

(19)



**Евразийское
патентное
ведомство**

(11) **019953**(13) **B1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента
2014.07.30

(51) Int. Cl. **C12Q 1/68** (2006.01)

(21) Номер заявки
200970891

(22) Дата подачи заявки
2008.03.26

**(54) ГЕНЕТИЧЕСКИЕ ВАРИАНТЫ CHR2 И CHR16 В КАЧЕСТВЕ МАРКЕРОВ
ДЛЯ ПРИМЕНЕНИЯ ПРИ ОЦЕНКЕ РИСКА, ДИАГНОСТИРОВАНИИ,
ПРОГНОЗИРОВАНИИ И ЛЕЧЕНИИ РАКА МОЛОЧНОЙ ЖЕЛЕЗЫ**

(31) 8625; 8648

(32) 2007.03.26; 2007.05.25

(33) IS

(43) 2010.04.30

(86) PCT/IS2008/000009

(87) WO 2008/117314 2008.10.02

(71)(73) Заявитель и патентовладелец:
ДЕКОУД ДЖЕНЕТИКС ЕХФ (IS)

(72) Изобретатель:
**Стейси Симон, Сулем Патрик,
Манолеску Андрей (IS)**

(74) Представитель:
**Липатова И.И., Новоселова С.В.,
Рыбаков В.М., Дощечкина В.В.,
Хмара М.В. (RU)**

(56) WO-A2-2006121991

Severine BEAUCOURT. Etude de 2 genes cibles de la proteine P53: siah (seven in absentia human homologue) et isap 6 (tumor suppressor activated pathway 6). THESE pour obtenir le grade de docteur de l'universite Paris XI, 25 Octobre 2005, с. 75, 79

WO-A1-2006/050573

Bohdan Gorski et al. A common missense variant in BRCA2 predisposes to early onset breast cancer. Breast Cancer Research 2005, 7: R1023-R1027 (Doi 10.1186/bcr1338), реферат

(57) Изобретение касается определенных генетических вариантов на Chr2q14, Chr2q35 и Chr16q12, являющихся маркерами подверженности раку молочной железы. Описываются способы оценки риска и диагностики повышенной и/или пониженной подверженности раку молочной железы с применением таких вариантов. Изобретение также относится к наборам для диагностики подверженности раку молочной железы.

B1

019953

019953

B1

Область изобретения

Рак молочной железы представляет собой наиболее распространенную форму рака, которым страдают женщины во всем мире. В настоящее время общая заболеваемость составляет более 1151000 новых случаев ежегодно (Parkin et al., 2005). Заболеваемость раком молочной железы является наиболее высокой в развитых странах, особенно среди населения, имеющего этническое происхождение из северной Европы, и неуклонно повышается. В Соединенных Штатах Америки ежегодная заболеваемость, стандартизованная по возрасту, составляет приблизительно 131 случай на 100000 населения, что более чем в три раза превышает среднее мировое значение. Показатели в странах северной Европы являются столь же высокими. По оценкам 2006 года в США ожидается ежегодное выявление 214650 новых случаев инвазивного рака молочной железы и гибель от заболевания 41430 человек [Jemal, et al., (2006), CA Cancer J Clin, 56, 106-30]. К этим цифрам необходимо добавить дополнительно 61980 диагностированных случаев протоковой и дольковой карциномы *in-situ*, ожидаемых в 2006 г. В индивидуальной перспективе вероятность развития рака молочной железы составляет 13,2% у женщин в США (т.е. у 1 из 8 женщин в течение жизни разовьется рак молочной железы). Как при большинстве форм рака, важными факторами являются ранняя диагностика и соответствующее лечение. В общем, вероятность прожить следующие 5 лет при раке молочной железы составляет 88%.

Однако, у лиц, имеющих местную инвазию или заболевание с метастазами, этот показатель снижается до 81 и 26% соответственно [Jemal, et al., (2006), CA Cancer J. Clin, 56, 106-30].

Все большее и большее значение придается выявлению индивидов, имеющих высокий риск первичного или рецидивирующего рака молочной железы.

С такими индивидами может проводиться более интенсивный скрининг, превентивная химиотерапия, гормональная терапия и, в случае чрезвычайно высокого риска, превентивное оперативное вмешательство. Программы массового скрининга ложатся огромным экономическим бременем на систему здравоохранения, тогда как превентивная терапия ассоциирована с риском и снижением качества жизни.

Генетическая предрасположенность к раку молочной железы

Две основные группы известных факторов риска рака молочной железы - это эндокринные и генетические факторы. Что касается последних, приблизительно 12% пациентов с раком молочной железы имеют одного или более 5 родственников первой степени с раком молочной железы [(2001), Lancet, 358, 1389-99]. Хорошо известные доминантные гены, предрасполагающие к раку молочной железы, BRCA1 и BRCA2 обуславливают значительно повышенный риск рака молочной железы у носителей, с оценками пенетрантности на протяжении всей жизни, варьирующими от 40 до 80%. Наличие мутаций BRCA1 и BRCA2 обнаруживается у большинства семей с 6 или более случаями рака молочной железы и в большой части семей, в которых встречается рак молочной железы и яичников или рак молочной железы у мужчин. Однако такие семьи в действительности встречаются очень редко. Мутации BRCA1 и BRCA2 обнаруживаются гораздо реже в семьях с меньшим количеством случаев рака или в семьях, где встречается только рак молочной железы. В общей сложности мутации BRCA1 и BRCA2 отвечают за 15-20% риска семейного рака молочной железы. Если было бы возможно детектировать все часто встречающиеся мутации BRCA в популяциях, не имеющих "эффекта основателя", от 2 до 3% пациентов с раком молочной железы имели бы мутацию [Gorski, et al., (2005), Breast Cancer Res Treat, 92, 19-24; (2000), Br J Cancer, 83, 1301-8]. Такая низкая вероятность обнаружения мутации препятствует грамотному проведению тестирования на наличие мутации BRCA в семьях, не имеющих явной генетической предрасположенности (Anon[(2003), J. Clin. Oncol., 21, 2397-406]). Известно, что редкие мутации с высокой пенетрантностью возникают в генах TP53 и PTEN, однако вместе они отвечают менее чем за 5% всего наследственно-обусловленного риска рака молочной железы [Easton, (1999), Breast Cancer Res, 1, 14-7]. В исследованиях по сцеплению не удалось выявить другие широко распространенные мутации, обуславливающие риск рака молочной железы [Smith, et al., (2006), Genes Chromosomes Cancer, 45, 646-55].

Проведенные недавно эпидемиологические исследования указали на то, что большинство случаев рака молочной железы возникает среди предрасположенного или подверженного к нему меньшинства населения [Antoniou, et al., (2002), Br J Cancer, 86, 76-83; Pharoah, et al., (2002), Nat Genet, 31, 33-6]. Данные исследований на близнецах и наблюдение постоянной высокой заболеваемости контралатеральным раком молочной железы пациентов, выживших после первичного рака молочной железы, указывают на то, что значительная часть неохарактеризованного риска рака молочной железы относится к эндогенным факторам, вероятнее всего, генетическим [Lichtenstein et al., (2000), N Engl J. Med, 343, 78-85; Peto и Mack, (2000), Nat Genet, 26, 411-4]. Сведения о генетических факторах, лежащих в основе этого широко распространенного риска, очень ограничены. Сегрегационный анализ позволяет предположить, что неохарактеризованный генетический риск рака молочной железы вероятнее всего имеет полигенную природу, а связанные с риском аллели обуславливают степень риска от низкой до средней и могут взаимодействовать между собой и гормональными факторами риска. Тем не менее, данные исследования прогнозируют почти 40-кратное различие в относительных рисках между верхним и нижним квинтилем в распределении, которое можно выявить при генетическом профилировании, охватывающем подобные аллели низкого и умеренного риска [Antoniou et al., (2002), Br J Cancer, 86, 76-83; Pharoah, et al., (2002), Nat Genet, 31, 33-6]. Предполагается, что 88% всех случаев рака молочной железы возникают среди 50%

населения, у которых имеется предрасположенность, а на долю 12% населения, имеющего наиболее высокий риск, приходится 50% всех случаев рака молочной железы (Pharoah et al., (2002), *Nat Genet*, 31, 33-6; Pharoah, (2003), *Recent Results Cancer Res*, 163, 7-18; discussion 264-6). Особое внимание, таким образом, сосредоточено на выявлении таких индивидов, имеющих генетическую предрасположенность, и разработку для них стратегий индивидуального лечения.

Нами и другими исследователями было показано, что существует значительный наследственный риск рака молочной железы в Исландии, который распространяется, по меньшей мере, на родственников 5 степени [Amundadottir, et al., (2004), *PLoS Med*, 1, e65; Tulinius, et al., (2002), *J Med Genet*, 39, 457-62]. Предполагалось, что вклад мутаций BRCA1 в наследственный риск в Исландии минимален [Arason, et al., (1998), *J Med Genet*, 35, 446-9; Bergthorsson et al., (1998), *Hum Mutat*, Suppl 1, S195-7]. Единственная мутация, связанная с эффектом основателя в гене BRCA2 (999del5), присутствует в общей исландской популяции с частотой носительства 0,6-0,8% и среди женщин с раком молочной железы с частотой 7,7-8,6% (Thorlacius, et al., (1997), *Am J Hum Genet*, 60, 1079-84; Gudmundsson, et al., (1996), *Am J Hum Genet*, 58, 749-56). Считается, что эта единственная мутация отвечает приблизительно за 40% наследуемого риска рака молочной железы у родственников первой-третьей степени (Tulinius et al., (2002), *J Med Genet*, 39, 457-62). Несмотря на то, что данная оценка выше, чем уровень наследственного риска, приписываемого всем мутациям BRCA 1 и 2 в сочетании в популяциях, не имеющих "эффекта основателя", остается еще 60% наследственного риска рака молочной железы в Исландии, которое требуется объяснить. Родственники первой степени пациентов, у которых результаты тестов на наличие BRCA2 999del5 отрицательны, имеют риск рака молочной железы в 1,72 раза выше популяционного риска (95% CI 1.49-1.96) (Tulinius, et al., (2002), *J Med Genet*, 39, 457-62).

Сведения о том, какие генетические факторы отвечают за резидуальный генетический риск рака молочной железы, очень ограничены. Были получены точные подтверждения, что варианты двух генов, CHEK2 и ATM, являются генами риска рака молочной железы с низкой пенетрантностью (Renwick et al., (2006), *Nat Genet*, 38, 873-5; (2004), *Am J Hum Genet*, 74, 1175-82). Изучалась роль многих других генов, однако их вклад в риск рака молочной железы не был подтвержден в анализе с использованием очень больших выборок [Breast Cancer Association, (2006), *J Natl Cancer Inst*, 98, 1382-96].

Генетический риск обуславливается небольшими различиями в генах индивидов в составе популяции. Чаще всего различия в генах между индивидами обусловлены однонуклеотидными полиморфизмами (ОНП), хотя и другие вариации также имеют важное значение. В человеческом геноме ОНП располагаются в среднем в каждой 1000 пар оснований. Соответственно, обычный человеческий ген, содержащий 250000 пар оснований, может содержать 250 различных ОНП. Только небольшая часть ОНП располагается в экзонах и приводит к изменению аминокислотной последовательности белка, кодируемого геном. Большинство ОНП оказывают небольшое влияние или не оказывают влияния на функцию генов, в то время как некоторые могут изменять транскрипцию, сплайсинг, трансляцию или стабильность кодируемой геном мРНК. Кроме того, генетический полиморфизм в человеческом геноме бывает связан с инсерциями, делециями, транслокациями или инверсиями либо коротких, либо длинных участков ДНК. Генетические полиморфизмы, обуславливающие риск заболеваний, таким образом, могут напрямую изменять аминокислотную последовательность белков, могут приводить к повышению количества белка, производимого геном, или могут вызывать уменьшение количества белка, производимого геном.

Поскольку генетические полиморфизмы, обуславливающие риск заболеваний, раскрыты, для клинической медицины становится важным генетическое тестирование на наличие подобных факторов риска. Примерами являются тестирование аполипопротеина Е для идентификации носительства полиморфизма гена апоЕ4 среди пациентов с деменцией для дифференциальной диагностики болезни Альцгеймера и тестирование на фактор V Лейден для выяснения предрасположенности к тромбозу глубоких вен. Более того, при лечении рака определение генетических вариантов опухолевых клеток используется для выбора наиболее подходящего режима лечения для конкретных пациентов. При раке молочной железы генетическая вариация в экспрессии эстрогеновых рецепторов или экспрессии тирозин киназы рецептора херегулина 2 типа определяет, будут ли включены в план лечения анти-эстрогеновые препараты (тамоксифен) или антитела anti-Her2 (герцептин). При хроническом миелоидном лейкозе (ХМЛ) обнаружение Филадельфийской хромосомы, появляющейся вследствие генетической транслокации, приводящей к объединению генов, кодирующих тирозиновые киназы Bcr и Abl указывает на то, что в лечении рака должен использоваться препарат Гливек (Gleevec, STI571), специфический ингибитор киназ Bcr-Abl. У пациентов с ХМЛ, имеющих подобное генетическое нарушение, ингибирование киназ Bcr-Abl приводит к быстрой элиминации опухолевых клеток и ремиссии лейкоза.

В настоящее время не существует универсального способа, позволяющего проводить успешную профилактику или лечение рака молочной железы. Ведение больных с раком молочной железы в настоящее время основано на сочетании первичной профилактики, ранней диагностики, соответствующего лечения и вторичной профилактики. Существуют четкие клинические показания для включения генетического тестирования по всем перечисленным направлениям. Выявление генов, предрасполагающих к развитию рака, может также обнаружить ключевые молекулярные механизмы, которыми можно управлять (например, с применением препаратов с низким или высоким молекулярным весом) и может обес-

печить более эффективное лечение.

Сущность изобретения

Данное изобретение касается способов оценки риска развития рака молочной железы. Они включают способы определения повышенной подверженности индивида к развитию рака молочной железы, а также способы определения пониженной подверженности к развитию рака молочной железы или диагностирования устойчивости индивида к раку молочной железы путем определения конкретных маркеров или гаплотипов, для которых обнаружена ассоциация с раком молочной железы, по данному описанию.

В первом аспекте изобретение касается способа определения подверженности к развитию рака молочной железы у человеческого индивида, включающего определение наличия или отсутствия по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера в образце нуклеиновой кислоты, взятом у индивида или в массиве данных генотипа, полученном у индивида, где по меньшей мере один полиморфный маркер выбран из полиморфных маркеров, указанных в любой из табл. 10, 15 и 19, и маркеров, находящихся в неравновесии по сцеплению с ними, и где наличие по меньшей мере одного аллеля указывает на подверженность индивида к развитию рака молочной железы. В одном воплощении способ может касаться определения наличия или отсутствия по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера в образце нуклеиновой кислоты, взятом у индивида. В другом воплощении способ касается определения наличия или отсутствия по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера в массиве данных генотипа, полученном у индивида. Массив данных генотипа может, в определенных воплощениях, быть получен у конкретного индивида на основе информации, содержащейся в массиве данных генотипа, относящегося к матрице определенной нуклеиновой кислоты образца, где образец нуклеиновой кислоты содержит нуклеиновую кислоту одного индивида.

Определение наличия конкретного аллеля подверженности полиморфного маркера напрямую указывает на существование определенной подверженности, обуславливаемой этим аллелем. С другой стороны, определение отсутствия такого аллеля подверженности указывает на отсутствие определенной подверженности у индивида, у которого получен образец генетического материала или массив данных генотипа. В конкретных случаях, когда у полиморфного маркера, например, ОНП или полиморфизма вызванного инсерцией/делецией, существует два возможных аллеля, выявление отсутствия конкретного аллеля означает наличие у индивида двух копий варианта аллеля (за исключением случаев, когда определенный регион генома индивида содержит делецию или дупликацию, когда может существовать либо единственная копия или более двух копий определенного региона генома индивида).

Во втором аспекте данное изобретение касается способа диагностирования подверженности человеческого индивида к развитию рака молочной железы, включающего определение наличия или отсутствия по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера в образце нуклеиновой кислоты, взятом у индивида, где по меньшей мере один полиморфный маркер выбран из группы маркеров, ассоциированных с блоком неравновесия по сцеплению rs4848543, блоком неравновесия по сцеплению rs3803662 и блоком неравновесия по сцеплению rs13387042, где наличие по меньшей мере одного аллеля указывает на подверженность к развитию рака молочной железы. В одном воплощении маркеры, ассоциированные с блоком неравновесия по сцеплению rs4848543, блоком неравновесия по сцеплению rs3803662 и/или блоком неравновесия по сцеплению rs13387042 LD являются маркерами, находящимися в неравновесии по сцеплению по меньшей мере с одним маркером в составе одного или более из этих блоков.

В следующем аспекте изобретение касается способа диагностирования подверженности человеческого индивида к развитию рака молочной железы, включающего определение наличия или отсутствия по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера в образце нуклеиновой кислоты, взятом у индивида, где по меньшей мере один полиморфный маркер выбран из группы маркеров, состоящей из маркеров, приведенных в табл. 10, табл. 15 и табл. 19, и маркеров в неравновесии по сцеплению с ними. В одном конкретном воплощении маркеры, находящиеся в неравновесии по сцеплению с маркерами, указанными в табл. 10, табл. 15 и табл. 19, являются маркерами со значениями r^2 выше 0,2. В другом воплощении маркеры, находящиеся в неравновесии по сцеплению с маркерами, указанными в табл. 10, табл. 15 и табл. 19, являются маркерами со значениями r^2 выше 0,2 в европеоидной популяции, такой как популяции НарМар СЕРН.

В другом аспекте изобретение относится к способу выявления маркера для использования в оценке подверженности к развитию рака молочной железы, способ включает:

а) выявление по меньшей мере одного полиморфного маркера, находящегося в неравновесии по сцеплению по меньшей мере с одним из маркеров в составе геномного сегмента, последовательность которого приведена в SEQ ID NO:4, SEQ ID NO:5 и SEQ ID NO:6;

б) определение генотипического статуса в выборке индивидов с диагностированным раком молочной железы или подверженными к нему; и

в) определение генотипического статуса в выборке контрольных индивидов, где достоверная разница в частоте по меньшей мере одного аллеля по меньшей мере одного полиморфизма у индивидов с диагностированным раком молочной железы или подверженных к нему, по сравнению с частотой по

меньшей мере одного аллеля в контрольной выборке указывает на то, что по меньшей мере один полиморфизм может применяться для оценки подверженности к развитию рака молочной железы. В другом воплощении пониженная частота по меньшей мере одного аллеля по меньшей мере одного полиморфизма у индивидов с диагностированным раком молочной железы или подверженных к нему, по сравнению с частотой по меньшей мере одного аллеля в контрольной выборке указывает на то, что полиморфизм может применяться для оценки пониженной подверженности, или устойчивости к раку молочной железы. В одном воплощении по меньшей мере один полиморфный маркер, находящийся в неравновесии по сцеплению по меньшей мере с одним полиморфным маркером в составе геномных сегментов, имеющих последовательность, приведенную в SEQ ID NO:4, SEQ ID NO:5 и SEQ ID NO:6, выбран из rs4848543 (SEQ ID NO:1), rs3803662 (SEQ ID NO:3) и rs13387042 (SEQ ID NO:2).

Изобретение также касается способа генотипирования образца нуклеиновой кислоты, взятого у человеческого индивида, имеющего риск развития или с диагностированным раком молочной железы, включающего определение наличия или отсутствия по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера в образце, где по меньшей мере один маркер выбран из группы, состоящей из маркеров, приведенных в табл. 10, табл. 15 и табл. 19, и маркеров, находящихся в неравновесии по сцеплению с ними, и где наличие или отсутствие по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера указывает на подверженность к раку молочной железы. В одном воплощении по меньшей мере один маркер выбран из rs4848543 (SEQ ID NO:1), rs3803662 (SEQ ID NO:3) и rs13387042 (SEQ ID NO:2). В другом воплощении генотипирование включает амплификацию сегмента нуклеиновой кислоты, включающего по меньшей мере один полиморфный маркер, в ходе полимеразной цепной реакции (ПЦР) с применением пары нуклеотидных праймеров, фланкирующих по меньшей мере один полиморфный маркер. В другом воплощении генотипирование проводится посредством процесса, выбранного из аллель-специфической гибридизации зондов, аллель-специфического удлинения праймеров, аллель-специфической амплификации, секвенирования нуклеиновых кислот, расщепления 5'-экзонуклеазами, использования "молекулярных маячков" (молекулярных беконов), лигирования олигонуклеотидов, анализа размера фрагментов и анализа одностречевой конформации. В одном предпочтительном воплощении процесс включает аллель-специфическую гибридизацию зондов. В другом предпочтительном воплощении процесс включает секвенирование ДНК. В особо предпочтительном воплощении способ включает этапы:

1) инкубации копий нуклеиновой кислоты с детектирующим олигонуклеотидным зондом и усиливающим олигонуклеотидным зондом в условиях для специфической гибридизации олигонуклеотидного зонда с нуклеиновой кислотой; где

а) детектирующий олигонуклеотидный зонд имеет длину 5-100 нуклеотидов и специфически гибридизуется с первым сегментом нуклеиновой кислоты, нуклеотидная последовательность которой представлена в SEQ ID NO:4, SEQ ID NO:5 or SEQ ID NO:6;

б) детектирующий олигонуклеотидный зонд содержит детектируемую метку на своем 3' конце и гасящий компонент на своем 5' конце;

в) усиливающий олигонуклеотид имеет длину 5-100 нуклеотидов и является комплементарным второму сегменту нуклеотидной последовательности, расположенному в положении 5' по отношению к олигонуклеотидному зонду, так что усиливающий олигонуклеотид располагается в положении 3' по отношению к детектирующему олигонуклеотидному зонду, когда оба олигонуклеотида гибридизованы с нуклеиновой кислотой; и

г) между первым сегментом и вторым сегментом существует пробел в одно основание, так что когда олигонуклеотидный зонд и усиливающий олигонуклеотидный зонд оба гибридизованы с нуклеиновой кислотой, между олигонуклеотидами существует пробел в одно основание;

2) обработки нуклеиновой кислоты эндонуклеазой, которая будет отщеплять детектируемую метку от 3' конца детектирующего зонда с высвобождением свободной детектируемой метки при гибридизации детектирующего зонда с нуклеиновой кислотой; и

3) измерения свободной детектируемой метки, когда присутствие свободной детектируемой метки указывает на то, что детектирующий зонд специфически гибридизован с первым сегментом нуклеиновой кислоты и указывает на то, что последовательность полиморфного сайта комплементарна детектирующему зонду

Следующий аспект изобретения касается способа оценки у индивида возможности ответа на терапевтический агент для лечения рака молочной железы, включающему: определение наличия или отсутствия по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера в образце нуклеиновой кислоты, взятом у индивида, где по меньшей мере один полиморфный маркер выбран из группы, состоящей из полиморфных маркеров, приведенных в табл. 10, табл. 15 и табл. 19, и маркеров в неравновесии по сцеплению с ними, где определение наличия по меньшей мере одного аллеля по меньшей мере одного маркера указывает на вероятность благоприятного ответа на терапевтический агент. В одном воплощении терапевтический агент представляет собой химиотерапевтический агент или агент для гормональной терапии. В одном воплощении агент для гормональной терапии представляет собой селективный модулятор эстрогеновых рецепторов или ингибитор ароматазы. В предпочтительных воплоще-

ниях селективный модулятор эстрогеновых рецепторов выбран из тамоксифена и ралоксифена. В определенных воплощениях ингибитор ароматазы выбран из экземестана, анастрозола и летроззола. В других воплощениях селективный модулятор эстрогеновых рецепторов вводят в комбинации с ингибитором ароматазы.

Следующий аспект изобретения касается способа определения прогноза у индивида с диагностированным раком молочной железы, способ включает определение наличия или отсутствия по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера в образце нуклеиновой кислоты, взятом у индивида, где по меньшей мере один полиморфный маркер выбран из группы, состоящей из полиморфных маркеров, приведенных в табл. 10, 15 и 19, и маркеров, находящихся в неравновесии по сцеплению с ними, где наличие по меньшей мере одного аллеля указывает на неблагоприятный прогноз рака молочной железы у индивида. Прогноз у индивида может в принципе касаться любого характера прогрессирования заболевания, включая формирование вторичной опухоли, быстрого распространения опухоли, степени злокачественности опухоли (от стадии 0 до стадии IV) и рецидивирования опухоли, но не ограничиваться ими.

Изобретение также касается способа мониторинга за ходом лечения у индивида, которому проводится лечение рака молочной железы, способ включает определение наличия или отсутствия по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера в образце нуклеиновой кислоты, взятом у индивида, где по меньшей мере один полиморфный маркер выбран из маркеров, приведенных в табл. 10, табл. 15 и табл. 19, и маркеров в неравновесии по сцеплению с ними, где наличие по меньшей мере одного аллеля указывает на исход лечения у индивида. Лечение может представлять собой хирургическое лечение, химиотерапевтическое лечение, радиотерапию, генную терапию или иммунотерапию. Любой из этих способов лечения может применяться самостоятельно или в комбинации. Например, после хирургического лечения может следовать химиотерапевтическое лечение и/или радиотерапевтическое лечение. Исход лечения индивида касается прогрессирования опухоли у индивида в конце лечения. Показатели такого исхода включают степень или возможность рецидивирования опухоли, распространение опухоли и формирование вторичной опухоли, но не ограничиваются ими.

Следующий аспект изобретения относится к набору для оценки подверженности к раку молочной железы у человеческого индивида, набор включает реагенты для селективной детекции по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера в геноме индивида, где полиморфный маркер выбран из маркеров, приведенных в табл. 10, табл. 15 и табл. 19, и маркеров в неравновесии по сцеплению с ними, и где наличие по меньшей мере одного аллеля указывает на подверженность к раку молочной железы. В одном воплощении по меньшей мере один полиморфный маркер представляет собой rs4848543 (SEQ ID NO:1), rs3803662 (SEQ ID NO:3) или rs13387042 (SEQ ID NO:2). В другом воплощении реагенты включают по меньшей мере один примыкающий нуклеотид, гибридизующийся с фрагментом генома индивида, содержащим по меньшей мере один полиморфный маркер, буфер и детектируемую метку. В другом воплощении реагенты включают по меньшей мере одну пару олигонуклеотидов, гибридизующихся с противоположными нитями сегмента геномной нуклеиновой кислоты, полученной у субъекта, где каждая пара олигонуклеотидных праймеров создана для селективной амплификации фрагмента генома индивида, включающего один полиморфный маркер, и где фрагмент имеет размер по меньшей мере 30 пар оснований. В следующем воплощении по меньшей мере один олигонуклеотид является полностью комплементарным геному индивида. В другом воплощении олигонуклеотид имеет размер от 18 до 50 нуклеотидов. В другом воплощении олигонуклеотид имеет размер от 20 до 30 нуклеотидов. В предпочтительном воплощении набор содержит:

- а) детектирующий олигонуклеотидный зонд, имеющий длину 5-100 нуклеотидов;
- б) усиливающий олигонуклеотидный зонд, имеющий длину 5-100 нуклеотидов; и
- в) фермент эндонуклеазу;

где детектирующий олигонуклеотидный зонд специфически гибридизуется с первым сегментом нуклеиновой кислоты, нуклеотидная последовательность которой приведена в SEQ ID NO:4, SEQ ID NO:5 или SEQ ID NO:6, и

где детектирующий олигонуклеотидный зонд содержит детектируемую метку на своем 3' конце и гасящий компонент на своем 5' конце;

где усиливающий олигонуклеотид имеет длину 5-100 нуклеотидов и является комплементарным второму сегменту нуклеотидной последовательности,

расположенному в положении 5' по отношению к олигонуклеотидному зонду, так что усиливающий олигонуклеотид располагается в положении 3' по отношению к детектирующему олигонуклеотидному зонду, когда оба олигонуклеотида гибридизованы с нуклеиновой кислотой;

где между первым сегментом и вторым сегментом существует пробел в одно основание, так что когда олигонуклеотидный зонд и усиливающий олигонуклеотидный зонд оба гибридизованы с нуклеиновой кислотой, между олигонуклеотидами существует пробел в одно основание; и

где при обработке нуклеиновой кислоты эндонуклеазой от 3' конца детектирующего зонда будет отщепляться детектируемая метка, так что высвобождение свободной детектируемой метки будет иметь место при гибридизации детектирующего зонда с нуклеиновой кислотой.

Следующий аспект изобретения касается устройств для определения у человеческого индивида генетического индикатора рака молочной железы, включающих

компьютерную память и

программы, хранящиеся в компьютерной памяти;

где программы адаптированы для выполнения на процессорах для анализа информации о маркерах и/или гаплотипах по меньшей мере для одного человеческого индивида в отношении по меньшей мере одного полиморфного маркера, выбранного из маркеров, приведенных в табл. 10, табл. 15 и табл. 19, и маркеров в неравновесии по сцеплению с ними, и для генерирования выходных данных на основании информации о маркере или гаплотипе, где выходные данные включают показатель риска по меньшей мере по одному маркеру или гаплотипу в качестве генетического индикатора рака молочной железы у человеческого индивида. В одном воплощении программа также включает показатель риска развития рака молочной железы, ассоциированный с по меньшей мере одним аллелем маркера и/или гаплотипом, где показатель риска основан на сравнении частоты по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера и/или гаплотипа у большинства индивидов с диагностированным раком молочной железы и индикатора частоты по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера и/или гаплотипа у большинства референсных индивидов, и где индивидуальный риск для человеческого индивида основан на сравнении статуса носительства индивидом по меньшей мере одного аллеля маркера и/или гаплотипа и показателя риска по меньшей мере для одного аллеля маркера и/или гаплотипа. В одном воплощении по меньшей мере один полиморфный маркер выбран из rs4848543 (SEQ ID NO:1), rs3803662 (SEQ ID NO:3) или rs13387042 (SEQ ID NO:2) и маркеров, находящихся в неравновесии по сцеплению с ними.

Изобретение также касается применения олигонуклеотидного зонда в производстве реагента для диагностирования и/или оценки подверженности к развитию рака молочной железы у человеческого индивида, где зонд гибридизуется с сегментом нуклеиновой кислоты, нуклеотидная последовательность которого приведена в SEQ ID NO:4, SEQ ID NO:5 или SEQ ID NO:6, где зонд имеет длину 15-500 нуклеотидов.

Изобретение также касается компьютерных носителей, на которых хранится:

а) идентификатор по меньшей мере для одного полиморфного маркера;

б) индикатор частоты по меньшей мере одного аллеля по меньшей мере одного указанного полиморфного маркера у большинства индивидов с диагностированным раком молочной железы и

в) индикатор частоты по меньшей мере одного аллеля по меньшей мере одного указанного полиморфного маркера у большинства референсных индивидов;

где по меньшей мере один полиморфный маркер выбран из полиморфных маркеров, приведенных в табл. 10, табл. 15 и табл. 19, и маркеров в неравновесии по сцеплению с ними.

В определенных воплощениях изобретения также оценивается частота по меньшей мере одного гаплотипа у индивида, где гаплотип включает по меньшей мере два маркера, и где наличие по меньшей мере одного гаплотипа указывает на подверженность к развитию рака молочной железы. В одном воплощении гаплотип представляет собой гаплотип, связанный с риском рака молочной железы, т.е. гаплотип обуславливает повышенный риск развития рака молочной железы. В одном воплощении гаплотип выбран из группы гаплотипов, указанных в табл. 7, табл. 8, табл. 9, табл. 13, табл. 14 и табл. 18.

В другом воплощении изобретения по меньшей мере один аллель по меньшей мере одного полиморфного маркера присутствует в составе гаплотипа, выбранного из группы состава гаплотипов, определенных гаплотипами, указанными в табл. 9 и 13. В другом воплощении способ по данному изобретению касается также выполнения этапа оценки высоко пенетрантных генетических факторов, связанных с раком молочной железы. Таким высоко пенетрантным генетическим фактором в одном воплощении является BRCA2 999del5. Недавно полученные свидетельства позволяют предположить, что риск, ассоциированный с вариантами BRCA1 и BRCA2, в некоторых случаях модифицирован другими генетическими факторами или факторами окружающей среды, которые концентрируются в семьях (Antoniou A.C., et al. Am J Hum Genet 2008 Mar 18 (Epub ahead of print)). Так, в определенных воплощениях, другие мутации с высокой пенетрантностью, связанные с раком молочной железы, такие как, обнаруженные в генах BRCA1 и BRCA2, которые хорошо документированы и известны специалистам (см, например, Breast Cancer Mutation Data Base at: <http://research.nhgri.nih.gov/bic/>; см. также Fackenthal, J.D. & Olopade, O.I., Nature Reviews Cancer 7:937-48 (2007) и ссылки, приводимые здесь) могут также оцениваться и использоваться в сочетании с вариантами по данному описанию ассоциированными с раком молочной железы.

В одном воплощении изобретения также выполняется этап оценки статуса эстрогеновых рецепторов или прогестероновых рецепторов индивида. В одном таком воплощении положительный эстрогеновый статус или положительный прогестероновый статус ассоциирован с повышенным риском, ассоциированным с аллелем A rs13387042 и аллелем T rs3803662 и ассоциированными с ними маркерами, например, маркерами, приведенными в табл. 15 и 19.

В другом аспекте данное изобретение касается способа оценки риска развития по меньшей мере второй первичной опухоли у индивида с диагностированным ранее раком молочной железы, способ включает определение наличия или отсутствия по меньшей мере одного аллеля по меньшей мере одного

полиморфного маркера в образце нуклеиновой кислоты, взятом у индивида, где по меньшей мере один полиморфный маркер выбран из группы маркеров, ассоциированных с блоком неравновесия по сцеплению rs4848543, блоком неравновесия по сцеплению rs13387042 и блоком неравновесия по сцеплению rs3803662, где наличие по меньшей мере одного аллеля указывает на риск развития по меньшей мере второй первичной опухоли. В одном воплощении по меньшей мере один полиморфный маркер выбран из маркеров, приведенных в любой из табл. 10, 15 и 19. В другом воплощении по меньшей мере один маркер ассоциирован с блоком неравновесия по сцеплению rs4848543. В одном воплощении по меньшей мере один полиморфный маркер выбран из группы, состоящей из полиморфных маркеров, указанных в табл. 10, 15, 19, 20, 21 и 22, и маркеров в неравновесии по сцеплению с ними. В другом воплощении по меньшей мере один полиморфный маркер выбран из группы маркеров, находящихся в сильном неравновесии по сцеплению, характеризуемом значениями r^2 выше 0,2, с маркером rs4848543. В следующем воплощении по меньшей мере один полиморфный маркер выбран из группы маркеров, указанных в табл. 7, 8 и 10. В следующем воплощении по меньшей мере один полиморфный маркер выбран из группы маркеров, указанных в табл. 10. В следующем воплощении по меньшей мере один полиморфный маркер ассоциирован с геном STEAP3/TSAP6. В предпочтительном воплощении маркером является rs4848543. В других воплощениях выполняются дополнительные этапы оценки частоты по меньшей мере одного гаплотипа индивида. Этот по меньшей мере один гаплотип в одном воплощении выбран из группы гаплотипов, указанных в табл. 7, 8 и 9. В других воплощениях также проводится оценка факторов риска рака молочной железы, имеющих высокую пенетрантность, таких как BRCA2 999del5.

Фенотип рака молочной железы, по изобретению, согласно данному описанию, в определенных воплощениях может быть выбран из всех форм рака молочной железы всех форм, первично-множественного рака молочной железы, рака молочной железы с ранним началом заболевания или другого принятого в медицине определения рака молочной железы. В другом воплощении фенотипом, ассоциированным с раком молочной железы, является суммарный семейный анамнез (от англ. summed family history).

В конкретных воплощениях способов, применений, наборов или устройств по изобретению, по данному описанию, человеческим индивидом, которому проводится оценка, является женщина.

Все маркеры по данному описанию, подходящие для оценки предрасположенности к раку молочной железы, могут применяться в различных способах, наборах, устройствах и применениях по данному описанию. Таким образом, в определенных воплощениях по меньшей мере один полиморфный маркер, подходящий для осуществления изобретения, может быть выбран из группы, состоящей из полиморфных маркеров, указанных в табл. 20, 21 и 22, и маркеров в неравновесии по сцеплению с ними. В другом воплощении по меньшей мере один полиморфный маркер выбран из группы маркеров, указанных в табл. 10, 15 и 19, и маркеров в неравновесии по сцеплению с ними. В определенных воплощениях по меньшей мере один полиморфный маркер расположен в составе геномного сегмента с последовательностью, приведенной в любой из SEQ ID NO:4, SEQ ID NO:5 и SEQ ID NO:6. В одном предпочтительном воплощении по меньшей мере один маркер в составе SEQ ID NO:4 выбран из маркеров, приведенных в табл. 20. В другом предпочтительном воплощении по меньшей мере один маркер в составе SEQ ID NO:5 выбран из маркеров, приведенных в табл. 21. В следующем предпочтительном воплощении по меньшей мере один маркер в составе SEQ ID NO:6 выбран из маркеров, приведенных в табл. 22. В следующем воплощении по меньшей мере один полиморфный маркер выбран из группы маркеров, указанных в табл. 7, 8, 10, 14, 15, 18 и 19. В другом воплощении по меньшей мере один полиморфный маркер выбран из группы маркеров, находящихся в сильном неравновесии по сцеплению, характеризуемом значениями r^2 выше 0,2, с маркером rs4848543 (SEQ ID NO:1), rs3803662 (SEQ ID NO:3) или rs13387042 (SEQ ID NO:2). В предпочтительном воплощении по меньшей мере один полиморфный маркер выбран из rs4848543 (SEQ ID NO:1), rs3803662 (SEQ ID NO:3) и rs13387042 (SEQ ID NO:2). В другом предпочтительном воплощении по меньшей мере один полиморфный маркер представляет собой rs4848543 (SEQ ID NO:1). В другом предпочтительном воплощении по меньшей мере один полиморфный маркер представляет собой rs3803662 (SEQ ID NO:3). В следующем предпочтительном воплощении по меньшей мере один полиморфный маркер представляет собой rs13387042 (SEQ ID NO:2). В следующем воплощении по меньшей мере один полиморфный маркер ассоциирован (т.е. находится в неравновесии по сцеплению) с геном STEAP3/TSAP6, геном LOC643714 и/или геном TNRC9. В одном таком воплощении по меньшей мере один полиморфный маркер расположен в составе гена STEAP3/TSAP6, гена LOC643714 и/или гена TNRC9. В другом предпочтительном воплощении по меньшей мере один полиморфный маркер выбран из rs4848543, rs13387042, rs3803662, rs12922061, rs4784227 и rs17271951.

В определенных способах, применениях, наборах или устройствах по изобретению наличие по меньшей мере одного аллеля или гаплотипа указывает на повышенную предрасположенность к раку молочной железы. В других способах по изобретению наличие по меньшей мере одного гаплотипа указывает на пониженную предрасположенность к раку молочной железы. Повышенная предрасположенность к раку молочной железы в одном воплощении обусловлена составом гаплотипа, выбранного из группы состава гаплотипов, состоящей из:

- (i) rs8955398 аллель 2, rs4848543 аллель 1, rs6759589 аллель 1, rs838066 аллель 2, rs838100 аллель 3,

rs838086 аллель 1, rs12711924 аллель 1 и rs3731603 аллель 4;

(ii) rs8955398 аллель 4, rs4848543 аллель 1, rs6759589 аллель 1, rs838066 аллель 2, rs838100 аллель 3, rs838086 аллель 1, rs12711924 аллель 1 и rs3731603 аллель 4;

(iii) rs8955398 аллель 4, rs4848543 аллель 1, rs6759589 аллель 1, rs838066 аллель 2, rs838100 аллель 3, rs838086 аллель 2, rs12711924 аллель 1 и rs3731603 аллель 4;

(iv) rs8955398 аллель 2, rs4848543 аллель 1, rs6759589 аллель 1, rs838066 аллель 2, rs838100 аллель 3, rs838086 аллель 2, rs12711924 аллель 1 и rs3731603 аллель 3;

(v) rs8955398 аллель 4, rs4848543 аллель 1, rs6759589 аллель 1, rs838066 аллель 2, rs838100 аллель 3, rs838086 аллель 1, rs12711924 аллель 1 и rs3731603 аллель 3;

(vi) rs8955398 аллель 2, rs4848543 аллель 1, rs6759589 аллель 1, rs838066 аллель 2, rs838100 аллель 3, rs838086 аллель 1, rs12711924 аллель 1 и rs3731603 аллель 4.

В другом воплощении повышенная предрасположенность к раку молочной железы обусловлена составом гаплотипа, выбранного из группы состава гаплотипов, состоящей из:

(i) rs10191184 аллель 3, rs6435957 аллель 4, rs10171745 аллель 3, rs6716542 аллель 1, rs4491709 аллель 4, rs12621130 аллель 1, rs6735174 аллель 1, rs6435959 аллель 2, rs2372943 аллель 3, rs13387042 аллель 1, rs10490444 аллель 1, rs13011060 аллель 3;

(ii) rs10191184 аллель 3, rs6435957 аллель 4, rs10171745 аллель 3, rs6716542 аллель 3, rs4491709 аллель 4, rs12621130 аллель 1, rs6735174 аллель 2, rs6435959 аллель 4, rs2372943 аллель 3, rs13387042 аллель 1, rs10490444 аллель 3, rs13011060 аллель 3;

(iii) rs10191184 аллель 3, rs6435957 аллель 2, rs10171745 аллель 3, rs6716542 аллель 3, rs4491709 аллель 4, rs12621130 аллель 1, rs6735174 аллель 2, rs6435959 аллель 4, rs2372943 аллель 3, rs13387042 аллель 1, rs10490444 аллель 3, rs13011060 аллель 3;

(iv) rs10191184 аллель 3, rs6435957 аллель 4, rs10171745 аллель 3, rs6716542 аллель 3, rs4491709 аллель 4, rs12621130 аллель 1, rs6735174 аллель 1, rs6435959 аллель 2, rs2372943 аллель 3, rs13387042 аллель 1, rs10490444 аллель 3, rs13011060 аллель 3;

(v) rs10191184 аллель 3, rs6435957 аллель 4, rs10171745 аллель 1, rs6716542 аллель 3, rs4491709 аллель 2, rs12621130 аллель 1, rs6735174 аллель 1, rs6435959 аллель 2, rs2372943 аллель 3, rs13387042 аллель 1, rs10490444 аллель 1, rs13011060 аллель 3;

(vi) rs10191184 аллель 3, rs6435957 аллель 4, rs10171745 аллель 3, rs6716542 аллель 3, rs4491709 аллель 4, rs12621130 аллель 1, rs6735174 аллель 2, rs6435959 аллель 4, rs2372943 аллель 3, rs13387042 аллель 1, rs10490444 аллель 3, rs13011060 аллель 1;

(vii) rs10191184 аллель 3, rs6435957 аллель 4, rs10171745 аллель 3, rs6716542 аллель 1, rs4491709 аллель 4, rs12621130 аллель 1, rs6735174 аллель 1, rs6435959 аллель 2, rs2372943 аллель 3, rs13387042 аллель 1, rs10490444 аллель 1, rs13011060 аллель 1;

(viii) rs10191184 аллель 3, rs6435957 аллель 2, rs10171745 аллель 3, rs6716542 аллель 1, rs4491709 аллель 4, rs12621130 аллель 1, rs6735174 аллель 1, rs6435959 аллель 2, rs2372943 аллель 3, rs13387042 аллель 1, rs10490444 аллель 1, rs13011060 аллель 3;

(ix) rs10191184 аллель 3, rs6435957 аллель 4, rs10171745 аллель 3, rs6716542 аллель 1, rs4491709 аллель 4, rs12621130 аллель 1, rs6735174 аллель 1, rs6435959 аллель 2, rs2372943 аллель 3, rs13387042 аллель 1, rs10490444 аллель 3, rs13011060 аллель 1;

(x) rs10191184 аллель 3, rs6435957 аллель 2, rs10171745 аллель 3, rs6716542 аллель 3, rs4491709 аллель 4, rs12621130 аллель 1, rs6735174 аллель 1, rs6435959 аллель 2, rs2372943 аллель 3, rs13387042 аллель 1, rs10490444 аллель 1, rs13011060 аллель 3.

В предпочтительном воплощении по меньшей мере один аллель или гаплотип, обуславливающий повышенный риск рака молочной железы, представляет собой аллель 1 rs4848543. В следующих воплощениях повышенный риск характеризуется значениями относительного риска или отношения шансов по меньшей мере 1,2, включая риск по меньшей мере 1,25, риск по меньшей мере 1,3, риск по меньшей мере 1,4, риск по меньшей мере 1,5, риск по меньшей мере 1,6, риск по меньшей мере 1,7 и риск по меньшей мере 2,0.

В других определенных воплощениях изобретения наличие по меньшей мере одного аллеля или гаплотипа указывает на пониженную предрасположенность (пониженный риск) к раку молочной железы. По меньшей мере один аллель или гаплотип может, например, быть выбран из группы аллелей маркеров и гаплотипов, указанных в табл. 8, 9, 13, 14 и 18, имеющих значения относительного риска (RR) или отношения шансов (OR) менее одного. В одном воплощении по меньшей мере один аллель указывающий на пониженную предрасположенность к раку молочной железы, представлен в составе гаплотипов, выбранном из группы состава гаплотипов, состоящей из:

(i) rs895398 аллель 2, rs4848543 аллель 2, rs6759589 аллель 3, rs838066 аллель 4, rs838100 аллель 1, rs838086 аллель 1, rs12711924 аллель 3 и rs3731603 аллель 3;

(ii) rs895398 аллель 2, rs4848543 аллель 2, rs6759589 аллель 3, rs838066 аллель 2, rs838100 аллель 1, rs838086 аллель 1, rs12711924 аллель 3 и rs3731603 аллель 3;

(iii) rs895398 аллель 2, rs4848543 аллель 2, rs6759589 аллель 1, rs838066 аллель 2, rs838100 аллель 3, rs838086 аллель 2, rs12711924 аллель 3 и rs3731603 аллель 3;

(iv) rs895398 аллель 2, rs4848543 аллель 2, rs6759589 аллель 3, rs838066 аллель 4, rs838100 аллель 3, rs838086 аллель 2, rs12711924 аллель 3 и rs3731603 аллель 3;

(v) rs895398 аллель 2, rs4848543 аллель 2, rs6759589 аллель 3, rs838066 аллель 2, rs838100 аллель 3, rs838086 аллель 2, rs12711924 аллель 3 и rs3731603 аллель 3;

(vi) rs895398 аллель 2, rs4848543 аллель 2, rs6759589 аллель 1, rs838066 аллель 2, rs838100 аллель 3, rs838086 аллель 1, rs12711924 аллель 1 и rs3731603 аллель 4;

(vii) rs895398 аллель 2, rs4848543 аллель 2, rs6759589 аллель 3, rs838066 аллель 4, rs838100 аллель 3, rs838086 аллель 1, rs12711924 аллель 1 и rs3731603 аллель 4;

(viii) rs895398 аллель 2, rs4848543 аллель 2, rs6759589 аллель 1, rs838066 аллель 2, rs838100 аллель 3, rs838086 аллель 1, rs12711924 аллель 3 и rs3731603 аллель 3; и

(ix) rs895398 аллель 2, rs4848543 аллель 2, rs6759589 аллель 3, rs838066 аллель 2, rs838100 аллель 1, rs838086 аллель 1, rs12711924 аллель 1 и rs3731603 аллель 4.

В другом воплощении пониженная предрасположенность к раку молочной железы обусловлена составом гаплотипа, выбранного из группы состава гаплотипов, состоящей из:

(i) rs10191184 аллель 3, rs6435957 аллель 2, rs10171745 аллель 3, rs6716542 аллель 3, rs4491709 аллель 2, rs12621130 аллель 3, rs6735174 аллель 1, rs6435959 аллель 2, rs2372943 аллель 3, rs13387042 аллель 3, rs10490444 аллель 1, rs13011060 аллель 3;

(ii) rs10191184 аллель 3, rs6435957 аллель 4, rs10171745 аллель 3, rs6716542 аллель 3, rs4491709 аллель 4, rs12621130 аллель 1, rs6735174 аллель 1, rs6435959 аллель 4, rs2372943 аллель 3, rs13387042 аллель 3, rs10490444 аллель 3, rs13011060 аллель 1;

(iii) rs10191184 аллель 1, rs6435957 аллель 4, rs10171745 аллель 3, rs6716542 аллель 3, rs4491709 аллель 4, rs12621130 аллель 1, rs6735174 аллель 1, rs6435959 аллель 4, rs2372943 аллель 1, rs13387042 аллель 3, rs10490444 аллель 3, rs13011060 аллель 1;

(iv) rs10191184 аллель 3, rs6435957 аллель 2, rs10171745 аллель 3, rs6716542 аллель 3, rs4491709 аллель 2, rs12621130 аллель 3, rs6735174 аллель 1, rs6435959 аллель 2, rs2372943 аллель 3, rs13387042 аллель 3, rs10490444 аллель 1, rs13011060 аллель 1;

(v) rs10191184 аллель 1, rs6435957 аллель 2, rs10171745 аллель 1, rs6716542 аллель 3, rs4491709 аллель 2, rs12621130 аллель 1, rs6735174 аллель 1, rs6435959 аллель 2, rs2372943 аллель 3, rs13387042 аллель 3, rs10490444 аллель 3, rs13011060 аллель 3;

(vi) rs10191184 аллель 3, rs6435957 аллель 4, rs10171745 аллель 3, rs6716542 аллель 3, rs4491709 аллель 4, rs12621130 аллель 1, rs6735174 аллель 1, rs6435959 аллель 4, rs2372943 аллель 1, rs13387042 аллель 3, rs10490444 аллель 3, rs13011060 аллель 3;

(vii) rs10191184 аллель 1, rs6435957 аллель 2, rs10171745 аллель 3, rs6716542 аллель 3, rs4491709 аллель 4, rs12621130 аллель 1, rs6735174 аллель 1, rs6435959 аллель 4, rs2372943 аллель 1, rs13387042 аллель 3, rs10490444 аллель 3, rs13011060 аллель 1;

(viii) rs10191184 аллель 3, rs6435957 аллель 4, rs10171745 аллель 3, rs6716542 аллель 3, rs4491709 аллель 4, rs12621130 аллель 1, rs6735174 аллель 1, rs6435959 аллель 2, rs2372943 аллель 3, rs13387042 аллель 3, rs10490444 аллель 1, rs13011060 аллель 3;

(ix) rs10191184 аллель 3, rs6435957 аллель 2, rs10171745 аллель 3, rs6716542 аллель 3, rs4491709 аллель 2, rs12621130 аллель 1, rs6735174 аллель 1, rs6435959 аллель 2, rs2372943 аллель 3, rs13387042 аллель 3, rs10490444 аллель 1, rs13011060 аллель 3;

(x) rs10191184 аллель 1, rs6435957 аллель 2, rs10171745 аллель 3, rs6716542 аллель 3, rs4491709 аллель 2, rs12621130 аллель 3, rs6735174 аллель 1, rs6435959 аллель 2, rs2372943 аллель 3, rs13387042 аллель 3, rs10490444 аллель 1, rs13011060 аллель 3;

(xi) rs10191184 аллель 3, rs6435957 аллель 4, rs10171745 аллель 3, rs6716542 аллель 3, rs4491709 аллель 4, rs12621130 аллель 1, rs6735174 аллель 1, rs6435959 аллель 4, rs2372943 аллель 3, rs13387042 аллель 3, rs10490444 аллель 3, rs13011060 аллель 1.

В определенных воплощениях пониженная подверженность характеризуется значениями риска (относительного риска или отношения шансов) менее 0,9, включая риск менее 0,8, а риск менее 0,7, риск менее 0,6 и риск менее 0,5.

В принципе, оценка маркеров и гаплотипов, для которых здесь была показана ассоциация с раком молочной железы, может использоваться в комбинации с анализом образца, содержащего геномную ДНК человеческого индивида или массива данных генотипа, полученного у человеческого индивида, на наличие или отсутствие по меньшей мере одного связанного с риском аллеля по меньшей мере одного варианта, связанного с риском рака молочной железы, не находящегося в неравновесии по сцеплению с каким-либо из маркеров, приведенных в табл. 10, 15 и 19. Другими словами, генетические факторы риска, не сцепленные с какими-либо факторами риска, описываемыми здесь, могут использоваться в комбинации с факторами риска по данному изобретению для получения оценки совокупного риска индивида на основании множества факторов риска. Более того, анализ более чем одного фактора риска, ассоциированного по данному описанию с раком молочной железы, может использоваться в комбинации для получения совокупного общего риска. В одном таком воплощении анализ rs4848543 (SEQ ID NO:1), rs3803662 (SEQ ID NO:3) и rs13387042 (SEQ ID NO:2) выполняется для индивида или образца, взятого у

индивида и производится анализ совокупного риска. В другом воплощении производится анализ rs3803662 (SEQ ID NO:3) и rs13387042 (SEQ ID NO:2).

Способы, применения, наборы и устройства по изобретению могут в некоторых воплощениях также включать анализирование информации, не относящейся к генетической, для оценки риска, диагностики или определения прогноза у индивида. Подобная информация, не относящаяся к генетической, в определенных воплощениях может включать возраст, пол, этническую принадлежность, социоэкономический статус, предшествующий диагноз заболевания, медицинский анамнез субъекта, семейный анамнез рака молочной железы, биохимические показатели и/или клинические показатели. Оценка общего риска по генетическим и негенетическим факторам риска может осуществляться с применением способов, известных специалистам.

Неравновесие по сцеплению (НС) между сегментами генома (например, маркерами) в определенных воплощениях изобретения характеризуется определенными значениями показателя неравновесия по сцеплению. Согласно описанию ниже, неравновесие по сцеплению может характеризоваться конкретными численными значениями показателей НС, r^2 и $|D'|$. В одном предпочтительном воплощении неравновесие по сцеплению характеризуется значениями r^2 выше 0,1. В другом предпочтительном воплощении неравновесие по сцеплению характеризуется значениями r^2 выше 0,2. Также возможны другие пороговые значения r^2 , включая 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 0,95, 0,96, 0,97, 0,98, 0,99, но не ограничиваясь ими. В другом предпочтительном воплощении неравновесие по сцеплению характеризуется значениями $|D'|$ выше 0,5. В другом предпочтительном воплощении неравновесие по сцеплению характеризуется значениями $|D'|$ выше 0,8. Также возможны другие пороговые значения $|D'|$, включая 0,2, 0,3, 0,4, 0,6, 0,7, 0,8, 0,9, 0,95, 0,96, 0,97, 0,98 и 0,99, но не ограничиваясь ими. В определенных воплощениях неравновесие по сцеплению характеризуется численными пороговыми значениями либо для $|D'|$, либо для r^2 . В одном таком воплощении неравновесие по сцеплению характеризуется численными пороговыми значениями выше 0,8 для $|D'|$ и выше 0,2, для r^2 или обоими. В определенных воплощениях НС определяется в конкретной популяции. В некоторых таких воплощениях популяция выбрана из европеоидной, китайской, японской и африканской популяций. В одном воплощении популяцией является европеоидная популяция CEPH.

В некоторых других воплощениях способов, применений, устройств или наборов по изобретению индивид имеет определенное происхождение. В одном воплощении происхождение выбрано из черного африканского происхождения, европеоидного происхождения и китайского происхождения. В другом воплощении происхождение представляет собой происхождение от черных африканцев. В другом воплощении происхождение представляет собой происхождение от афроамериканцев. В другом воплощении происхождение представляет собой европеоидное происхождение. В определенном воплощении происхождение указывается самостоятельно индивидом, проходящим генетический анализ или генотипирование. В других воплощениях происхождение определяется с помощью генетического тестирования, включающего определение по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера в образце нуклеиновой кислоты индивида, где наличие или отсутствие аллеля указывает на происхождение индивида.

Краткое описание графических материалов

Описанные выше и другие объекты, признаки и преимущества данного изобретения будут очевидны из следующего далее более подробного описания предпочтительных воплощений данного изобретения.

Фиг. 1 отображает геномную структуру региона хромосомы 2, включающего маркер rs4848543, который, как здесь показано, ассоциирован с раком молочной железы. Маркеры в неравновесии по сцеплению с rs4848543 также находятся в составе этого региона, обозначаемого здесь как блок неравновесия по сцеплению rs4848543. Регион характеризуется выраженным неравновесием по сцеплению (высокое НС), а его границы характеризуются участками с высокой рекомбинацией. Маркеры в неравновесии по сцеплению с rs4848543, характеризуемые значениями показателя НС r^2 выше 0,2 в европеоидной популяции CEU HarMap, охватывают позиции 120,023,583 - 120,117,062 на хромосоме 2 в NCBI Build 34, и позиции 119,644,908-119,738,387 пн в NCBI Build 36. В обоих этих формах последовательностей, регион, охватываемый этими маркерами, составляет 93,479 пн.

Фиг. 2 отображает геномную структуру региона на хромосоме 2, включающего маркер rs13387042. Маркеры в неравновесии по сцеплению с rs13387042 также находятся в составе данного региона, обозначаемого здесь как блок неравновесия по сцеплению rs13387042. Регион характеризуется выраженным неравновесием по сцеплению (высокое НС), а его границы характеризуются участками с высокой рекомбинацией. Маркеры в неравновесии по сцеплению с rs13387042, характеризуемые значениями показателя НС r^2 выше 0,2 в европеоидной популяции CEU HarMap, охватывают позиции 218,059,508 - 218,141,061 на хромосоме 2 в NCBI Build 34, и позиции 217,565,211-217,646,764 пн в NCBI Build 36. В обоих этих формах последовательностей, регион охватывает 81,553 пн.

Фиг. 3 отображает геномную структуру региона хромосомы 16, включающего маркер rs3803662. Маркеры в неравновесии по сцеплению с rs3803662 также находятся в составе данного региона, обозначаемого здесь как блок неравновесия по сцеплению rs3803662. Регион характеризуется выраженным не-

равновесием по сцеплению (высокое НС), а его границы характеризуются участками с высокой рекомбинацией. Маркеры в неравновесии по сцеплению с rs3803662, характеризуются значениями показателя НС r^2 выше 0,2 в европеоидной популяции CEU HarMap, охватывают позиции 52,314,403 - 52,413,602 на хромосоме 16 в NCBI Build 34 и позиции 51,093,311-51,192,501 пн в NCBI Build 36. В обеих этих формах последовательностей регион охватывает 99,190 пн.

Сведения, подтверждающие возможность осуществления изобретения

Ниже следует описание предпочтительных воплощений изобретения.

Данное изобретение описывает полиморфные варианты и гаплотипы, для которых была обнаружена ассоциация с раком молочной железы. Было установлено, что определенные аллели конкретных полиморфных маркеров (например, маркера rs4848543, маркера rs13387042 и маркера rs3803662, и маркеров в неравновесии по сцеплению с ними, маркеров, ассоциированных с блоком неравновесия по сцеплению rs4848543, блоком неравновесия по сцеплению rs13387042 и блоком неравновесия по сцеплению rs3803662, маркеры по табл. 10, 15, 19, 20, 21 и 22, например, маркеры, указанные в таблице 10, 15 и 19) и гаплотипы, включающие такие аллели, ассоциированы с раком молочной железы. Такие маркеры и гаплотипы могут применяться для контроля за риском рака молочной железы, как будет более подробно описано далее. Следующие воплощения изобретения включают наборы для оценки наличия или отсутствия определенных аллелей данных маркеров.

Определения

Последовательности нуклеиновых кислот изображены слева направо в направлении от 5' к 3', если не указано иначе. Численные диапазоны, перечисленные в спецификации, включают значения, определяющие диапазон, и включают каждое целое значение или дробное значение в составе определенного диапазона. Если не указано иначе, все используемые здесь технические и научные термины имеют то же значение, в котором они обычно используются специалистами в области, к которой относится изобретение.

Следующие термины будут в настоящем контексте иметь указанные значения:

"Полиморфный маркер," иногда обозначаемый как "маркер" в данном описании, относится к полиморфному сайту генома. Каждый полиморфный маркер имеет по меньшей мере две разновидности последовательности, характерные для определенных аллелей полиморфного сайта. Таким образом, генетическая ассоциация с полиморфным маркером подразумевает, что существует ассоциация по меньшей мере с одним специфическим аллелем этого конкретного полиморфного маркера. Маркер может включать любой аллель любого вида варианта, обнаруженного в геноме, включая ОНП, микросателлиты, инсерции, делеции, дупликации и транслокации.

"Аллель" относится к нуклеотидной последовательности заданного локуса (позиции) на хромосоме. Аллель полиморфного маркера таким образом относится к составу (т.е. последовательности) маркера на хромосоме. Геномная ДНК индивида содержит два аллеля (например, аллель-специфичные последовательности) для любого заданного полиморфного маркера, репрезентативного для каждой копии маркера на каждой хромосоме. Коды нуклеотидов, используемые здесь в последовательностях, следующие: A = 1, C = 2, G = 3, T = 4. Для аллелей микросателлитов в качестве референсной использована выборка CEPH (Centre d'Etudes du Polymorphisme Humain, генобанк, выборка CEPH 1347-02), наиболее короткий аллель каждого микросателлита в данной выборке обозначен как 0, а все другие аллели в других выборках пронумерованы относительно данной последовательности. Так, например, аллель 1 является на 1 пн длиннее, чем самый короткий аллель в выборке CEPH, аллель 2 является на 2 пн длиннее, чем самый короткий аллель в выборке CEPH, аллель 3 является на 3 пн длиннее, чем самый короткий аллель в выборке CEPH, и т.д., а аллель -1 является на 1 пн короче, чем самый короткий аллель в выборке CEPH, аллель -2 является на пн короче, чем самый короткий аллель в выборке CEPH и т.д.

Возможные вариации в последовательности, приведенные здесь, находятся в соответствии с предложенными IUPAC-IUB. Данные коды совместимы с кодами, используемыми в базах данных EMBL, GenBank и PIR.

код IUB	Значение
A	Аденозин
C	Цитидин
G	Гуанин
T	Тимидин
R	G или A
Y	T или C
K	G или T
M	A или C
S	G или C

W	A или T
B	C G или T
D	A G или T
H	A C или T
V	A C или G
N	A C G или T (любое основание)

Позиция нуклеотида, по которой возможно существование более одной последовательности в популяции (либо в естественной популяции, либо в искусственной популяции, например библиотеке синтетических молекул), обозначается здесь как "полиморфный сайт".

"Однонуклеотидный полиморфизм" или ОНП представляет собой вариацию последовательности ДНК, возникающую, когда единичный нуклеотид в определенном положении в геноме отличается между представителями вида или между парами хромосом у индивида. Большинство ОНП имеют два аллеля. Каждый индивид в этом отношении является либо гомозиготным по одному аллелю полиморфизма (т.е. обе хромосомные копии индивида имеют одинаковые нуклеотиды в месте ОНП) или индивид является гетерозиготным (т.е. две сестринские хромосомы индивида содержат разные нуклеотиды). Номенклатура ОНП по данному описанию соответствует официальным идентификационным кодам (от англ. Reference SNP (rs) ID identification tag), приписанным каждому уникальному ОНП Национальным Центром по Биотехнологической Информации (от англ. National Center for Biotechnological Information, NCBI).

"Вариант" по данному описанию относится к сегменту ДНК, отличающемуся от референсной ДНК. "Маркер" или "полиморфный маркер" по данному описанию представляет собой вариант. Аллели, отличающиеся от референсных, обозначаются здесь как аллели "вариантов".

"Микросателлит" является полиморфным маркером, имеющим множественные малые повторы длиной от 2 до 8 нуклеотидов (например, повторы CA) в конкретном сайте, у которых число повторяющихся мотивов варьирует в общей популяции.

"Индели" представляют собой распространенную форму полиморфизма, включающую небольшую инсерцию или делецию, которая обычно имеет длину всего несколько нуклеотидов.

"Гаплотип" в данном описании относится к сегменту геномной ДНК, в пределах одной ната-ДНК, характеризующемуся специфической комбинацией последовательно расположенных аллелей. У диплоидных организмов, таких как человек, гаплотип включает одного представителя из пары аллелей по каждому полиморфному маркеру или локусу. В определенном воплощении гаплотип может включать два или более аллелей, три или более аллелей, четыре или более аллелей, или пять или более аллелей. Гаплотипы описываются здесь в контексте названия маркера и аллеля маркера в данном гаплотипе, например "1 rs4848543" или "1-rs4848543" означает 1 аллель (аллель A) маркера rs4848543, входящего в гаплотип, и является эквивалентом "аллель 1 rs4848543" или "аллель A rs4848543). Кроме того, коды аллелей в гаплотипе являются такими же, как в отдельных маркерах, т.е. 1 = A, 2 = C, 3 = G и 4 = T.

Термин "подверженность" по данному описанию относится к индивиду (или группе индивидов), склонных к развитию рака молочной железы или способных в меньшей степени противостоять развитию рака молочной железы по сравнению со среднестатистическим индивидом. Термин включает как повышенную подверженность, так и пониженную подверженность. Так, определенные полиморфные маркеры и/или гаплотипы по изобретению в данном описании могут быть характерны для повышенной подверженности (т.е. повышенного риска) раку молочной железы, характеризующей относительным риском (ОР) более одного или отношением шансов (ОШ) более одного. Возможно, маркеры и/или гаплотипы по изобретению характерны для пониженной подверженности (т.е. пониженного риска) сердечно-сосудистых заболеваний, характеризующих относительным риском менее одного.

Термин "и/или" в данном контексте следует трактовать как указание на то, что подразумеваются либо один, либо оба из пунктов, соединенных таким термином. Другими словами, данный термин здесь нужно воспринимать в значении "один или другой, или оба".

Термин "ассоциированный с" в данном описании сцепления генетических элементов (генов и/или маркеров) означает элементы, находящиеся в неравновесии по сцеплению. Предпочтительно термин должен использоваться для обозначения элементов, находящихся в неравновесии по сцеплению, характеризующимся значениями показателя неравновесия по сцеплению r^2 выше 0,2.

Термин "переводная таблица" в данном описании означает таблицу, приводящую в соответствие одну форму данных с другой или одну или более форм данных с прогнозируемым исходом, к которому относятся данные, например фенотип или признак. Например, переводная таблица может включать соответствие между аллельными данными по меньшей мере для одного маркера и определенным признаком или фенотипом, например диагнозом определенного заболевания, которое вероятно проявится у индивида, имеющего определенные аллельные данные или которое может проявиться у него с большей вероятностью, чем у индивида, не имеющего определенные аллельные данные. Переводные таблицы могут быть многомерными, т.е. они могут содержать информацию одновременно по нескольким аллелям отдельных маркеров или они могут содержать информацию о множестве маркеров, и они также могут включать другие факторы, например подробности о диагнозе заболеваний, информацию о расовой при-

надлежности, биомаркерах, биохимических показателях, способах терапии или лекарственных препаратах и т.д.

"Компьютерные носители" представляют собой носители для хранения информации, к которой можно получить доступ при помощи компьютера с использованием коммерчески доступных или созданных по заказу программ. Примеры компьютерных носителей включают блоки памяти (например, RAM, ROM, флэш-память, и т.д.), оптические носители информации (например, CD-ROM), магнитные носители (например, компьютерные жесткие диски, гибкие диски и т.д.), перфокарты или другие коммерчески доступные носители. Для хранения или доступа к хранящейся информации, информация может быть перенесена между системой, представляющей интерес, и носителем, между компьютерами, либо между компьютерами и компьютерными носителями. Такое перенесение может быть электронным или с помощью других доступных методов, например ИК-связи, беспроводных соединений и т.д.

"Образец нуклеиновой кислоты" представляет собой образец, полученный у индивидов и содержащий нуклеиновую кислоту. В определенных воплощениях, т.е. для определения конкретных полиморфных маркеров и/или гаплотипов образец нуклеиновой кислоты содержит геномную ДНК. Такой образец нуклеиновой кислоты может быть получен из любого источника, содержащего геномную ДНК, включая как образец крови, образец амниотической жидкости, образец спинномозговой жидкости или образец ткани кожи, мышцы, буккальной слизистой или конъюнктивы, плаценты, желудочно-кишечного тракта или других органов.

Термин "препарат, используемый при раке молочной железы" в данном описании означает агент, который может применяться для облегчения или предупреждения симптомов, ассоциированных с раком молочной железы.

Термин "нуклеиновая кислота, ассоциированная с раком молочной железы" в данном описании означает нуклеиновую кислоту, которая, как было показано, ассоциируется с раком молочной железы. Термин включает описанные здесь маркеры и гаплотипы, а также маркеры и гаплотипы в сильном неравновесии по сцеплению (НС) с ними, но не ограничивается ими. В одном воплощении нуклеиновая кислота, ассоциированная с раком молочной железы, относится к определенному региону генома, например, блоку неравновесия по сцеплению, для которого обнаружена ассоциация с раком молочной железы по меньшей мере по одному полиморфному маркеру, расположенному в составе региона или ассоциированному (т.е. находящемуся в неравновесии по сцеплению) с регионом.

Термин "все формы рака молочной железы" или "все формы РМЖ" относится к индивидам с диагностированным раком молочной железы, независимо от конкретного подтипа рака молочной железы.

Термин "умеренная предрасположенность" к раку молочной железы или рак молочной железы "УмерПред" (от англ. "Medium Predisposition") относится к подтипу рака молочной железы. Чтобы фенотип подходил под это определение, требуется соответствие пробанда по меньшей мере одному из следующих критериев:

1. Пробанд является представителем группы случаев рака молочной железы, объединяющей 3 или более пораженных родственников, находящихся на генетическом расстоянии не более 3 мейотических делений (3М).

2. Пробанд является одним из пары пораженных родственников, находящихся на расстоянии не более 3М, у одного из которых диагноз был поставлен в возрасте 50 лет или младше.

3. Пробанд является одним из пары пораженных родственников, находящихся на расстоянии не более 3М, у одного из которых была диагностирована вторая первичная опухоль любого типа.

4. У пробанда была диагностирована вторая первичная опухоль любого типа.

Из 1600 исландских пациентов, вошедших в представленное здесь исследование, 653 отвечают критериям "УмерПред" (40,8%).

Термин "первично-множественные опухоли молочной железы", или "ПМ ОМЖ" в данном описании означает случаи, когда наряду с постановкой первого диагноза рака молочной железы диагностируется по меньшей мере еще одна первичная опухоль, и как клинически, так и гистологически подтверждается, что две опухоли являются независимыми первичными опухолями, развивающимися одновременно или последовательно с первоначально диагностированной опухолью молочной железы, и возникают контралатерально или ипсилатерально.

Термин "индекс семейного анамнеза" (ИСА) в данном описании определяется по количеству имеющих рак молочной железы родственников у пробанда с заболеванием. Для каждого пробанда индекс, равный 1, присваивается за каждого пораженного родственника первой степени, 0,5 за каждого пораженного родственника второй степени и 0,25 за каждого родственника третьей степени. Общая сумма, полученная таким образом, по всем пораженным родственникам, представляет суммированный индекс семейного анамнеза, или ИСА.

Термин "блок НС rs4848543" или "блок неравновесия по сцеплению rs4848543" в данном описании означает регион генома на хромосоме 2 между позициями 119,987,002 и 120,129,001 по NCBI (National Center for Biotechnology Information) Build 34, и между позициями 119,608,327 и 120,129,001 по NCBI Build 36. Блок НС rs4848543 имеет размер 141,999 пн.

Термин "ген STEAP3" или "ген TSAP6" в данном описании означает ген 6-трансмембранного эпи-

телиального антигена простаты 3 (от англ. Six-transmembrane Epithelial Antigen of the Prostate или STEAP3), также обозначаемый геном пути активации опухолевых супрессоров 6 (от англ. Tumor Suppressor Activated Pathway 6 или TSAP6). Ген располагается на хромосоме 2q14.2 и охватывает позиции с 120076561 по 120118373 (NCBI Build 34).

Термин "блок НС rs13387042", или "блок неравновесия по сцеплению rs13387042" в данном описании означает блок неравновесия по сцеплению (НС) на хромосоме 2 между позициями 218,062,001 и 218,141,002 по NCBI (National Center for Biotechnology Information) Build 34. Блок НС охватывает 79,001 пар оснований (Build 34).

Термин "блок НС rs3803662", или "блок неравновесия по сцеплению rs3803662" в данном описании означает блок неравновесия по сцеплению (НС) на хромосоме 16 между позициями 52,291,041 и 52,436,127 по NCBI (National Center for Biotechnology Information) Build 34. Блок НС охватывает 145,086 пар оснований (Build 34).

Термин "гаплотип Daly" в данном описании означает 2-маркерные гаплотипы в составе заданного блока НС (регион высокого НС), которые идентифицированы в качестве хороших суррогатов ($r^2 > 0,8$) для дополнительного набора ОНП, типированных в проекте НарМар. Эти дополнительные ОНП имеют частоту минорного аллеля $> 5\%$ в выборках УТАН СЕРН (CEU) НарМар, и они не представлены в составе чипа Нар300 SNP и также они не маркируются ОНП, имеющимися на чипе [Pe'er, et al., (2006), Nat Genet, 38, 663-7].

Термин "рак молочной железы, позитивный по эстрогеновым рецепторам", или "ЭР-позитивный рак молочной железы", в данном описании относится к образцу ткани опухоли молочной железы, который считается положительным по экспрессии эстрогенового рецептора. Экспрессия эстрогенового рецептора может, например, определяться в радиоиммунном анализе или методом иммуногистохимии. Положительным считается уровень ≥ 10 фмоль/мг, определенный радиоиммунометрически или $\geq 10\%$ ядер, оцененных положительными при иммуногистохимическом определении.

Термин "рак молочной железы, позитивный по прогестероновым рецепторам" или "ПР-позитивный рак молочной железы" в данном описании относится к образцу ткани опухоли молочной железы, который считается положительным по экспрессии прогестеронового рецептора. Экспрессия прогестеронового рецептора может, например, определяться в радиоиммунном анализе или методом иммуногистохимии. Положительным считается уровень > 10 фмоль/мг, определенный радиоиммунометрически или $> 10\%$ ядер, оцененных положительными при иммуногистохимическом определении.

В ходе ассоциативного анализа популяции индивидов с диагностированным раком молочной железы было обнаружено, что конкретные аллели конкретных полиморфных маркеров ассоциированы с раком молочной железы. Общегеномный анализ вариантов, ассоциированных с раком, выявил ассоциацию рака молочной железы с двумя различными регионами на хромосоме 2 в составе сегментов Chr2q14 и Chr2q35, а также в составе региона на хромосоме 16 (Chr16q12). Для определенных маркеров и гаплотипов в этих регионах была обнаружена ассоциация с повышенным риском рака молочной железы.

Как показано в табл. 1, было обнаружено, что аллель А маркера rs4848543 на хромосоме 2q14.2 (также обозначаемый аллель А rs4848543 или 1 аллель rs4848543 или А-rs4848543) ассоциирован с повышенным риском рака молочной железы. При анализе фенотипа рака молочной железы "УмерПред" у 1598 пациентов в сравнении с 4477 популяционными контролями, не имеющими рака, оцениваемое значение относительного риска (RR) составило 1,42, при значении $p = 8,3 \times 10^{-8}$. После коррекции на множественность сравнений (317 089 ОНП маркеров, протестированных при помощи чипа для общегеномного сканирования, использованного в эксперименте, значение p составило 0,026, т.е. ассоциация является достоверной на общегеномном уровне. Ассоциация с более общим фенотипом "все формы РМЖ" несколько слабее, значение RR составляет 1,16 (см. Табл. 1). Данные результаты были воспроизведены на независимой исландской когорте лиц с раком молочной железы.

Маркер rs4848543 расположен в составе региона, обозначаемого здесь как блок НС rs4848543. Данный маркер ОНП и маркеры, сцепленные с данным маркером (например, маркеры, указанные в табл. 10), могут применяться в способах по данному изобретению. Благодаря существованию паттерна локального неравновесия по сцеплению в человеческом геноме, существует множество полиморфных маркеров, находящихся в сильном НС с rs4848543, для которых здесь была продемонстрирована ассоциация с раком. Эти сцепленные маркеры (например, маркеры, указанные в табл. 10), включая известные ОНП или другие полиморфные маркеры, например, микросателлиты или индели, а также другие сцепленные ОНП или другие полиморфные маркеры могут, таким образом, применяться поодиночке или в комбинации, в качестве суррогатных маркеров для выявления ассоциации с раком молочной железы, по данному описанию. В частности, ожидается, что другие полиморфные маркеры, расположенные в составе блока НС rs4848543, могут применяться в качестве суррогатных маркеров в способах по изобретению.

Второй регион на хромосоме 2 (2q35) был идентифицирован в ходе общегеномного ассоциативного исследования. Было обнаружено, что аллель А маркера rs13387042 (также обозначаемый аллель А rs13387042 или 1 аллель rs13387042 или А-rs13387042) ассоциирован с повышенным риском рака молочной железы (см. табл. 11). Данный ОНП расположен на хромосоме 2q35, на определенном расстоянии от

локуса STEAP3/TSAP6. Анализ 2181 индивида с раком молочной железы ("все формы РМЖ") и 12 441 популяционного контроля, показал, что аллель A rs13387042 обуславливает риск, равный 1,19, со значением $p = 4,0 \times 10^{-5}$. Риск (RR) при раке молочной железы "УмерПред" оказался сравнимым с таковым при Any BC. Суррогатные маркеры, которые также могли применяться для выявления ассоциации с rs13387042, указаны в табл. 15.

Данные результаты были воспроизведены на второй независимой выборке из населения Исландии (583 "случая" и 7966 "контролей"), и была получена оценка значения относительного риска RR для аллеля A rs13387042, ассоциированного с фенотипом "все формы РМЖ" в этой второй группе, равная 1,20 (значение $P_{\text{corr}} = 3,8 \times 10^{-3}$). Таким образом, исходные результаты были достоверно воспроизведены на независимой исландской выборке с очень близкой точечной оценкой относительного риска RR, и совокупным значением $P_{\text{corr}} = 2,0 \times 10^{-7}$, что очень близко к уровню значимости после коррекции Бонферрони на 317089 протестированных ОНП. Наблюдаемая частота и относительный риск для варианта аллели A rs13387042 соответствует оцененному популяционному атрибутивному риску в исландской популяции, равному 15,6%.

Исследование по воспроизведению на испанской и шведской выборках подтвердило данные результаты. При анализе 446 случаев с раком молочной железы и 977 контрольных образцов из Испании было получено значение RR, равное 1,21 (значение $P = 1,8 \times 10^{-2}$; табл. 11). Частота варианта A-rs13387042 была выше в контрольных образцах Испании, что позволило предположить, что более широкая распространенность варианта способствует большему генетическому грузу в отношении рака молочной железы в популяциях испанского происхождения. Кроме того, были проанализированы две шведские когорты "шведская семейственная" и "шведская консеквативная". В "шведской консеквативной" когорте был продемонстрирован значимый RR рака молочной железы, составлявший 1,31 (значение $P = 2,0 \times 10^{-4}$), тогда как в "шведской семейственной" когорте оценка относительного риска составила 1,11, что было статистически недостоверно. Тем не менее, в совокупности на объединенных шведских когортах была получена значимая оценка относительного риска, равная 1,22 (значение $P = 8,1 \times 10^{-4}$).

Оценки у исландской, испанской и шведской когорт были объединены при совместном анализе, по методу Мантеля-Хензеля. В результате была получена оценка относительного риска, равная 1,20, и значение P , равное $3,8 \times 10^{-11}$. Это значительно ниже порогового уровня значимости на уровне генома при использовании метода Бонферрони для коррекции на 317 089 протестированных ОНП. Таким образом, было сделано заключение, что аллель A rs13387042 обуславливает достоверный и воспроизводимый риск рака молочной железы в нескольких популяциях европейского происхождения. Совокупный популяционный атрибутивный риск был оценен как равный 16,4%.

Последующий общегеномный анализ ОНП по данным, полученным для 1600 пациентов с раком молочной железы и 11563 контролей, обнаружил, что аллель T ОНП rs3803662 (T-rs3803662; rs3803662 аллель T) обуславливает 1,23-кратное оцениваемое повышение риска для Any BC (табл. 16). Данные результаты были подтверждены на второй независимой когорте 594 исландских пациентов с раком молочной железы и 1433 контролей. При объединении данных по этим двум исландским выборкам была получена оценка относительного риска 1,23 и значение $P = 2,8 \times 10^{-7}$ при коррекции на родство между индивидами. Это соответствовало оцениваемому популяционному атрибутивному риску, равному 10,1% (табл. 16).

Суррогатные маркеры, которые могут применяться для выявления ассоциации с rs3803662 указаны в табл. 19.

Был проведен анализ воспроизводимости данных результатов в шведской и испанской когортах, описанных выше, а также анализ 558 случаев с раком молочной железы и 1384 контролей из г. Неймеген, Голландия. Как показано в табл. 16, достоверно повышенный риск для аллеля T rs3803662 наблюдался во всех трех из перечисленных дублирующих когорт. При совместном анализе этих трех не-исландских дублирующих когорт оценка совокупного относительного риска составила 1,35, а значение $P = 5,1 \times 10^{-12}$. При совместном анализе исландской и дублирующих когорт оценка совокупного относительного риска составила 1,28 при значении $P = 2,7 \times 10^{-17}$. Это значительно ниже порогового уровня значимости на уровне генома при коррекции на количество протестированных ОНП. Соответствующая оценка совокупного популяционного атрибутивного риска составила 13,4% (табл. 16).

Была изучена ассоциация между аллелем A rs13387042 и аллелем T rs3803662, статусом по эстрогеновым рецепторам (ЭР) и прогестероновым рецепторам (ПР). Значительный риск рака молочной железы, связанный с аллелем A rs13387042 и аллелем T rs3803662, наблюдался в пределах группы пациентов с диагностированными ЭР-позитивными опухолями и разница между значениями отношения шансов OR для ЭР-позитивных и ЭР-негативных опухолей также была достоверной (табл. 25). Также наблюдалась тенденция к существованию риска рака молочной железы преимущественно среди пациентов с диагностированными ПР-позитивными опухолями.

Биология STEAP3/TSAP6

ДНК STEAP3/TSAP6 была впервые получена из одной из мРНК, индуцированных активацией p53 в клетках линии миеломы грызунов LTR6, содержащих температуро-чувствительный ген p53 [Amson, et al., (1996), Proc Natl Acad Sci USA, 93, 3953-7]. Данная форма у грызунов получила название гена пути

активации опухолевых супрессоров (от англ. Tumour Suppressor Activated Pathway 6 или TSAP6). Позднее было подтверждено, что данный ген включается при активации p53 в клетках LTR6 [Passer et al., (2003), *Proc Natl Acad Sci USA*, 100, 2284-9]. Его аналог у крыс, обозначенный pHyde, был выделен из клеточной линии опухоли простаты в ходе экспериментов, проводимых с целью выявления генов, по-разному экспрессируемых в двух клеточных линиях с различной склонностью к метастазированию [Rinaldy и Steiner, (1999), *DNA Cell Biol*, 18, 829-36]. Было показано, что кДНК pHyde вызывает апоптоз в клеточных линиях простаты крысы и человека и ксенотрансплантатах [Rinaldy et al., (2000), *Gan To Kagaku Ryoho*, 27 Suppl 2, 215-22; Steiner et al., (2000), *Cancer Res*, 60, 4419-25]. Позже было показано, что апоптоз, вызванный pHyde, запускается через путь каспазы-3 [Zhang et al., (2001), *Oncogene*, 20, 5982-90]. Подавление STEAP3/TSAP6 также было связано с прогрессированием гепатоцеллюлярной карциномы [Coulouarn, et al., (2005), *J Hepatol*, 42, 860-9].

Passer et al. выделили кДНК человеческого аналога STEAP3/TSAP6 и показали, что мРНК можно было индуцировать при активации p53 в клетках линии карциномы молочной железы MCF7 [Passer et al., (2003), *Proc Natl Acad Sci USA*, 100, 2284-9]. Индукция p53 происходит на уровне транскрипции и опосредуется консервативным p53-отвечающим элементом в промоторе STEAP3/TSAP6 между нуклеотидами -478 и -357 (последовательность получила кодом доступа AY214461) [Passer et al., (2003), *Proc Natl Acad Sci USA*, 100, 2284-9]. Белок был охарактеризован как состоящий из 488 аминокислот, имеющий размер 50-55 кДа, 6-кратно пронизывающий мембрану трансмембранный белок [Passer et al., (2003), *Proc Natl Acad Sci USA*, 100, 2284-9]. Антисенсовая РНК, прицельно воздействующая на STEAP3/TSAP6, ингибировала p53-зависимый апоптоз, что также свидетельствовало в пользу того, что STEAP3/TSAP6 действует как эффектор апоптоза [Passer et al., (2003), *Proc Natl Acad Sci USA*, 100, 2284-9]. STEAP3/TSAP6 физически взаимодействует с двумя белками, задействованными в регуляции клеточного цикла и апоптоза; Nix и Myt1 [Passer et al., (2003), *Proc Natl Acad Sci USA*, 100, 2284-9]. Nix (также обозначаемый BNIP3L) является митохондриальным проапоптотическим белком, связанным с Bcl2. Nix и STEAP3/TSAP6 усиливают проапоптотические эффекты друг друга [Passer et al., (2003), *Proc Natl Acad Sci USA*, 100, 2284-9]. Myt1 обладает двойной специфичностью (Ser/Thr и Tyr) киназы, блокирующей клеточный цикл в переходе G2/M путем фосфорилирования и таким образом, ингибирования циклин-зависимой киназы p34^{cdc2}. Взаимодействие между Myt1 и TSAP6 способствует Myt1-зависимому фосфорилированию p34^{cdc2} путем поддержания Myt1 в его фосфорилированном активированном состоянии [Passer et al., (2003), *Proc Natl Acad Sci USA*, 100, 2284-9].

Несколько неожиданно было обнаружено, что STEAP3/TSAP6 является геном, отвечающим за микроцитарную гипохромную анемию у мышей мутантной линии pt1054 [Ohgami et al., (2005), *Nat Genet*, 37, 1264-9]. Внеклеточное железо переносится трансферрином (Tf) из кишечника, ретикулоэндотелиальной системы и печени ко всем пролиферирующим клеткам организма. Tf-связанное железо поступает внутрь клетки посредством эндоцитоза, опосредованного рецептором к трансферрину (TfR1). Железо высвобождается из комплекса с Tf в эндосомах при окислении и затем доставляется в цитоплазму переносчиком двухвалентных ионов металлов Dmt1. Tf и TfR1 затем возвращаются во внеклеточное пространство. В окружающей среде и в связанном с Tf виде железо преимущественно находится в окисленном состоянии Fe³⁺ (трехвалентном) и должно быть восстановлено до Fe²⁺ (двухвалентного) состояния для транспортировки через цитоплазматическую мембрану. Благодаря наличию ассоциации с анемией, STEAP3/TSAP6 был идентифицирован в качестве основной редуктазы железа, осуществляющей эту реакцию в клетках эритроидного ряда, что подразумевает его важное значение в поступлении железа [Ohgami et al., (2005), *Nat Genet*, 37, 1264-9]. STEAP3/TSAP6 также обладает способностью восстанавливать ионы меди, и такая металлоредуктазная активность свойственна и другим представителям семейства STEAP: STEAP2 и STEAP4 [Ohgami, et al., (2006), *Blood*, 108, 1388-94].

Было также показано, что STEAP3/TSAP6 участвует в стимуляции секреции белков альтернативным путем. Классическая секреция белков опосредована через NH₂-концевую сигнальную последовательность на белках, подлежащих секреции. Сигнальная последовательность направляет белок через эндоплазматический ретикулум/аппарат Гольджи с последующим транспортом через плазматическую мембрану в секреторных везикулах. После слияния секреторной везикулы с плазматической мембраной секретированный белок высвобождается во внеклеточное пространство. Секреция белков может также осуществляться альтернативным путем, в ходе которого внутриклеточные везикулы отпочковываются в просвет эндосом, образуя так называемые мультивезикулярные тельца. Слияние мультивезикулярных телец с плазматической мембраной приводит к выбросу таких мембрано-инкапсулированных везикул во внеклеточное пространство. Такие везикулы известны как экзосомы. Было показано, что STEAP3/TSAP6 стимулирует образование экзосом p53-зависимым способом [Amzallag et al., (2004), *J Biol Chem*, 279, 46104-12; Yu et al., (2006), *Cancer Res*, 66, 4795-801]. Экзосомы могут быть вовлечены в противоопухолевый ответ, имеющий важное значение. Одним из белков, секретируемых в экзосомах в ответ на активацию пути p53-STEAP3/TSAP6, является маспин (Maspin, от англ. mammary serine protease inhibitor), обладающий ингибиторным эффектом в отношении ангиогенеза, опухолевой инвазии и метастазирования клеток опухоли молочной железы [Sheng et al., (1996), *Proc Natl Acad Sci USA*, 93, 11669-74; Shi, et al., (2001), *Cancer Res*, 61, 6945-51; Zhang et al., (1997), *Mol. Med.*, 3, 49-59; Zou et al., (1994), *Science*, 263,

526-9]. Экзосомы могут содержать провоспалительные факторы, такие как опухолевый белок, регулируемый на уровне трансляции/фактор высвобождения гистамина (от англ. Translationally Controlled Tumour Protein (TCTP)/Histamine Releasing Factor) и опухоли

специфичные антигены, такие как Her2/Neu и Marti [Amzallag, et al., (2004), J Biol Chem, 279, 46104-12; Ure et al., (2002), Lancet, 360, 295-305]. Предполагается, что STEAP3/TSAP6-опосредованное формирование эндосом может отвечать за так называемый «эффект постороннего», при котором клетки, в которых индуцирована экспрессия p53, могут вызывать остановку клеточного цикла, проапоптотический ответ и гибель соседних клеток [Yu et al., (2006), Cancer Res, 66, 4795-801].

Почти ничего не известно о взаимоотношениях между возможными генетическими вариантами STEAP3/TSAP6 и риском рака. Скрининг мутаций в клеточных линиях карциномы простаты, ксенотрансплантатах и образцах опухолевой ткани позволил обнаружить 2 миссенс-мутации (Ala184Thr и Ile305Thr) в ксенотрансплантатах при исследовании всего 4 клеточных линий, 8 ксенотрансплантатов и 56 образцов опухолей (всего 68 образцов). Авторы пришли к выводу, что STEAP3/TSAP6 не является классическим геном супрессии опухолей при раке простаты. До настоящего времени не было описано других исследований генетических вариантов STEAP3/TSAP6 при какой-либо другой форме рака.

Возможность применения пути STEAP3/TSAP6 для разработки усовершенствованных способов лечения рака молочной железы

Наше наблюдение позволяет предположить, что генетический вариант блока HC STEAP3/TSAP6 представляет собой распространенный фактора риска развития карциномы молочной железы. Таким образом, способы лечения, направленно воздействующие на STEAP3/TSAP6, его гомологи и другие компоненты пути STEAP3/TSAP6, могут считаться кандидатами для лечения или предупреждения рака молочной железы. Такие мишени включают: сам STEAP3/TSAP6 (OMIM# 609671); его гомологи STEAP1 (OMIM# 604415), STEAP2/TIARP/STAMP1 (OMIM# 605094), и STEAP4 (NM_024636); взаимодействующие с ним белки Nix/BNIP3L (OMIM# 605368), Myt1 (OMIM# 602474), TCTP/фактор высвобождения гистамина (OMIM# 600763); белки пути p34^{cdc2} (OMIM# 116940), циклин B1 (OMIM# 123836), HER2/Neu (OMIM# 164870), маспин (OMIM# 154790); и другие представители пути STEAP3/TSAP6, которые будут идентифицированы или о вовлечении которых станет известно. Способы лечения могут включать малые молекулы, являющиеся агонистами или антагонистами STEAP3/TSAP6, его гомологов или компонентов пути; макромолекулярные агенты, являющиеся агонистами или антагонистами STEAP3/TSAP6, его гомологов или компонентов пути; или применение STEAP3/TSAP6, гомологов или компонентов пути в качестве трансгенов в протоколах генной терапии или иммунотерапии. STEAP3/TSAP6, его гомологи или компоненты пути и/или молекулы, воздействующие на их активность, могут также быть использованы в качестве адъювантов для цитостатической или цитотоксической терапии (химиотерапии, радиотерапии, генной или иммунной терапии) посредством потенциации их "эффекта постороннего".

Оценка маркеров и гаплотипов

Когда проводится сравнение индивидов, геномная последовательность в популяциях оказывается не идентичной. Скорее, геном отражает вариабельность последовательности между индивидами во многих участках генома. Такие вариации в последовательности обычно обозначаются как полиморфизмы, и в геноме существует множество таких сайтов. Например, человеческий геном имеет вариации последовательности, которые появляются в среднем на каждые 500 пар оснований. Наиболее распространенный вариант последовательности представляет собой вариации оснований по одной позиции в геноме, и такие варианты последовательности, или полиморфизмы, обычно называют однонуклеотидными полиморфизмами (ОНП). Полагают, что данные ОНП появились в результате единственной мутации, таким образом, обычно существует два возможных аллеля в каждом сайте ОНП; исходный аллель и мутировавший аллель. Благодаря естественному генетическому дрейфу и, возможно, также благодаря давлению отбора исходная мутация превратилась в полиморфизм, характеризующийся определенной частотой его аллелей в любой заданной популяции. В человеческом геноме найдены многие другие виды вариантов последовательностей, включая микросателлиты, инсерции, делеции, инверсии и вариации числа копий. Полиморфный микросателлит имеет множество коротких повторов оснований (таких как CA повторы, или TG на комплементарной цепи) в определенном участке, число повторяющихся мотивов варьирует в общей популяции. В общих терминах, каждая версия последовательности полиморфного сайта представляет собой определенный аллель полиморфного сайта. Такие варианты последовательностей могут все обозначаться полиморфизмами, возникающими в определенных полиморфных сайтах, характерных для обсуждаемого варианта последовательности. В общих терминах полиморфизмы могут включать любое число определенных аллелей. Так, в одном воплощении изобретения полиморфизм характеризуется наличием двух или более аллелей в любой заданной популяции. В другом воплощении полиморфизм характеризуется наличием трех или более аллелей. В других воплощениях полиморфизм характеризуется четырьмя или более аллелями, пятью или более аллелями, шестью или более аллелями, семью или более аллелями, девятью или более аллелями или десятью или более аллелями. Все такие полиморфизмы могут применяться в способах и наборах по данному изобретению, и таким образом находятся в рамках изобретения.

В некоторых случаях делается ссылка на разные аллели в полиморфных сайтах без выбора рефе-

ренсного аллеля. В других случаях для конкретного полиморфного сайта может быть обозначена референсная последовательность. Референсный аллель в некоторых случаях обозначается как аллель "дикого типа" и обычно выбирается либо как первый секвенированный аллель или как аллель "не пораженного" индивида (например, индивида, не имеющего признаков или фенотипа заболевания).

Аллели маркеров ОНП в данном описании означают основания А, С, G или Т, обнаруживаемые в полиморфных сайтах в ходе детекции ОНП. В данном описании использовали следующие коды аллелей ОНП: 1 = А, 2=С, 3=G, 4=Т. Специалисту очевидно, что при исследовании или чтении противоположной цепи ДНК можно в каждом случае определить комплементарный аллель. Следовательно, для полиморфного сайта (полиморфного маркера), характеризуемого полиморфизмом А/G, проводимый анализ может быть использован для специфической детекции наличия как одного из двух возможных оснований, так и обоих, т.е. А и G. Также возможно, что в ходе анализа, разработанного для детекции противоположной цепи ДНК-матрицы, может быть определено наличие комплементарных оснований Т и С. В количественном отношении (например, в терминах относительного риска) идентичные результаты будут получены при определении любой из цепей ДНК (+ цепи или - цепи).

Обычно референсная последовательность относится к определенной последовательности. Аллели, отличающиеся от референсного, иногда обозначаются как "варианты" аллелей. Вариант последовательности в данном описании означает последовательность, отличающуюся от референсной последовательности, но в остальном существенно схожую. Аллели полиморфных генетических маркеров, описываемые здесь, являются вариантами. Другие варианты могут включать изменения, затрагивающие полипептид. Отличия в последовательностях при сравнении с референсной нуклеотидной последовательностью могут включать инсерции или делеции единичного нуклеотида, или более чем одного нуклеотида, приводящие к сдвигу рамки считывания; замену по меньшей мере одного нуклеотида, приводящую к замене кодируемой аминокислоты; замену по меньшей мере одного нуклеотида, приводящую к образованию преждевременного стоп-кодона; делецию нескольких нуклеотидов, приводящую к делеции одной или более аминокислот, кодируемых нуклеотидами; инсерции одного или нескольких нуклеотидов, такие как при неравной рекомбинации или конверсии генов, приводящие к прерыванию кодирующей последовательности рамки считывания; дубликации всей или части последовательности; транспозиции; или перестройки нуклеотидной последовательности. Такие изменения последовательности могут изменять полипептид, кодируемый нуклеиновой кислотой. Например, если изменения в последовательности нуклеиновых кислот вызывают сдвиг рамки считывания, сдвиг рамки считывания может привести к изменению кодируемых аминокислот и/или привести к образованию преждевременного стоп-кодона, вызывая образование укороченного полипептида. Возможно, что полиморфизм, ассоциированный с заболеванием или его признаками, может быть молчащей заменой одного или нескольких нуклеотидов (т.е. замена, которая не приводит к изменению аминокислотной последовательности). Такой полиморфизм может, например, изменять сайты сплайсинга, влиять на стабильность транспортной мРНК или иначе влиять на транскрипцию или трансляцию кодируемого полипептида. Он может также изменять ДНК, так что увеличивается вероятность появления на соматическом уровне таких структурных изменений, таких как амплификации или делеции. Полипептид, кодируемый референсной нуклеотидной последовательностью является "референсным" полипептидом с определенной референсной последовательностью аминокислот, а полипептиды, кодируемые вариантами аллелей обозначаются "вариантными" полипептидами с вариантами последовательности аминокислот.

Гаплотип означает сегмент ДНК, характеризующийся специфической комбинацией последовательно расположенных аллелей. У диплоидных организмов, таких как человек, гаплотип включает одного представителя из представителя пары аллелей по каждому полиморфному маркеру или локусу. В определенном воплощении гаплотип может включать два или более аллелей, три или более аллелей, четыре или более аллелей, или пять или более аллелей, каждый аллель соответствует определенному полиморфному маркеру в составе сегмента. Гаплотипы могут включать комбинацию различных полиморфных маркеров, например, ОНП и микросателлиты, имеющие конкретные аллели в полиморфных сайтах. Гаплотипы таким образом включают комбинацию аллелей различных генетических маркеров.

Детекция специфических полиморфных маркеров и/или гаплотипов может выполняться известными способами детекции последовательностей в полиморфных сайтах. Например, могут использоваться такие стандартные способы генотипирования наличия ОНП и/или микросателлитных маркеров, как флуоресцентные методики (Chen X. et al., *Genome Res.* 9(5): 492-98 (1999)), применяющие ПЦР, ЛЦР, гнездовой ПЦР и другие методики амплификации нуклеиновых кислот. Определенные методики, используемые для генотипирования ОНП, включают генотипирующие системы TaqMan и платформы SNPlex (Applied Biosystems), масс-спектрометрию (например, систему MassARRAY от Sequenom), методы минисеквенирования, ПЦР в реальном времени, систему Bio-Plex (BioRad), системы CEQ и SNPstream (Beckman), технологию Molecular Inversion Probe array (например, Affymetrix GeneChip), и технологии BeadArray (например, Illumina GoldenGate и Infinium assays), но не ограничиваются ими. С помощью этих или других методов, используемых специалистами, можно идентифицировать один или более аллелей полиморфных маркеров, включая микросателлиты, ОНП или другие виды полиморфных маркеров.

В некоторых из описанных здесь методов индивид с повышенной подверженностью (т.е. с повы-

шенным риском) к любому исследуемому определенному заболеванию или его признакам (например, раком молочной железы), является индивидом, у которого идентифицирован по меньшей мере один специфический аллель одного или более полиморфных маркеров или гаплотипов, обуславливающих повышенную подверженность к заболеванию или его признакам (т.е. связанный с риском аллель маркера или гаплотип). В одном аспекте, связанный с риском маркер или гаплотип является тем, который обуславливает достоверно повышенный риск (или подверженность) к раку молочной железы. В одном воплощении достоверность, ассоциированная с маркером или гаплотипом, измеряется относительным риском (RR). В другом воплощении достоверность, ассоциированная с маркером или гаплотипом, измеряется/выражается отношением шансов (OR). В следующем воплощении достоверность выражается в процентах. В одном воплощении достоверно повышенный риск выражается как риск (относительный риск и/или отношение шансов) равный по меньшей мере 1,2, что включает, но не ограничивается значениями по меньшей мере 1,2, по меньшей мере 1,3, по меньшей мере 1,4, по меньшей мере 1,5, по меньшей мере 1,6, по меньшей мере 1,7, по меньшей мере 1,8, по меньшей мере 1,9, по меньшей мере 2,0, по меньшей мере 2,5, по меньшей мере 3,0, по меньшей мере 4,0 и по меньшей мере 5,0. В конкретном воплощении достоверным является риск (относительный риск и/или отношение шансов) по меньшей мере 1,2. В другом конкретном воплощении достоверным является риск по меньшей мере 1,3. В еще одном воплощении достоверным является риск по меньшей мере 1,4. В следующем воплощении достоверным является относительный риск по меньшей мере около 1,5. В следующем воплощении достоверным увеличением риска является по меньшей мере 1,7. Однако также рассматриваются и другие пороговые значения, например, по меньшей мере 1,15, 1,25, 1,35, и так далее, и такие пороговые значения также находятся в рамках данного изобретения. В других воплощениях достоверным является увеличение риска по меньшей мере около 20%, включая, но не ограничиваясь значениями 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 300 и 500%. В одном конкретном воплощении достоверным является увеличение риска по меньшей мере на 20%. В других воплощениях достоверным является увеличение риска по меньшей мере на 30%, по меньшей мере на 40%, по меньшей мере на 50%, по меньшей мере на 60%, по меньшей мере на 70%, по меньшей мере на 80%, по меньшей мере на 90% и по меньшей мере на 100%. Другие пороговые значения или диапазоны, которые специалисты сочтут подходящими для характеристики изобретения, также рассматриваются, и они также находятся в рамках данного изобретения. В определенных воплощениях достоверное увеличение риска характеризуется значением p , например значением p менее 0,05, менее 0,01, менее 0,001, менее 0,0001, менее 0,00001, менее 0,000001, менее 0,0000001, менее 0,00000001 или менее 0,000000001.

Связанный с риском полиморфный маркер или гаплотип по данному изобретению представляет собой по меньшей мере один аллель по меньшей мере одного маркера или гаплотип, встречающийся более часто у индивида с риском рака молочной железы ("пораженного"), по сравнению с частотой его встречаемости в группе сравнения (в контроле), и где наличие маркера или гаплотипа указывает на подверженность к заболеванию или его признакам. Контрольная группа в одном воплощении может быть популяционной выборкой, т.е. случайной выборкой из общей популяции. В другом воплощении, контрольная группа представлена группой индивидов, не имеющих заболевания. Такие не имеющие заболевания контроли могут в одном воплощении характеризоваться отсутствием одного или более специфических симптомов, ассоциированных с заболеванием. В другом воплощении контрольная группа лиц, не имеющих заболевания, характеризуется отсутствием одного или более факторов риска, специфических для заболевания. Такие факторы риска в одном воплощении представляют собой по меньшей мере один фактор риска окружающей среды. Примерами факторов окружающей среды являются натуральные продукты, минералы или химические вещества, для которых известна или предполагается способность влиять на риск развития определенного заболевания или его признаки. Другие факторы риска окружающей среды являются факторами риска, относящимися к образу жизни, включая особенности питания и питья, географическое расположение места проживания, и профессиональные факторы риска, но не ограничиваются ими. В другом воплощении факторы риска представляют собой по меньшей мере один генетический фактор риска.

Примером простого теста на корреляцию может быть точный тест Фишера в таблице два на два. Для заданной когорты хромосом таблица два на два строится исходя из числа хромосом, которые включают оба маркера или гаплотипа, один из маркеров или гаплотипов, но не хромосом, которые не включают ни один из маркеров или гаплотипов или хромосом, которые включают другие маркеры или гаплотипы. Другие статистические тесты на ассоциацию, известные специалистам, также рассматриваются и находятся в рамках изобретения.

В других воплощениях по изобретению индивид с пониженной подверженностью (т.е. со сниженным риском) к заболеванию или его признакам является индивидом, у которого идентифицирован по меньшей мере один специфический аллель одного или более полиморфного маркера или гаплотипа, обуславливающий пониженную предрасположенность к заболеванию или его признакам. Аллели маркеров и/или гаплотипы, обуславливающие пониженный риск, также называются протективными. В одном аспекте протективный маркер или гаплотип - это тот, который обуславливает достоверно пониженный риск (или подверженность) заболевания или его признаков. В одном воплощении достоверно понижен-

ный риск выражается как относительный риск менее, чем 0,9, включая, но не ограничиваясь значениями менее чем 0,9, менее чем 0,8, менее чем 0,7, менее чем 0,6, менее чем 0,5, менее чем 0,4, менее чем 0,3, менее чем 0,2 и менее чем 0,1. В одном конкретном воплощении достоверно пониженный риск составляет менее 0,7. В другом воплощении достоверно пониженный риск составляет менее 0,5. В еще одном воплощении достоверно пониженный риск составляет менее 0,3. В другом воплощении снижением риска (или подверженности) является снижение по меньшей мере на 25%, по меньшей мере на 30%, по меньшей мере на 35%, по меньшей мере на 40%, по меньшей мере на 45%, по меньшей мере на 50%, по меньшей мере на 55%, по меньшей мере на 60%, по меньшей мере на 65%, по меньшей мере на 70%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по меньшей мере на 95% и по меньшей мере на 98%. В одном конкретном воплощении достоверным снижением риска является снижение по меньшей мере около 30%. В другом воплощении достоверным является снижение риска по меньшей мере около 50%. В другом воплощении снижением риска является снижение по меньшей мере около 70%. Другие пороговые значения или диапазоны, которые специалисты сочтут подходящими для характеристики изобретения, также находятся в рамках данного изобретения.

Специалистам в данной области известно, что для маркеров, присутствующих в исследуемой популяции в двух аллелях, при обнаружении одного аллеля в популяции в группе индивидов с признаками заболевания или заболеванием с повышенной частотой по сравнению с контролем, другой аллель или маркер будет обнаруживаться в группе индивидов с признаками заболевания или заболеванием с пониженной частотой по сравнению с контролем. В таком случае один аллель маркера (тот, что обнаруживается с повышенной частотой среди индивидов с признаками заболевания или заболеванием) будет аллелем, связанным с риском, а другой аллель будет протективным аллелем.

Генетический вариант, ассоциированный с заболеванием или признаком заболевания (например, раком молочной железы), может применяться самостоятельно для прогнозирования риска заболевания по данному фенотипу. Для биаллельных маркеров, таких как ОНП, существует 3 возможных генотипа: гомозиготный по связанному с риском варианту, гетерозиготный и не несущий варианта, связанного с риском. Риск, ассоциируемый с вариантами по множеству локусов, может применяться для оценки совокупного риска. При множестве вариантов ОНП возможное число генотипов k определяется формулой $k = 3^n \times 2^p$; где n является числом аутосомных локусов, а p является числом гоносомных локусов (на половой хромосоме). При оценке совокупного риска, как правило, принимается, что относительные риски по разным генетическим вариантам мультиплицируются, т.е. совокупный риск (например, ОР или ОШ), ассоциированный с определенной комбинацией генотипа, является производным от значений риска для генотипа по каждому локусу. Если представленный риск является относительным риском для одного лица или для определенного генотипа лица, в сравнении с референсной популяцией, сопоставимой по полу и возрасту, то общий риск является производным значений риска, характерных для определенных локусов, и соответствует оценке совокупного риска в сравнении с популяцией. Если риск для одного лица основан на сравнении с лицами, не являющимися носителями связанного с риском аллеля, то общий риск соответствует оценке, которая сравнивает это лицо с заданной комбинацией генотипов по всем локусам с группой индивидов, которые не имеют связанных с риском вариантов по какому-либо из этих локусов. Группа лиц, не являющихся носителями какого-либо из связанных с риском вариантов, имеет наименьший оцениваемый риск и имеет общий риск равный 1,0 при сравнении с самой собой (т.е. с группой не являющихся носителями), но имеет совокупный риск менее 1,0 при сравнении с популяцией. Следует отметить, что группа лиц, не являющихся носителями, потенциально может быть очень малой, особенно по большому числу локусов, и в этом случае ее значимость, соответственно, небольшая.

Мультипликативная модель представляет собой простую модель, которая, как правило, достаточно хорошо описывает данные по сложным признакам. Отклонения от мультипликативности для часто встречающихся вариантов для распространенных заболеваний описываются редко, а если описываются, то, как правило, являются только предположительными, т.к. для демонстрации статистически достоверных взаимосвязей между локусами требуются очень большие размеры выборок.

Для рассмотрения одного примера примем, что с раком простаты была описана ассоциация всего с 8 вариантами (Gudmundsson J. et al., Nat Genet 39: 631-7 (2007), Gudmundsson J. et al., Nat Genet 39:977-83 (2007); Yeager M. et al, Nat Genet 39: 645-49 (2007), Amundadottir L. et al., Nat Genet 38:652-8 (2006); Haiman C.A. et al., Nat Genet 39:638-44 (2007)). Семь из этих локусов находятся на аутосомах, а оставшийся один локус находится на хромосоме X. Общее число теоретически возможных генотипических комбинаций составляет $3^7 \times 2^1 = 4374$. Некоторые из этих классов генотипов являются очень редкими, но все же возможными и должны учитываться при оценке совокупного риска. Вероятно, что мультипликативная модель, применяемая в случае множества генетических вариантов, будет также валидна при сочетании с вариантами, не являющимися генетическими, при предположении, что генетические варианты не имеют четкой корреляции с факторами "окружающей среды". Другими словами, при мультипликативной модели для оценки общего риска должны оцениваться как генетические, так и негенетические варианты, связанные с риском, при предположении, что между негенетическими и генетическими факторами риска

отсутствует взаимодействие.

Используя аналогичный количественный подход, можно оценивать общий или совокупный риск, ассоциированный с множеством вариантов, ассоциированных с раком молочной железы. В одном из таких воплощений, исследуются маркеры rs4848543, rs13387042 и rs3803662 или суррогатные маркеры, находящиеся в неравновесии по сцеплению с ними, и подсчитывается риск для различных комбинаций генотипов ($3^3 = 27$ возможных комбинаций). В другом воплощении совокупный риск определяется для комбинации генотипов двух маркеров (например, маркеров rs13387042 и rs3803662). В других воплощениях исследуются дополнительные маркеры, известные способностью предрасполагать к раку молочной железы, (например, такие факторы риска с высокой пенетрантностью, как BRCA1, BRCA2, BARD1) в сочетании с одним или более факторами риска, описываемыми здесь.

Неравновесие по сцеплению

Естественный феномен рекомбинации, происходящей в каждой паре хромосом в среднем один раз во время каждого мейоза, представляет один из способов, с помощью которого природой обеспечиваются вариации последовательностей (и вследствие этого, биологической функции). Обнаружено, что рекомбинация в геноме происходит не случайным образом, а скорее существуют большие вариации в частотах рекомбинации, что обуславливает существование небольших участков с высокой частотой рекомбинации (также называемых горячими точками рекомбинации) и больших участков с низкой частотой рекомбинации, обычно обозначаемыми блоками неравновесия по сцеплению (Myers S. et al., *Biochem Soc. Trans* 34: 526-530 (2006); Jeffreys A.J. et al., *Nature Genet* 29: 217-222 (2001); May C.A. et al., *Nature Genet* 31: 272-275(2002)).

Неравновесие по сцеплению (НС) относится к неслучайному распределению двух генетических элементов. Например, если конкретный генетический элемент (например, "аллели" полиморфного маркера) встречается в популяции с частотой 0,50 (50%) и другой элемент встречается с частотой 0,50 (50%), то расчетная встречаемость обоих элементов у человека составит 0,25 (25%), учитывая случайное распределение элементов. Однако, если обнаружится, что оба элемента встречаются вместе с частотой выше, чем 0,25, то говорят, что элементы находятся в неравновесии по сцеплению, т.к. они имеют тенденцию наследоваться совместно с более высокой частотой, чем можно ожидать при независимости частот их встречаемости. Грубо говоря, НС обычно коррелирует с частотой рекомбинации между двумя элементами. Частота аллеля или гаплотипа в популяции может быть определена путем генотипирования индивидов в популяции и определения встречаемости в популяции каждого аллеля или гаплотипа. В популяциях диплоидных организмов, например, человеческой популяции, индивиды обычно имеют по два аллеля для каждого генетического элемента (например, маркера, гаплотипа или гена).

Для оценки силы неравновесия по сцеплению (НС) были предложены различные показатели. Большинство из них оценивают силу ассоциации между парами диаллельных сайтов. Двумя важными попарными показателями НС являются r^2 (иногда изображаемый как Δ^2) и $|D'|$. Оба показателя изменяются от 0 (отсутствие неравновесия) до 1 ("полное" неравновесие), но их интерпретация несколько различается. $|D'|$ определяется таким образом, что он равен 1, только если существуют два или три из возможных гаплотипов и он <1 если существуют все четыре возможных гаплотипа. Таким образом, значение $|D'| <1$ указывает на то, что в ходе эволюции между двумя сайтами могла произойти рекомбинация (повторная мутация также может обуславливать значения $|D'| <1$, но для однонуклеотидных полиморфизмов (ОМП) это обычно считается менее вероятным, чем рекомбинация). Показатель r^2 отражает статистическую корреляцию между двумя сайтами и принимает значение 1, только если существуют два гаплотипа.

Показатель r^2 , возможно, является наиболее подходящим показателем картирования сцепления, поскольку существует простое обратное отношение между r^2 и размером выборки, необходимым для обнаружения ассоциации между подверженностью к заболеванию и ОМП. Эти показатели определяются для двух сайтов, но в некоторых случаях может требоваться определение того, насколько сильным является НС по всему участку, содержащему множество полиморфных сайтов (например, определение достоверно ли отличается сила НС между локусами или между популяциями, или является ли НС в участке более сильным или слабым по сравнению с предсказанным по определенной модели). Измерение НС в участке не является простым, одним из подходов является применение разработанного в популяционной генетике показателя r . Грубо говоря, r показывает, какой должна быть рекомбинация в конкретной популяционной модели для возникновения НС, наблюдаемого в результатах. Данный способ дает возможность использования статистического подхода для оценки того, что данные НС демонстрируют свидетельства существования горячих точек рекомбинации. Для описанных здесь способов значимой величиной r^2 может быть по меньшей мере 0,1, например, по меньшей мере 0,1, 0,15, 0,2, 0,25, 0,3, 0,35, 0,4, 0,45, 0,5, 0,55, 0,6, 0,65, 0,7, 0,75, 0,8, 0,85, 0,9, 0,91, 0,92, 0,93, 0,94, 0,95, 0,96, 0,97, 0,98 или по меньшей мере 0,99. В одном предпочтительном воплощении значимой величиной r^2 может быть по меньшей мере 0,2. Кроме того, неравновесие по сцеплению в данном описании означает неравновесие по сцеплению, характеризующее значениями $|D'|$ по меньшей мере 0,2, например, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,85, 0,9, 0,95, 0,96, 0,97, 0,98 или по меньшей мере 0,99. Таким образом, неравновесие по сцеплению отражает корреляцию между аллелями отдельных маркеров. Оно измеряется коэффициентом корреляции или $|D'|$ (r^2 до 1,0 и $|D'|$ до

1,0). В определенных воплощениях неравновесие по сцеплению определяется в терминах значений обоих показателей, как r^2 так и $|D'|$. В одном из таких воплощений достоверное неравновесие по сцеплению определяется значениями $r^2 > 0,1$ и $|D'| > 0,8$. В другом воплощении достоверное неравновесие по сцеплению определяется значениями $r^2 > 0,2$ и $|D'| > 0,9$. Другие сочетания и перестановки значений r^2 и $|D'|$ для определения неравновесия по сцеплению также рассматриваются и находятся в рамках изобретения. Неравновесие по сцеплению может быть определено в одной человеческой популяции, согласно данному описанию, или оно может быть определено в нескольких выборках, составленных из индивидов, происходящих более чем из одной человеческой популяции. В одном воплощении изобретения НС определяется в выборке из одной или более популяций HarMap (европеоидной, африканской, японской, китайской), согласно описанию (<http://www.harmap.org>). В одном из таких воплощений НС определяется в популяции CEU выборки HarMap. В другом воплощении НС определяется в популяции YRI. В еще одном воплощении НС определяется в выборках исландской популяции.

Если бы все полиморфизмы в геноме были идентичными на уровне популяции, то в ассоциативных исследованиях необходимо было бы изучать каждый из них. Однако, благодаря существованию неравновесия по сцеплению между полиморфизмами, между тесно связанными полиморфизмами наблюдается сильная корреляция, что сокращает число полиморфизмов, которые требуется изучить в ассоциативном исследовании для обнаружения достоверной ассоциации. Другим следствием существования НС является то, что между многими полиморфизмами прослеживается ассоциативный сигнал, поскольку между этими полиморфизмами существует сильная корреляция.

Были разработаны генетические карты НС в геноме, и такие карты НС были предложены для использования в качестве основы для картирования генов, связанных с заболеваниями (Risch N. & Merikangas K, *Science* 273:1516-1517 (1996); Maniatis N. et al., *Proc Natl Acad Sci USA* 99: 2228-2233 (2002); Reich DE et al., *Nature* 411:199-204 (2001)).

На данный момент установлено, что многие части человеческого генома могут быть разбиты на серии отдельных блоков гаплотипов, содержащих несколько общих гаплотипов; в этих блоках данные неравновесия по сцеплению не демонстрируют значительных свидетельств рекомбинации (см., например, Wall, J.D. и Pritchard J.K., *Nature Reviews Genetics* 4:587-597 (2003); Daly M. et al., *Nature Genet.* 29: 229-232 (2001); Gabriel S.B. et al., *Science* 296: 2225-2229 (2002); Patil N. et al., *Science* 294: 1719-1723 (2001); Dawson E. et al., *Nature* 478: 544-548 (2002); Phillips M.S. et al., *Nature Genet.* 33:382-387 (2003)). Существует два основных способа определения таких блоков гаплотипов: блоки могут быть определены как участки ДНК, имеющие ограниченное разнообразие по гаплотипам (см., например, Daly, M. et al., *Nature Genet.* 29:229-232 (2001); Patil N. et al., *Science* 294: 1719-1723 (2001); Dawson E. et al., *Nature* 478: 544-548 (2002); Zhang K. et al., *Proc. Natl. Acad. Sci. USA* 99:7335-7339 (2002)), или как участки между зонами с интенсивным темпом обмена, демонстрирующими значительную рекомбинацию в истории человечества, на что указывает неравновесие по сцеплению (см., например, Gabriel S.B. et al., *Science* 296: 2225-2229 (2002); Phillips M.S. et al., *Nature Genet.* 33:382-387 (2003); Wang N. et al., *Am. J. Hum. Genet.* 77: 1227-1234 (2002); Stumpf M.P. и Goldstein D.B., *Curr. Biol.* 73:1-8 (2003)). Недавно была разработана подробная карта частот рекомбинации и соответствующих горячих точек в человеческом геноме (Myers S. et al., *Science* 310:321-32324 (2005); Myers S. et al., *Biochem Soc Trans* 34:526530 (2006)). Карта отражает огромную вариацию по рекомбинации в геноме, с частотами рекомбинации, достигающими до 10-60 сМ/Мб в горячих точках и приближающимися к 0 в участках между ними, что таким образом отражает участки ограниченного разнообразия по гаплотипам и высокое НС. Следовательно, карта может быть использована для определения блоков гаплотипов/блоков НС как участков, ограниченных горячими точками рекомбинации. В данном описании термин "блок гаплотипа" или "блок неравновесия по сцеплению" включает блоки, определенные согласно любым из вышеописанных характеристик или другими альтернативными способами, применяемыми специалистами для определения таких участков.

Некоторые репрезентативные методы для идентификации блоков гаплотипов приведены, например, в опубликованных патентных заявках U.S. №№ 20030099964, 20030170665, 20040023237 и 20040146870. Блоки гаплотипов могут использоваться для картирования ассоциации между фенотипом и генотипическим статусом, с применением одиночных маркеров или гаплотипов, включающих множество маркеров. В каждом блоке гаплотипов могут быть выделены основные гаплотипы, а затем может быть выбран ряд ОНП или маркеров, выполняющих роль "ярлычков" (наименьший набор ОНП или маркеров, необходимый для идентификации гаплотипов). Такие ОНП или маркеры - "ярлычки" могут затем применяться для оценки образцов, полученных у групп индивидов, для выявления ассоциации между фенотипом и гаплотипом. При желании соседствующие блоки гаплотипов могут оцениваться параллельно, поскольку между блоками гаплотипов может также существовать неравновесие по сцеплению.

Известно, что неравновесие по сцеплению может варьировать в различных популяциях из-за различных степеней расхождения генов в ходе эволюции в различных популяциях. В некоторых воплощениях данного изобретения неравновесие по сцеплению относится к неравновесию по сцеплению, определенному в европеоидных популяциях. В конкретном воплощении неравновесие по сцеплению определяется в европеоидных выборках CEPH консорциума HarMap (согласно описанию на

<http://www.hapmap.org>). В других воплощениях НС определяется в африканской популяции, африкано-американской популяции, латиноамериканской популяции, японской популяции, китайской популяции. В некоторых воплощениях НС определяется в выборках HarMap из Китая, Японии, Африки, согласно описанию (<http://www.hapmap.org>).

Таким образом, очевидно, что при наличии какой-либо ассоциации с полиморфным маркером генома вероятно, что ассоциация обнаружится и между дополнительными маркерами генома. Это является естественным следствием неравного распределения НС по геному, что выражается в виде большой вариации частот рекомбинации. Маркеры, использованные для выявления ассоциации, таким образом являются "метками" для участка генома (т.е. блока гаплотипа или блока НС), для которого обнаружена ассоциация с заданным заболеванием или признаком заболевания. Внутри участка, имеющего ассоциацию с заболеванием или признаком заболевания, может существовать один или более причинных (функциональных) вариантов или мутаций. Такие варианты могут обуславливать более высокий относительный риск (ОР) или отношение шансов (ОШ) по сравнению с маркерами - "ярлычками", использованными для обнаружения ассоциации. Данное изобретение таким образом касается маркеров, применяемых для обнаружения ассоциации с заболеванием, по данному описанию, а также маркеров в неравновесии по сцеплению с маркерами. Так, в определенных воплощениях изобретения маркеры, находящиеся в НС с маркерами и/или гаплотипами по изобретению, согласно данному описанию могут использоваться как суррогатные маркеры. Суррогатные маркеры в одном воплощении имеют значения относительного риска (ОР) и/или отношения шансов (ОШ) меньше, чем для маркеров или гаплотипов, для которых исходно была обнаружена ассоциативная связь с заболеванием, по данному описанию. В других воплощениях суррогатные маркеры имеют показатели ОР или ОШ выше тех, что были исходно обнаружены для маркеров, для которых исходно была обнаружена ассоциативная связь с заболеванием, по данному описанию. Примером такого воплощения может быть редкий или относительно редкий (например, с частотой аллеля в популяции < 10%) вариант в НС с более распространенным вариантом (частота встречаемости в популяции > 10%), для которого исходно была обнаружена ассоциативная связь с заболеванием, например, с описанными здесь вариантами. Идентификация и использование таких маркеров для выявления ассоциации, обнаруженной авторами изобретения, согласно данному описанию, может выполняться с помощью рутинных методов, хорошо известных специалистам, и, таким образом, находится в рамках данного изобретения.

Определение частоты гаплотипа

Частоты гаплотипов в контрольной группе и группе пациентов могут быть оценены с применением алгоритма максимизации ожидания (Dempster A. et al., J. R. Stat. Soc. B, 39:1-38 (1977)). Применение данного алгоритма позволяет работать при отсутствующих генотипах и неопределенности гаплотипа. При нулевой гипотезе принимается, что пациенты и контроли имеют одинаковые частоты. С помощью метода правдоподобия проверяется альтернативная гипотеза, при которой допускается, что связанный с риском кандидатный гаплотип, который может включать описанные здесь маркеры, может иметь более высокую частоту у пациентов по сравнению с контролями, в то время как соотношения частот других гаплотипов в обеих группах считаются одинаковыми. Максимизация правдоподобия проводится раздельно для обеих гипотез, и статистическая достоверность оценивается с помощью соответствующей статистики отношения правдоподобия с 1 степенью свободы.

Для рассмотрения связанных с риском и протективных маркеров и гаплотипов в составе региона сцепления, например, исследуется ассоциация всех возможных комбинаций генотипированных маркеров, при условии, что эти маркеры охватывают настоящий регион. Объединенные группы пациентов и контролей могут быть случайным образом поделены на две группы, равные по размеру исходным группам пациентов и контролей. Затем повторяется анализ маркеров и гаплотипов и определяется наиболее статистически достоверное значение Р из всех определенных. Такая процедура рандомизации может быть повторена, например, более 100 раз для получения эмпирического распределения значений Р. В предпочтительном воплощении значение $P < 0,05$ указывает на достоверную ассоциацию маркера и/или гаплотипа.

Анализ гаплотипов

Один из общих подходов к анализу гаплотипов включает их выведение, основанное на правдоподобии, которое применяется в NEstED MOdels (Gretarsdottir S. et al., Nat. Genet. 35:131-38 (2003)). Данный способ используется в программе NEMO, позволяющей анализировать множество полиморфных маркеров, ОНП и микросателлитов. Способ и программное обеспечение специально разработаны для исследований случай-контроль, целью которых является идентификация групп гаплотипов, которые обуславливают различный риск. Он также является инструментом для изучения структур НС. В NEMO оценки максимального правдоподобия, отношения правдоподобия и значения Р подсчитываются напрямую при обработке имеющихся данных с помощью алгоритма максимизации правдоподобия (от англ. Expectation Maximization, EM), позволяющего решить проблему недостающих данных.

Несмотря на то, что тест отношения правдоподобия, основанный на правдоподобии, подсчитанном напрямую для наблюдаемых данных, зафиксировавших потерю информации из-за неопределенности

гаплотипов и пропусках генотипов, достаточно надежен для получения достоверных значений P , все же представляет интерес узнать, какое количество информации было потеряно из-за неполноты информации. Измерение количества информации для анализа гаплотипа описано у Nicolae и Kong (Technical Report 537, Department of Statistics, University of Statistics, University of Chicago; Biometrics, 60(2): 368-75 (2004)) как естественное обобщение определения мер количества информации, используемого при анализе сцепления, и применяется в NEMO.

В случае ассоциации с заболеванием одного маркера для расчета двусторонних значений P для каждого отдельного аллеля может применяться точный тест Фишера. Обычно, при отсутствии специальных оговорок, все значения P представлены без коррекции на множественность сравнений. Представленные частоты (для микросателлитов, ОНП и гаплотипов) являются частотами аллелей, а не частотами носителей. Для уменьшения возможной систематической ошибки, обусловленной родством пациентов, рекрутированных для анализа сцепления целыми семьями, из списка пациентов могут быть исключены родственники первой и второй степени. Кроме этого, тест может быть повторен для корректирования ассоциации, обусловленной возможно оставшимися родственными связями между пациентами, путем применения процедуры уточнения дисперсии, описанной Risen, N. & Teng, J. (Genome Res., 8:1273-1288 (1998)) при пулировании ДНК (там же) родных братьев и сестер, для родственных связей вообще и получения как скорректированных, так и нескорректированных значений P для сравнения. Как и предполагается, разница обычно бывает очень малой. Для оценки значимости ассоциации одного маркера, скорректированной на большое количество тестов, можно провести рандомизирующий тест с использованием тех же генотипических данных. Когорты пациентов и контролей могут быть рандомизированы и ассоциативный анализ проведен заново множество раз (например, до 500000 раз), тогда значение P будет долей реплик, давших значение P для некоторого аллеля маркера, меньшее или равное значению P , наблюдавшемуся в исходных когортах пациентов и контролей.

Как для анализа одиночных маркеров, так и гаплотипов можно подсчитать относительный риск и (ОР) популяционный атрибутивный риск (ПАР) с применением мультипликативной модели (модели относительного риска гаплотипов) (Terwilliger J.D. & Ott J., Hum. Hered. 42:337-46 (1992) и Falk C.T. & Rubinstein P., Ann. Hum. Genet. 51 (Pt 3): 227-33 (1987)), т.е. о том, что при носительстве двух аллелей/гаплотипов, риск мультиплицируется. Например, если риск "А" по сравнению с "а" равен RR , то риск у гомозиготного "АА" человека будет в RR раз больше по сравнению с гетерозиготой "Аа" и в RR^2 раз больше по сравнению с гомозиготой "аа". Мультипликативная модель имеет допущение, которое упрощает анализ и подсчеты - о том, что гаплотипы независимы, т.е. соответствуют равновесию Харди-Вайнберга как в популяции "пораженных", так и в контрольной популяции. В результате, частоты гаплотипов "пораженных" и контролей имеют мультиномиальное распределение, но при альтернативных гипотезах частоты гаплотипов различаются. В частности, для двух гаплотипов, h_i и h_j , $\text{риск}(h_i)/\text{риск}(h_j) = (f_i/p_i)/(f_j/p_j)$, где f и p обозначают, соответственно частоты в популяции "пораженных" и в популяции "контролей". Хотя в случае, если истинная модель не является мультипликативной, происходит некоторое снижение мощности, это снижение оказывается умеренным, за исключением некоторых крайних случаев. Наиболее важно то, что значения P всегда валидны, поскольку они подсчитаны для нулевой гипотезы.

Оценка неравновесия по сцеплению с применением NEMO

НС между парами маркеров может быть рассчитано с применением стандартного определения D' и r^2 (Lewontin R., Genetics 49:49-67 (1964); Hill W.G. & Robertson, A. Theor. Appl. Genet. 22:226-231 (1968)). С применением NEMO частоты комбинаций аллелей двух маркеров оцениваются по максимальному правдоподобию, а отклонение от равновесия по сцеплению оценивается по тесту отношения правдоподобия. Определения D' и r^2 обобщены для включения микросателлитов путем усреднения значений всех возможных комбинаций аллелей двух маркеров, взвешенных предельными вероятностями аллелей. При изображении всех комбинаций маркеров для выяснения структуры НС в определенном участке, мы откладываем D' в верхнем левом углу, а значения P в правом нижнем углу. На картах НС при желании маркеры могут быть нанесены равноудалено, а не исходя из их физического расположения.

Оценка риска и диагностика

В составе любой заданной популяции существует абсолютный риск развития заболевания или признака, определяемый как вероятность развития у лица определенного заболевания или признака за определенный промежуток времени. Например, для женщины абсолютный пожизненный риск развития рака молочной железы составляет 1 из 9. То есть, у одной женщины из 9 в какой-то момент их жизни разовьется рак молочной железы. Риск обычно определяется при рассмотрении очень большого количества людей, а не отдельных индивидов. Риск чаще всего выражается в терминах Абсолютного Риска (АР) и Относительного Риска (ОР). Относительный риск используется для сравнения рисков, ассоциированных с двумя вариантами, или рисков по двум различным группам людей. Например, он может применяться для сравнения группы людей с конкретным генотипом и другой группы людей, имеющих отличный генотип. В случае заболевания относительный риск, равный 2, означает, что в одной группе вероятность развития заболевания в два раза выше, чем в другой группе. Представленный риск, как правило, является относительным риском для лица, или для определенного генотипа лица, в сравнении с популяцией, со-

поставимой по полу и возрасту. Риск для двух индивидов одного пола и возраста можно сравнить простым способом. Например, если при сравнении с популяцией первый индивид имеет относительный риск 1,5, а второй индивид имеет относительный риск 0,5, то риск первого индивида по сравнению со вторым индивидом составляет $1,5/0,5 = 3$.

Согласно данному описанию, определенные полиморфные маркеры и гаплотипы, включающие такие маркеры, могут применяться для оценки риска рака молочной железы. Оценка риска может включать применение маркеров для диагностики подверженности раку молочной железы. Определенные аллели полиморфных маркеров обнаруживаются у индивидов с раком молочной железы чаще, чем у индивидов без диагностированного рака молочной железы. Следовательно, эти аллели маркеров имеют прогностическую значимость для выявления у индивида рака молочной железы или подверженности раку молочной железы. Маркеры - "ярлычки" в составе блоков гаплотипов или блоков НС, включающие связанные с риском маркеры, такие как маркеры по данному изобретению, могут применяться в качестве суррогатов для других маркеров и/или гаплотипов в составе блока гаплотипов или блока НС. Маркеры со значениями χ^2 равными 1 являются идеальными суррогатами для связанных с риском вариантов, т.е. генотипы одного маркера позволяют идеально спрогнозировать генотипы другого. Маркеры со значениями χ^2 менее 1 также могут служить суррогатами для связанных с риском вариантов, или, кроме того, представлять варианты с такими же высокими значениями относительного риска или возможно даже более высокими, чем у связанного с риском варианта. Идентифицированный связанный с риском вариант сам по себе может не быть функциональным вариантом, но в этом случае может находиться в неравновесии по сцеплению с истинно функциональным вариантом. Данное изобретение включает оценку таких суррогатных маркеров для маркеров по данному описанию. Такие маркеры описаны, картированы и приведены в общедоступных базах данных, известных специалистам, или могут быть легко идентифицированы путем секвенирования участка или его части, идентифицированного маркерами по данному изобретению у группы индивидов, и идентифицировать полиморфизмы в полученной группе последовательностей. В результате, специалисты могут легко и без чрезмерных усилий генотипировать суррогатные маркеры, находящиеся в неравновесии по сцеплению с маркерами и/или гаплотипами по данному описанию. Маркеры - "ярлычки", или суррогатные маркеры, находящиеся в НС с обнаруженными связанными с риском вариантами в составе гаплотипа или блока НС, также имеют прогностическое значение для выявления ассоциации с раком молочной железы или подверженности индивида раку молочной железы. Такие маркеры - "ярлычки", или суррогатные маркеры, находящиеся в НС с маркерами по данному изобретению, могут также включать и другие маркеры, которые характеризуют гаплотипы, поскольку они также имеют прогностическое значение для выявления подверженности раку молочной железы.

Данное изобретение в некоторых воплощениях может осуществляться посредством оценки образца, содержащего геномную ДНК индивида, на наличие вариантов, согласно данному описанию ассоциированных с раком молочной железы. Такая оценка включает этапы определения наличия или отсутствия по меньшей мере одного аллеля по меньшей мере одного маркера, с помощью способов, хорошо известных специалистам и описываемых здесь, и определения на основе такой оценки существования повышенного или пониженного риска (повышенной или пониженной подверженности) раку молочной железы у индивида, от которого получен образец. Кроме того, изобретение может осуществляться с применением массива данных, содержащего информацию о генотипическом статусе по меньшей мере одного полиморфного маркера, согласно данному описанию ассоциированного с раком молочной железы (или маркерами, находящимися в неравновесии по сцеплению с по меньшей мере одним маркером, который, как здесь показано, ассоциирован с раком молочной железы). Другим словами, массив данных, содержащий информацию о таком генетическом статусе, например, в форме частот генотипа по определенному полиморфному маркеру или множеству маркеров (например, указания на наличие или отсутствие определенных связанных с риском аллелей), или реальные генотипы по одному маркеру или более, могут быть исследованы на наличие или отсутствие определенных связанных с риском аллелей определенных полиморфных маркеров, для которых авторами изобретения продемонстрирована ассоциация с раком молочной железы. Положительный результат по варианту (например, аллелю маркера), ассоциированному с раком молочной железы согласно данному описанию, указывает на то, что индивид, для которого получен массив данных, имеет повышенную подверженность (повышенный риск) раку молочной железы.

В определенных воплощениях изобретения корреляция полиморфного маркера с раком молочной железы выявляется при сравнении данных генотипа полиморфного маркера с переводной таблицей, показывающей корреляцию между по меньшей мере одним аллелем полиморфизма и раком молочной железы. В некоторых воплощениях таблица показывает корреляцию по одному полиморфизму. В других воплощениях таблица показывает корреляцию по множеству полиморфизмов. В обоих случаях, по ссылкам на переводную таблицу, указывающую на наличие корреляции между маркером и раком молочной железы, для индивида, у которого получен образец, можно определить риск развития рака молочной железы или подверженность раку молочной железы. В некоторых воплощениях корреляция отражена в виде статистического показателя.

Статистический показатель может быть выражен в виде показателя риска, например, относительного риска (ОР), абсолютного риска (АР) или отношения шансов (ОШ).

Маркеры и гаплотипы по изобретению, например, полиморфные маркеры и гаплотипы на хромосоме 2q14.2, хромосоме 2q35 и хромосоме 16q12, например, маркеры, представленные в табл.10, 15, 19, 20, 21 и 22 и маркеры в неравновесии по сцеплению с ними, например маркеры rs4848543, rs13387042 и rs3803662 могут применяться для оценки риска и для диагностических целей, как по-отдельности, так и в комбинации. Так, даже в случаях, когда увеличение риска по отдельным маркерам относительно умеренное, т.е. порядка 10-30%, ассоциация может иметь существенное значение. Так, относительно распространенные варианты могут вносить существенный вклад в совокупный риск (иметь высокий популяционный атрибутивный риск), либо комбинация маркеров может применяться для выделения групп индивидов, которые в случае комбинации риска по маркерам, будут иметь значительный суммарный риск развития заболевания.

Так, в одном воплощении изобретения для оценки совокупного риска применяется несколько вариантов (маркеров и/или гаплотипов). В одном воплощении эти варианты выбирают из вариантов по данному описанию. Другие воплощения включают применение вариантов по данному изобретению в комбинации с другими вариантами, применяющимися в диагностике подверженности раку молочной железы. В таких воплощениях у индивида определяется генотипический статус нескольких маркеров и/или гаплотипов и статус индивида сравнивается с популяционной частотой ассоциированных вариантов или с частотой вариантов у клинически здоровых лиц, например, лиц сопоставимых по полу и возрасту. Известные способы, например, многофакторный анализ или анализ общего риска, могут применяться для определения совокупного риска, обусловленного генотипическим статусом по множественным локусам. Оценка риска на основании такого анализа может впоследствии использоваться в способах и наборах по изобретению, согласно данному описанию.

Согласно вышеописанному структура блока гаплотипа человеческого генома имеет особенность, что большое число вариантов (маркеров и/или гаплотипов) в неравновесии по сцеплению с вариантом, для которого исходно была отмечена ассоциация с заболеванием или признаком заболевания, может использоваться в качестве суррогатных маркеров для оценки ассоциации с заболеванием или признаком заболевания. Количество таких суррогатных маркеров будет зависеть от таких факторов, как интенсивность рекомбинации в участке в ходе истории человечества, частота мутаций в участке (т.е. число полиморфных сайтов или маркеров в участке) и степень НС (размером блока НС) в участке. Эти маркеры обычно располагаются внутри физических границ изучаемого блока НС или блока гаплотипов, что определяется с помощью описываемых здесь способов или с помощью других способов, известных специалистам. Однако иногда ассоциация маркеров и гаплотипов оказывается выходящей за пределы физических границ блока гаплотипов согласно определению. Такие маркеры и/или гаплотипы могут в таких случаях использоваться также в качестве суррогатных маркеров и/или гаплотипов для маркеров и/или гаплотипов, физически находящихся в составе блока гаплотипов согласно определению. Вследствие этого маркеры и гаплотипы в НС (обычно характеризующиеся r^2 более 0,1, например, r^2 более 0,2, включая r^2 более 0,3, а также включая r^2 более 0,4) с маркерами и гаплотипами по данному изобретению, также находятся в рамках данного изобретения, даже если они физически располагаются вне границ блока гаплотипов согласно определению. Сюда входят описанные здесь маркеры (например, в табл. 10, 15 и 19), но могут также входить и другие маркеры, находящиеся в сильном НС (характеризующемся r^2 более 0,1 или r^2 более 0,2 и/или $|D'| > 0,8$) с одним маркером или более, приведенными в табл. 10, 15 и 19.

В случае описываемых здесь маркеров ОНП, аллель противоположный аллелю, частота которого превышена у пациентов (связанному с риском аллелю), обнаруживается с пониженной частотой у пациентов с раком молочной железы. Следовательно, такие маркеры и гаплотипы в НС и/или включающие такие маркеры являются протективными в отношении рака молочной железы, т.е. они обуславливают пониженный риск или подверженность индивидов, несущих такие маркеры и/или гаплотипы, к развитию рака молочной железы.

Определенные варианты по данному изобретению, в том числе определенные гаплотипы, включают, в некоторых случаях комбинации нескольких генетических маркеров, например, ОНП и микросателлиты. Таким образом, детектирование гаплотипов может проводиться способами определения последовательностей полиморфных сайтов, известными в данной области техники, и/или описываемыми здесь. Кроме того, корреляция между определенными гаплотипами или группами маркеров и фенотипом заболевания может быть верифицирована с помощью стандартных методик. Репрезентативным примером простого теста на корреляцию является точный тест Фишера в таблице два на два.

В определенных воплощениях аллель маркера или гаплотип, ассоциированный с раком молочной железы (например, аллели, маркеров, приведенные в табл. 10, 15, 19, 20, 21 и 22 и маркеры в неравновесии по сцеплению с ними), представляет собой аллель маркера или гаплотип, встречающийся у индивидов с риском развития рака молочной железы ("пораженных") чаще по сравнению с частотой его встречаемости у здоровых индивидов (контроль), и наличие аллеля маркера или гаплотипа указывает на рак молочной железы или подверженность раку молочной железы. В других воплощениях связанные с риском маркеры, находящиеся в неравновесии по сцеплению с одним или более маркерами, ассоциированными с раком молочной железы (например, маркеры, приведенные в табл. 10, 15, 19, 20, 21 и 22 и маркеры в неравновесии по сцеплению с ними), являются маркерами-ярлычками, встречающимися у индиви-

дов с риском развития рака молочной железы ("пораженных") чаще по сравнению с частотой их встречаемости у здоровых индивидов (контроль), и наличие маркеров-ярлычков указывает на повышенную подверженность раку молочной железы. В следующем воплощении связанные с риском аллели маркеров (т.е. обуславливающие повышенную подверженность), находящиеся в неравновесии по сцеплению с одним или более маркерами, ассоциированными с раком молочной железы (например, аллели маркеров, приведенные в табл. 10, 15, 19, 20, 21 и 22 и маркеры в неравновесии по сцеплению с ними), являются маркерами, включающими один или более аллелей, которые встречаются у индивидов с риском развития рака молочной железы чаще по сравнению с частотой их встречаемости у здоровых индивидов (контроль), и наличие маркеров указывает на повышенную подверженность раку молочной железы.

Исследуемая популяция

Исландская популяция представляет собой европеоидную популяцию северо-европейского происхождения. За последние годы было опубликовано большое число исследований, представляющих результаты генетического сцепления и ассоциации в Исландской популяции. Многие исследования демонстрируют воспроизведение вариантов, первоначально идентифицированных на исландской популяции как ассоциированных с определенным заболеванием, и в других популяциях (Stacey S.N. et al., *Nat Genet.* May 27 2007 (Epub ahead of print; Helgadóttir A., et al., *Science* 316:1491-93 (2007); Steinthorsdóttir V., et al., *Nat Genet.* 39:770-75 (2007); Gudmundsson J., et al. *Nat Genet.* 39:631-37 (2007); Amundadóttir L.T. et al., *Nat Genet.* 38:652-58 (2006); Grant, S.F., et al., *Nat Genet.* 38:320-23 (2006)). Таким образом, генетические находки в исландской популяции в общем были воспроизведены в других популяциях, включая популяции из Африки и Азии.

Считается, что маркеры по данному изобретению, для которых была обнаружена ассоциация с раком молочной железы, демонстрируют схожую ассоциацию в других человеческих популяциях. Определенные воплощения, включающие отдельные человеческие популяции, также рассматриваются и находятся в рамках изобретения. Такие воплощения касаются человеческих субъектов, происходящих более чем из одной человеческой популяции, включая европеоидные популяции, популяции Европы, популяции Америки, популяции Евразии, популяции Азии, популяции Центральной и Южной Азии, популяции Восточной Азии, популяции Среднего Востока, популяции Африки, популяции Латинской Америки и популяции Океании, но не ограничиваются ими. Популяции Европы включают шведскую, норвежскую, финскую, русскую, датскую, исландскую, ирландскую, кельтскую, английскую, шотландскую, голландскую, бельгийскую, французскую, немецкую, испанскую, португальскую, итальянскую, польскую, болгарскую, славянскую, сербскую, боснийскую, чешскую, греческую и турецкую популяции, но не ограничиваются ими. Кроме того, в других воплощениях изобретение может осуществляться на определенных человеческих популяциях, включая банту, манденка, иоруба, сан, пигмеев мбути, жителей оркнейских островов, адыгов, русских, сардинцев, тосканцев, мозабитов, бедуинов, друзов, палестинцев, белуджей, брагуев, макрани синдхов, пуштунов, буришей, хазарейцев, уйгуров, калашей, хань, дай, дауров, хэцже, лаху, мяо, орохонов, шэ, туцзя, ту, сибо, и, монголов, наси, камбоджийцев, японцев, якутов, меланезийцев, папуасов, каритиана, суруи, колумбийцев, майя и пима.

В одном предпочтительном воплощении изобретение касается популяций, имеющих происхождение от черных африканцев, например, популяций, включающих лиц с африканским происхождением или родословной. Происхождение от черных африканцев может определяться при обозначении себя африкано-американцами, афро-американцами, черными американцами, принадлежности к черной расе или принадлежности к негроидной расе. Например, африкано-американцы или черные американцы - это люди, живущие в Северной Америке и происходящие из любой группы черной расы из Африки. В другом примере лица, сообщившие о происхождении от черных африканцев, могут иметь по меньшей мере одного родителя черного африканского происхождения или по меньшей мере дедушку или бабушку черного африканского происхождения.

Расовый вклад индивидов может также определяться с помощью генетического анализа. Генетический анализ происхождения можно провести с помощью несцепленных микросателлитных маркеров, например, как описано у Smith et al. (*Am J Hum Genet* 74, 1001-13 (2004)).

В определенных воплощениях изобретение относится к маркерам и/или гаплотипам, идентифицированным в конкретных популяциях, согласно описанию выше. Специалистам известно, что показатели неравновесия по сцеплению (НС) могут дать различные результаты на разных популяциях. Это объясняется различиями в популяционной истории у разных человеческих популяций, а также различным давлением отбора, которое могло привести к различиям в НС в определенных участках генома. Специалистам также хорошо известно, что определенные маркеры, например маркеры ОНП, являются полиморфными в одной популяции, но не в другой. Специалисты будут использовать доступные и рекомендуемые здесь способы для применения данного изобретения в любой заданной человеческой популяции. Сюда может входить оценка полиморфных маркеров в участке НС по данному изобретению для идентификации тех маркеров, которые имеют наиболее сильную ассоциацию в определенной популяции. Так, связанные с риском варианты по данному изобретению в различных человеческих популяциях могут встречаться в различных гаплотипах и с различной частотой. Однако, используя способы, известные в данной области техники, и маркеры по данному изобретению, изобретение можно применять в любой данной человеческой популяции.

ской популяции.

Применение генетического тестирования

Специалистам известно и понятно, что описанные здесь варианты, как правило, сами по себе не позволяют абсолютно точно идентифицировать индивидов, у которых разовьется рак молочной железы. Тем не менее, описанные здесь варианты указывают на повышенную и/или пониженную вероятность того, что у индивидов, имеющих связанные с риском или протективные варианты по изобретению, разовьется рак молочной железы. Такая информация чрезвычайно ценна сама по себе, как это будет подробно продемонстрировано ниже, поскольку она может применяться для контролирования заболевания и выбора соответствующих вариантов лечения.

Знание генетических вариантов, обуславливающих риск развития заболевания (например, рака молочной железы), дает возможность применения генетических тестов для дифференцирования между индивидами с повышенным риском развития заболевания (т.е. носителей, связанных с риском вариантов) и тех, у кого риск развития заболевания понижен (т.е. носителей протективных вариантов). Основными достоинствами генетического тестирования индивидов, относящихся к обеим вышеупомянутым группам, является возможность диагностирования заболевания или предрасположенности к заболеванию на ранней стадии и получение клиницистом информации о прогнозе/агрессивности заболевания для возможности применения наиболее подходящего лечения.

Модели, позволяющие прогнозировать риск рака молочной железы, обусловленный наследственностью

Оценка риска рака молочной железы осуществляется для того, чтобы обосновать разработку персонализированного подхода в лечении всех женщин с целью увеличения продолжительности и качества жизни у женщин, имеющих высокий риск, при одновременной минимизации расходов, ненужных вмешательств и беспокойности у женщин, имеющих более низкий риск. В моделях, прогнозирующих риск, производится попытка оценить риск рака молочной железы у индивида, имеющего определенный набор наследственных характеристик риска (например, семейный анамнез, имевшееся ранее доброкачественное поражение молочной железы, имевшаяся ранее опухоль молочной железы). Наиболее часто применяющиеся в клинической практике модели оценки риска рака молочной железы, учитывают наследственные факторы риска, принимая во внимание семейный анамнез. Оценки риска основываются на обнаружении повышенного риска у индивидов, имеющих одного или более близких родственников с ранее диагностированным раком молочной железы. Они не учитывают сложные структуры родословных. Еще один недостаток таких моделей - это невозможность отличить носителей генов с мутациями, предрасполагающими к раку молочной железы, от лиц, не являющихся носителями.

Более сложные модели оценки риска имеют усовершенствованные подходы, позволяющие учитывать определенный семейный анамнез, и способны принимать в расчет статус носительства мутаций BRCA1 и BRCA2. Например, алгоритм Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA) (Antoniou et al., 2004) учитывает семейный анамнез, основанный на индивидуальных структурах родословных, с помощью программы анализа родословных MENDEL. Кроме того, учитывается и информация об известном статусе по BRCA1 и BRCA2. Основными недостатками BOADICEA и других моделей оценки риска рака молочной железы, применяемых в настоящее время, является то, что они не включают информацию по другим генам предрасположенности. Существующие в настоящее время модели в значительной степени используют семейный анамнез в качестве суррогата для того, чтобы компенсировать отсутствие сведений о генетических детерминантах риска, отличных от BRCA. Таким образом, существующие модели привязаны к ситуациям, когда известен семейный анамнез заболевания. Гены, предрасполагающие к раку молочной железы, имеющие более низкую пенетрантность, могут быть довольно распространенными в популяции, но могут не обладать такой способностью вызывать семейную концентрацию, как гены BRCA1 и BRCA2. Пациенты с относительно большим генетическим грузом аллелей предрасположенности могут иметь слабо выраженный семейный анамнез заболевания или не иметь его. Таким образом, существует необходимость создания моделей, которые бы учитывали данные о наследственной предрасположенности, полученные непосредственно при генетическом тестировании. Кроме того, что такой подход позволит сделать модели более точными, он также уменьшит зависимость от параметров семейного анамнеза и поможет определять риск в приложении к более широкой популяции, в которых семейный анамнез не является ключевым фактором.

Внедрение усовершенствованных моделей оценки генетического риска при проведении первичной профилактики рака молочной железы в клинической практике

Существующие в клинике варианты первичной профилактики можно подразделить на проведение превентивной химиотерапии (или гормональной терапии) и превентивного хирургического вмешательства. Пациентам, у которых выявлен высокий риск, могут быть предписаны длительные курсы превентивной химиотерапии. Такой подход широко применяется в области сердечно-сосудистой медицины, но только начинает внедряться в клинической онкологии. Наиболее широко применяемым в онкологии агентом для превентивной химиотерапии является Тамоксифен, селективный модулятор эстрогеновых рецепторов (СМЭР). Используемый изначально в качестве адъювантной терапии, направленной против рецидивирования рака молочной железы, в настоящее время тамоксифен доказал свою эффективность в

качестве превентивного агента при раке молочной железы (Cuzick et al. 2003; Martino et al. 2004). Агентство по контролю за продуктами питания и лекарственными средствами FDA одобрило применение тамоксифена в качестве агента для превентивной химиотерапии у некоторых женщин с высоким риском.

К сожалению, длительное применение тамоксифена повышает риск развития рака эндометрия приблизительно в 2,5 раза, риск венозного тромбоза приблизительно в 2 раза. Риск развития легочной эмболии, инсульта и катаракты также повышаются (Cuzick et al., 2003). Следовательно, благоприятный эффект от применения тамоксифена для снижения заболеваемости раком молочной железы не обуславливает соответствующего снижения общей смертности. Другой селективный модулятор эстрогеновых рецепторов, ралоксифен, может оказаться более эффективным для профилактики, и он не связан с таким риском рака эндометрия. Однако у пациентов, принимающих длительное время ралоксифен, риск тромбоза остается повышенным (Cuzick et al. 2003; Martino et al. 2004). Более того, с приемом тамоксифена и ралоксифена связаны проблемы, касающиеся качества жизни. Для проведения разумного анализа соотношения риска и пользы применения СМЭР в качестве превентивной химиотерапии, в клинике существует необходимость выявления индивидов, имеющих наиболее высокий риск рака молочной железы. Учитывая тот факт, что генетическая составляющая риска рака молочной железы является весьма значительной, в клинической практике существует явная необходимость проведения генетического тестирования для измерения подобного риска у индивидов. Можно предвидеть, что аналогичные проблемы возникнут в отношении любой химиотерапии для предупреждения рака, которая может появиться в будущем, например, в отношении ингибиторов ароматазы. Более того, по мере того как превентивная химиотерапия становится более безопасной, нарастает необходимость выявления пациентов, имеющих генетическую предрасположенность, но не имеющих значительно повышенный риск, ассоциированный с носительством BRCA1 & 2.

Для пациентов, у которых выявлен высокий риск развития рака молочной железы, рассматривается возможность проведения превентивного хирургического вмешательства; двусторонней мастэктомии или овариэктомии, или и того, и другого. Очевидно, что подобные решительные меры рекомендуются только тем пациентам, у которых риск оценивается как чрезвычайно высокий. На практике, в настоящее время такой риск выявляется только у индивидов, имеющих мутации в генах BRCA1, BRCA2 или генах, задействованных в редких синдромах с предрасположенностью к раку молочной железы, например, p53 при синдроме Ли-Фромена, PTEN при синдроме Коудена.

Оценки пенетрантности мутаций BRCA1 и BRCA2 оказываются более высокими, когда их получают для семей, в которых имеется несколько "случаев" рака молочной железы, чем оценки, полученные для популяции. Это объясняется тем, что в различных семьях с носительством мутаций отмечается различная пенетрантность в отношении рака молочной железы (для примера см. Thorlacius et al., 1997). Одним из основных факторов, вносящих вклад в подобную вариацию, является действие до сих пор не известных генов предрасположенности, эффект которых модифицирует пенетрантность мутаций BRCA1 и BRCA2. Следовательно, абсолютный риск для индивида, имеющего мутации в генах BRCA1 и BRCA2, невозможно определить точно при отсутствии сведений о существовании и действии модифицирующих генов. Поскольку варианты лечения для носителей BRCA1 и BRCA2 могут быть агрессивными, важно проводить количественную оценку рисков у отдельных носителей BRCA с максимально возможной точностью. Таким образом, существует необходимость выявления генов предрасположенности, эффект которых изменяет степень пенетрантности рака молочной железы у носителей BRCA1 и BRCA2, и разработки усовершенствованных моделей оценки риска, основанных на использовании этих генов.

Кроме того, существуют индивиды, у которых риск развития рака молочной железы расценивается как очень высокий, возможно из-за наличия выраженного семейного анамнеза рака молочной железы, но у которых не удалось выявить мутаций в известных генах предрасположенности. В таких случаях сложно рассматривать возможность проведения превентивного хирургического вмешательства, поскольку невозможно провести тестирование индивида для выяснения, унаследовала ли она ген предрасположенности, имеющий высокую пенетрантность. Соответственно, невозможно с точностью оценить риск индивида. Таким образом, в клинической практике существует очевидная необходимость идентификации генов предрасположенности с высокой пенетрантностью, которые остаются неизвестными, и разработки связанных с ними генетических тестов для применения в целях первичной профилактики.

Ранняя диагностика

Скрининг рака молочной железы в клинической практике большинства западных стран представляет собой периодическое клиническое обследование молочной железы и маммографию с помощью рентгеновского излучения. Существуют веские доказательства того, что клиническое обследование молочной железы имеет мало дополнительных преимуществ при использовании в сочетании с хорошей программой маммографического скрининга. В Великобритании женщинам в возрасте от 50 до 70 лет предлагается пройти скрининговую маммографию каждые три года. В Соединенных Штатах ситуация различается в зависимости от организаций, обеспечивающих медицинское обслуживание, но, тем не менее, Американское Общество Рака (American Cancer Society) рекомендует проводить ежегодный маммографический скрининг начиная с 40 лет. Доказана эффективность маммографического скрининга в снижении смертности среди прошедших скрининг женщин в возрасте старше 50 лет.

Маловероятно, что генетическое тестирование когда-либо будет применяться в качестве средства, сокращающего доступ к существующим программам маммографического скрининга. Однако маммографический скрининг не лишен недостатков, и подразумевается, что генетическое тестирование должно применяться для отбора людей, нуждающихся в более интенсивных программах скрининга. Одним из недостатков маммографического скрининга является то, что до сих пор не удалось продемонстрировать значимый эффект на повышение выживаемости у женщин, проходящих скрининг в возрасте до 50 лет.

Одной из причин, объясняющей меньшую эффективность маммографии у женщин в возрасте до 50 лет, может быть более высокая плотность тканей молочной железы у молодых женщин, что затрудняет выявление опухолей при маммографии. Тем не менее, рак молочной железы у предрасположенных индивидов имеет тенденцию развиваться в более молодых возрастных группах и существует четкая зависимость между высокой плотностью ткани молочной железы и риском рака молочной железы. Следовательно, существуют затруднения с простой интенсификацией маммографического скрининга у индивидов с высокой предрасположенностью, поскольку к ним будет применяться технология, показывающая не самые оптимальные результаты в группе наиболее высокого риска. Последние исследования показали, что магнитно-резонансная маммография с контрастным усилением (МРМ-КУ) является более чувствительным методом и позволяет детектировать опухоли на более ранних стадиях в данной группе высокого риска по сравнению с маммографическим скринингом (Warner et al., 2004; Leach et al., 2005). Методика МРМ-КУ дает наилучшие результаты в комбинации с обыкновенной маммографией, использующей рентгеновское излучение (Leach et al., 2005). Поскольку для проведения МРМ-КУ требуются специализированные центры, что связано с высокими затратами, скрининг индивидов моложе 50 лет может быть ограничен теми, у кого риск наиболее высокий. В проводимых в настоящее время испытаниях МРМ-КУ набор пациентов ограничен индивидами, имеющими мутации BRCA1, BRCA2 или p53, либо имеющими очень выраженный семейный анамнез заболевания. Применению данной скрининговой тактики к более широкому кругу пациентов, имеющих высокий риск, будет значительно способствовать создание средств для определения профиля риска, основанных на генетической информации.

Существуют убедительные доказательства в пользу точки зрения, что рак молочной железы с ранним началом и рак, возникающий у женщин с генетической предрасположенностью, прогрессирует быстрее, чем у женщин старшего возраста и с меньшей предрасположенностью. Об этом свидетельствует обнаружение более высоких показателей межскринингового рака у молодых женщин, т.е. рак, развивающийся в промежутках между скрининговыми визитами в популяциях, где проводится хороший скрининг, чаще наблюдается у молодых женщин. Вследствие этого предлагается сокращение интервалов скрининга, проводимого любым способом, для молодых женщин. В этой связи возникает парадокс, поскольку оказывается, что наиболее частый скрининг с применением более дорогих методик необходимо проводить в возрастной группе, в которой совокупные показатели рака молочной железы сравнительно низкие. В клинической практике существует очевидная необходимость выявления тех молодых индивидов, у которых имеется наибольшая предрасположенность к раннему развитию заболевания, и направления их в более дорогие и подробные скрининговые программы.

Лечение

В настоящее время в лечении рака молочной железы используются хирургическое вмешательство, адъювантная химиотерапия, радиотерапия, с последующей длительной гормональной терапией. Зачастую применяются комбинации трех или четырех способов лечения.

Пациенты с раком молочной железы, имеющие одинаковые стадии заболевания, могут по-разному отвечать на адъювантную химиотерапию, что приводит к большой вариации общего исхода лечения. Для определения того, подходит ли пациентам с раком молочной железы лечение с помощью адъювантной химиотерапии, были разработаны консенсусные рекомендации (критерии Сен-Галлен и Национального института здоровья). Однако даже наиболее выраженные клинические и гистологические предикторы метастазирования не позволяют с точностью предсказать клинический ответ при наличии опухолей молочной железы (Goldhirsch et al., 1998; Eifel et al., 2001). Химиотерапия или гормональная терапия сокращает риск метастазирования приблизительно только на 1/3, однако 70-80% пациентов, получающих такое лечение, выжили бы и без него. Таким образом, большинство пациентов с раком молочной железы в настоящее время получают лечение, которое является либо неэффективным, либо ненужным. В клинической практике существует очевидная необходимость разработки усовершенствованных прогностических показателей, которые позволили бы клиницистам назначать лечение, предназначенное для тех, кому оно принесет наибольшую пользу. Следует ожидать, что определение генетической предрасположенности у индивидов может выявить информацию, имеющую значение для исхода лечения и, таким образом, способствовать рациональному планированию лечения.

Примерами данной концепции могут служить несколько предыдущих исследований: пациенты с раком молочной железы, являющиеся носителями мутации BRCA, демонстрируют лучшие степени клинического ответа и выживаемость при применении адъювантной химиотерапии [Chappuis et al., (2002), J Med Genet, 39, 608-10; Goffin et al., (2003), Cancer 97, 527-36]. Носители мутации BRCA отвечают на химиотерапевтическое лечение рака яичников препаратами платины лучше чем лица, не являющиеся носителями [Cass et al., (2003), Cancer 97, 2187-95]. Аналогичные соображения могут касаться пациентов,

имеющих предрасположенность, у которых задействованные при этом гены не известны. Например, известно, что при инфильтрующей долевой карциноме молочной железы (ИДК МЖ) существует значимый наследственный компонент, но задействованные в нем генетические варианты до сих пор не идентифицированы. Пациенты с ИДК МЖ хуже отвечают на общепринятые режимы химиотерапии [Mathieu et al., (2004), *Eur J Cancer*, 40, 342-51].

Модели генетической предрасположенности могут не только способствовать выбору стратегий лечения, но могут играть важную роль в разработке таких стратегий. Например, было обнаружено, что клетки опухолей при мутациях BRCA1 и BRCA2 очень чувствительны к ингибиторам поли-(АДФ-рибоза)-полимеразы (от англ. poly (ADP-ribose) polymerase, или PARP) вследствие дефекта их пути репарации ДНК [Farmer, et al., (2005), *Nature*, 434, 917-21]. Это способствовало разработке препаратов малых молекул, направленно воздействующих на PARP с целью их специфического применения у пациентов с носительством BRCA. Из данного примера следует, что информация о генетической предрасположенности может идентифицировать мишени для лекарственных препаратов, что позволит разработать индивидуальные режимы химиотерапии для применения в сочетании с определением профиля генетического риска.

Химиотерапия рака имеет хорошо известные ограничивающие дозировку побочные эффекты на нормальные ткани, особенно на интенсивно пролиферирующие гемопоэтические и эпителиальные клетки отделов кишечника. Можно ожидать, что в чувствительности нормальных тканей к цитотоксическим препаратам существуют индивидуальные различия, обусловленные генетически. Понимание этих факторов может способствовать рациональному планированию лечения и разработке лекарственных препаратов, предназначенных для защиты нормальных тканей от побочных эффектов химиотерапии.

Изучение генетического профиля может также способствовать совершенствованию методов радиотерапии: в составе группы пациентов с раком молочной железы, которым радиотерапия проводится в стандартном режиме, часть пациентов будет иметь побочные эффекты в ответ на дозы радиации, которые обычно переносятся. Острые реакции включают эритему, влажную десквамацию, отек и радиационную пневмонию. Отдаленные реакции, включая телеангиэктазии, отек, фиброз, легочный фиброз и фиброз молочной железы могут развиваться много лет спустя после радиотерапии. Как острые, так и отдаленные реакции поражают значительный процент пациентов и могут вызывать летальный исход. В одном исследовании у 87% пациентов были обнаружены некоторые отрицательные побочные эффекты на радиотерапию, тогда как 11% имели серьезные побочные реакции (LENT/SOMA Grade 3-4); [Hoeller et al., (2003), *Int J Radiat Oncol Biol Phys*, 55, 1013-8]. Вероятность развития неблагоприятной реакции в ответ на радиотерапию обусловлена прежде всего конститутивными индивидуальными различиями в реакциях нормальной ткани и существует подозрение, что они имеют существенный генетический компонент. Ряд известных генов предрасположенности к раку молочной железы (например, BRCA1, BRCA2, ATM) затрагивают пути репарации повреждений двойной цепи ДНК. Повреждения двойной цепи ДНК представляют собой первичные цитотоксические повреждения, вызываемые радиотерапией. Это вызвало беспокойство по тому поводу, что индивиды, генетически предрасположенные к раку молочной железы за счет носительства вариантов генов, принадлежащих этим путям, могут также иметь повышенный риск избыточного повреждения нормальных тканей при радиотерапии.

Существование в популяции индивидов с конститутивной чувствительностью к облучению означает, что для большей части пациентов дозы радиации при радиотерапии должны быть ограничены для поддержания частоты побочных реакций на приемлемом уровне. Таким образом, в клинической практике существует потребность в надежных тестах, которые позволили бы выявлять индивидов, имеющих повышенный риск развития побочных реакций в ответ на радиотерапию. Такие тесты позволили бы назначать консервативное или альтернативное лечение индивидам с чувствительностью к облучению и в то же время увеличивать дозы при радиотерапии большинству пациентов, являющимся относительно устойчивыми к облучению. Согласно расчетам, увеличение дозы, возможное на основании простого разделения пациентов на категории чувствительных к облучению, устойчивых к облучению и имеющих промежуточную чувствительность, позволило бы добиться улучшения локального контроля над опухолями на 35% и в результате повысить показатели выживаемости (Burnet et al., 1996).

Воздействие ионизирующей радиации является известным фактором, способствующим онкогенезу молочной железы (Dumitrescu и Cotarla 2005). Известные гены предрасположенности раку молочной железы кодируют компоненты пути клеточного ответа на повреждение ДНК, индуцированное радиацией (Narod и Foulkes 2004). В этой связи существует обеспокоенность по поводу того, что риск развития второй первичной опухоли молочной железы может повышаться при воздействии радиации на нормальные ткани в ходе радиотерапии. По-видимому, не существует какого-либо измеримого повышения риска для носителей BRCA при радиотерапии, однако риск развития у них второй первичной опухоли уже чрезвычайно высок. Существуют основания полагать, что риск развития второй первичной опухоли повышен среди носителей аллелей генов ATM и CHEK2, предрасполагающих к раку молочной железы, которым проводится радиотерапия (Bernstein et al., 2004; Broeks et al., 2004). Предполагается, что риск развития вторых первичных опухолей на фоне радиотерапии (и, возможно, на фоне интенсивного маммографического скрининга) будет возможно лучше рассчитать при получении точного профиля генетического рис-

ка у пациентов на этапе планирования лечения.

Вторичная профилактика

Приблизительно у 30% пациентов, у которых была диагностирована 1 или 2 стадия рака молочной железы возникает либо локо-региональный, либо дистантный метастатический рецидив их первоначальной опухоли. Пациенты, имевшие первичный рак молочной железы, имеют также значительно повышенный риск обнаружения второй первичной опухоли либо в контралатеральной молочной железе, либо в ипсилатеральной молочной железе в случае проведения органно-сохраняющей операции. Вторичная профилактика касается способов, применяемых для предупреждения рецидивов или развития вторых первичных опухолей. Способы, применяемые в настоящее время, включают долгосрочное лечение тамоксифеном или другим СМЭР, либо отдельно, либо при чередовании с ингибитором ароматазы, снижающую риск мастэктомии контралатеральной молочной железы и снижающую риск овариэктомии (у пациентов, имеющих риск развития семейного рака молочной железы и яичников). Мнения насчет применения тамоксифена были изложены выше. При проведении снижающего риск хирургического вмешательства очевидно, что для выполнения квалифицированного анализа соотношения затрат и пользы, необходимо максимально точно проводить количественную оценку риска.

Существуют некоторые данные, указывающие на то, что качество жизни у пациентов, имеющих генетическую предрасположенность к раку молочной железы, оказывается хуже, чем у основной массы пациентов. Пациенты с носительством варианта 1100delC гена CHEK2 имеют риск развития дистантных метастазов, в 2,8 раза выше и риск рецидивирования заболевания в 3,9 раза выше, чем при отсутствии носительства (de Bock et al., (2004), J Med Genet, 41, 731-5). Пациенты с носительством BRCA1 с опухолями без поражения лимфатических узлов имеют риск метастазирования выше, чем такие же пациенты, не являющиеся носителями мутации BRCA1 (Goffin et al., (2003), Cancer, 97, 527-36; Moller et al., (2002), Int J Cancer, 101, 555-9; Eerola et al., (2001), Int J Cancer, 93, 368-72). Определение генетического профиля, таким образом, может применяться для оценки риска местного рецидивирования и метастазирования, таким образом, направляя выбор мер вторичной профилактики.

В целом, пациенты с диагнозом первичной опухоли имеют риск развития второй первичной опухоли с ежегодной частотой возникновения новых случаев 0,7% (Peto и Mack 2000). Пациенты с мутациями BRCA имеют значительно более высокий риск развития вторых первичных опухолей, чем большинство пациентов с раком молочной железы, абсолютный риск у них находится в пределах 40-60% (Easton 1999). Носители мутаций BRCA имеют значительно более высокий риск развития вторых первичных опухолей (Stacey et al., (2006), PLoS Med, 3, e217; Metcalfe et al., (2004), J Clin Oncol, 22, 2328-35). Пациенты с мутациями гена CHEK2 имеют оцениваемый риск рака противоположной молочной железы в 5,7 раза выше (de Bock, et al., (2004), J Med Genet, 41, 731-5). Носители варианта BARD1 Cys557Ser имеют в 2,7 более высокую вероятность диагностирования второй первичной опухоли (Stacey et al., (2006), PLoS Med, 3, e217). Определение профиля генетического риска может применяться для оценки риска развития второй первичной опухоли у пациентов и принятия взвешенного решения о том, насколько агрессивными должны быть меры профилактики.

Способы по изобретению

Здесь описываются способы оценки риска рака молочной железы и они находятся в рамках изобретения. Изобретение также включает способы оценки возможности ответа индивида на терапевтический агент, применяемый при раке молочной железы, а также способы прогнозирования эффективности терапевтического агента, применяемого при раке молочной железы. Наборы для изучения образца, взятого у пациента для определения подверженности к сердечно-сосудистому заболеванию, также находятся в рамках изобретения.

Диагностические способы и скрининговые тесты

В определенных воплощениях данное изобретение относится к способам диагностики или помощи в диагностике рака молочной железы или подверженности раку молочной железы путем детектирования конкретных аллелей генетических маркеров, которые наиболее часто встречаются у субъектов, имеющих рак молочной железы или субъектов, предрасположенных к раку молочной железы. В конкретном воплощении изобретение представляет собой способ диагностики подверженности раку молочной железы путем детектирования по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера (например, описанных здесь маркеров). Данное изобретение описывает способы, при которых детектирование конкретных аллелей конкретных маркеров или гаплотипов указывает на подверженность раку молочной железы. Такие прогностические или прогнозирующие тесты могут также проводиться для выбора мер профилактики для пациента до того, как у него возникнут симптомы рака молочной железы. Данное изобретение в некоторых воплощениях относится к способам клинического применения диагностики, например диагностике, проводимой медицинскими сотрудниками. В других воплощениях изобретение относится к способам диагностики или определения подверженности, проводимой неспециалистом. Последние технологические достижения в технологиях генотипирования, включая высокопроизводительное генотипирование маркеров ОНП, например, технология Molecular Inversion Probe array (например, в системе GeneChip компании Affymetrix) и технологии BeadArray (например, в системах Golden Gate и Infinium компании Illumina) обеспечили возможность проведения индивидам оценки их генома

одновременно по одному миллиону ОНП по относительно небольшой стоимости. Полученная информация о генотипе, выдаваемая пациенту, может сравниваться с информацией по риску заболевания или признаках заболевания, ассоциированном с различными ОНП, находящейся в открытом доступе в научной литературе. Диагностическую информацию по ассоциированным с заболеванием аллелям по данному описанию, таким образом, могут получать либо индивиды, путем анализа данных его/ее генотипа, либо медицинские сотрудники на основании результатов клинического теста. Другими словами, диагностика, или оценка подверженности, основанная на генетическом риске, может выполняться медицинскими сотрудниками, генетическими консультантами или неспециалистами на основе информации о генотипе человека и публикаций по различным факторам риска. В данном контексте термин "диагностика", "диагностировать подверженность" и "определить подверженность" используются в отношении любого диагностического способа, включая те, которые были упомянуты выше.

Кроме того, в других конкретных воплощениях данное изобретение относится к способам диагностики, или помощи в диагностике пониженной подверженности раку молочной железы путем детектирования определенных аллелей генетических маркеров или гаплотипов, которые встречаются у пациентов с диагностированным раком молочной железы реже, чем у индивидов, у которых не диагностирован рак молочной железы, либо в общей популяции.

Согласно данному описанию и примерам, определенные аллели маркеров или гаплотипы (например, маркеры, расположенные в составе блока НС rs4848543, блока НС rs13387042, блока НС rs3803662, гена STEAP3/TSAP6, маркеры и гаплотипы, приведенные в табл. 10, 15, 19, 20, 21 и 22, и маркеры в неравновесии по сцеплению с ними) ассоциированы с раком молочной железы (например, рак молочной железы с фенотипом "все формы РМЖ" и/или "УмерПред"). В одном воплощении аллель маркера или гаплотип обуславливает существенный риск или подверженность раку молочной железы. В другом воплощении изобретение касается способа диагностики подверженности раку молочной железы у человеческого индивида, способ включает определение наличия или отсутствия по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера в образце нуклеиновой кислоты, взятого у индивида или массиве данных генотипа, полученного для индивида, где по меньшей мере один полиморфный маркер выбран из группы, состоящей из полиморфных маркеров, расположенных в составе блока НС rs4848543, блока НС rs13387042, блока НС rs3803662, гена STEAP3/TSAP6, маркеров и гаплотипов, приведенных в табл. 10, 15, 19, 20, 21 и 22, и маркеров в неравновесии по сцеплению с ними (например, характеризуемому $r^2 > 0,2$). В другом воплощении изобретение относится к способам диагностики подверженности раку молочной железы у человеческого индивида путем скрининга по меньшей мере одного аллеля маркера или гаплотипа, например, маркеров, расположенных в составе блока НС rs4848543, блока НС rs13387042, блока НС rs3803662, гена STEAP3/TSAP6, маркеров и гаплотипов, приведенных в табл. 10, 15, 19, 20, 21 и 22, и маркеров в неравновесии по сцеплению с ними. В другом воплощении аллель маркера или гаплотип встречаются у лиц, имеющих или подверженных раку молочной железы ("пораженных", например, с фенотипов РМЖ "все формы РМЖ" и/или "УмерПред") чаще по сравнению с частотой их встречаемости у здоровых лиц (в контроле, например, в популяционном контроле). В другом воплощении изобретение относится к способу диагностики подверженности раку молочной железы у человеческого индивида, способ включает определение наличия или отсутствия по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера в образце нуклеиновой кислоты, взятом у индивида, где по меньшей мере один полиморфный маркер выбран из маркера rs4848543, маркера rs13387042, маркера rs3803662 и маркеров в неравновесии по сцеплению с ними. В определенных воплощениях достоверность ассоциации по меньшей мере одного аллеля маркера или гаплотипа характеризуется значением $P < 0,05$. В других воплощениях достоверность ассоциации характеризуется меньшими значениями P , такими как $< 0,01$, $< 0,001$, $< 0,0001$, $< 0,00001$, $< 0,000001$, $< 0,0000001$ или $< 0,00000001$.

В данных воплощениях наличие по меньшей мере одного аллеля маркера или гаплотипа указывает на подверженность раку молочной железы. Данные диагностические способы включают детекцию наличия или отсутствия по меньшей мере одного аллеля маркера или гаплотипа, ассоциированного с раком молочной железы. Гаплотипы по данному описанию включают комбинации аллелей по разным генетическим маркерам (например, ОНП, микросателлитам). Детекция конкретных аллелей генетических маркеров, составляющих определенные гаплотипы, может выполняться при помощи разнообразных методов, описываемых здесь и/или известных в данной области техники. Например, генетические маркеры можно детектировать на уровне нуклеиновых кислот (например, с помощью прямого секвенирования нуклеотидов или с помощью других средств, известных специалистам) или на уровне аминокислот, если генетический маркер затрагивает кодирующую последовательность для белка, кодируемого нуклеиновой кислотой, ассоциированной с раком молочной железы (например, путем секвенирования белка или с помощью иммунологических тестов с применением антител, распознающих такой белок). Аллели маркеров или гаплотипы по данному изобретению соответствуют фрагментам последовательности геномной ДНК, ассоциированной с раком молочной железы. Такие фрагменты включают последовательность ДНК обобщаемого полиморфного маркера или гаплотипа, но также могут включать сегменты ДНК в сильном НС (неравновесии по сцеплению) с маркером или гаплотипом (например, характеризуемому значениями

r^2 выше 0,2 и/или $|D'| > 0,8$).

В одном воплощении диагностика подверженности раку молочной железы может осуществляться с помощью методов гибридизации, например Саузерн анализу, Нозерн анализу, и/или гибридизации *in situ* (см. *Current Protocols in Molecular Biology*, Ausubel F. et al., eds., John Wiley & Sons, включая все приложения). Биологический образец тестируемого субъекта или индивида («тестируемый образец») геномной ДНК, РНК или кДНК получают у субъекта с подозрением на наличие, предрасположенность или подверженность раку молочной железы ("тестируемого субъекта"). Субъект может быть взрослым, ребенком или плодом. Тестируемый образец может быть получен из любого источника, содержащего геномную ДНК, например, образца крови, образца амниотической жидкости, образца цереброспинальной жидкости, или образца ткани кожи, мышцы, буккальной слизистой или конъюнктивы, плаценты, желудочно-кишечного тракта или других органов. Тестируемый образец ДНК из фетальных клеток или тканей может быть получен с помощью соответствующих способов, например, амниоцентеза или обследования хорионических ворсинок. Затем проводится тестирование образца ДНК, РНК или кДНК. Наличие конкретного аллеля маркера может быть обнаружено при помощи специфической гибридизации зонда нуклеиновой кислоты, специфического в отношении определенного аллеля. Наличие более одного аллеля специфического маркера или специфического гаплотипа может быть обнаружено с помощью нескольких последовательность-специфических зондов нуклеиновых кислот, каждый из которых специфичен в отношении конкретного аллеля. В одном воплощении гаплотип может быть идентифицирован с помощью единственного зонда нуклеиновой кислоты, специфичного в отношении определенного гаплотипа (т.е. специфически гибридизующегося с нитью ДНК, содержащей аллели определенных маркеров, характерных для гаплотипа). Зонд, специфичный в отношении последовательности, может быть предназначен для гибридизации с геномной ДНК, РНК или кДНК. "Зонд нуклеиновой кислоты" в данном описании может представлять собой зонд ДНК или зонд РНК, который гибридизуется с комплементарной последовательностью. Специалистам известно, как создать такой зонд, чтобы гибридизация, специфичная в отношении последовательности, имела место только при наличии конкретного аллеля в геномной последовательности из тестируемого образца.

Для диагностики подверженности раку молочной железы гибридизованный образец формируется при инкубации тестируемого образца, содержащего нуклеиновую кислоту, ассоциированную с раком молочной железы, например, образца геномной ДНК, с по меньшей мере одним зондом нуклеиновой кислоты. Нелимитирующим примером зонда для детекции мРНК или геномной ДНК является меченый зонд нуклеиновой кислоты, способный гибридизоваться с описанными здесь последовательностями мРНК или геномной ДНК. Зонд нуклеиновой кислоты может быть, например, целой молекулой нуклеиновой кислоты или ее частью, например, олигонуклеотидом с длиной по меньшей мере 15, 30, 50, 100, 250 или 500 нуклеотидов, достаточной для специфической гибридизации с соответствующей мРНК или геномной ДНК в строгих условиях. Например, зонд нуклеиновой кислоты может включать нуклеотидную последовательность, приведенную в SEQ ID NO:4, SEQ IN NO:5 or SEQ ID NO:6, или ее часть, нуклеотидную последовательность блока NC rs4848543, блока NC rs13387042, блока NC rs3803662 и/или гена STEAP3/TSAP или ее часть, возможно, включая по меньшей мере один аллель маркера по данному описанию или по меньшей мере один маркер или гаплотип по данному описанию (например, маркеры и гаплотипы, приведенные в табл. 10, 15, 19, 20, 21 и 22, и маркеров в неравновесии по сцеплению с ними) или по меньшей мере один аллель, входящий в состав гаплотипов по данному описанию, или зонд может быть комплементарной последовательностью к такой последовательности. Здесь описываются и другие зонды, подходящие для применения в диагностических тестах по изобретению. Гибридизация может проводиться с помощью способов, хорошо известных специалистам (см., например, *Current Protocols in Molecular Biology*, Ausubel F. et al., eds., John Wiley & Sons, включая все приложения). В одном воплощении гибридизация означает специфическую гибридизацию, т.е. гибридизацию без несовпадений (точную гибридизацию). В одном воплощении условия гибридизации при специфической гибридизации являются строгими.

Специфическая гибридизация, если она имеет место, определяется с помощью стандартных методов. Если специфическая гибридизация происходит между зондом нуклеиновой кислоты и ассоциированной с раком молочной железы нуклеиновой кислотой в тестируемом образце, значит образец содержит аллель, комплементарный нуклеотиду, присутствующему в зонде нуклеиновой кислоты. Процесс можно повторить для других маркеров по данному изобретению или маркеров, составляющих гаплотип по данному изобретению, или несколько зондов можно использовать параллельно для детекции аллелей более чем одного маркера одновременно. Возможно также создать единый зонд, содержащий аллели более чем одного маркера определенного гаплотипа (например, зонд, содержащий аллели, комплементарные к 2, 3, 4, 5 или всем маркерам, составляющим конкретный гаплотип). Детектирование конкретных маркеров гаплотипа в образце указывает на то, что источник, у которого был взят образец, имеет конкретный гаплотип (например, гаплотип) и таким образом является подверженным раку молочной железы.

В одном предпочтительном воплощении способ, в котором применяется детектирующий олигонуклеотидный зонд, включающий флуоресцентный компонент или группу на своем 3' конце и гаситель на

своем 5' конце, и усиливающий олигонуклеотид, применяется согласно описанию у Kutuyavin et al. (Nucleic Acid Res. 34:e128 (2006)). Флуоресцентным компонентом может служить Gig Harbor Green или Yakima Yellow, или другой подходящий флуоресцентный компонент. Детектирующий зонд создается для гибридизации с короткой нуклеотидной последовательностью, содержащей однонуклеотидный полиморфизм, подлежащий определению. Предпочтительно, однонуклеотидный полиморфизм находится в любом положении от концевой остатка до -6 остатка с 3' конца детектирующего зонда. Усилитель представляет собой короткий олигонуклеотидный зонд, который гибридизуется с ДНК-матрицей в 3' положении относительно детектирующего зонда. Зонды создаются таким образом, что, когда детектирующий зонд и усиливающий нуклеотидный зонд оба связаны с матрицей, между ними существует однонуклеотидный пробел. Этот пробел приводит к появлению искусственного сайта в виде остатка дезоксирибозы, лишённого азотистого основания, распознаваемого эндонуклеазой, например, эндонуклеазой IV. Фермент отщепляет краситель от полностью комплементарного детектирующего зонда, но не может расщепить детектирующий зонд, содержащий несовпадение. Таким образом, путем измерения флуоресценции высвобожденного флуоресцентного компонента может проводиться оценка наличия конкретного аллеля, определяемого нуклеотидной последовательностью детектирующего зонда.

Детектирующий зонд может быть любого подходящего размера, хотя предпочтительно зонды бывают относительно короткими. В одном воплощении зонд имеет длину 5-100 нуклеотидов. В другом воплощении зонд имеет длину 10-50 нуклеотидов, а в другом воплощении зонд имеет длину 12-30 нуклеотидов. Другие размеры зондов также возможны и находятся в рамках компетенции специалиста в данной области.

В предпочтительном воплощении ДНК-матрица, содержащая однонуклеотидный полиморфизм, перед детекцией амплифицируется в ходе полимеразной цепной реакции (ПЦР). В таком воплощении амплифицируемая ДНК служит матрицей для детектирующего зонда и усиливающего зонда.

Определенные воплощения детектирующего зонда, усиливающего зонда, и/или праймеров, используемых для амплификации матрицы в ПЦР, включают применение модифицированных оснований, включая модифицированные А и модифицированные G. Применение модифицированных оснований может использоваться для подгонки температур плавления нуклеотидной молекулы (зонда и/или праймера) и ДНК-матрицы, например, для повышения температуры плавления в участках с низким процентным содержанием оснований G или C, в которых могут применяться модифицированные А, со способностью формировать три водородные связи со своими комплементарными Т, или для понижения температуры плавления в участках с высоким процентным содержанием оснований G или C, например, при применении модифицированных оснований G, формирующих только две водородные связи с комплементарными им основаниями C в двунитевой молекуле ДНК. В предпочтительном воплощении модифицированные основания применяются для создания детектирующего нуклеотидного зонда. Любое модифицированное основание, известное специалистам, может быть выбрано в этих методах, и выбор подходящих оснований находится в рамках компетенции специалиста и основывается на приведенных здесь сведениях и имеющихся в распоряжении специалиста основаниях, доступных из коммерческих источников.

При другом способе гибридизации, для идентификации наличия полиморфизма, ассоциированного с раком молочной железы, используется Нозерн анализ (см. Current Protocols in Molecular Biology, Ausubel F. et al., eds., John Wiley & Sons, supra). Для Нозерн-анализа тестируемый образец РНК получают у субъекта соответствующими способами. Согласно данному описанию, специфическая гибридизация зонда нуклеиновой кислоты с РНК субъекта указывает на наличие конкретного аллеля, комплементарного к зонду. Репрезентативные примеры применения зондов нуклеиновых кислот приведены, например, в US Patent Nos. 5288611 и 4851330.

Кроме того, в описываемых здесь гибридизационных методах в дополнение к зонду нуклеиновой кислоты или вместо него можно использовать зонд пептидной нуклеиновой кислоты (ПНК). ПНК представляет собой ДНК, имитирующую наличие пептидоподобного неорганического скелета, например, единиц N-(2-аминоэтил)глицина с органическими основаниями (A, G, C, T или U), присоединенных к атому азота глицина через метиленовый карбонильный линкер (см, например, Nielsen P. et al., Bioconjug. Chem. 5:3-7 (1994)). Зонд ПНК может быть специально сконструирован для специфической гибридизации в образце с молекулой, предположительно содержащей одну или более аллелей маркера или гаплотипа, ассоциированных с раком молочной железы. Таким образом, гибридизация зонда ПНК будет диагностичной в отношении рака молочной железы или подверженности раку молочной железы.

В одном воплощении изобретения у субъекта берется тестируемый образец, содержащий геномную ДНК, и для амплификации фрагмента, содержащего один или более маркеров или гаплотипов по данному изобретению применяется полимеразная цепная реакция (ПЦР). Согласно данному описанию, идентификация конкретного аллеля маркера или гаплотипа, ассоциированного с раком молочной железы, может проводиться с помощью разнообразных методов (например, анализа последовательности, анализа с помощью рестрикционного расщепления, специфической гибридизации, анализа конформационного полиморфизма однонитевой ДНК (от англ. Single Stranded Conformation Polymorphism, SSCP), электрофоретического анализа и т.д. В другом воплощении диагностика выполняется путем анализа экспрессии с помощью количественной ПЦР (кинетический термоциклер). Для такой методики, например, могут

применяться коммерческие доступные системы, такие как TaqMan® (Applied Biosystems, Foster City, CA). Методика позволяет оценить наличие изменения в экспрессии или составе полипептида или варианта (вариантов) сплайсинга, кодируемых нуклеиновой кислотой, ассоциированной с раком молочной железы. Кроме того, экспрессия варианта (вариантов) может расцениваться как физическое или функциональное отличие.

В другом способе по изобретению для детектирования определенного аллеля может применяться рестрикционный анализ, в том случае, если аллель обуславливает образование или исчезновение рестрикционного сайта референсной последовательности. Может быть проведен анализ полиморфизма длины рестрикционных фрагментов (от англ. Restriction Fragment Length Polymorphism, RFLP), например, согласно описанию в *Current Protocols in Molecular Biology*, supra. Профиль расщепления соответствующего фрагмента ДНК указывает на присутствие или отсутствие конкретного аллеля в образце.

Анализ последовательности может также применяться для детекции определенных аллелей полиморфных сайтов, ассоциированных с раком молочной железы (например, полиморфных маркеров и гаплотипов по табл. 10, 15, 19, 20, 21 и/или 22 и маркеров в неравновесии по сцеплению с ними). Так, в одном воплощении определение наличия или отсутствия конкретных аллелей маркеров или гаплотипов включает анализ последовательности тестируемого образца ДНК или РНК, полученного у субъекта или индивида. Для амплификации части нуклеиновой кислоты, ассоциированной с раком молочной железы, может применяться ПЦР или другой соответствующий метод, а наличие определенного аллеля затем может определяться непосредственно путем секвенирования полиморфного сайта (или множества полиморфных сайтов в гаплотипе) геномной ДНК в образце.

Аллель-специфические олигонуклеотиды могут также применяться для детекции наличия конкретного аллеля нуклеиновой кислоты, ассоциированной с раком молочной железы (например, полиморфных маркеров и гаплотипов по табл. 10, 15, 19, 20, 21 и/или 22 и маркеров в неравновесии по сцеплению с ними), с помощью дот-блот гибридизации амплифицированных олигонуклеотидов с аллель-специфическими олигонуклеотидными зондами (от англ. Allele-Specific Oligonucleotide, ASO) (см. для примера Saiki R. et al., *Nature*, 324:163-166 (1986)). "Аллель-специфический олигонуклеотид" (также обозначаемый здесь как "аллель-специфический олигонуклеотидный зонд") представляет собой олигонуклеотид состоящий примерно из 10-50 пар оснований или примерно из 15-30 пар оснований, который специфически гибридизуется с нуклеиновой кислотой, ассоциированной с раком молочной железы, содержащей определенный аллель полиморфного сайта (например, маркер или гаплотип, согласно данному описанию). Аллель-специфический олигонуклеотидный зонд, специфичный к одной или более конкретной нуклеиновой кислоте, ассоциированной с раком молочной железы, может быть создан с применением стандартных способов (см., например, *Current Protocols in Molecular Biology*, supra). Для амплификации желаемого участка может использоваться ПЦР. Можно провести дот-блоттинг ДНК, содержащей амплифицированный участок, с помощью стандартных методов (см., например, *Current Protocols in Molecular Biology*, supra), и блот может быть инкубирован с олигонуклеотидным зондом. Затем можно определить наличие специфической гибридизации зонда с амплифицированным участком. Специфическая гибридизация аллель-специфического олигонуклеотидного зонда с ДНК субъекта указывает на наличие определенного аллеля полиморфного сайта, ассоциированного с раком молочной железы (см., например, Gibbs R. et al., *Nucleic Acids Res.*, 77: 2437-2448 (1989) и WO 93/22456).

С добавлением таких аналогов, как закрытые нуклеиновые кислоты (от англ. Locked Nucleic Acids, LNAs), размер праймеров и зондов может быть уменьшен всего до 8 оснований. LNAs являются новым классом бициклических аналогов ДНК, в которых 2' и 4' положения фуранозного кольца связаны через О-метиленовые (окси-LNA), S-метиленовые (тио-LNA), или amino-метиленовые (амино-LNA) группировки. Общим для всех этих вариантов LNA является аффинность к комплементарным нуклеиновым кислотам, которая является наиболее высокой среди всех известных аналогов ДНК. Например, было показано, что определенные нонамеры окси-LNA имеют температуры плавления (T_m) 64 и 74°C будучи в комплексе, соответственно, с комплементарной ДНК или РНК, в противоположность 28°C для ДНК и РНК в случае соответствующего нонамера ДНК. Существенное повышение T_m также происходит, когда мономеры LNA используются в комбинации с мономерами стандартной ДНК или РНК. T_m для праймеров и зондов может быть существенно повышена в зависимости от того, куда включаются мономеры LNA (например, на 3' конце, на 5' конце или в середине).

В другом воплощении, панели олигонуклеотидных зондов, комплементарных к целевым сегментам последовательности нуклеиновой кислоты субъекта, могут применяться для идентификации полиморфизмов в нуклеиновой кислоте, ассоциированной с раком молочной железы (например, полиморфные маркеры и гаплотипы по табл. 10, 15, 19, 20, 21 и/или 22 и маркеры в неравновесии по сцеплению с ними).

Например, могут быть использованы нуклеотидные панели. Олигонуклеотидные панели обычно содержат множество различных олигонуклеотидных зондов, связанных с поверхностью субстрата в разных известных положениях. Такие олигонуклеотидные чипы, также обозначаемые как "Genechips™", широко описаны в данной области техники (см. например, U.S. Patent No. 5143854, PCT Patent Publication

Nos. WO 90/15070 и 92/10092). Эти чипы обычно производятся с помощью механических методов синтеза, включающих сочетание фотолитографических методов и методов твердофазного синтеза олигонуклеотидов, или с помощью других методов, известных специалистам (см., например, Fodor S. et al., *Science*, 251: 767-773 (1991); Pirrung et al., U.S. Patent No. 5143854 (см. также опубликованную заявку PCT No. WO 90/15070); и Fodor S. et al., опубликованная заявка PCT No. WO 92/10092 и U.S. Patent No. 5424186, описания которых включены сюда во всей их полноте путем ссылок). Технологии синтеза этих чипов с применением методов механического синтеза описаны, например, в U.S. Patent No. 5,384,261; описания которых включены сюда во всей их полноте путем ссылок. В другом примере могут применяться линейные чипы.

После получения олигонуклеотидного чипа производится гибридизация нуклеиновой кислоты, представляющей интерес, с чипом. Детекция гибридизации представляет собой детекцию определенного аллеля нуклеиновой кислоты, представляющей интерес. Гибридизация и сканирование как правило проводятся при помощи способов, описываемых здесь или при помощи других способов, известных специалистам, например, согласно описанию в опубликованных патентных заявках WO 92/10092 и WO 95/11995, и патенте U.S. Patent No. 5424186 описания которых включены сюда во всей их полноте путем ссылок. Коротко, последовательность нуклеиновой кислоты - мишени, включающая один или более идентифицированных ранее полиморфных маркеров, амплифицируют при помощи хорошо известных методов амплификации (например, ПЦР). Обычно при этом применяют праймеры, последовательность которых комплементарна двум цепочкам последовательности - мишени, как выше, так и ниже полиморфного сайта. Также могут применяться методики асимметричной ПЦР. Амплифицированная мишень, как правило, инкорпорирующая метку, затем гибридизуется с чипом при соответствующих условиях, при которых возможна последовательность-специфическая гибридизация. После окончания гибридизации и отмывки чипа, чип сканируют для определения участка чипа, с которым гибридизуется последовательность - мишень. Данные о гибридизации, полученные при сканировании, обычно бывают представлены в форме интенсивностей флуоресценции, указанных для позиций чипа.

Несмотря на то, что чипы изначально были описаны для одиночного детектируемого блока, например, для детекции одиночного полиморфного сайта, они могут включать множество детектируемых блоков, и таким образом позволять анализировать множественные специфические полиморфизмы (например, множественные полиморфизмы определенного гаплотипа). Возможны компоновки, в которых детектируемые блоки могут быть сгруппированы в пределах одного чипа или нескольких отдельных чипов таким образом, что можно варьировать оптимальные условия для гибридизации мишени с чипом. Например, зачастую бывает желательно провести детекцию тех полиморфизмов, которые приходятся на участки геномной последовательности, богатые G-C, в отдельности от тех полиморфизмов, которые приходятся на сегменты, богатые A-T. Таким образом оказывается возможным отдельно оптимизировать условия гибридизации для каждой ситуации.

Дополнительные описания применений олигонуклеотидных чипов для детекции полиморфизмов можно найти, например, в патентах США US 5858659 и 5837832, описания которых включены сюда во всей их полноте путем ссылок.

Для детекции конкретного аллеля полиморфного сайта, ассоциированного с раком молочной железы (например, полиморфных маркеров по табл. 10, 15, 19, 20, 21 и 22 и маркеров в неравновесии по сцеплению с ними), могут применяться другие методы анализа нуклеиновых кислот, доступные специалистам. Репрезентативные методы включают, например, прямое секвенирование вручную (Church и Gilbert, *Proc. Natl. Acad. Sci. USA*, 81: 1991-1995 (1988); Sanger, F. et al., *Proc. Natl. Acad. Sci. USA*, 74: 5463-5467 (1977); Beavis et al., патент США US 5,288,644); автоматизированное секвенирование с использованием флуоресценции; анализ конформационного полиморфизма однонитевой ДНК (single stranded conformation polymorphism assays, SSCP); электрофорез в денатурирующем геле с фиксацией GC (clamped denaturing gel electrophoresis, CDGE); электрофорез в градиентном денатурирующем геле (denaturing gradient gel electrophoresis, DGGE) (Sheffield V. et al., *Proc. Natl. Acad. Sci. USA*, 86: 232-236 (1989)), анализ по сдвигу электрофоретической подвижности (mobility shift analysis) (Orita M. et al., *Proc. Natl. Acad. Sci. USA*, 86: 2766-2770 (1989)), рестрикционный анализ (Flavell R. et al., *Cell*, 15:25-41 (1978); Geever, R., et al., *Proc. Natl. Acad. Sci. USA*, 78:5081-5085 (1981)); гетеродуплексный анализ; химическое расщепление некомплементарных сайтов (chemical mismatch cleavage, CMC) (Cotton R. et al., *Proc. Natl. Acad. Sci. USA*, 85: 4397-4401 (1988)); анализ с помощью защиты от рибонуклеазы (RNase protection assays) (Myers R. et al., *Science*, 230: 1242-1246 (1985)); применение полипептидов, распознающих нуклеотидные несовпадения, таких, как белок E. coli mutS; и аллель-специфическую ПЦР.

В другом воплощении изобретения диагностика рака молочной железы или подверженности раку молочной железы может проводиться при изучении экспрессии и/или состава полипептида, кодируемого нуклеиновой кислотой, ассоциированной с раком молочной железы в тех случаях, когда генетический маркер (маркеры) или гаплотип (гаплотипы) по данному изобретению приводят к изменению состава или экспрессии полипептида. Так, диагностика подверженности раку молочной железы может проводиться при изучении экспрессии и/или состава одного из указанных полипептидов или другого полипептида, кодируемого нуклеиновой кислотой, ассоциированной с раком молочной железы в тех случаях, когда

генетический маркер или гаплотип по данному изобретению приводит к изменению состава или экспрессии полипептида. Гаплотипы и маркеры по данному изобретению, ассоциированные с раком молочной железы, могут оказывать эффект на один или более близкорасположенных генов (например, ген PITX2). Возможный механизм влияния на эти гены включает, например, влияние на транскрипцию, влияние на сплайсинг РНК, изменения в относительном количестве форм альтернативного сплайсинга мРНК, влияние на стабильность РНК, влияние на транспорт из ядра в цитоплазму и влияние на эффективность и точность трансляции.

Так, в другом воплощении варианты (маркеры или гаплотипы) по изобретению, демонстрирующие ассоциацию с раком молочной железы, влияют на экспрессию близкорасположенного гена. Известно, что регуляторные элементы, влияющие на экспрессию генов, могут располагаться за десятки и даже сотни килобаз от участка промотора гена. Оценивая наличие или отсутствие по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера по данному изобретению, возможно оценить уровень экспрессии таких близрасположенных генов. Предполагается, что детекция маркеров или гаплотипов по данному изобретению может применяться для оценки экспрессии одного или более генов STEAP3/TSAP6, LOC643714 и TNRC9.

Для определения уровня экспрессии белков могут применяться различные методы, включая иммуноферментный анализ, Вестерн блоттинг, иммунопреципитацию и иммунофлуоресценцию. Тестируемый образец субъекта оценивается на наличие изменения в экспрессии и/или изменения в составе полипептида, кодируемого нуклеиновой кислотой, ассоциированной с раком молочной железы. Изменения в экспрессии полипептида, кодируемого нуклеиновой кислотой, ассоциированной с раком молочной железы могут, например, представлять собой изменения в количестве экспрессируемого полипептида (т.е. количество продуцируемого полипептида). Изменения в составе полипептида, кодируемого нуклеиновой кислотой, ассоциированной с раком молочной железы представляют собой качественные изменения в экспрессии полипептида (например, экспрессия мутантного полипептида или другого сплайсингового варианта). В одном воплощении диагностика подверженности раку молочной железы производится путем детекции определенного сплайсингового варианта, кодируемого нуклеиновой кислотой, ассоциированной с раком молочной железы, или определенного профиля сплайсинговых вариантов.

Могут иметь место оба вида таких нарушений (количественные и качественные). "Изменение" в экспрессии или составе полипептида, по данному описанию, относится к изменению экспрессии или состава тестируемого образца, по сравнению с экспрессией или составом полипептида, кодируемого нуклеиновой кислотой, ассоциированной с раком молочной железы, в контрольном образце. Контрольным образцом является образец, соответствующий тестируемому образцу (например, происходит из того же типа клеток) и забирается у субъекта, который не поражен или который не подвержен раку молочной железы. В одном воплощении контрольный образец берется у субъекта, который не имеет аллеля маркера или гаплотипа по данному описанию. Также наличие одного или более различных вариантов сплайсинга в тестируемом образце или присутствие значительно различающегося количества разных вариантов сплайсинга в тестируемом образце по сравнению с контрольным образцом может указывать на подверженность раку молочной железы. Изменения экспрессии или состава полипептида в тестируемом образце по сравнению с контрольным образцом может указывать на наличие определенного аллеля в случае, когда аллель приводит к изменению сайта сплайсинга относительно референсного в контрольном образце. Могут применяться различные способы изучения экспрессии или состава полипептида, кодируемого нуклеиновой кислотой, ассоциированной с раком молочной железы, включая спектроскопию, колориметрию, электрофорез, изоэлектрическое фокусирование и иммунологические методы (например, David et al., U.S. Pat. No. 4376110), такие как иммуноблоттинг (см., например, Current Protocols in Molecular Biology, особенно главу 10, supra).

Например, в одном воплощении может использоваться антитело (например, антитело с детектируемой меткой), способное связываться с полипептидом, кодируемым нуклеиновой кислотой, ассоциированной с раком молочной железы. Антитела могут быть поликлональными или моноклональными. Может использоваться интактное антитело или его фрагмент (например, Fv, Fab, Fab', F(ab')₂). Термин "меченое" в отношении зонда или антитела применяется для обозначения прямого мечения зонда путем сшивания (т.е. физического связывания) детектируемого вещества с зондом или антителом, а также непрямого мечения зонда или антитела по его способности реагировать с другим реагентом, меченым напрямую. Примеры непрямого мечения включают детекцию первичного антитела с помощью меченого вторичного антитела (например, вторичное антитело с флуоресцентной меткой) и концевое мечение ДНК-зонда биотином, что позволяет его детектировать флуоресцентно-меченым стрептавидином.

В одном воплощении данного метода уровень или количество полипептида, кодируемого нуклеиновой кислотой, ассоциированной с раком молочной железы (например, нуклеиновой кислотой, кодирующей гены STEAP3/TSAP6, LOC643714 или TNRC9), в тестируемом образце сравнивается с уровнем или количеством полипептида в контрольном образце. Количество или уровень полипептида в контрольном образце, которое выше или ниже уровня или количества полипептида в контрольном образце настолько, что различие статистически достоверно, указывает на изменение экспрессии полипептида, кодируемого нуклеиновой кислотой, и является диагностическим в отношении конкретного аллеля или

гаплотипа, ответственного за различие в экспрессии. В альтернативном варианте, состав полипептида в тестируемом образце сравнивается с составом полипептида в контрольном образце. В другом воплощении могут оцениваться как уровень или количество, так и состав полипептида в тестируемом образце и в контрольном образце.

В другом воплощении диагностика подверженности раку молочной железы проводится по детектированию по меньшей мере одного маркера или гаплотипа по данному изобретению (например, ассоциированных аллелей маркеров и гаплотипов, приведенных в табл. 10, 15, 19, 20, 21 и 22 и маркеров в неравновесии по сцеплению с ними) в сочетании с дополнительным исследованием белков, РНК или ДНК. Способы по изобретению могут также использоваться в сочетании с анализом семейного анамнеза субъекта и факторов риска (например, факторов риска внешней среды, факторов риска образа жизни).

Наборы

Наборы, применяемые в способах по изобретению, включают компоненты, применяемые в любых описываемых здесь способах, включая, например, гибридизационные зонды, рестрикционные ферменты (например, для анализа RFLP), аллель-специфические олигонуклеотиды, антитела, связывающиеся с измененными полипептидами, кодируемыми нуклеиновой кислотой по изобретению согласно данному описанию (например, сегментом генома, содержащим по меньшей мере один полиморфный маркер и/или гаплотип по данному изобретению) или с неизменным (нативным) полипептидом, кодируемым нуклеиновой кислотой по изобретению согласно данному описанию, средства для амплификации нуклеиновой кислоты, ассоциированной с раком молочной железы, средства для анализа нуклеотидной последовательности нуклеиновой кислоты, ассоциированной с раком молочной железы, средства для анализа аминокислотной последовательности полипептида, кодируемого нуклеиновой кислотой, ассоциированной с раком молочной железы и т.д. Наборы могут, например, содержать необходимые буферы, праймеры нуклеиновых кислот для амплификации нуклеиновых кислот по изобретению (например, одного или более полиморфных маркеров по данному описанию) и реагенты для аллель-специфической детекции амплифицируемых фрагментов с применением таких праймеров и необходимых ферментов (например, ДНК-полимеразы). Кроме того, наборы могут иметь реагенты для тестов, для применения в сочетании со способами по данному изобретению, например, реагенты для применения в тестах для диагностики рака молочной железы.

В одном воплощении изобретение представляет собой набор для тестирования образца, взятого у субъекта, для детекции наличия у субъекта рака молочной железы или подверженности раку молочной железы, когда набор включает реагенты, необходимые для избирательного детектирования по меньшей мере одного аллеля по меньшей мере одного полиморфизма по данному изобретению в геноме индивида. В конкретном воплощении реагенты включают по меньшей мере один примыкающий олигонуклеотид, гибридизующийся с фрагментом в геноме индивида, содержащим по меньшей мере один полиморфизм по данному изобретению. В другом воплощении реагенты включают по меньшей мере одну пару олигонуклеотидов, гибридизующихся с противоположными нитями сегмента генома, полученного у субъекта, когда каждая пара олигонуклеотидных праймеров создана для селективной амплификации фрагмента генома индивида, содержащего по меньшей мере один полиморфизм, когда полиморфизм выбран из группы, состоящей из полиморфизмов, приведенных в табл. 10, 15, 19, 20, 21 и 22 и полиморфных маркеров в неравновесии по сцеплению с ними. В другом воплощении фрагмент имеет длину по меньшей мере 20 пар оснований. Такие олигонуклеотиды или нуклеиновые кислоты (например, олигонуклеотидные праймеры) могут быть созданы с применением части последовательностей нуклеиновых кислот, фланкирующих полиморфизмы (например, ОНП или микросателлиты), указывающие на рак молочной железы. В другом воплощении набор включает одну или более меченых нуклеиновых кислот, пригодных для аллель-специфической детекции одного или более определенных полиморфных маркеров или гаплотипов, ассоциированных с раком молочной железы, и реагентами для детекции метки. Подходящие метки включают, например, радиоизотоп, флуоресцентную метку, ферментную метку, метку в виде кофактора фермента, магнитную метку, спин-метку, эпиген-метку.

В конкретных воплощениях полиморфный маркер или гаплотип, подлежащий определению реагентами по набору, включает один или более маркеров, два или более маркеров, три или более маркеров, четыре или более маркеров или пять или более маркеров, выбранных из группы, состоящей из маркеров по табл. 10, 15, 19, 20, 21 и 22. В другом воплощении маркер, подлежащий определению, выбран из маркеров, приведенных в табл. 10, 15 или 19. В одном воплощении маркер, подлежащий определению, выбран из маркеров, приведенных в табл. 10. В другом воплощении маркер, подлежащий определению, выбран из маркеров, приведенных в табл. 15. В другом воплощении маркер, подлежащий определению, выбран из маркеров, приведенных в табл. 19. В предпочтительном воплощении маркер, подлежащий определению, является rs4848543, rs13387042 и/или rs3803662. Другие предпочтительные воплощения включают те, что включают реагенты для детекции маркера rs4848543, те, что включают реагенты для детекции маркера rs13387042 и те, что включают реагенты для детекции маркера rs3803662. В другом воплощении маркер или гаплотип, подлежащий определению, включает по меньшей мере один маркер из группы маркеров в сильном неравновесии по сцеплению, которое описывается значениями r^2 выше 0,2, с по меньшей мере одной из групп маркеров, состоящей из маркеров, приведенных в табл. 10, 15, 19,

20, 21 и 22. В другом воплощении маркер или гаплотип, подлежащий определению, включает по меньшей мере один маркер, выбранный из группы маркеров, состоящей из маркеров rs4848543, rs13387042 и rs3803662, и маркеров в неравновесии по сцеплению с ними.

В одном таком воплощении наличие маркера или гаплотипа указывает на подверженность (повышенную подверженность или пониженную подверженность) раку молочной железы. В другом воплощении маркером является аллель A rs4848543, аллель A rs13387042 и/или аллель T rs3803662, наличие которого указывает на повышенный риск рака молочной железы (например, с фенотипами РМЖ "все формы РМЖ" или "УмерПред"). В следующем воплощении наличие маркера или гаплотипа указывает на ответ на терапевтический агент, применяемый при раке молочной железы. В другом воплощении наличие маркера или гаплотипа указывает на прогноз рака молочной железы. В еще одном воплощении наличие маркера или гаплотипа указывает на успешность лечения рака молочной железы. Такое лечение может включать хирургическое вмешательство, радиотерапию или другие способы (например, изменение образа жизни).

Диагностика рака молочной железы, ассоциированного с полиморфизмами по изобретению

Несмотря на то, что способы диагностики были в общем описаны в контексте диагностики подверженности раку молочной железы, способы могут также использоваться для диагностики рака молочной железы, ассоциированного с полиморфными маркерами по данному изобретению. Например, индивид, имеющий рак молочной железы или факторы риска, ассоциированные с раком молочной железы, может обследоваться для определения, может ли наличие у индивида полиморфизма или гаплотипа по данному изобретению являться фактором, способствующим развитию у индивида рака молочной железы. В одном воплощении выявление рака молочной железы, ассоциированного с маркерами и/или гаплотипами по данному изобретению, облегчает планирование тактики лечения. Например, могут проводиться профилактические мероприятия для минимизации возникновения у индивида рака молочной железы. Такие превентивные мероприятия также включают определение, является ли индивид гетерозиготным или гомозиготным в отношении связанных с риском вариантов по изобретению. В других воплощениях по изобретению может быть разработано лечение и выбраны препараты для целенаправленного воздействия на соответствующий ген или белок, ассоциированный с полиморфизмом и/или гаплотипом по данному изобретению.

В других воплощениях изобретение касается способа диагностики и выявления у субъекта рака молочной железы, ассоциированного с полиморфизмами по изобретению, путем выявления наличия полиморфного маркера или гаплотипа по изобретению, согласно подробному описанию здесь. Например, описываемые здесь полиморфные маркеры и/или гаплотипы чаще обнаруживаются у субъектов с раком молочной железы, чем у субъектов, не пораженных раком молочной железы. Таким образом, данные маркеры и/или гаплотипы имеют прогностическое значение для диагностики рака молочной железы. В одном воплощении маркеры или гаплотипы, имеющие прогностическое значение для выявления рака молочной железы, включают один или более маркеров, выбранных из группы, состоящей из маркеров по табл. 10, 15, 19, 20, 21 и 22. В другом воплощении маркеры, имеющие прогностическое значение для диагностики рака молочной железы, выбраны из маркеров rs4848543, rs13387042 и rs3803662 и маркеров в неравновесии по сцеплению с ними.

В другом воплощении аллели маркеров, имеющих прогностическое значение для диагностики рака молочной железы, включают аллель A rs4848543, аллель A rs13387042 и аллель T rs3803662 и маркеры в неравновесии по сцеплению с ними.

Таким образом, описываемые здесь способы могут применяться для обследования образца, взятого у субъекта, на наличие или отсутствие определенного аллеля маркера или гаплотипа; наличие или отсутствие определенного аллеля маркера или гаплотипа указывает на подверженность раку молочной железы.

В одном воплощении изобретения диагностика рака молочной железы, ассоциированного с полиморфизмами и/или гаплотипами по данному изобретению, осуществляется путем детекции полиморфизмов по данному изобретению. Здесь описываются конкретные полиморфизмы (см, например, табл. 10, 15, 19, 20, 21 и 22 и маркеры в неравновесии по сцеплению с ними). Тестируемый образец геномной ДНК, РНК или кДНК получают у субъекта, имеющего рак молочной железы, для определения, связано ли заболевание с одним или более полиморфизмами по изобретению. Затем образец ДНК, РНК или кДНК исследуется для определения, присутствует ли в образце определенный аллель полиморфизма или определенный гаплотип по изобретению. Если обнаруживается, что образец нуклеиновой кислоты содержит определенный аллель полиморфизма или гаплотип, то наличие аллеля или гаплотипа указывает на то, что рак молочной железы ассоциирован с полиморфизмом и/или гаплотипом.

Для детекции полиморфизма могут применяться способы, известные специалистам, включающие методы гибридизации, такие как Саузерн-анализ, Нозерн-анализ, количественную ПЦР, гибридизацию *in situ*, рестрикционное расщепление или анализ последовательности, но не ограничивающиеся ими. Диагностика рака молочной железы, ассоциированного с полиморфизмами по изобретению, может также проводиться путем изучения экспрессии и/или состава полипептида, кодируемого нуклеиновой кислотой, ассоциированной с полиморфизмами по изобретению, при помощи различных способов, включая

иммуноферментный анализ (ИФА), Вестерн блоттинг, иммунопреципитацию и иммунофлуоресценцию, или другие способы, известные специалистам.

Терапевтические агенты

Варианты по данному изобретению (например, маркеры и/или гаплотипы по изобретению, например, маркеры, приведенные в любой из табл. 10, 15 и 19), могут применяться для идентификации новых терапевтических мишеней при раке молочной железы. Например, гены, содержащие варианты или находящиеся в неравновесии по сцеплению с вариантами (маркерами и/или гаплотипами), ассоциированными с раком молочной железы, или их продукты (например, один или более из генов STEAP3/TSAP6, LOC643714 и/или TNRC9), а также гены или их продукты, прямо или косвенно регулируемые этими вариантами генов или их продуктами, или взаимодействующие с ними, могут быть мишенями для разработки терапевтических агентов для лечения рака молочной железы. Терапевтические агенты могут включать одну или более молекул, например, малых молекул, не относящихся к белкам и нуклеиновым кислотам, белков, пептидов, белковых фрагментов, нуклеиновых кислот (ДНК, РНК), ПНК (пептидных нуклеиновых кислот) или их производных или миметиков, способных модулировать функцию и/или уровень генов-мишеней или их генных продуктов.

Нуклеиновые кислоты и/или варианты по изобретению или нуклеиновые кислоты, содержащие комплементарную им последовательность, могут применяться как анти-сенс конструкции для контроля экспрессии генов в клетках, тканях или органах. Методики, связанные с техникой анти-сенс, хорошо известны специалистам и их описание и обзор приводятся в *AntisenseDrug Technology: Principles, Strategies, u Applications*, Crooke, ed., Marcel Dekker Inc., New York (2001). В общем, анти-сенс молекулы нуклеиновых кислот создаются комплементарными к участку мРНК, экспрессируемой геном, так что анти-сенс молекула гибридизуется с мРНК и таким образом блокирует трансляцию белка мРНК. Специалистам известно несколько классов анти-сенс олигонуклеотидов, включая расщепляющие и блокирующие. Первые связываются с сайтами-мишенями РНК, активируют внутриклеточные нуклеазы (например, Рнказу Н или Рнказу L), которые расщепляют РНК-мишень. Блокаторы связываются с РНК-мишенью, ингибируют трансляцию белка путем стерического блокирования рибосом. Примерами блокаторов могут быть нуклеиновые кислоты, морфолиновые соединения, закрытые нуклеиновые кислоты и метилфосфонаты (Thompson, *Drug Discovery Today*, 7:912-917 (2002)). Анти-сенс олигонуклеотиды применяются непосредственно в качестве терапевтических агентов, а также применяются для определения и подтверждения функции гена, например, в экспериментах с нокаутом гена или нокдауном гена. Техника анти-сенс также описана Lavery et al., *Curr. Opin. Drug Discov. Devel.* 6: 561-569 (2003), Stephens et al., *Curr. Opin. Mol. Ther.* 5:118-122 (2003), Kurreck, *Eur. J. Biochem.* 270: 1628-44 (2003), Dias et al., *Mol. Cancer Ter.* 1: 347-55 (2002), Chen, *Methods Mol. Med.* 75:621-636 (2003), Wang et al., *Curr. Cancer Drug Targets* 1: 177-96 (2001) и Bennett, *Antisense Nucleic Acid Drug. Dev.* 12: 215-24 (2002).

Описываемые здесь варианты могут использоваться для выбора и создания анти-сенс реагентов, специфических для конкретных вариантов. Используя информацию об описываемых здесь вариантах, можно создать анти-сенс олигонуклеотиды или другие анти-сенс молекулы, специфически взаимодействующие с молекулами мРНК, содержащими один или более вариантов по данному изобретению. В этом случае экспрессия молекул мРНК, содержащих один или более вариантов по данному изобретению (маркеры и/или гаплотипы), может быть подавлена или блокирована. В одном воплощении анти-сенс молекулы созданы для специфического связывания определенной аллельной формы (т.е. одного или нескольких вариантов (аллелей и/или гаплотипов)) нуклеиновой кислоты - мишени, и, таким образом, подавления трансляции продукта, производимого этим определенным аллелем или гаплотипом, без связывания других или альтернативных вариантов в специфических полиморфных сайтах молекулы нуклеиновой кислоты - мишени.

Поскольку молекулы анти-сенс могут применяться для инактивации мРНК для ингибирования экспрессии генов и, следовательно, экспрессии белков, молекулы могут использоваться для лечения рака молочной железы. Методика может включать расщепление с помощью рибозимов, содержащих нуклеотидные последовательности, комплементарные к одному или более участкам мРНК, что изменяет способность мРНК к трансляции. Такие участки мРНК включают, например, белок-кодирующие участки, в частности белок-кодирующие участки, соответствующие каталитической активности, сайтам связывания субстрата и/или лиганда, или другим функциональным доменам белка.

Феномен РНК интерференции (RNAi) активно изучался на протяжении последних 10 лет с момента его первоначального открытия у *C. elegans* (Fire et al., *Nature* 391: 806-11 (1998)), и в последние годы активно обсуждалось его возможное применение в лечении заболеваний человека (см. обзор Kim & Rossi, *Nature Rev. Genet.* 8: 173-204 (2007)). РНК интерференция (RNAi), также обозначаемая сайленсингом генов, основана на применении двунитевых молекул РНК (dsRNA) для выключения определенных генов. В клетке молекулы цитоплазматической двунитевой РНК (dsRNA) подвергаются процессингу клеточными комплексами в малые интерферирующие РНК (siRNA, от англ. small interfering RNA). siRNA направляют комплексы протеин-РНК к специфическим сайтам на мРНК-мишени, приводя к расщеплению мРНК (Thompson, *Drug Discovery Today*, 7:912-917 (2002)). Молекулы siRNA обычно бывают длиной 20, 21, 22 или 23 нуклеотида. Таким образом, один аспект изобретения касается изолированных молекул

нуклеиновых кислот и применения этих молекул для РНК интерференции, т.е. в качестве малых интерферирующих молекул РНК (siRNA). В одном воплощении изолированные молекулы нуклеиновых кислот имеют размер 18-26 нуклеотидов, предпочтительно 19-25 нуклеотидов, более предпочтительно 20-24 нуклеотида, и более предпочтительно 21, 22 или 23 нуклеотида.

Другой путь RNAi-опосредованного подавления генов берет начало от эндогенно кодируемых первичных транскриптов микро РНК (primary microRNA, pri-miRNA), которые подвергаются процессингу в клетках и образуют предшественник miRNA (precursor, pre-miRNA). Эти молекулы miRNA транспортируются из ядра в цитоплазму, где они подвергаются процессингу и образуют зрелые молекулы miRNA (miRNA), которые управляют ингибированием процесса трансляции путем распознавания сайтов-мишеней в 3' нетранслируемых участках мРНК, и последующей деградации мРНК Р-тельцами (см. обзор Kim & Rossi, *Nature Rev. Genet.* 8: 173-204(2007)).

Клинические применения RNAi включают получение искусственно синтезированных дуплексов siRNA, которые предпочтительно имеют размер 20-23 нуклеотидов и предпочтительно имеют выступающие 3' концы в 2 нуклеотида. Нокдаун экспрессии генов вызывают за счет создания последовательности, специфической к мРНК-мишени. Специалистам известно несколько коммерческих сайтов для оптимизации разработки и синтеза таких молекул.

Другие применения связаны с более длинными молекулами siRNA (обычно 25-30 нуклеотидов длиной, предпочтительно около 27 нуклеотидов), а также малыми РНК, образующими шпильку (small hairpin RNAs, shRNAs; обычно около 29 нуклеотидов длиной). Последние экспрессируются естественным путем, что описано Amarzguioui et al. (*FEBS Lett.* 579:5974-81 (2005)). Химически синтезированные siRNAs и shRNAs являются субстратами для процессинга *in vivo* и в некоторых случаях обеспечивают более сильное подавление генов по сравнению с более короткими аналогами (Kim et al., *Nature Biotechnol.* 23:222-226 (2005); Siolas et al., *Nature Biotechnol.* 23:227-231 (2005)). Как правило, siRNAs обеспечивают транзитное подавление экспрессии генов, поскольку их внутриклеточная концентрация снижается при последующих делениях клетки. Напротив, экспрессируемые shRNAs опосредуют длительно существующий и стабильный нокдаун транскриптов-мишеней, поскольку имеет место транскрипция shRNA (Marques et al., *Nature Biotechnol.* 23: 559-565 (2006); Brummelkamp et al., *Science* 296: 550-553 (2002)).

Поскольку функционирование молекул RNAi, включая siRNA, miRNA и shRNA является зависимым от последовательности, варианты по данному изобретению (например, маркеры или гаплотипы, приведенные в табл. 10, 15 и 19) могут быть использованы для создания реагентов RNAi, распознающих определенные молекулы нуклеиновых кислот, содержащих определенные аллели и/или гаплотипы (например, аллели и/или гаплотипы по данному изобретению), при этом не распознающих молекулы нуклеиновых кислот, содержащих другие аллели или гаплотипы. Эти реагенты RNAi могут таким образом распознавать и разрушать молекулы нуклеиновых кислот - мишеней. Как и в случае реагентов анти-сенс, реагенты RNAi могут применяться в качестве терапевтических агентов (например, для выключения генов, ассоциированных с заболеванием или вариантов генов, ассоциированных с заболеванием), но могут также применяться для характеристики или подтверждения функции гена (например, в экспериментах с нокаутом или нокдаун генов).

Введение RNAi в клетку может осуществляться с помощью ряда методик, известных специалистам. Способы с применением невирусного внесения включают холестерин, стабилизированные микрочастицы, содержащие комплекс коротких интерференционных рибонуклеиновых кислот и липидов (SNALP, от англ. stable nucleic-acid-lipid particle), фрагменты тяжелых цепей антител (Fab), аптамеры и наночастицы. Способы вирусного внесения включают применение лентивирусов, аденовирусов и аденоассоциированных вирусов. Молекулы siRNA в некоторых воплощениях являются химически модифицированными для повышения их стабильности. Сюда входят модификации по положению 2' рибозы, включая 2'-0-метилпурины и 2'-флуоропиримидины, которые обеспечивают устойчивость к рибонуклеазной активности. Возможны и другие химические модификации, и они известны специалистам.

Обзор информации по RNAi и возможностям направленного воздействия на специфические гены с помощью RNAi можно получить в следующих источниках: Kim & Rossi, *Nat. Rev. Genet.* 8: 173-184 (2007), Chen & Rajewsky, *Nat. Rev. Genet.* 8: 93-103 (2007), Reynolds et al., *Nat. Biotechnol.* 22: 326-330 (2004), Chi et al., *Proc. Natl. Acad. Sci. USA* 100: 6343-6346 (2003), Vickers et al., *J. Biol. Chem.* 278: 7108-7118 (2003), Agami *Curr. Opin. Chem. Biol.* 6: 829-834 (2002), Lavery et al., *Curr. Opin. Drug Discov. Devel.* 6:561-569 (2003), Shi, *Trends Genet.* 19:9-12 (2003), Shuey et al., *Drug Discov. Today* 7:1040-46 (2002), McManus et al., *Nat. Rev. Genet.* 3:737-747 (2002), Xia et al., *Nat. Biotechnol.* 20:1006-10 (2002), Plasterk et al., *curr. Opin. Genet. Dev.* 10:562-7 (2000), Boshier et al., *Nat. Cell Biol.* 2: E31-6 (2000) и Hunter, *Curr. Biol.* 9: R440-442(1999).

Генетический дефект, приводящий к повышенной предрасположенности или риску развития рака молочной железы, или генетический дефект, вызывающий рак молочной железы, может подвергаться постоянной коррекции путем введения субъекту, являющемуся носителем дефекта, фрагмента нуклеиновой кислоты, несущей репарирующую последовательность, доставляющей нормальный нуклеотид (нуклеотиды), т.е. нуклеотид (нуклеотиды) дикого типа к участку генетического дефекта. Такая сайт-специфическая репарирующая последовательность может содержать олигонуклеотид РНК/ДНК, который

способствует эндогенному исправлению повреждений геномной ДНК субъекта. Введение репарирующей последовательности может проводиться с применением соответствующего носителя, например, в комплексе с полиэтиленгликолем, в инкапсулированном в анионные липосомы виде, с помощью вирусного вектора, например, аденовирусного вектора, или других фармацевтических композиций, пригодных для облегчения поступления внутрь клетки вводимой нуклеиновой кислоты. В этом случае генетический дефект может быть исправлен, поскольку химерные олигонуклеотиды запускают инкорпорирование нормальной последовательности в геном субъекта, приводя к экспрессии нормального генетического продукта (дикого типа). Замещение распространяется, таким образом обеспечивая постоянное исправление дефекта и облегчение симптомов, ассоциирующихся с заболеванием или состоянием.

Данное изобретение обеспечивает способы идентификации соединений или агентов, которые могут использоваться для лечения и/или предупреждения рака молочной железы. Так, варианты по изобретению могут применяться как мишени для поиска и/или разработки терапевтических агентов. В определенных воплощениях такие способы включают оценку способности агента или соединения модулировать активность и/или экспрессию нуклеиновой кислоты, включающей по меньшей мере один из вариантов (маркеров и/или гаплотипов) по данному изобретению, или продукт, кодируемый нуклеиновой кислотой, например, один или более из генов STEAP3/TSAP6, LOC643714 и/или TNRC9. Это в свою очередь может быть использовано для поиска агентов или соединений, подавляющих или нарушающих нежелательную активность или экспрессию продукта, кодируемого нуклеиновой кислотой. Как известно специалистам, тесты для проведения таких экспериментов могут проводиться в клеточных системах или бесклеточных системах. Клеточные системы включают клетки, которые естественным образом экспрессируют молекулы нуклеиновой кислоты, представляющей интерес, или рекомбинантные клетки, генетически модифицированные для экспрессии молекул определенной требуемой нуклеиновой кислоты.

Экспрессия варианта гена у пациента может оцениваться по экспрессии вариант-содержащей последовательности нуклеиновой кислоты (например, гена, содержащего по меньшей мере один вариант по данному изобретению, который может транскрибироваться в РНК, содержащую по меньшей мере один вариант и, в свою очередь, транслироваться в белок), или по изменению экспрессии нормальной последовательности (дикого типа) нуклеиновой кислоты из-за влияния вариантов на уровень или профиль экспрессии нормальных транскриптов, например, вариантов в регуляторных или контролируемых участках гена. Исследования экспрессии генов включают прямые исследования нуклеиновых кислот (мРНК), исследования уровня экспрессируемых белков, или исследования побочных продуктов, задействованных в пути, например, в сигнальном пути. Кроме того, можно оценивать экспрессию генов, которая претерпевает положительную или отрицательную регуляцию в ответ на сигнальный путь. Одно воплощение включает оперативное связывание репортерного гена, например, люциферазы с регуляторным участком гена (генов), представляющего интерес.

Модуляторы экспрессии генов в одном воплощении могут идентифицироваться, при контактировании клетки с соединением или агентом-кандидатом, и определении экспрессии мРНК. Уровень экспрессии мРНК в присутствии соединения или агента-кандидата сравнивается с уровнем экспрессии в отсутствии соединения или агента. На основании такого сравнения можно идентифицировать кандидатные соединения или агенты для лечения и/или предупреждения таких нарушений, как фибрилляция предсердий, трепетание предсердий и инсульт, по их способности модулировать генетическую экспрессию варианта гена. Если экспрессия мРНК или кодируемого белка статистически достоверно выше в присутствии соединения или агента-кандидата, чем в его отсутствие, то соединение или агент-кандидат считаются стимуляторами или положительными регуляторами экспрессии нуклеиновой кислоты. Если экспрессия нуклеиновой кислоты или уровень белка статистически достоверно ниже в присутствии соединения или агента - кандидата, чем в его отсутствие, то соединение или агент-кандидат считаются ингибиторами или отрицательными регуляторами экспрессии нуклеиновой кислоты.

Изобретение также обеспечивает способы лечения с применением соединения, идентифицированного при скрининге лекарств (соединений и/или агентов) в качестве модулятора гена (т.е. стимулятора и/или ингибитора экспрессии гена).

Способы оценки вероятности ответа на терапевтические агенты, способы мониторинга хода лечения и способы лечения рака молочной железы

Как известно, индивиды могут иметь различные ответы на определенное лечение (например, на терапевтический агент или способ лечения согласно данному описанию). Различия в ответах могут быть частично основаны на генетической детерминированности. Фармакогеномика занимается изучением того, как генетические вариации (например, варианты (маркеры и/или гаплотипы) по данному изобретению) влияют на ответ на лекарственные препараты, по причине изменения фармакокинетики и/или нарушения или изменения действия лекарственного препарата. Следовательно, в своей основе различия в ответах частично могут быть генетически детерминированы. Клинические исходы благодаря наличию генетических вариаций, оказывающих влияние на ответ на лекарства, могут выражаться в виде токсического эффекта препарата у некоторых индивидов (например, у носителей или лиц, не являющихся носителями генетических вариантов по данному изобретению), либо терапевтической неэффективности препарата. Следовательно, варианты по данному изобретению могут детерминировать характер воздействия

терапевтического агента на организм или способ метаболизма терапевтического агента в организме.

Соответственно, в одном воплощении наличие определенного аллеля полиморфного сайта или гаплотипа указывает на различную степень ответа на определенный способ лечения. Это означает, что пациент с диагностированным раком молочной железы и являющийся носителем определенного аллеля полиморфного сайта или гаплотипа по данному изобретению (например, связанного с риском аллеля и/или гаплотипа по изобретению) будет отвечать лучше или хуже на определенное терапевтическое средство, лекарственный препарат и/или другое терапевтическое мероприятие, используемое для лечения заболевания. Таким образом, наличие или отсутствие аллеля маркера или гаплотипа может помочь в принятии решения, какое лечение нужно применить для пациента. Например, для вновь диагностированного пациента можно оценить наличие маркера или гаплотипа по данному изобретению (например, по анализу ДНК, полученной из образца крови, согласно данному описанию). Если пациент позитивен в отношении аллеля маркера или гаплотипа (что означает наличие по меньшей мере одного определенного аллеля маркера или гаплотипа), то врач назначает один способ лечения (например, один определенный терапевтический агент или комбинацию терапевтических агентов), а если пациент негативен в отношении по меньшей мере одного аллеля маркера или гаплотипа, то может быть назначен другой курс лечения (который может подразумевать рекомендации вести мониторинг прогрессирования заболевания без начала незамедлительного лечения). Таким образом, статус носительства пациента может помочь в определении того, следует ли применять конкретный способ лечения. Ценность такого подхода заключается в возможности диагностировать заболевание на ранней стадии, выбрать наиболее подходящее лечение и дать врачу информацию о прогнозе/агрессивности заболевания для выбора наиболее подходящего лечения.

Согласно следующему ниже описанию существующие в клинике варианты профилактики рака молочной железы в основном представляют собой хемотрепивную терапию (химиотерапию или гормональную терапию) и превентивную хирургию. Наиболее распространенными хемотрепивными агентами являются тамоксифен и ралоксифен; другие варианты включают ингибиторы ароматазы. Варианты лечения также включают радиотерапию, при которой часть пациентов испытывает неблагоприятные симптомы. Маркеры по изобретению согласно данному описанию могут применяться для оценки ответа на данные варианты лечения или для прогнозирования хода лечения при помощи одного из таких вариантов лечения. Таким образом, определение генетического профиля может применяться для выбора соответствующей тактики лечения, основанной на генетическом статусе индивида, или может применяться для прогнозирования исхода определенного варианта лечения и, таким образом, применяться при стратегическом выборе вариантов лечения или комбинации доступных вариантов лечения.

Данное изобретение также касается способов мониторинга эффективности лечения рака молочной железы. Это может быть осуществлено на основании генотипического и/или гаплотипического статуса маркеров и гаплотипов по данному изобретению, т.е. путем оценки отсутствия или наличия по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера согласно данному описанию или путем мониторинга экспрессии генов, ассоциированных с вариантами (маркерами и гаплотипами) по данному изобретению. мРНК связанного с риском гена или кодируемый полипептид могут быть измерены в образце ткани (например, образце периферической крови или биопсийном образце). Уровни экспрессии и/или уровни мРНК могут, таким образом, быть определены до и во время лечения для мониторинга его эффективности. Кроме того, до и во время лечения для мониторинга его эффективности проводится определение генотипического и/или гаплотипического статуса по меньшей мере одного варианта, связанного с риском рака молочной железы, как представлено здесь.

Кроме того, биологические цепочки и метаболические пути, относящиеся к маркерам и гаплотипам по данному изобретению, можно отслеживать путем определения уровня мРНК и/или полипептидов. Это может быть осуществлено, например, путем отслеживания уровня экспрессии или уровня полипептидов по нескольким генам, принадлежащим цепочке и/или пути, в образцах, взятых до и во время лечения. Кроме того, до и во время лечения можно определять метаболиты, принадлежащие биологической цепочке или метаболическому пути. Эффективность лечения определяется путем сравнения наблюдаемых во время лечения изменений в уровнях экспрессии/уровнях метаболитов с соответствующими показателями для здоровых субъектов.

В дальнейшем аспекте маркеры по данному изобретению могут использоваться для повышения статистической чувствительности и эффективности клинических испытаний. Так, индивиды, являющиеся носителями связанных с риском вариантов по данному изобретению, т.е. индивиды, являющиеся носителями по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера, обуславливающего повышенный риск развития рака молочной железы, могут с большей вероятностью отвечать на определенные способы лечения. В одном воплощении индивиды, являющиеся носителями связанных с риском вариантов гена (генов) пути и/или метаболической цепочки, на которые направлено определенное лечение (например, препарат малых молекул), с большей вероятностью будут отвечать на лечение. В другом воплощении индивиды, являющиеся носителями связанных с риском вариантов гена, экспрессия и/или функция которого изменена связанным с риском вариантом, с большей вероятностью будут отвечать на способ лечения, направленный на данный ген, его экспрессию или его генный продукт.

В дальнейшем аспекте маркеры и гаплотипы по данному изобретению могут применяться для прицельного выбора фармацевтических агентов для отдельных индивидов. Персонализированный выбор схем лечения, изменения образа жизни или комбинация обоих могут выполняться с учетом связанных с риском вариантов по данному изобретению. Таким образом, знание статуса индивида по определенным маркерам по данному изобретению может использоваться для выбора вариантов лечения, направленно воздействующих на гены или генные продукты, находящиеся под влиянием связанных с риском вариантами по изобретению. Определенные комбинации вариантов могут подходить для выбора одних способов лечения, тогда как на другие комбинации вариантов генов могут обуславливать другие способы лечения. Подобные комбинации вариантов могут включать один вариант, два варианта, три варианта или четыре или более вариантов, в зависимости от необходимости выбора схемы лечения с клинически надежной точностью.

Аспекты, связанные с компьютерными технологиями

Данное изобретение также относится к связанному с компьютерными технологиями применению полиморфных маркеров и гаплотипов, которые по данному описанию ассоциированы с раком молочной железы. Данные приложения могут применяться для хранения, обработки или же анализа данных генотипа, что может использоваться в способах по изобретению, согласно данному описанию. Один пример относится к хранению информации о генотипе, полученной у индивида, на компьютерных носителях, таким образом, что информация о генотипе становится доступна третьим лицам (например, индивиду) или получения информации о данных генотипа, например, путем сравнения данных генотипа с информацией о генетических факторах риска, которые способствуют повышенной подверженности раку молочной железы и представления результатов, основанных на подобном сравнении.

Один из таких аспектов относится к компьютерным носителям. В общих терминах подобные носители способны хранить (1) информацию, позволяющую выявить по меньшей мере один полиморфный маркер или гаплотип; (2) индикатор частоты по меньшей мере одного аллеля по меньшей мере одного указанного маркера или частоты гаплотипа у индивида с раком молочной железы; и индикатор частоты по меньшей мере одного аллеля по меньшей мере одного указанного маркера или частоты гаплотипа в референсной популяции. Референсной популяцией может быть популяция индивидов, не имеющих заболевания. Кроме того, референсной популяцией может быть случайная выборка из общей популяции, таким образом являющаяся репрезентативной для популяции в целом. Индикатором частоты может быть расчетная частота, частота аллелей и/или копий гаплотипов, или нормализованные или иным образом обработанные значения реальных частот, соответствующие определенному носителю.

Дополнительная информация об индивидах может храниться на носителях, например, информация о происхождении, информация о поле, физических признаках или характеристиках (включая рост и вес), биохимических показателях, проводимом лечении, исходе лечения, лекарственных препаратах или другой важной информации, которую желательно хранить или обрабатывать в контексте генотипического статуса определенного индивида.

Изобретение также относится к устройствам, которые могут использоваться для определения или обработки генетических данных с целью определения подверженности человеческого индивида раку молочной железы. Подобные устройства могут включать компьютерную память, программы для обработки данных, хранящихся в компьютерной памяти и программы для генерирования выходных данных, включающих показатели генетических данных. Подобные показатели могут включать такие значения, как частоты аллелей или гаплотипов, частоты генотипов, пол, возраст, информацию о фенотипе, значения отношения шансов (ОШ) или относительного риска (ОР), популяционный атрибутивный риск (ПАР) или иную важную информацию, которая представляет собой либо статистическую обработку исходных данных генотипа или основывается на расчетах с использованием генетических данных.

Маркеры и гаплотипы, которые, как здесь показано, ассоциированы с повышенной подверженностью (например, повышенным риском) раку молочной железы, в некоторых воплощениях могут применяться для интерпретации и/или анализа данных генотипа. Так, в некоторых воплощениях выявление аллеля или гаплотипа, связанного с риском рака молочной железы, как здесь показано, или аллеля полиморфного маркера или гаплотипа, находящегося в неравновесии по сцеплению с любым из маркеров и/или гаплотипов, которые, как здесь показано, ассоциированы с раком молочной железы, указывает на то, что индивид, у которого получены генетические данные, имеет повышенный риск рака молочной железы. В одном подобном воплощении генетические данные получены по меньшей мере для одного полиморфного маркера, который, как здесь показано, ассоциирован с раком молочной железы, или маркера, находящегося в неравновесии по сцеплению с ним. Генетические данные впоследствии становятся доступными для третьей стороны, например, для индивида, у которого получены эти данные, например, через пользовательский интерфейс, к которому можно получить доступ через Интернет, наряду с интерпретацией генетических данных, например в форме показателей риска (например, абсолютного риска (АР), относительного риска (ОР) или отношения шансов (ОШ)) заболевания (например, рака молочной железы). В другом воплощении проводится оценка связанных с риском маркеров, идентифицированных в массиве данных генотипа, полученного у индивида, и результаты оценки риска, обусловленного наличием в массиве данных подобных вариантов, связанных с риском, становятся доступными индивиду,

например, через надежный веб-интерфейс или с применением иных средств коммуникации. Результаты подобной оценки риска могут быть представлены в численном виде (например, в виде значений риска, таких как абсолютный риск, относительный риск и/или отношение шансов или в виде процентного увеличения риска по сравнению с референсным значением), в графическом виде или в другом виде, подходящем для иллюстрации риска для индивида, у которого получены генетические данные. В конкретных воплощениях результаты оценки риска становятся доступными для третьих лиц, например, врачей, других медицинских сотрудников или генетических консультантов.

Маркеры, применяемые в различных аспектах изобретения

Все вышеописанные способы и приложения могут осуществляться с маркерами и гаплотипами по изобретению, которые были более подробно описаны здесь в общих терминах в качестве подходящих для оценки подверженности раку молочной железы. Так, подобные приложения могут в целом осуществляться на практике с применением маркеров, приведенных в любой из табл. 10, 15, 19, 20, 21 и 22, и маркеров, находящихся в неравновесии по сцеплению с ними. В определенных воплощениях маркеры или гаплотипы присутствуют в составе сегментов генома, последовательности которых приведены в SEQ ID NO: 4, SEQ ID NO: 5 или SEQ ID NO: 6. В определенных воплощениях маркеры находятся в составе блока HC rs4848543, блока HC rs13387042, или блока HC rs3803662. В определенных воплощениях маркеры выбраны из маркеров, приведенных в табл. 10. В других воплощениях маркеры выбраны из маркеров, приведенных в табл. 15. В других воплощениях маркеры выбраны из маркеров, приведенных в табл. 19. В некоторых воплощениях маркеры выбраны из маркеров, приведенных в табл. 7 и 8. В некоторых других воплощениях маркеры выбраны из маркеров, приведенных в табл. 14. В некоторых других воплощениях маркеры выбраны из маркеров, приведенных в табл. 18. В других воплощениях маркеры являются одним из rs4848543 (SEQ ID NO:1), rs13387042 (SEQ ID NO:2) или rs3803662 (SEQ ID NO:3), возможно включая маркеры в неравновесии по сцеплению с ними. В одном воплощении маркеры являются одним из rs4848543 (SEQ ID NO:1), rs13387042 (SEQ ID NO:2) или rs3803662 (SEQ ID NO:3). В определенных предпочтительных воплощениях маркером является rs4848543 (SEQ ID NO:1). В других предпочтительных воплощениях маркером является rs13387042 (SEQ ID NO:2). В других предпочтительных воплощениях маркером является rs3803662 (SEQ ID NO:3). В определенных воплощениях неравновесие по сцеплению определяется численными значениями r^2 выше 0,2. В другом воплощении маркеры или гаплотипы включают по меньшей мере один маркер, выбранный из аллеля A rs4848543, аллеля A rs13387042 и аллеля T rs3803662.

Нуклеиновые кислоты и полипептиды

Описываемые здесь нуклеиновые кислоты и полипептиды могут применяться в способах и наборах данного изобретения, как описано выше. "Выделенная" молекула нуклеиновой кислоты в данном описании представляет собой молекулу нуклеиновой кислоты, отделенную от нуклеиновых кислот, обычно фланкирующих ген или нуклеотидную последовательность (как в геномных последовательностях) и/или полностью или частично очищенную от других транскрибируемых последовательностей (например, как в библиотеке РНК). Например, выделенная нуклеиновая кислота по изобретению может быть фактически выделена из сложной клеточной среды, в которой она появляется естественным путем, или из культуральной среды, где она производится с помощью рекомбинантных технологий, или из химических предшественников или других химических соединений, если проводится ее химический синтез. В некоторых случаях выделенный материал может быть составной частью смеси (например, грубого экстракта, содержащего другие соединения), буферной системы или смеси реагентов. В других случаях материал может быть очищен до необходимой гомогенности, что, например, определяется при электрофорезе в полиакриламидном геле (ПААГ) или колоночной хроматографии (например, ВЭЖХ). Выделенная молекула нуклеиновой кислоты по определению может включать по меньшей мере около 50%, по меньшей мере около 80% или по меньшей мере около 90% (в молярном отношении) всех присутствующих видов макромолекул. В отношении геномной ДНК термин «выделенная» может также означать молекулы нуклеиновой кислоты, отделенные от хромосомы, с которой естественным образом связывается геномная ДНК. Например, выделенная молекула нуклеиновой кислоты может содержать меньше чем примерно 250, 200, 150, 100, 75, 50, 25, 10, 5, 4, 3, 2, 1, 0,5 или 0,1 кб нуклеотидов, фланкирующих молекулу нуклеиновой кислоты в геномной ДНК клетки, из которой получена молекула нуклеиновой кислоты.

Молекула нуклеиновой кислоты может быть соединена с другими кодирующими или регуляторными последовательностями, но тем не менее считается выделенной. Так, рекомбинантная ДНК, содержащаяся в векторе, подходит под определение "выделенная" согласно данному описанию. Кроме того, молекулы выделенной нуклеиновой кислоты включают молекулы рекомбинантной ДНК в гетерологичных клетках хозяина или гетерологичных организмах, а также частично очищенные или практически очищенные молекулы ДНК в растворе. Молекулы "выделенной" нуклеиновой кислоты также включают РНК-транскрипты молекул ДНК по данному изобретению, полученные *in vivo* и *in vitro*. Молекулы выделенной нуклеиновой кислоты или нуклеотидной последовательности могут включать молекулы нуклеиновой кислоты или последовательности нуклеотидов, синтезированные химическим путем или с помощью рекомбинантных технологий. Такие выделенные нуклеотидные последовательности могут применяться, например, в производстве кодируемых полипептидов, в качестве зондов для выделения гомо-

логичных последовательностей (например, из других видов млекопитающих), для генного картирования (например, с помощью *in situ* гибридизации с хромосомами) или для определения экспрессии генов в тканях (например, в человеческих тканях), например, с помощью Нозерн анализа или других гибридизационных технологий.

Изобретение также относится к молекулам нуклеиновой кислоты, которые гибридизуются в жестких условиях гибридизации, например, как в случае селективной гибридизации, с описываемыми здесь нуклеотидными последовательностями (например, молекулам нуклеиновой кислоты, специфически гибридизующимся с нуклеотидными последовательностями, содержащими полиморфный сайт, ассоциированный с описываемым здесь маркером или гаплотипом). Такие молекулы нуклеиновой кислоты могут быть детектированы и/или выделены путем аллель- или последовательность-специфической гибридизации (например, в жестких условиях). Жесткие условия гибридизации и способы гибридизации нуклеиновых кислот хорошо знакомы специалистам (см., например, *Current Protocols in Molecular Biology*, Ausubel F. et al., John Wiley & Sons, (1998) и Kraus M. и Aaronson S., *Methods Enzymol.*, 200: 546-556 (1991), описания которых включены сюда во всей полноте путем ссылки.

Процент идентичности двух нуклеотидных или аминокислотных последовательностей может быть определен при выравнивании последовательностей с целью удобства сравнения (например, в изображении первой последовательности можно вставить пробелы). Затем сравниваются нуклеотиды или аминокислоты в соответствующих позициях и процент идентичности между двумя последовательностями рассчитывается как функция количества идентичных позиций для обеих последовательностей (т.е. % идентичности = количество идентичных позиций/общее количество позиций $\times$ 100). В некоторых воплощениях длина последовательностей, выровненных с целью их сравнения, может быть по меньшей мере 30%, по меньшей мере 40%, по меньшей мере 50%, по меньшей мере 60%, по меньшей мере 70%, по меньшей мере 80%, по меньшей мере 90% или по меньшей мере 95% от длины референсной последовательности. Реальное сравнение двух последовательностей может выполняться при помощи общеизвестных способов, например, с применением математического алгоритма. Нелимитирующий пример такого математического алгоритма описан Karlin S. и Altschul S., *Proc. Natl. Acad. Sci. USA*, 90: 5873-5877 (1993). Такой алгоритм встроен в программы NBLAST и XBLAST (версия 2.0), как описано Altschul S. et al., *Nucleic Acids Res.*, 25: 3389-3402 (1997). При использовании программ BLAST и Gapped BLAST могут использоваться параметры по умолчанию соответствующих программ (например, NBLAST). См. Веб-сайт в интернете по адресу ncbi.nlm.nih.gov. В одном воплощении параметры для сравнения последовательностей могут быть зафиксированы: score=100, wordlength=12, или варьироваться (e.g., W=5 or W=20).

Другие примеры включают алгоритм Myers и Miller, CABIOS (1989), ADVANCE и ADAM, как описано у Torellis, A. и Robotti, C, *Comput. Appl. Biosci.* 10:3-5 (1994); и FASTA, как описано у Pearson, W. и Lipman, D., *Proc. Natl. Acad. Sci. USA*, 85:2444-48 (1988). В другом воплощении процент идентичности между двумя аминокислотными последовательностями может быть определен с помощью программы GAP в пакете программного обеспечения GCG (Accelrys, Cambridge, UK).

Данное изобретение также относится к молекулам выделенной нуклеиновой кислоты, содержащим фрагмент или часть, гибридизующуюся в жестких условиях с нуклеиновой кислотой, которая включает или состоит из нуклеотидной последовательности, содержащей полиморфные маркеры, приведенные в табл. 10, 15, 19, 20, 21 и 22, и маркеры в неравновесии по сцеплению с ними, и нуклеотидной последовательности генов STEAP3/TSAP6, LOC643714 и TNRC9 или нуклеотидной последовательности, которая включает или состоит из нуклеотидной последовательности, комплементарной нуклеотидной последовательности, включающей полиморфные маркеры, приведенные в табл. 10, 15, 19, 20, 21 и 22, и маркеры в неравновесии по сцеплению с ними и нуклеотидной последовательности генов STEAP3/TSAP6, LOC643714 и TNRC9 или их фрагментов, где нуклеотидная последовательность включает по меньшей мере один полиморфный аллель, содержащийся в описываемых здесь маркерах и гаплотипах. Фрагменты нуклеиновых кислот по изобретению имеют размер по меньшей мере около 15, по меньшей мере около 18, 20, 23 или 25 нуклеотидов, и могут быть размером 30, 40, 50, 100, 200, 500, 1000, 10000 нуклеотидов или более.

Фрагменты нуклеиновых кислот по изобретению применяются как зонды или праймеры в описываемых здесь тестах. "Зонды" или "праймеры" являются олигонуклеотидами, которые гибридизуются специфическим по отношению к основаниям образом с комплементарной цепью молекулы нуклеиновой кислоты. Кроме ДНК и РНК такие зонды и праймеры включают полипептидные нуклеиновые кислоты (ПНК), как описано у Nielsen P. et al., *Science* 254:1497-1500 (1991). Зонд или праймер имеет участок нуклеотидной последовательности, который гибридизуется по меньшей мере приблизительно с 15, обычно с 20-25, а в некоторых воплощениях приблизительно с 40, 50 или 75 последовательными нуклеотидами в молекуле нуклеиновой кислоты. В одном воплощении зонд или праймер включает по меньшей мере один аллель по меньшей мере одного полиморфного маркера или по меньшей мере одного гаплотипа, описываемых здесь, или комплементарную им последовательность. В конкретных воплощениях зонд или праймер могут включать 100 или менее нуклеотидов, например, в некоторых воплощениях от 6 до 50 нуклеотидов, или, например, от 12 до 30 нуклеотидов. В других воплощениях зонд или праймер идентичен по меньшей мере на 70%, по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на

90% или по меньшей мере на 95% примыкающей нуклеотидной последовательности или последовательности, комплементарной к примыкающей нуклеотидной последовательности. В другом воплощении зонд или праймер способны селективно гибридизоваться с примыкающей нуклеотидной последовательностью или с последовательностью, комплементарной примыкающей нуклеотидной последовательности. Зачастую зонд или праймер также включает метку, например, радиоизотоп, флуоресцентную метку, ферментативную метку, метку в виде кофактора фермента, магнитную метку, спин-метку, эпитоп-метку.

Молекулы нуклеиновых кислот по изобретению, такие как описываемые здесь, могут быть идентифицированы и выделены с помощью стандартных методов молекулярной биологии, хорошо известных специалистам. Амплифицированная ДНК может быть помечена (например, радиоактивно помечена) и использована в качестве зонда для скрининга библиотеки кДНК, полученной из клеток человека. кДНК может быть синтезирована на основе мРНК и содержаться в подходящем векторе. Могут быть выделены соответствующие клоны, ДНК может быть получена с помощью эксцизии *in vivo*, и клонированная вставка может быть секвенирована в одном или в обоих направлениях с помощью общеизвестных способов для определения точной рамки считывания, кодирующей полипептид соответствующего молекулярного веса. С помощью данных способов или аналогичных им, можно выделить полипептид и кодирующую полипептид ДНК, секвенировать их и охарактеризовать более подробно.

Как правило, последовательности выделенных нуклеиновых кислот по изобретению могут применяться в качестве маркеров молекулярного веса для Саузерн гелей и в качестве хромосомных маркеров, меченых для картирования позиций ассоциированных генов. Последовательности нуклеиновых кислот также могут применяться для сравнения с эндогенными последовательностями ДНК пациентов для выявления рака молочной железы или подверженности раку молочной железы, и в качестве зондов, например, для гибридизации и обнаружения ассоциированных последовательностей ДНК или для вычитания известных последовательностей из образца (например, вычитающая гибридизация). Последовательности нуклеиновых кислот могут быть в дальнейшем использованы для получения праймеров для генетической идентификации, для выработки антител к полипептидам с помощью методов иммунизации и/или в качестве антигенов для выработки антител к ДНК или для развития иммунного ответа.

Антитела

Поликлональные антитела и/или моноклональные антитела, специфически связывающие одну форму генного продукта, но не связывающие другую форму генного продукта, также рассматриваются в изобретении. Рассматриваются антитела, связывающие часть варианта или референсного генного продукта, содержащего полиморфный сайт или сайты. Используемый здесь термин "антитело" относится к молекулам иммуноглобулинов и иммунологически активных частей иммуноглобулиновых молекул, т.е. молекулам, содержащим антиген-связывающие сайты, специфически связывающие антиген. Молекула, специфически связывающая полипептид по изобретению, представляет собой молекулу, связывающуюся с этим полипептидом или его фрагментом, но не связывающуюся существенно с другими молекулами в образце, например, в биологическом образце, который естественным образом содержит полипептид. Примеры иммунологически активных частей молекул иммуноглобулина включают F(ab) и F(ab')₂, которые могут быть получены при обработке антитела таким ферментом, как пепсин. В изобретении также рассматриваются поликлональные и моноклональные антитела, связывающиеся с полипептидом по изобретению. Термин "моноклональное антитело" или "композиция моноклональных антител" в данном описании относится к группе молекул антител, содержащих только один вид антиген-связывающего сайта, способного к иммунному взаимодействию с определенным эпитопом полипептида по изобретению. Композиция моноклональных антител, таким образом, обычно проявляет одинаковую аффинность к определенному полипептиду по изобретению, с которым она взаимодействует.

Поликлональные антитела могут быть получены согласно описанию выше путем иммунизации подходящего организма необходимым иммуногеном, например, полипептидом по изобретению или его фрагментом. Титр антител в иммунизированном организме можно отслеживать во времени с помощью стандартных методик, таких как иммуноферментный анализ (ELISA) с применением иммобилизованных полипептидов. При желании из организма млекопитающего (например, из крови) могут быть выделены молекулы антител, направленные против полипептида, и затем очищены для получения фракции IgG с помощью общеизвестных методик, например, с помощью хроматографии на протеине А. Через определенное время после иммунизации, например, когда титр антител наиболее высокий, антитело-продуцирующие клетки могут быть выделены из организма и использованы для получения моноклональных антител с помощью стандартных методик, например, гибридомной технологии, первоначально описанной Kohler и Milstein, *Nature* 256:495-497 (1975), технологии человеческих В-клеточных гибридом (Kozbor et al., *Immunol. Today* 4: 72 (1983)), технологии EBV-гибридом (Cole et al., *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, 1985, Inc., pp. 77-96) или триомной технологии. Технология получения гибридом хорошо известна (см. *Current Protocols in Immunology* (1994) Coligan et al., (eds.) John Wiley & Sons, Inc., New York, NY). Кратко, бессмертная клеточная линия (обычно, миелома) подвергается слиянию с лимфоцитами (обычно, спленоцитами) млекопитающего, иммунизированного иммуногеном, как описано выше, и проводится скрининг супернатантов культур полученных клеток гибридомы для выделения гибридом, продуцирующих моноклональное антитело, связывающее полипептид по изобретению.

Для получения моноклональных антител к полипептидам по изобретению могут применяться любые из множества общеизвестных протоколов, используемых для слияния лимфоцитов и бессмертных клеточных линий (см, например, *Current Protocols in Immunology*, supra; Galfre et al., *Nature* 266: 55052 (1977); R.H. Kenneth, in *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York, New York (1980); и Lerner, *Yale J. Biol. Med.* 54:387-402 (1981)). Кроме того, любому специалисту понятно, что существует множество вариаций таких методик, которые тоже могут использоваться.

Кроме получения гибридом, секретирующих моноклональные антитела, идентификация и выделение моноклонального антитела к полипептиду по изобретению может быть проведено путем скрининга рекомбинантной комбинаторной библиотеки иммуноглобулинов (например, библиотеки фаговых частиц, экспонирующих антитела) при помощи полипептидов для выделения таким образом представителей библиотеки иммуноглобулинов, связывающих полипептид. Наборы для генерирования и скрининга фаговых библиотек имеются на рынке (например, *Recombinant Phage Antibody System*, компании Pharmacia, каталоговый номер 27-9400-01; и *SurfZAP™ Phage Display Kit*, компании Stratagen, каталоговый номер 240612). Кроме того, примеры методик и реагентов, удобных для использования при генерировании и скрининге библиотеки антител, можно найти, например, в U.S. Patent No. 5223409; PCT Publication No. WO 92/18619; PCT Publication No. WO 91/17271; PCT Publication No. WO 92/20791; PCT Publication No. WO 92/15679; PCT Publication No. WO 93/01288; PCT Publication No. WO 92/01047; PCT Publication No. WO 92/09690; PCT Publication No. WO 90/02809; Fuchs et al., *BioTechnology* 9: 1370-1372 (1991); Hay et al., *Hum. Antibod. Hybridomas* 3: 81-85 (1992); Huse et al., *Science* 246: 1275-1281 (1989); и Griffiths et al, *EMBO J.* 12:725-734(1993).

Кроме того, рекомбинантные антитела, такие как химерные и гуманизированные моноклональные антитела, содержащие как человеческие, так и не принадлежащие человеку участки, которые могут быть получены с помощью стандартных технологий рекомбинантной ДНК, также находятся в рамках изобретения. Такие химерные и гуманизированные моноклональные антитела могут быть получены с помощью известных специалистам технологий рекомбинантной ДНК.

В общих чертах, антитела по изобретению (например, моноклональные антитела) могут использоваться для выделения полипептидов по изобретению с помощью стандартных методик, таких как аффинная хроматография или иммунопреципитация. Антитело, специфическое в отношении полипептида, может облегчить очистку естественного полипептида из клеток и полипептида, произведенного рекомбинантным способом, экспрессируемого в клетках хозяина. Кроме того, антитело, специфическое в отношении полипептида по изобретению, может применяться для детекции полипептида (например, в клеточном лизате, клеточном супернатанте или образце ткани) для оценки содержания и характера экспрессии полипептида. Антитела могут применяться в диагностических целях для мониторингирования уровня белков в тканях, как часть процесса клинического тестирования, например, для определения эффективности применяемого режима лечения. Антитело может быть связано с детектируемым соединением для облегчения его детекции. Примеры детектируемых соединений включают различные ферменты, простетические группы, флуоресцентные вещества, люминесцентные вещества, биолюминесцентные вещества и радиоактивные вещества. Примеры подходящих ферментов включают пероксидазу хрена, щелочную фосфатазу, бета-галактозидазу или ацетилхолинэстеразу; примеры подходящих комплексов простетических групп включают стрептавидин/биотин и авидин/биотин; примеры подходящих флуоресцентных материалов включают умбеллиферон, флуоресцеин, флуоресцеин изотиоцианат, родамин, дихлоротриазиниламин флуоресцеин, дансил хлорид или фикоэритрин; примеры люминесцентных веществ включают люминол; примеры биолюминесцентных веществ включают люциферазу, люциферин и экворин, а примеры подходящих радиоактивных веществ включают ^{125}I , ^{131}I , ^{35}S или ^3H .

Антитела могут также применяться для фармакогеномного анализа. В таких воплощениях антитела против полиморфных вариантов белков, кодируемых нуклеиновыми кислотами по изобретению, например полиморфных вариантов белков, кодируемых нуклеиновыми кислотами, содержащими по меньшей мере один полиморфный маркер по изобретению, могут применяться для идентификации индивидов, нуждающихся в изменении схем лечения.

Кроме того, антитела могут применяться для оценки экспрессии полиморфных вариантов белков при заболеваниях, например, активных стадий заболевания или у индивидов с предрасположенностью к заболеванию, связанной с функцией белка, в частности, к раку молочной железы. Антитела, специфические в отношении полиморфного варианта белка по данному изобретению, кодируемого нуклеиновой кислотой, включающей по меньшей мере один полиморфный маркер или гаплотип согласно данному описанию, могут применяться для скрининга наличия полиморфного варианта белка, например, скрининга предрасположенности к раку молочной железы, на что будет указывать наличие полиморфного варианта белка.

Антитела могут применяться и в других методах. Так, антитела, могут использоваться в качестве диагностических средств для оценки белков, например, вариантов белков по изобретению, в ходе анализа по электрофоретической подвижности, изоэлектрической точки, расщепления трипсином или другими протеазами, или для применения в других физических тестах, известных специалистам. Антитела могут

также применяться для типирования тканей. В одном воплощении определенный вариант белка коррелирует с экспрессией в определенном типе ткани, и антитела, специфические для варианта белка, могут применяться для идентификации определенного типа тканей.

С помощью антител можно также определять субклеточную локализацию белков, включая полиморфные варианты белков, что может использоваться для определения аномальной субклеточной локализации белков в клетках различных тканей. Такое применение может использоваться в генетическом тестировании, а также мониторинге определенных схем лечения. В случае, когда лечение нацелено на коррекцию уровня экспрессии или наличия полиморфного варианта белка или аномального распределения в тканях или экспрессии в процессе дифференцировки полиморфного варианта белка, антитела, специфические в отношении полиморфного варианта белка или его фрагментов, могут использоваться для мониторинга эффективности терапии.

Антитела также применяются для ингибирования функции полиморфного варианта белка, например, путем блокирования связывания варианта белка со связывающей молекулой или партнером. Подобные применения могут также использоваться в терапевтическом контексте, когда лечение включает ингибирование функции полиморфного варианта белка. Например, антитело может применяться для блокирования или конкурентного ингибирования связывания, таким образом модулируя (т.е. выполняя роль агониста или антагониста) активности белка. Антитела можно получать против фрагментов определенного белка, содержащих сайты, необходимые для определенной функции или против интактного белка, ассоциированного с клеткой или клеточной мембраной. Для применения *in vivo*, антитело может быть связанным с дополнительной функциональной частью с терапевтическими свойствами, например, радионуклидом, ферментом, иммуогенным эпитопом или цитотоксическим агентом, включая бактериальные токсины (дифтерийный или растительные токсины, например, рицин). Время полужизни *in vivo* антитела или его фрагмента может быть увеличено при помощи пегилирования путем конъюгирования с полиэтиленгликолем.

Данное изобретение также относится к наборам для применения антител в описываемых здесь способах. Они включают наборы для детектирования наличия полиморфного варианта белка в тестируемом образце, но не ограничиваются ими. Одно предпочтительное воплощение включает антитела, такие как меченые или способные быть помеченными антитела, а также соединения или агенты для детектирования полиморфных вариантов белка в биологическом образце, средств для определения количества или наличия и/или отсутствия полиморфного варианта белка в образце и средств для сравнения количества полиморфного варианта белка в образце со стандартом, а также инструкции по применению набора.

Ниже будут описаны примеры осуществления изобретения, не являющиеся лимитирующими.

Описание примеров осуществления изобретения

Выбор пациентов и контролей.

Данное исследование было согласовано с Национальным Комитетом Исландии по Биоэтике и Комитетом Исландии по защите Данных. Медицинская документация по пациентам с диагнозом рака молочной железы была получена из Исландского Регистра по Раку (от англ. Icelnic Cancer Registry, ICR). В документации описывались все случаи инвазивных опухолей молочной железы и протоковой или дольковой карциномы *in-situ*, выявленные в Исландии с 1 января 1955 до 31 декабря 2005. В ICR имелась документация по 4603 индивидов, которым за этот период был поставлен диагноз. Преобладающая часть пациентов, включающая всех живущих пациентов (приблизительно 2840), соответствовала критериям включения в исследование. Информированное согласие и образец крови и клинические данные были получены у 2210 пациентов, что составило приблизительно 78%. Исходная группа сравнения состояла из всех индивидов, участвующих в проводимых в deCODE Genetics общегеномных ассоциативных исследованиях и прошедших генотипирование с помощью чипов Illumina SNP на тот момент, когда были генотипированы пациенты из группы с раком молочной железы. Индивиды, включенные в исландский реестр по раку с диагнозом рака молочной железы, рака простаты или колоректального рака на конец 2005 г, были исключены из контрольной группы, в которую в результате вошли 4477 "контролей". Независимая дублирующая группа контролей состояла дополнительно из 7406 индивидов, участвовавших в проектах, не связанных с раком молочной железы, типированных после группы из 4477 человек. В этой группе из 7406 человек пациенты с раком простаты и колоректальным раком не исключались, поскольку были получены предварительные данные, указывающие на то, что частоты изучаемых ОНП в этих группах пациентов не отличались от контрольных. Ни в одной группе контролей не наблюдалось различий между полами по частотам ОНП, приведенных ниже в табл. 8, 9 и 14. Таким образом, данные контрольные группы были достаточно репрезентативными в отношении популяционных частот изучаемых ОНП.

Генеалогическая база данных.

В deCODE Genetics существует компьютеризированная база данных генеалогии исландцев. Имеются данные практически по всем индивидам, родившимся в Исландии на протяжении последних двух столетий и для 95% рождений за этот период известна информация о родителях [Sigurdardottir, et al., (2000), *Am J Hum Genet*, 66, 1599-609]. Кроме того, для большинства индивидов указывается округ, в котором он проживает, по данным переписи и церковно-приходским записям. Информация содержится в реляционной базе данных с закодированными персональными идентификаторами, соответствующими тем, ко-

которые используются для биологических образцов и документации в ICR, что позволяет сопоставлять генотипы и фенотипы исследуемых пациентов с их генеалогией.

Обработка образцов.

Образцы крови хранили в ЭДТА при -20°C. ДНК выделяли из цельной крови при помощи колоночного метода выделения Qiagen (<http://www.qiagen.com>). ДНК хранили при 4°C.

Генотипирование.

Впоследствии с помощью микрочипов Illumina Infinium Harp300 SNP (Illumina, San Diego, CA, USA), содержащих 317 511 ОНП, установленных в ходе I фазы международного проекта HarpMap, были генотипированы образцы 1600 пациентов, выбранных случайным образом. Образцы "контролей", не имеющих рака, были генотипированы с помощью той же платформы. Из всех ОНП, представленных на чипе, для 170 не удалось вывести генотип, а у других 24 ОНП выход был ниже 80%. Другие 61 ОНП имели высокосзначимое отклонение от равновесия Харди-Вайнберга среди "контролей" ($p < 1 \times 10^{-10}$). 104 ОНП были мономорфными и 58 ОНП были близки к тому, чтобы считаться мономорфными (т.е. частота минорного аллеля в объединенной когорте пациентов и контролей составляла менее 0,001). Наконец, после изучения определенных регионов и возможных сигналов в ходе нескольких общегеномных ассоциативных исследований, проводящихся в deCODE, оказалось, что 5 маркеров генотипировать невозможно. Все проблематичные ОНП были исключены из анализа. Таким образом, в итоговом анализе, описываемом в тексте, было использовано 317 089 ОНП. Любые чипы с уровнем сигнала ниже 98% также исключались из анализа.

Ниже представлены последовательности трех ключевых ОНП, изучавшихся в данном исследовании:

rs4848543 (SEQ ID NO:1):

```
caggataagattattctgaggctgtccatcagtggtctgtgagctggcccatgtattcagccagtcagtaggcc
atcaagcatgcatctacgcaaggcactagaagggttacatacaatgggaaactgaggcccagaaggaggaaagatga
agtgcccaaaagtaaggagataaaagcagaagtgagggaagagaaggagtattcatattactgcaggaatccaaattct
acaagaacatacgtgttcttaaggaggccctcgcttaagcattattgccgatcaatacc[A/C]attaataattacagg
acatgggcaaatattctagcaaggctcctgaggtctctagatgtgccccaaataacaaggactcccccgagggtggca
ccaaaagtcggaggatgtgtggaagctgacctccccacctgtccctgcctcaccacagtcacccacccctctaggatgt
aatgtgtaatcattcaacacatgggccaccttggaagtcgcttaacctctctccctttcacgttataaaatgggcatagtcatt
acacaaatgcataaatcaacctcaattttctcatct
```

rs13387042 (SEQ ID NO:2):

```
ctgacatagccgctgtgcagggttttgggaccagggtcgtccatctgtattcccaggctgggctggagcacctgag
gctactatggttacagaaaggaaacctggcatcgaaggtataggctcaaatcccagctctcatgattgctagctttgaggttc
agggtcatttctaagtgtctgagcttgggtccctctctcaaatgggaatgaggagcatgtctagctccctgtttgtgag
tgaatagctacatgaagaacagctaaccagaacagaaagaaggcaaatgga[A/G]gctacagaaaccaaggatttc
ctgttgaatcgaatctcctcaatctcctcaccacactagtggtatcctctgtgggagggtgtgagagtgctccgtgtttttt
tggtttttttttttttttgtgatggagtctcactctgcacccagactggaatgcagtggtgcgactctcggtcactgcaacctctgc
ctccagggtcaagtgattctcctgcctcagctcccaagtagccaggactacaggcataatgcatcacaccagcctaattttt
gtatttttagtag
```

rs3803662 (SEQ ID NO:3):

```
gctagtccttggtgtgtgatcaaatgcaattgtttccatttctaagaacagaatctttatagactcagttctttt
cttggttattttacaattttattgttctatggttattaaaaataacatgtcatataaattaagcctagaacatggatgtttatcaaa
gaaaggattgtcatccaaagcaccaactatgagagatatctatgtgcaatggtatatagatctgtcatagaagggttaattat
atctgcctaataatgttctccttaatgcctctatagctgtc[C/T]cttagcgaagaataaaactgtggactgacccccacca
tttgcgaagaagtagtggtctcagcttctcagcgggtgtgtgttgacacaccagggggccctcctgtgtaag
actgtctggactcctgattttcactcctgtataaacccctccccattactaccttgtctcctgggaaaaaaatgttttcac
ttaagcaaaataaactaccagattttgttgaggagaaattgtaaacctcattgtacattctcctggccttgatgcttatcg
```

Были разработаны и апробированы способы анализа для определения количества ОНП в блоке НС rs4848543. Данные способы анализа основаны на технологии анализа ОНП на платформе Centaurus, описанной в [Kutyavin, et al., (2006), Nucleic Acids Res, 34, e128]. Были использованы следующие анализы и праймеры для них:

SG02S733

rs895398

(Build 34) chr2:120,058,180

Прямой праймер: TTGGAACCTCCCACTGCCACA

Обратный праймер: ACAGCCCAGGTTTCAGGTTGGCAT

VIC-Зонд: AGCGACACTG

FAM-Зонд: GCGACACCG

Энхансер: CCTG*AGATCCAGGCA

SG02S738

rs4848543

(Build 34) chr2:120,061,096

Прямой праймер: AGGGAGCCCTCGCTTTAAGCATTA

Обратный праймер: TTTGGGCACATCTAGGAAGACCTC

VIC-Зонд: GAT*CAATACCC

FAM-Зонд: CGAT*CAATACCA

Энхансер: TTAATAATTACAGGACATG

SG02S739

rs6759589

(Build 34) chr2:120,064,976

Прямой праймер: TCAGCAGCCTCTTGTCTCACTAAT

Обратный праймер: CTTCTTGCCCAGGCAGGACAGA

VIC-Зонд: TAGACTGAGCC

FAM-Зонд: AAT*AGACTAAGCC

Энхансер: CCATCCCCAACTAA

SG02S753

rs895397

(Build 34) chr2:120,072,400

Прямой праймер: CACTACTAGATCATGGGAGATGCCT

Обратный праймер: GTTTCCTCAGTTGTAATGTAGACC

VIC-Зонд: TT*T*A*ACCCTT

FAM-Зонд: GTTTTAACCCCT

Энхансер: GCTGCTCGGAAG

SG02S740

rs838102

(Build 34) chr2:120,079,878

Прямой праймер: GGTTTTGCATCCTTGCTCACTCACAT

Обратный праймер: GTGCCCTCTGCGTACTTGCGAAT

VIC-Зонд: TACTGCCACC
 FAM-Зонд: CTTACTGCCACT
 Энхансер: TGTCTCCTCACTC

SG02S741
 rs838100
 (Build 34) chr2:120,083,710
 Прямой праймер: ACTTACCAACAGCAGGCTGGTG
 Обратный праймер: CCAGAAAGGGTGGCTCACCTCA
 VIC-Зонд: GTCTGAGGCT
 FAM-Зонд: GGTCTGAGGTT
 Энхансер: TCAGGTCACTCG

SG02S734
 rs12711924
 (Build 34) chr2:120,107,316
 Прямой праймер: AACGCTTATCACAGAGCCAGGTAG
 Обратный праймер: GGAAAACCATGAACACTGTGGCAA
 VIC-Зонд: CCCGCCGT
 FAM-Зонд: GCCCGCCAT
 Энхансер: TTGTCTTTCTTGTGTC

SG02S742
 rs3731603
 (Build 34) chr2:120,117,062
 Прямой праймер: AACA*CTCTAGGGAGTAAAGCTC
 Обратный праймер: GAGGA*CATTGGAGTCCCCAAT
 VIC-Зонд: GACTGTCTTCAA
 FAM-Зонд: ACTGT*CT*TCAC
 Энхансер: AAGTCAGTCCTGAG

SG02S728
 rs13387042
 (Build 34) chr2:218,108,374
 Прямой праймер: CCACTAGTGTGGTGA*AGGAAGATT
 Обратный праймер: GCTACATGA*AGAA*CAGCTAAACC

VIC-Зонд: TTTCTGTAGCCT
 FAM-Зонд: TT*T*CTGTAGCT*T
 Энхансер: CATTTGCCTTCTTTC

*указывает модифицированное основание, согласно описанию [Kutyavin et al., (2006), Nucleic Acids Res, 34, e128].

Единственная основополагающая мутация гена BRCA2 gene (999del5) представлена с частотой носительства 0,6-0,8% в общей исландской популяции и 7,7-8,6% среди женщин с раком молочной железы [Thorlacius et al., (1997), Am J Hum Genet, 60, 1079-84; Gudmundsson et al., (1996), Am J Hum Genet, 58, 749-56]. Возможность наличия мутации BRCA2 999del5 определяли у 1499 "случаев" с применением анализа по технологии ПЦР-анализа микросателлитов, обозначенного DG13S3727. Использовали следующие праймеры:

Прямой: TGTGAAAAGCTATTTTCCAATC
 Обратный: ATCACGGGTGACAGAGCAA

Последовательность, амплифицированная в DG13S3727, представляла собой:

Расположение DG13S3727 =chr13:30703058-30703261

```
tgtgaaaagctatttttccaatcatgatgaaagtctgaagaaaatgata
gatttatcgcttctgtgacagacagtgaaaacacaaatcaaagagaagct
gcaagtcattggaagtcctctgtttagtgaactacaggtttttttgttg
ttgtgttttgatttttttttttgaggtggagtcttgcctctgtcacccg
tgat
```

Статистические методы.

Относительный риск (RR) для аллеля ОНП был подсчитан с применением мультипликативной модели следующим образом: $RR = [p/(1-p)]/[s/(1-s)]$, где p и s представляют собой частоты варианта у пациентов и контролей, соответственно. Соответствующие значения P для RR были рассчитаны на основании стандартной статистики хи-квадрат отношения правдоподобия. Доверительные интервалы рассчитывали, допуская, что оценки RR имеют логарифмически нормальное распределение.

Частоты гаплотипов оценивали при помощи максимального правдоподобия, а анализ различий между "случаями" и "контролями" выполняли при помощи обобщенного теста отношения правдоподобия. Для подсчета результатов для гаплотипов применяли программу анализа гаплотипов с названием NEMO, от англ. NEsted Models. Подчеркивается, что для решения проблемы неопределенности гаплотипов и пропусков генотипов в NEMO не применяется традиционный двухэтапный способ исследования ассоциации, когда вначале оценивают частоты гаплотипов, возможно с применением алгоритма максимизации правдоподобия, а затем выполняются тесты, в которых оценки частот принимают за истинные значения частот, поскольку данный способ иногда может вызывать затруднения и требовать рандомизации для того, чтобы корректно оценить статистическую достоверность. Вместо этого, в NEMO оценки максимального правдоподобия, отношения правдоподобия и значения P подсчитываются напрямую для имеющихся данных и, следовательно, потеря информации из-за неопределенности гаплотипов и пропусков генотипов автоматически фиксируется в отношении правдоподобия.

Некоторые пациенты и контроли из Исландии приходятся друг другу родственниками, как внутри групп, так и между группами. Для внесения поправки на наличие родства в результаты ассоциативного анализа, была проведена симуляция генотипов по распределению аллелей в генеалогии исландцев. По каждой симуляции тест на наличие ассоциации у случаев и контролей выполняли таким же образом, как и для реальных генотипов, т.е. тест хи-квадрат отношения правдоподобия. Для N числа симуляций было подсчитано среднее значение статистик хи-квадрат (что аналогично вариации соответствующего Z -значения (Z -score)), которое при нулевой гипотезе должно равняться 1 для индивидов, не состоящих в родстве, а для родственников должно быть несколько выше. Затем поправку на наличие родства учитывали путем деления статистики хи-квадрат истинных показателей ассоциации на среднее значение, полученное при симуляциях.

Совместный анализ множества групп случай-контроль проводили с применением метода Мантеля-Хензеля, при котором предполагается, что у групп могут быть различные популяционные частоты аллелей, гаплотипов и генотипов, но они имеют одинаковый относительный риск.

Все приведенные значения P являются двусторонними.

Анализ первично-множественных опухолей молочной железы

Документация по пациентам с диагнозом первично-множественного рака молочной железы (ПМ РМЖ) была получена в ICR. При диагностировании первичных опухолей после первоначального диагностирования рака молочной железы подтверждали как клинически, так и гистологически, что они являются независимыми первичными опухолями, развивающимися одновременно или последовательно с первоначально диагностированными опухолями молочной железы, и возникают контралатерально или ипсилатерально. Опухоли, классифицированные клинически или гистологически как рецидив первоначальной опухоли или как отдельные мультифокальные первичные опухоли, исключали из анализа. Пациентов с диагностированными двумя или более независимыми первичными опухолями, зарегистрированных в реестре на конец 2005 г включительно, считали имеющими ПМ РМЖ. Пациентов, не имеющих диагностированной второй первичной опухоли согласно ICR на 31 декабря 2005 г., считали случаями с одиночными первичными опухолями молочной железы (ОП ОМЖ). Риск ПМ РМЖ по отношению к ОП ОМЖ определяли при помощи логистической регрессии, с учетом количества лет, прошедших после постановки первого диагноза рака молочной железы, в качестве ковариаты. Возраст при начале заболевания сравнивали с помощью линейной регрессии с применением программы NEMO.

Анализ семейного анамнеза.

Для каждого пораженного пробанда определяли индекс семейного анамнеза с применением генеалогической базы данных и документации ICR. Круг родственников пробанда от первой до третьей степени определяли используя генеалогическую базу данных. Затем у родственников определяли статус поражения заболеванием согласно записям в ICR. Пробандам присваивали индекс семейного анамнеза, равный 1 для каждого пораженного родственника первой степени, 0,5 для каждого пораженного родственника второй степени и 0,25 для каждого пораженного родственника третьей степени. Затем для каждого

пробанда подсчитывали общий индекс. Потенциальное взаимоотношение между генотипом ОНП маркера и индексом семейного анамнеза затем определяли при помощи критерия Вилкоксона, рассчитываемого в программе JMP v4 (S.A.S. Institute Inc.).

Ассоциативный анализ фенотипов:

Для ассоциативного анализа использовали два родственных фенотипа. Первый фенотип включал всех 1600 индивидов с диагностированным раком молочной железы. Данный фенотип был обозначен "Все формы рака молочной железы" ("все формы РМЖ"). Ко второму фенотипу были отнесены индивиды с предрасположенностью от умеренной до высокой, обозначавшейся как "Умеренная предрасположенность" ("УмерПред").

Чтобы фенотип подходил под это определение, требовалось соответствие пробанда по меньшей мере одному из следующих критериев:

1. Пробанд является представителем группы случаев рака молочной железы, объединяющей 3 или более пораженных родственников, находящихся на генетическом расстоянии не более 3 мейотических делений (3М).

2. Пробанд является одним из пары пораженных родственников, находящихся на расстоянии не более 3М, у одного из которых диагноз был поставлен в возрасте 50 лет или младше.

3. Пробанд является одним из пары пораженных родственников, находящихся на расстоянии не более 3М, у одного из которых была диагностирована вторая первичная опухоль любого типа.

4. У пробанда была диагностирована вторая первичная опухоль любого типа.

Из 1600 исландских пациентов, вошедших в представленное здесь исследование, 653 отвечают критериям "УмерПред" (40,8%).

Результаты

Раздел 1: rs4848543 на хромосоме 2q14.2

Аллель А rs4848543 ассоциирован с повышенным риском рака молочной железы.

Общегеномный ассоциативный анализ показал, что аллель А rs4848543 был ассоциирован с раком молочной железы "УмерПред". Данные по этому ОНП были получены с помощью чипа Illumina для 1598 пациентов (из которых 653 соответствовали критериям фенотипа "УмерПред") и 4477 контролей, не имеющих рак. Оценки RR для аллеля А rs4848543 в случае ассоциации с раком молочной железы "УмерПред" составили 1,42 при значении Р, равном $8,3 \times 10^{-8}$ после коррекции на наличие родства между индивидами в группах случаев и контролей (табл. 1). После применения поправки Бонферрони с учетом 317089 протестированных ОНП "суммарное" скорректированное значение Р составило 0,026. Для аллеля А rs4848543 также была обнаружена допустимо значимая ассоциация с фенотипом "все формы РМЖ", с оценкой относительного риска, равной 1,16 (табл. 1). Для подтверждения этих результатов на независимой выборке из Исландской популяции была разработана и апробирована технология анализа ОНП на платформе Centaurus, SG02S738, для исследования ОНП rs4848543. Анализ SG02S738 был проведен дополнительно у 573 пациентов с раком молочной железы, из которых 198 удовлетворяли определению фенотипа

"УмерПред". Этих пациентов сравнивали с независимой контрольной группой, включавшей 7406 индивидов, типированных в отношении rs4848543 при помощи чипа Illumina или анализа SG02S738. Оценки RR для аллеля А rs4848543 в случае ассоциации с раком молочной железы "УмерПред" во второй группе составили 1,26 (значение $p=3,4 \times 10^{-2}$). Таким образом, исходные результаты были воспроизведены с допустимым уровнем значимости на независимой исландской выборке. Результаты совместного анализа, при котором объединили 1598 пациентов из первоначальной группы и 573 пациентов из дублирующей группы и провели сравнение с 11883 контролями, показаны в табл. 1. Совместный анализ показал допустимый уровень значимости риска для фенотипа "все формы РМЖ" и риск, значимый на общегеномном уровне для фенотипа "УмерПред". Суммарное значение Р для рака молочной железы "УмерПред" после поправки Бонферрони составило 0,0095 (коррекция на 317089 ОНП). Данный показатель является значимым на общегеномном уровне, даже с учетом того, что были протестированы два фенотипа "все формы РМЖ" и "УмерПред".

Поскольку документацию в ICR начали вести с 1955 г, некоторые из пациентов, вошедших в исследование, имели большую продолжительность выживания после обнаружения рака. Если пациенты с носительством вариантного аллеля А rs4848543 имели бы иную вероятность длительного выживания, чем лица, не несущие его, то частота варианта среди распространенных случаев могла бы различаться. Для изучения этой возможности была выделена подгруппа пациентов, состоявшая из 833 индивидов, которым диагноз был поставлен после 1 января 2000 г и у которых время от постановки диагноза до включения в исследование составило менее 5 лет. В этой когорте пациентов, которым диагноз был поставлен недавно, частота аллеля А rs4848543 составила 0,372945. Также была выделена группа из 1338 пациентов, которым диагноз был поставлен после 1 января 2000 г и которые дожили до включения в исследование. Частота аллеля А rs4848543 в этой группе составила 0,372749, что достоверно не отличалось от частоты среди недавно диагностированных пациентов (значение $P = 0,99$). Таким образом, нет очевидных доказательств того, что различная выживаемость оказывает влияние на оценки риска, приведенные в

табл. 1.

Кроме того, было отмечено, что у пациентов с раком молочной железы, не имеющих фенотипа "УмерПред", не обнаруживалось повышенной частоты аллеля A rs4848543 (табл. 1). Значение этого наблюдения будет обсуждаться ниже. Были также изучены взаимоотношения между аллелем A rs4848543 и различными слагаемыми определения фенотипа "УмерПред":

Аллель A rs4848543 не имеет достоверной ассоциации с ранним возникновением рака молочной железы.

Поскольку молодой возраст при постановке диагноза был одним из критериев, использованных в определении фенотипа "УмерПред", была рассмотрена ассоциация между возрастом при постановке диагноза и генотипом rs4848543. Для 1598 пациентов, которым был проведен анализ с помощью чипов Illumina, не было обнаружено достоверной ассоциации между генотипом rs4848543 и возрастом при постановке диагноза (табл. 2).

Риск, обусловленный аллелем A rs4848543, ассоциирован с семейным анамнезом рака молочной железы.

Индекс семейного анамнеза подсчитывали для каждого пробанда, включенного в исследование. С применением генеалогической базы данных для каждого пробанда был определен круг родственников от первой до третьей степени. Затем, при обращении к данным ICR, были идентифицированы все родственники, у которых также был диагностирован рак молочной железы. Затем для каждого пробанда был подсчитан суммарный индекс семейного анамнеза (ИСА), когда за каждого пораженного родственника первой степени присваивали 1 балл, за каждого пораженного родственника второй степени присваивали 0,5 балла и за каждого родственника третьей степени присваивали 0,25 балла. Затем 1598 пробандов, типированных с применением чипов Illumina, исследовали на наличие ассоциации между ИСА и генотипом rs4848543. Как показано в табл. 2, при носительстве аллеля A rs4848543 наблюдалась тенденция к увеличению ИСА. Данные результаты было бы невозможно ожидать только на том основании, что вариантный аллель A rs4848543 обуславливает относительный риск среди пациентов с фенотипом "все формы РМЖ", равный всего лишь 1,16; такой относительный риск сам по себе не мог бы обуславливать в семьях концентрации случаев с раком молочной железы [Stacey, et al., (2006), PLoS Med, 3, e217]. Тем не менее, аллель A rs4848543 был обнаружен у лиц с фенотипом "УмерПред", в определении которого важным слагаемым является наличие семейного анамнеза. Данные наблюдения позволяют предположить, что аллель A rs4848543 может находиться в неравновесии по сцеплению с менее распространенным вариантом, имеющим высокую пенетрантность. Возможно также, что аллель A rs4848543 может взаимодействовать с другими предрасполагающими факторами, имеющими высокую пенетрантность, повышая их пенетрантность и, следовательно, ассоциацию с семейным анамнезом рака молочной железы.

Ассоциация между риском, обусловленным аллелем A rs4848543 и семейным анамнезом рака молочной железы, частично (но не полностью) объясняется семейным носительством мутации BRCA2 999del5.

Для дальнейшего изучения взаимосвязи между аллелем A rs4848543 и семейным анамнезом рака молочной железы было исследовано, объясняется ли семейная концентрация, характерная для пациентов с носительством аллеля A rs4848543, существованием в семьях носительства хорошо охарактеризованной в Исландии мутации BRCA2 999del5. Из группы 1600 пациентов, 1499 было типировано на наличие мутации BRCA2 999del5. ИСА был пересчитан согласно описанию выше, на этот раз все пробанды, не имеющие мутации BRCA2 999del5, были исключены. Как показано в табл. 2, ассоциация между аллелем A rs4848543 и ИСА после исключения носителей мутации 999del5 оставалась достоверной, что позволяло предположить, что наличие ассоциации невозможно полностью отнести на счет семей с мутацией BRCA2 999del5. Однако значение P было существенно выше, чем без исключения носителей 999del5, что дало основания предположить, что BRCA2 999del5 может в некоторой степени обуславливать семейную концентрацию, характерную для аллеля A rs4848543.

Риск, обусловленный носительством аллеля A rs4848543, распространяется на носителей мутации BRCA2 999del5.

Поскольку мутация BRCA2 999del5 имеет столь важный эффект при семейном раке молочной железы в Исландии, была исследована ее взаимосвязь с вариантным аллелем A rs4848543. Одним из возможных сценариев было бы то, что вариантный аллель A rs4848543 обуславливает незначительный дополнительный риск носителям BRCA2 999del5, как это было предположено в случае взаимодействия между мутациями CHEK2 и BRCA [(2004), Am J Hum Genet, 74, 1175-82; Meijers-Heijboer, et al., (2002), Nat Genet, 31, 55-9]. В таком случае частота вариантного аллеля A rs4848543 среди пораженных носителей BRCA2 999del5 приблизительно соответствовала бы его частоте среди контролей. Напротив, если бы частота вариантного аллеля A rs4848543 у пораженных носителей BRCA2 999del5 была выше, чем среди популяционных контролей, это бы означало, что аллель A rs4848543 обуславливает у носителей BRCA2 риск, превышающий риск, обусловленный мутацией 999del5. Для проведения анализа на максимально большое число носителей BRCA2 999del5, группу исландских пациентов Illumina объединили с дублирующей группой. Изучение данных показало, что частота вариантного аллеля A rs4848543 среди носителей BRCA2 999del5, имеющих рак молочной железы, составила 0,463, что было достоверно выше, чем

частота среди контролей и соответствовало RR, равному 1,65 (табл. 3). Таким образом, у носителей BRCA2 999del5 (уже имеющих высокий риск рака молочной железы) риск мультиплицируется на коэффициент, равный по оценкам 1,65 для каждого имеющегося аллеля A rs4848543. Данных, указывающих на то, что генотип rs4848543 отклоняется от мультипликативной модели, получено не было. Данные наблюдения свидетельствуют о том, что повышенный риск рака молочной железы, обусловленный вариантным аллелем A rs4848543 распространяется на носителей BRCA2 999del5. Поскольку BRCA2 999del5 представляет собой ноль-мутацию, при которой образуется нефункциональный белок, эти результаты позволяют предположить, что риск, обусловленный аллелем A rs4848543, распространяется на всех носителей мутаций BRCA2, когда из-за мутации белок превращается в нефункциональный.

Аллель A rs4848543 может обладать синергичным действием с BRCA2 999del5.

Наблюдение за тем, что риск, обусловленный аллелем A rs4848543, распространяется на носителей BRCA2 999del5, заставляет задаться вопросом, обуславливает ли вариант такой же риск у носителей 999del5, как и у лиц, не являющихся носителями 999del5. Было отмечено, что оценки относительного риска, обусловленного аллелем A rs4848543 у носителей BRCA2 999del5 (1,65) были несколько выше, чем оценки относительного риска, обусловленного аллелем A rs4848543 у лиц, не являющихся носителями мутации BRCA2 (1,13, табл. 3). Поэтому было исследовано, отличается ли относительный риск, обусловленный аллелем A rs4848543 в составе генотипа BRCA2 999del5, от обусловленного в составе генотипа с отсутствием 999del5. Как показано в табл. 3, относительный риск, обусловленный аллелем A rs4848543, оценивался в 1,46 раза выше у носителей 999del5 по сравнению с лицами, не являющимися носителями. Данные наблюдения согласуются с мнением, согласно которому аллель A rs4848543 обладает синергичным действием с BRCA2 999del5. Поскольку BRCA2 999del5 представляет собой ноль-мутацию, при которой образуется нефункциональный белок, эти результаты позволяют предположить, что аллель A rs4848543 обладает синергизмом со всеми мутациями BRCA2, когда из-за мутации белок превращается в нефункциональный.

Возможное взаимодействие с BRCA2 999del5 было также исследовано путем оценки того, наблюдается ли у пациентов достоверная корреляция между количеством вариантных аллелей rs4848543 и 999del5. Как показано в табл. 4, у пациентов отмечалась достоверная корреляция между количеством имеющихся вариантных аллелей A rs4848543 и носительством аллелей BRCA2 999del5. Для исследования возможности того, что корреляция между аллелем A rs4848543 и аллелями BRCA2 999del5 у пациентов возникла вследствие стратификации популяции, существование корреляции было изучено в контрольной популяции. У 5938 контролей, прошедших тестирование на наличие обоих вариантов, корреляции не было отмечено. Из данной группы 5938 контролей были исключены пациенты с раком молочной железы. Поскольку BRCA2 999del5 является очень выраженным фактором риска рака молочной железы, может возникнуть предположение, что контрольная группа была истощена по носителям 999del5 (поскольку не страдать раком молочной железы могло относительно небольшое количество женщин-носителей). Поэтому эксперимент был повторно проведен на контрольной группе, состоявшей из 2925 мужчин, которые прошли тестирование на наличие обоих вариантов. В этой группе не было получено данных, свидетельствующих о корреляции между аллелем A rs4848543 и носительством 999del5. Таким образом, было сделано заключение о том, что свидетельств стратификации популяции, которая могла бы быть причиной совместного наследования у пациентов аллеля A rs4848543 и 999del5, обнаружено не было. Следовательно, наиболее вероятным объяснением является то, что аллель A rs4848543 и 999del5 обладают синергичным действием.

Затем мы вновь вернулись к рассмотрению того, что аллель A rs4848543 обнаруживал ассоциацию с семейным анамнезом после исключения из анализа пациентов с известным носительством BRCA2 999del5 (табл. 2). Было исследовано, обнаруживается ли повышенная частота аллеля A rs4848543 у пациентов с высоким риском, не являющихся носителями BRCA2 999del5, но соответствующих критериям фенотипа "УмерПред". У данных пациентов частота была выше, чем у контролей и соответствовала относительному риску 1,35 (табл. 3). Таким образом, ассоциация риска и аллеля A rs4848543 распространяется на пациентов, не являющихся носителями 999del5, но, тем не менее, имеющих характеристики "УмерПред".

Как показано в табл. 1, у пациентов, не соответствующих определению фенотипа "УмерПред", не было обнаружено риска, ассоциированного с аллелем A rs4848543. Это позволяет предположить, что аллель A rs4848543 обуславливает риск исключительно в составе фенотипа "УмерПред". Это может иметь место при существовании относительно редкой мутации с высокой пенетрантностью на фоне вариантного аллеля A rs4848543. Возможно также, что вариантный аллель A rs4848543 может указывать на существование другого более распространенного аллеля, влияющего на пенетрантность BRCA2, а также других, еще не обнаруженных факторов риска с высокой пенетрантностью, ассоциированных с наследственностью. Такие неизвестные факторы риска могут представлять собой варианты других генов с высокой пенетрантностью, полигенные сочетания вариантов, или факторы риска окружающей среды, концентрирующиеся в семьях. Для удобства подобные гипотетические факторы риска с высокой пенетрантностью будут обозначаться BRCAx. Для дальнейшего изучения возможности аллеля A rs4848543 действовать в качестве гена-Модификатора, было исследовано, обнаруживается ли аллель A rs4848543 среди

пораженных носителей BRCA2 999del5 с той же частотой, как и у пораженных пациентов с фенотипом "УмерПред", не являющихся носителями BRCA. Действительно, это подвергает сомнению представление о том, что аллель A rs4848543 обуславливает у носителей 999del5 такой же риск, как и у индивидов с BRCAх. Как показано в табл. 3, между этими группами не наблюдалось достоверной разницы в частотах вариантного аллеля A rs4848543. Тем не менее, точечная оценка относительного риска составляла 1,22, что позволяло предположить, что эффект может быть более выраженным у носителей 999del5.

Носители аллеля A rs4848543 и ряда других маркеров в составе блока HC rs4848543 (в неравновесии по сцеплению с rs4848543) имеют повышенный риск развития первично-множественного рака молочной железы:

Возникновение множественных первичных опухолей молочной железы является индикатором наследственной предрасположенности и имеет важное клиническое значение. Пациенты, у которых был выявлен рак молочной железы, имеют существенно больший риск развития второй первичной опухоли. Генетическая предрасположенность может являться значительным слагаемым такого повышенного риска. По этой причине было исследовано развивался ли первично-множественный рак молочной железы (ПМ РМЖ) с частотой, превышающей ожидаемую, у носителей аллеля A rs4848543 по сравнению с лицами, не являющимися носителями. Данные, полученные с помощью чипов Illumina, указывали на возможность прогнозирования ПМ РМЖ с помощью других маркеров в блоке HC rs4848543 по меньшей мере с таким же успехом, как и с помощью rs4848543. Поэтому была разработана и апробирована панель, состоящая из 8 тестов на маркеры в составе этого блока. При оценке риска ПМ РМЖ с помощью логистической регрессии и коррекции на продолжительность ведения пациентов, для 7 из 8 ОНП был продемонстрирован повышенный риск ПМ РМЖ у носителей при сравнении с пациентами, у которых были диагностированы только одиночные первичные опухоли молочной железы (ОП ОМЖ) (табл. 5). Было сделано заключение, что с помощью анализа данных 7 ОНП возможно определить, имеется ли у пациента повышенный риск ПМ РМЖ.

Несколько ОНП в составе одного блока неравновесия по сцеплению ассоциированы с фенотипами рака молочной железы "все формы РМЖ" и "УмерПред".

Аллель A rs4848543 находится в составе блока неравновесия по сцеплению (HC), который ограничивается координатами, приведенными в табл. 6, в соответствии с NCBI Build 34. Блок HC был определен как участок, располагающийся между горячими точками рекомбинации в соответствии с определением Группы Математической Генетики университета Оксфорда с применением теста отношения правдоподобия, описанного в [McVean, et al., (2004), Science, 304, 581-4; Winckler, et al., (2005), Science, 308, 107-11]. Данные, полученные с помощью чипа Illumina Har300 позволили обнаружить несколько ОНП, находящихся в том же блоке неравновесия по сцеплению, что и аллель A rs4848543, имевших допустимо значимый уровень ассоциации с двумя исследованными факторами, предрасполагающими к раку молочной железы. В табл. 7 и 8 приведен список всех ОНП в блоке HC, представленных в составе чипа Illumina Har300, и показатели ассоциации, полученные по ним для двух фенотипов. Кроме того, были рассмотрены 2-маркерные гаплотипы в составе блока, которые были идентифицированы в качестве хороших суррогатов ($r^2 > 0,8$) для дополнительной группы ОНП, типированных в ходе проекта HarMap (все ссылки на данные HarMap приведены в соответствии с 20 версией). Эти добавочные ОНП имели частоту минорного аллеля $> 5\%$ в выборках UTAH CEPH (CEU) HarMap, и они не были представлены ни в составе чипа Har300 SNP и не маркировались ОНП, представленными в составе чипа [Pe'er, et al., (2006), Nat Genet, 38, 663-7]. Результаты по так называемым гаплотипам "Daly" также приведены в табл. 7 и 8. В случае рака молочной железы "УмерПред" для 29 отдельных ОНП или 2-точечных гаплотипов "Daly" были показаны допустимо значимые значения p .

Риск, ассоциированный с аллелем A rs4848543, прослеживается в составе нескольких гаплотипов.

Для выяснения, маркирует ли аллель A rs4848543 относительно редкий вариант с высокой пенетрантностью или обладает свойствами распространенного модификатора других факторов с высокой пенетрантностью, было исследовано все разнообразие гаплотипов, в составе которых появляется аллель A rs4848543. Блок HC rs4848543 LD содержит участок с достаточно высокой частотой рекомбинации, дистальный конец которого располагается вблизи нуклеотида 120034174. ОНП, обуславливающие допустимо значимый риск рака молочной железы, расположены дистально (правее) от этого участка рекомбинации. Таким образом, анализ разнообразия гаплотипов был ограничен этой зоной в правой части блока. Для идентификации гаплотипов в составе указанного региона (между нуклеотидами 120034174 и 120129001) с применением ОНП и образцов, типированных с помощью чипов Illumina, применяли программу NEMO. Лишние ОНП, которые не были необходимыми для идентификации гаплотипов, были исключены, а затем для каждого полученного гаплотипа были оценены относительные риски и частоты. Результаты приведены в табл. 9. Очевидно, что во всех идентифицированных гаплотипах, несущих аллель A rs4848543, оценка относительного риска составляла более 1 и, в большинстве случаев, оценка относительного риска RR значительно превышала 1. Наоборот, ни один гаплотип, несущий аллель C rs4848543, не имел оценки относительного риска RR выше 1, и все значимые оценки были менее 1 (т.е. они являлись протективными). Таким образом, разделить риск и гаплотипы, несущие аллель A

rs4848543, оказалось невозможно. Это позволяет предположить, что патогенным вариантом (т.е. вариантом, который механистически обуславливает риск) является аллель А rs4848543 или вариант, находящийся в сильном неравновесии по сцеплению с аллелем А rs4848543. Наличие патогенного варианта в составе множества гаплотипов позволяет предположить, что его частота может приблизительно равняться частоте аллеля А rs4848543 (т.е. значение r^2 для аллеля А rs4848543 и патогенного варианта является высоким). Данное наблюдение свидетельствует в пользу точки зрения, что аллель А rs4848543 не указывает на существование редкого варианта с высокой пенетрантностью в составе блока НС rs4848543, а скорее маркирует достаточно распространенный модификатор факторов с высокой пенетрантностью, таких как BRCA2 и BRCAx, которому было дано определение выше.

Многочисленные маркеры могут быть ассоциированы с риском рака молочной железы при их корреляции с аллелем А rs4848543.

Ссылки на данные проекта НаРМар (20 выпуск) позволили обнаружить в выборке белой европейской популяции CEU ряд известных ОНП, сцепленных с rs4848543 со значениями r^2 выше 0,2. Поскольку эти ОНП имеют достоверную корреляцию с rs4848543, их можно использовать для определения риска рака молочной железы наряду с самим rs4848543. Данные ОНП приведены в табл. 10.

Ген STEAP3/TSAP6 является предполагаемым геном, обуславливающим риск рака молочной железы.

На настоящий момент известно три гена в регионе блока НС rs4848543.

Это:

1) ген 2 подобного суб-компоненту 1q компонента комплемента (от англ. Complement Component 1q Subcomponent-like 2 gene, или C1QI2), (NM_182528)

2) ген 6-трансмембранного эпителиального антигена простаты 3 (STEAP3), известный также под названиями гена пути активации опухолевых супрессоров 6 (TSAP6), pHyde и Dudulin-2. Известно три основных альтернативных транскрипта данного гена, обозначаемых в RefSeq под номерами NM_018234 (изоформа b), NM_182915 (изоформа a) и NM_001008410 (также обозначаемый AF262322). Был описан четвертый транскрипт этого гена, получивший идентификатор AK024163.

3) транскрипт AK127773, перекрывающийся с геном STEAP3 и транскрибируемый с противоположной цепи (левее). Этот ген был обозначен как Гипотетический белок FLJ45874.

Блок НС rs4848543 содержит регион с относительно высокой частотой рекомбинации, дистальный конец которого располагается вблизи нуклеотида 120034174. Большинство ОНП и 2-точечные гаплотипы Daly, обуславливающие допустимо значимый уровень риска рака молочной железы, расположены дистально (правее) от этого участка рекомбинации. Это позволяет предположить, что варианты в составе C1QI2 и вблизи него не могут быть источником ассоциативного сигнала. Варианты, в значительной мере ассоциированные с риском, расположены в положении 5' по отношению к гену STEAP3/TSAP6 и в его составе. Действительно, два из ОНП, имеющих допустимый уровень значимости, rs838100 и rs3731603, расположены в составе транскрибируемых последовательностей гена, а другие два, rs838102 и rs12711924, расположены в составе интронов STEAP3/TSAP6. Последний из этих двух ОНП расположен вблизи 3' сайта сплайсинга гена IVS4. С учетом данных наблюдений и сведений о биологии гена STEAP3/TSAP6 (см. ниже), геном, с наибольшей вероятностью задействованным в наблюдающемся повышении риска рака молочной железы, является STEAP3/TSAP6.

Раздел 2: rs13387042 на хромосоме 2q35

Аллель А rs13387042 ассоциирован с повышенным риском рака молочной железы в Исландии:

Общегеномный ассоциативный анализ показал, что аллель А rs13387042 ассоциирован с фенотипом Any BC (табл. 11). Этот ОНП расположен на хромосоме 2q35, на определенном расстоянии от локуса STEAP3/TSAP6. Данные по этому ОНП были получены с помощью чипов Illumina у 1598 пациентов и 4475 контролей, не имеющих рака. Оценки относительного риска RR для А-rs13387042 составили 1,19, а значение Р после коррекции на наличие родства между индивидами в группе "случаев" и "контролей" составило $4,0 \times 10^{-5}$. В группе рака молочной железы "УмерПред" относительный риск RR был аналогичным, а значение Р составило $4,4 \times 10^{-3}$.

Сами по себе данные результаты не достигали уровня значимости, который бы считался достоверным на уровне генома после применения поправки Бонферрони, с учетом 317 089 протестированных ОНП. Поэтому была предпринята попытка подтвердить результаты на независимой выборке из исландской популяции, как это было сделано для аллеля А rs4848543. Для исследования ОНП rs13387042 был разработан и апробирован анализ ОНП на платформе Centaurus, обозначенный SG02S728. Анализ SG02S728 был проведен дополнительно у 583 пациентов и независимой контрольной группе, состоявшей из 7966 индивидов, типированных по аллелю А rs13387042 либо с помощью чипов Illumina, либо с помощью методики Centaurus. В контрольную группу вошли 7966 индивидов с диагностированным раком простаты или колоректальным раком, но не индивиды с диагностированным раком молочной железы. Такой выбор был основан на предварительном наблюдении, свидетельствовавшем, что аллель А rs13387042 не обуславливал сколько-нибудь измеримого риска в отношении этих форм рака (данные не приводятся). Оценки относительного риска RR для аллеля Arsl3387042, ассоциированного с фенотипом

АпуВС в этой второй группе составили 1,20 (значение $P_{\text{корр}} = 3,8 \times 10^{-3}$). Таким образом, исходные результаты были достоверно воспроизведены на независимой исландской выборке, с очень близкими точечными оценками RR. Результаты совместного анализа, при котором 1598 исходных и 583 дублирующих образцов объединили и сравнили с 12441 контролями, приведены в табл. 11. Совместный анализ показал значимую оценку относительного риска, составившую 1,20 (значение $P_{\text{корр}} = 2,0 \times 10^{-7}$), что близко к уровню значимости после коррекции Бонферрони на 317 089 протестированных ОНП. Наблюдавшаяся частота и относительный риск для вариантного аллеля A rs13387042 соответствует оцененному популяционному атрибутивному риску в исландской популяции, равному 15,6%.

Поскольку документацию в ICR начали вести с 1955 г, некоторые из пациентов, вошедших в исследование, имели большую продолжительность выживания после обнаружения рака. Если пациенты с носительством вариантного аллеля A rs13387042 имели бы иную вероятность длительного выживания, чем лица, не несущие его, то частота варианта среди распространенных случаев могла бы различаться. Для изучения этой возможности была выделена подгруппа пациентов, состоявшая из 837 индивидов, которым диагноз был поставлен после 1 января 2000 г, и у которых время от постановки диагноза до включения в исследование составило менее 5 лет. В этой когорте пациентов, которым диагноз был поставлен недавно, частота аллеля A rs13387042 составила 0,498. Также была выделена группа из 1344 пациентов, которым диагноз был поставлен после 1 января 2000 г и которые дожили до включения в исследование. Частота аллеля A rs13387042 в этой группе составила 0,492, что достоверно не отличалось от частоты среди недавно диагностированных пациентов (значение $P = 0,71$). Таким образом, нет очевидных доказательств того, что различная выживаемость отражается на оценках риска, приведенных в табл. 11.

Аллель A rs13387042 ассоциирован с повышенным риском рака молочной железы в испанской выборке случай/контроль.

Для дальнейшего изучения роли аллеля A rs13387042 при раке молочной железы была исследована независимая серия образцов, полученных в Испании у "случаев" с раком молочной железы и "контролей". Испанская исследуемая популяция состояла из 446 "случаев" с раком молочной железы, рекрутированных в онкологическом отделении больницы Сарагоса в г. Сарагоса, Испания. Группа сравнения состояла из 977 индивидов, обратившихся в больницу университета Сарагоса, Испания, у которых было подтверждено отсутствие рака молочной железы до включения их в исследование. Все субъекты подписали информированное согласие.

Испанские "случаи" и "контроли" были генотипированы по rs13387042 при помощи анализа SG02S728 на платформе Centaurus. Результаты представлены в табл. 11. Для аллеля A rs13387042 в данной популяции была обнаружена достоверная ассоциация с риском рака молочной железы, с точечной оценкой относительного риска RR, равной 1,21 (значение P составило $1,8 \times 10^{-2}$). Частота вариантного аллеля A rs13387042 в испанских контрольных образцах была выше, что позволяло предположить, что вариант может иметь большую распространенность и, таким образом, обуславливать больший генетический груз в отношении рака молочной железы в популяциях испанского происхождения.

Аллель A rs13387042 ассоциирован с повышенным риском рака молочной железы в шведской выборке случай/контроль.

Для дальнейшего подтверждения роли вариантного аллеля A rs13387042 в отношении риска рака молочной железы, ОНП был типирован в двух шведских когортах: "шведская семейственная" когорта включала 346 пациентов с раком молочной железы, обратившихся в онкогенетическое консультационное отделение института Каролинска в Стокгольме, для исследования семейного анамнеза рака молочной железы. Каждый пациент относился к отдельной семье. "Шведская консеквативная" когорта включала 482 последовательно рекрутированных пациентов с раком молочной железы, посещавших клинику для пациентов с раком молочной железы с институте Каролинска. При отборе этой второй когорты семейный анамнез не принимался во внимание. Контролями были 1300 доноров крови в районе Стокгольма и 434 не имеющих рака супруга пациентов с колоректальным раком. Среди контролей не отмечалось достоверной разницы между полами по частоте A-rs13387042, так же как и достоверной разницы между группами доноров крови и не имеющих рака "контролей".

Как показано в табл. 11, в "шведской семейственной" был отмечен значимый относительный риск рака молочной железы, равный 1,31 (значение $P = 2,0 \times 10^{-4}$). В "шведской консеквативной" когорте оценка относительного риска, равная 1,11, не была значимой, возможно по причине того, что в этой группе могло иметь место носительство генов с высокой пенетрантностью, маскирующих эффект A-rs13387042. В сумме, в объединенной шведской когорте была получена значимая оценка относительного риска, составившая 1,22 (значение $P = 8,1 \times 10^{-4}$).

Оценки, полученные на исландской, испанской и шведской когортах, были объединены в ходе совместного анализа по методу Мантеля-Хензеля. В результате оцениваемый относительный риск составил 1,20 при значении P , равном $3,8 \times 10^{-11}$. Это существенно ниже порогового уровня общегеномной значимости после применения метода Бонферрони для коррекции на 317 089 исследованных ОНП. Таким образом, было сделано заключение, что A-rs13387042 обуславливает значимый и воспроизводимый риск рака молочной железы в нескольких популяционных выборках европейского происхождения. Совокуп-

ный популяционный атрибутивный риск по оценкам составил 16,4%.

Риск, ассоциированный с аллелем A rs13387042, прослеживается в составе нескольких гаплотипов.

ОНП rs13387042 находится в составе блока неравновесия по сцеплению (НС), который ограничивается координатами, приведенными в табл. 12, в соответствии с NCBI Build 34. Блок НС был определен как участок, расположенный между горячими точками рекомбинации, в соответствии с определением Группы Математической Генетики университета Оксфорда с применением теста отношения правдоподобия, описанного в [McVean et al., (2004), Science, 304, 581-4; Winckler et al., (2005), Science, 308, 107-11]. Для выяснения, маркирует ли аллель A rs13387042 относительно редкий высокопенетрантный вариант или более распространенный низкопенетрантный вариант, были исследованы различные составы гаплотипов, в составе которых появляется аллель A rs13387042. Для идентификации гаплотипов в составе блока НС с применением ОНП и образцов, типированных с помощью чипов Illumina, применяли программу NEMO. Лишние ОНП, которые не были необходимыми для идентификации гаплотипов, были исключены, а затем для каждого полученного гаплотипа были оценены относительные риски и частоты. Результаты приведены в табл. 13. Очевидно, что несколько идентифицированных гаплотипов, несущих аллель A rs13387042, обуславливают относительный риск, по оценкам составляющий более 1, и в нескольких случаях оценка относительного риска RR значительно превышала 1. Напротив, только для одного редкого гаплотипа, несущего аллель C rs13387042, оценка относительного риска RR превышала 1, а две допустимо значимых оценки относительного риска RR для гаплотипов, содержащих аллель C rs13387042, составили менее 1 (т.е. они являлись протективными). Для двух гаплотипов, несущих аллель A rs13387042, не были получены оценки относительного риска RR, превышающие 1, и одна из этих оценок была допустимо значимой. Вместе данные наблюдения позволяют предположить, что патогенный вариант (т.е. вариант, который механистически обуславливает риск) находится в сильном НС с аллелем A rs13387042, но может не являться самим аллелем A rs13387042.

Несколько ОНП в блоке НС rs13387042 ассоциированы с риском рака молочной железы.

Данные, полученные с помощью чипа Illumina Nap300 позволили обнаружить несколько ОНП, находящихся в том же блоке неравновесия по сцеплению, что и rs13387042. В табл. 14 приведен список всех ОНП в блоке НС, представленных в составе чипа Illumina Nap300, и показатели ассоциации, полученные для них в ходе анализа по одной точке и анализа по двум точкам гаплотипа "Daly". Можно заметить, что для 21 ОНП или гаплотипов "Daly" были зарегистрированы допустимо значимые значения P после коррекции на наличие родства среди субъектов.

Многочисленные маркеры могут быть ассоциированы с риском рака молочной железы при их сцеплении с аллелем A rs13387042:

Ссылки на данные проекта NapMap (20 выпуск) позволили обнаружить в выборке белой европейской популяции CEU ряд известных ОНП, сцепленных с rs13387042 со значениями r^2 выше 0,2. Поскольку эти ОНП имеют значимое сцепление с rs13387042, их можно использовать для определения риска рака молочной железы наряду с самим rs13387042. Данные ОНП приведены в табл. 15.

Раздел 3: rs3803662 на хромосоме 16q12

The T аллель of rs3803662 is associated with risk for Breast Cancer: При последующем общегеномном анализе ОНП с применением данных, полученных на 1600 пациентах с раком молочной железы и 11563 контролях, было установлено, что аллель T ОНП rs3803662 обуславливает 1,23-кратное повышение оцениваемого риска рака молочной железы для фенотипа Any BC (табл. 16). Данный результат был подтвержден на второй независимой когорте 594 исландских пациентов с раком молочной железы и 1433 контролей. Объединенные данные по этим двум исландским выборкам позволили получить оценку относительного риска, равную 1,23 и значение $P 2,8 \times 10^{-7}$ после коррекции на наличие родства между индивидами. Это соответствовало оцениваемому популяционному атрибутивному риску, равному 10,1% (табл. 16).

Ассоциация аллеля T rs3803662 с риском рака молочной железы воспроизводится на независимых шведской, испанской и голландской "когорте" с раком молочной железы: "контроли".

Для подтверждения ассоциации аллеля T rs3803662 с риском рака молочной железы была предпринята попытка воспроизвести результаты на трех независимых когортах из других стран. Были генотипированы индивиды из шведской и испанской когорт, описанных выше. Кроме того, были генотипированы 558 случаев с раком молочной железы и 1384 контролей из г. Неймеген, Голландия. Данная когорта была отобрана из популяционного ракового реестра, содержащего информацию о пациентах из восточной части Голландии с раком молочной железы, диагностированным в 2005-2006 гг. Всем пациентам, у которых за это период рак молочной железы был диагностирован в возрасте ранее 70 лет, было предложено принять участие в исследовании. Контрольная группа была набрана в ходе опроса, проведенного в 2002-2003 гг. (Неймегенское Биомедицинское Исследование), на основании случайной выборки из списков населения. Как показано в табл. 16, достоверно повышенный риск, обуславливаемый аллелем T rs3803662, наблюдался во всех трех дублирующих когортах. При совместном анализе этих трех дублирующих исландских когорт совокупная оценка относительного риска составила 1,35, а значение P составило $5,1 \times 10^{-12}$. При совместном анализе исландской и дублирующих когорт была получена суммарная оценка отно-

сительного риска, равная 1,28, и значение P , равное $2,7 \times 10^{-17}$. Эти показатели существенно ниже порогового уровня для достоверности на общегеномном уровне после коррекции на количество протестированных ОНП. Соответствующая оценка совокупного популяционного атрибутивного риска составила 13,4% (табл. 16).

Определение блока НС, содержащего rs3803662 и идентификация возможных кандидатных генов

ОНП rs3803662 находится в составе блока неравновесия по сцеплению (НС), который ограничивается координатами, приведенными в табл. 17, в соответствии с NCBI Build 34. Блок НС был определен как участок, располагающийся между горячими точками рекомбинации, в соответствии с определением Группы Математической Генетики университета Оксфорда с применением теста отношения правдоподобия, описанного в [McVean, et al., (2004), Science, 304, 581-4; Winckler et al., (2005), Science, 308, 107-11]. rs3803662 представлен в составе транскрипта гипотетического белка, обозначенного LOC643714. Этот гипотетический белок, таким образом, включает в себе возможный ген предрасположенности к раку. В блок также входит 5' конец известного гена TNRC9. TNRC9 является членом семейства ассоциированных с хроматином белков с высокой подвижностью (от англ. chromatin associated high mobility group, или HMG), группы белков, в которую входят факторы транскрипции и агенты, вызывающие ремоделирование хроматина. При раке молочной железы с метастазами наблюдалась пониженная экспрессия белка TNRC9 по сравнению с первичными опухолями молочной железы (Olendrowitz, C. 2006, Bachelor of Science Bioinformatics Thesis, Charite-Universitätsmedizin, Berlin). Выявление варианта ОНП вблизи 5' конца этого гена позволяет предположить, что вариант может находиться в неравновесии по сцеплению с генетическими факторами, изменяющими экспрессию или функцию белка TNRC9, что приводит к повышению риска рака молочной железы.

Несколько ОНП в блоке НС rs3803662 ассоциированы с риском рака молочной железы.

Данные, полученные с помощью чипа Illumina Nap300 позволили обнаружить несколько ОНП, находящихся в том же блоке неравновесия по сцеплению, что и rs3803662. В табл. 18 приведен список всех ОНП в блоке НС, представленных в составе чипа Illumina Nap300, и показатели ассоциации, полученные для них в ходе анализа по одной точке и анализа по двум точкам гаплотипа "Daly". Можно заметить, что для 80 ОНП или гаплотипов "Daly" были зарегистрированы допустимо значимые значения P .

Многочисленные маркеры могут быть ассоциированы с риском рака молочной железы при их корреляции с аллелем T rs3803662:

Ссылки на данные проекта NapMap (20 выпуск) позволили обнаружить в выборке белой европейской популяции CEU ряд известных ОНП, сцепленных с rs3803662 со значениями r^2 выше 0,2. Поскольку эти ОНП имеют значимое сцепление с rs3803662, их можно использовать для определения риска рака молочной железы наряду с самим rs3803662. Данные ОНП приведены в табл. 19.

Аллель T rs3803662 и/или связанные с ним ОНП обуславливают риск рака молочной железы в нескольких различных этнических группах.

Для исследования универсальности характера данных, полученных для аллеля T rs3803662, у пациентов с различным этническим происхождением было исследовано наличие ассоциации ОНП в выборке пациентов с раком молочной железы и контролей из мультиэтнического когортного исследования (от англ. Multiethnic Cohort study, или MEC). В MEC вошли более 215 000 мужчин и женщин из Гавайских островов и Лос Анжелеса (а также американцы африканского происхождения из других регионов Калифорнии), подробное описание представлено в [Kolonel et al., (2000), Am J Epidemiol, 151, 346-57]. Когорта состоит преимущественно из американцев африканского происхождения, коренных жителей Гавайских островов, американцев японского происхождения, латиноамериканцев и американцев европейского происхождения, которые были включены в исследование между 1993 и 1996 гг. после заполнения 26-страничного опросника, в котором требовалось указать подробные сведения о характере питания, демографических факторах, личном поведении, предшествующих заболеваниях в анамнезе, семейном анамнезе распространенных форм рака, и для женщин - о репродуктивной истории и применении экзогенных гормонов. При включении в исследование участники имели возраст от 45 до 75 лет. Случаи рака в MEC идентифицировали путем привязки когорты к популяционным раковым реестрам по выживаемости, эпидемиологии и конечным результатам (от англ. Surveillance, Epidemiology, and End Results, или SEER), охватывающим Гавайи и Лос Анжелес, а также Реестру по Раку Штата Калифорния, охватывающему всю Калифорнию. Начиная с 1994 образцы крови брали у пациентов с раком молочной железы и у случайной выборки участников MEC, служивших контрольным пулом для генетического анализа в когорте. Подходящие случаи в гнездовом исследовании случай-контроль были представлены женщинами с инвазивным раком, диагностированным после включения в MEC и до 31 декабря 2002. Контроли были участниками, не имеющими рака молочной железы, соответственно, до включения в когорту и без диагностированного рака молочной железы на 31 декабря 2002. Контроли имели сопоставимую частоту со случаями по расовой/этнической принадлежности и возрасту (в интервалах 5 лет). Исследование было согласовано с Институтскими Наблюдательными Комитетами университета Южной Калифорнии и Университета штата Гавайи.

Была предпринята попытка исследовать риск, ассоциированный с rs3803662 у американцев европейского происхождения, латиноамериканцев и американцев африканского происхождения из MEC. По-

сколькo корреляция между ОНП rs13387042 и другими ОНП могла различаться в различных этнических группах, для исследования на этой выборке был выбран ряд ОНП. ОНП, их значения r^2 и D' , полученные по данным проекта НарМар (20 выпуск) для американцев европейского происхождения и Йоруба (африканцев) приведены в табл. 23. Для этих ОНП были разработаны анализы на платформе Centaurus и для образцов из МЕС было проведено генотипирование. Для удобства сравнения результатов по ОНП был использован тест обобщенного отношения правдоподобия, учитывающий отсутствующие генотипы при анализе гаплотипов. Это позволило уравновесить количество случаев и контролей по различным ОНП, и позволило легче проводить сравнение между ними. Результаты приведены в табл. 24. Прежде всего, очевидно, что риск, обусловленный аллелем T rs3803662, достоверно воспроизводится у американцев европейского происхождения. Два из взаимосвязанных ОНП, аллель T rs4784227 и аллель C rs17271951, имеют более высокий относительный риск при допустимом уровне значимости и более низкие значения P, чем аллель T rs3803662 у американцев европейского происхождения. Этот пример свидетельствует в пользу концепции, согласно которой взаимосвязанные с исходно выявленным ОНП, могут обуславливать более высокий относительный риск и таким образом служить лучшими маркерами заболевания.

Сигнал, обусловленный аллелем T rs3803662, также с достоверностью воспроизводится среди латиноамериканцев. У американцев африканского происхождения результаты по аллелю T маркера rs3803662 с достоверностью не воспроизводились, в действительности, аллель T оказался достоверно протективным в выборке американцев африканского происхождения. Это позволяет предположить, что взаимоотношения между аллелем T rs3803662 и предполагаемой патогенной мутацией значительно отличаются у американцев африканского происхождения. Было отмечено, что три взаимосвязанных аллеля ОНП; аллель T rs12922061, аллель T rs4784227 и аллель C rs17271951 имеют сильную корреляцию с аллелем T rs3803662 у европейцев, однако данная корреляция не прослеживается у африканцев Йоруба (табл. 24). Однако, в отличие от аллеля T rs3803662, все аллели взаимосвязанных ОНП, аллель T rs12922061, аллель T rs4784227 и аллель C rs17271951, имеют повышенный относительный риск рака молочной железы при допустимом уровне значимости, и два из них достигают статистической достоверности на с 1-сторонними значениями P (табл. 24). Это показывает, что при использовании групп, взаимосвязанных ОНП, можно получить маркеры риска, которые можно применять в различных этнических группах.

Аллель A rs13387042 и аллель T rs3803662 обуславливают повышенный риск при раке молочной железы, позитивный по эстрогеновым рецепторам.

При наличии документации по пациентам, включенным в исследование, проводили ее анализ. На объединенных группах образцов была исследована ассоциация между аллелем A rs13387042 и аллелем T rs3803662, возрастом при постановке диагноза, статусом по эстрогеновым рецепторам (ЭР) и прогестероновым рецепторам (ПР). Ни для одного из вариантов не была обнаружена ассоциация с возрастом при постановке диагноза. Достоверный риск рака молочной железы, обусловленный аллелем A rs13387042 и аллелем T rs3803662, наблюдался исключительно у пациентов с диагностированными ЭР-позитивными опухолями и разница между отношениями шансов при ЭР-позитивных и ЭР-негативных опухолях была достоверной (табл. 24). Аналогично, наблюдалась тенденция к выявлению риска рака молочной железы преимущественно у пациентов с диагностированными ПР-позитивными опухолями, однако разница между отношениями шансов у ПР-позитивных и ПР-негативных была недостоверна. Это позволяет предположить, что ЭР-позитивные и негативные опухоли могут иметь в некоторой степени различную генетическую этиологию.

Таблицы

Таблица 1. Ассоциация аллеля A ОНП SG02S738 (rs4848543) с риском рака молочной железы

	p	p_корр*	r	95%CI	#пораж	Част поража	#конт	Част конт
«УмерПред» Illumina	6,70E-09	8,30E-08	1,42	(1,25,1,62)	653	0,423	4477	0,341
«УмерПред» Исландия дублир.	2,90E-02	3,40E-02	1,26	(1,02,1,55)	198	0,399	7406	0,346
«УмерПред» объедин.	9,70E-10	3,00E-08	1,37	(1,23,1,53)	851	0,418	11883	0,344
«Все формы РМЖ» Illumina	7,10E-04	1,10E-03	1,16	(1,06,1,26)	1598	0,374	4477	0,341
«Все формы РМЖ» Исландия дублир.	9,70E-02	1,00E-01	1,11	(0,98,1,26)	573	0,37	7406	0,346
«Все формы РМЖ» объедин.	2,20E-04	3,50E-04	1,14	(1,06,1,22)	2171	0,373	11883	0,344
не-«УмерПред» в сравн с Конт	9,80E-01		1		1320	0,344	11883	0,344

*после коррекции на наличие родства между индивидами

Таблица 2. Ассоциация ОНП SG02S738 (rs4848543) с возрастом при постановке диагноза и семейным анамнезом рака молочной железы

Сравнение	Значение P (Вилкоксона)
Генотип rs4848543 и Возраст при постановке диагноза	0,5139
Генотип rs4848543 и Индекс семейного анамнеза	<0,001
Генотип отличный от BRCA2 rs4848543 и Индекс семейного анамнеза	0,0044

Таблица 3. Ассоциация ОНП SG02S738 (rs4848543) с риском рака молочной железы среди носителей BRCA2 999del5

	р_корр*	р	г	95%CI	#пораж	Част поража	#конт	Част поража
Носители BRCA2 в сравн с Контр	3,67E-03	8,02E-04	1,65	(1,18,2,30)	94	0,463	9452	0,344
Без носительства BRCA2 в сравн с Контр	2,95E-03	2,19E-03	1,13	(1,04,1,22)	1621	0,371	9452	0,344
Носители BRCA2 в сравн с не имеющими носительства BRCA2	3,06E-02	1,29E-02	1,46	(1,04,2,05)	94	0,463	1621	0,371
Без носительства «УмерПред» BRCA2 в сравн с Контр	4,71E-06	5,92E-07	1,35	(1,19,1,55)	622	0,414	9452	0,344
Носители BRCA2 в сравн с не имеющими носительства «УмерПред» BRCA2	Не опр	2,08E-01	1,22	Не опр	94	0,463	622	0,414

*после коррекции на наличие родства между индивидами

Таблица 4. Корреляция между аллелем A rs4848543 и генотипом BRCA2 999del5

Корреляция генотипа BRCA2 - STEAP3			
(тест Пирсона)			
	Значение Р	Коеф коррел.	#индив
Пациенты с раком молочной железы	1,4E-02	0,059	1715
Контроли	8,0E-01	-0,0033	5938
Контроли мужчины	6,7E-01	-0,0079	2925

Таблица 5. Ассоциация ОНП с риском первично-множественного рака молочной железы

#ПМ РМЖ	Част ПМ РМЖ	#ОП ОМЖ	Част ОП ОМЖ	Р носит	RR носит	Аллель маркера	rs SNP	Полож в Seq ID 4
128	0,258	1566	0,189	0,0111	1,615	4 SG02S733	rs895398	34898
160	0,409	2010	0,370	0,0433	1,419	1 SG02S738	rs4848543	37814
160	0,547	2019	0,488	0,0438	1,498	1 SG02S739	rs6759589	41694
163	0,414	2001	0,371	0,0389	1,427	4 SG02S753	rs895397	49118
162	0,645	2020	0,575	0,0072	1,928	3 SG02S740	rs838102	56596
160	0,650	2014	0,575	0,0050	2,012	3 SG02S741	rs838100	60428
127	0,445	1565	0,380	0,0176	1,602	1 SG02S734	rs12711924	84034
159	0,387	2004	0,343	0,1741	1,258	4 SG02S742	rs3731603	93780

Таблица 6. Определение блока HC rs4848543 на Хромосоме 2

NCBI Build	Начало	Окончание
34	119,987,002	120,129,001
36	119,608,327	119,750,326

Таблица 7. Ассоциация ОНП маркеров и 2-маркерных гаплотипов в блоке HC rs4848543 с риском рака молочной железы (фенотип "все формы РМЖ"):

Приведены следующие коды аллелей: 1=A, 2=C, 3=G, 4=T.

Значение p	RR	#пораж	Част пораж	#конт	Част конт	Част H0	X2	инфо	аллель	Маркер	аллель	Маркер
0,0000825	1,1852	1557	0,377649	4626	0,33863	0,348455	15,5001	1	1	rs4848543		
0,00015	0,8545	1558	0,498176	4627	0,537411	0,527529	14,3718	0,999171	2	rs895398	3	rs6759589
0,0001842	1,1677	1558	0,500642	4626	0,461954	0,471701	13,9859	1	1	rs6759589		
0,0004919	1,2131	1558	0,18702	4627	0,159399	0,166396	12,1462	0,966913	2	rs4849765	4	rs895398
0,0006022	1,1539	1558	0,560738	4627	0,525225	0,534172	11,7693	0,994181	2	rs4849760	3	rs838102
0,0006816	1,1529	1558	0,587612	4625	0,552757	0,56154	11,5388	1	3	rs838102		
0,0007483	1,1517	1556	0,587082	4622	0,552466	0,561185	11,3653	1	3	rs838100		
0,0007856	0,8681	1558	0,401958	4627	0,436379	0,427706	11,2749	0,995483	3	rs6759589	1	rs838102
0,0009395	1,1929	1557	0,196853	4626	0,17045	0,177098	10,9433	1	4	rs895398		
0,0009835	1,1959	1556	0,204802	4619	0,177192	0,184711	10,8584	0,936107	2	rs4849765	4	rs2084154
0,0019878	1,1829	1558	0,187099	4626	0,162884	0,168984	9,5608	1	4	rs935361		
0,0039725	1,3609	1558	0,067214	4627	0,050286	0,054483	8,2963	0,668984	2	rs838086	1	rs12711924
0,0040463	1,1296	1558	0,387356	4627	0,358872	0,366047	8,1083	1	1	rs12711924		
0,0053777	1,1574	1558	0,202503	4627	0,179922	0,18561	7,7478	1	2	rs7564432	2	rs11680207
0,0076689	1,1238	1558	0,349807	4627	0,323752	0,330315	7,1092	1	4	rs3731603		
0,0092636	0,895	1558	0,58833	4626	0,614901	0,608199	6,7713	0,984119	2	rs6542504	2	rs12711923
0,0119574	1,1118	1557	0,418433	4626	0,392888	0,399321	6,3172	1	4	rs1530561		
0,0120093	1,1289	1558	0,253591	4627	0,231338	0,236949	6,3095	0,998425	1	rs745888	2	rs838087
0,0127242	1,1276	1558	0,253851	4625	0,231784	0,237344	6,2071	1	1	rs745888		
0,0140734	0,8975	1554	0,330746	4621	0,355095	0,348951	6,0289	0,987772	3	rs2084154	4	rs838066
0,0143097	0,8986	1558	0,337519	4627	0,361823	0,355701	5,9995	0,993263	3	rs6759589	4	rs838066
0,0158735	0,9024	1558	0,586151	4627	0,610816	0,60459	5,8168	0,983893	2	rs6542504	2	rs4849760
0,0166389	1,1057	1558	0,432285	4627	0,407824	0,413985	5,7341	1	3	rs838069		
0,0199778	1,1096	1558	0,332002	4626	0,309345	0,315063	5,4138	0,982361	4	rs6542504	2	rs12711923
0,02322	0,9	1558	0,716624	4627	0,737523	0,732257	5,152	0,998986	2	rs4849760	3	rs745888
0,0234152	0,9054	1558	0,339147	4627	0,361769	0,356092	5,1374	0,982618	2	rs895398	4	rs838066
0,0236887	1,1011	1555	0,397106	4610	0,374295	0,380049	5,1173	1	4	rs6542504		

Значение p	RR	#пораж	Част пораж	#конт	Част конт	Част H0	X2	инфо	аллель	Маркер	аллель	Маркер
0,0346288	1,0972	1522	0,658673	4559	0,63753	0,642822	4,4634	1	2	rs838066		
0,0376797	0,9131	1558	0,444559	4627	0,467115	0,461494	4,3194	0,903903	1	rs838086	1	rs838069
0,0612227	0,9009	1558	0,161601	4627	0,176246	0,172556	3,5039	0,98925	3	rs6759589	2	rs838066
0,0675678	1,0805	1558	0,403402	4627	0,384915	0,389572	3,3411	1	2	rs11680207		
0,0742724	1,1028	1558	0,20821	4627	0,192542	0,196469	3,186	0,887553	3	rs1867750	2	rs11680207
0,1275005	1,0674	1511	0,527134	3962	0,510853	0,515348	2,3227	1	4	rs2084154		
0,1817962	0,9326	1558	0,209889	4627	0,221696	0,218727	1,7829	0,931297	1	rs10496555	3	rs745888
0,2193745	1,056	1542	0,679313	4602	0,667319	0,670329	1,5085	1	2	rs4849765		
0,2349078	1,0589	1548	0,251292	4614	0,240681	0,243346	1,4109	1	2	rs838086		
0,2686289	0,9532	1558	0,456896	4627	0,468807	0,465537	1,2227	0,919405	3	rs7562894	3	rs2084154
0,2941036	0,9538	1558	0,331602	4627	0,342173	0,339281	1,1007	0,945087	4	rs2084154	2	rs895398
0,455931	1,0331	1558	0,354244	4627	0,346821	0,34869	0,5559	0,98416	2	rs330763	3	rs725346
0,4591499	1,033	1558	0,358115	4627	0,350681	0,352554	0,548	0,971987	3	rs1439945	3	rs725346
0,4793247	1,0562	1557	0,079641	4622	0,075725	0,076711	0,5004	1	1	rs12711923		
0,497656	1,03	1558	0,356544	4627	0,349799	0,351497	0,4599	0,989528	2	rs330763	3	rs10496556
0,5015258	0,9669	1558	0,216303	4627	0,222066	0,220614	0,4517	1	2	rs11680207	2	rs935361
0,5315142	1,0273	1553	0,370573	4610	0,364317	0,365893	0,3915	1	3	rs725346		
0,5977098	1,0229	1557	0,376365	4619	0,371076	0,372409	0,2785	1	3	rs10496556		
0,6233236	1,0256	1558	0,799101	4627	0,795007	0,796039	0,2412	1	2	rs7564432		
0,7140552	1,0158	1558	0,33668	4627	0,333192	0,334071	0,1343	1,054416	3	rs10496555	3	rs725346
0,7338904	0,9756	1558	0,084777	4627	0,086711	0,086224	0,1156	1,040852	3	rs12711924	3	rs838069
0,7373932	1,0483	1554	0,977479	4049	0,976416	0,976709	0,1124	1	2	rs838087		
0,7425392	1,0407	1558	0,970475	4627	0,96931	0,969604	0,1079	1	2	rs4849760		
0,8144482	1,01	1558	0,341581	4627	0,339335	0,339901	0,0551	1,052093	3	rs10496555	3	rs10496556
0,8352386	0,9855	1558	0,092358	4627	0,093587	0,093277	0,0433	1,03608	4	rs330773	3	rs745888
0,8666485	1,0069	1558	0,421023	4627	0,419347	0,419769	0,0262	1,049887	3	rs10496555	1	rs725346
0,8876881	1,0095	1558	0,106083	4627	0,10519	0,105415	0,0199	1,01195	1	rs1867750	2	rs7564432
0,915614	1,0149	1557	0,977521	4626	0,977194	0,977276	0,0112	1	3	rs1545360		
0,9646863	0,998	1558	0,700899	4627	0,701318	0,701213	0,002	1	3	rs1867750		
0,9767802	0,9959	1558	0,021916	4627	0,022003	0,021981	0,0008	1,023095	1	rs1545360	1	rs10165398

Таблица 8. Ассоциация ОНП маркеров и гаплотипов из 2 маркеров в блоке НС rs4848543 LD с риском рака молочной железы (фенотип "умеренная предрасположенность к РМЖ"):

Приведены следующие коды аллелей: 1=A, 2=C, 3=G, 4=T.

Значение p	RR	#пораж	Част пораж	#конт	Част конт	Част H0	X2	инфо	аллель	Маркер	аллель	Маркер
9,40E-10	1,4567	632	0,427215	4626	0,33863	0,349277	37,4465	1	1	rs4848543		
1,33E-06	0,748	632	0,464983	4627	0,537453	0,528745	23,3794	0,999257	2	rs895398	3	rs6759589
1,51E-06	1,3348	632	0,534019	4626	0,461954	0,470616	23,1368	1	1	rs6759589		
1,62E-06	1,3411	632	0,428798	4627	0,358872	0,367275	22,9946	1	1	rs12711924		
2,57E-06	1,4384	632	0,214003	4627	0,159157	0,165836	22,1127	0,970516	2	rs4849765	4	rs895398
4,01E-06	1,4104	632	0,224684	4626	0,17045	0,176968	21,2621	1	4	rs895398		
1,00E-05	1,3099	632	0,61788	4622	0,552466	0,560335	19,5116	1	3	rs838100		
0,0000125	1,3136	632	0,386076	4627	0,323752	0,331242	19,0859	1	4	rs3731603		
1,398E-05	1,3039	632	0,617089	4625	0,552757	0,560491	18,8713	1	3	rs838102		
0,0000142	0,7657	632	0,37216	4627	0,43636	0,428644	18,8419	0,995144	3	rs6759589	1	rs838102
1,579E-05	1,388	632	0,230115	4619	0,177183	0,184123	18,6401	0,943665	2	rs4849765	4	rs2084154
0,0000258	1,2911	632	0,588229	4627	0,525264	0,532825	17,7045	0,993478	2	rs4849760	3	rs838102
0,0004697	1,2366	631	0,444533	4626	0,392888	0,399087	12,2324	1	4	rs1530561		
0,0007986	0,7553	632	0,139112	4627	0,176242	0,171752	11,2444	0,995218	3	rs6759589	2	rs838066
0,0025521	1,2002	632	0,452532	4627	0,407824	0,413196	9,1029	1	3	rs838069		
0,0027736	1,2608	632	0,196994	4626	0,162884	0,166984	8,9507	1	4	rs935361		
0,0038501	1,5279	632	0,074928	4627	0,050342	0,053363	8,3532	0,690307	2	rs838086	1	rs12711924
0,0086964	0,8452	631	0,317499	4621	0,355011	0,350481	6,8841	0,989658	3	rs2084154	4	rs838066
0,0109182	1,1925	632	0,264011	4627	0,231241	0,235192	6,4786	0,998578	1	rs745888	2	rs838087
0,0116941	1,1903	632	0,264241	4625	0,231784	0,235686	6,3567	1	4	rs745888		
0,0125611	0,8469	632	0,704114	4627	0,737523	0,733507	6,2299	0,99995	2	rs4849760	3	rs745888
0,0140093	0,8557	632	0,326665	4627	0,361823	0,357582	6,037	0,997006	3	rs6759589	4	rs838066
0,0167713	1,1975	632	0,20807	4627	0,179922	0,183305	5,7202	1	2	rs7564432	2	rs11680207
0,0206458	0,8629	632	0,328408	4627	0,36172	0,357717	5,3665	0,986182	2	rs895398	4	rs838066
0,0250242	0,8675	632	0,431088	4627	0,466247	0,462051	5,0222	0,905031	1	rs838086	1	rs838069
0,0262532	0,772	632	0,068325	4627	0,086756	0,084542	4,9393	0,958473	3	rs12711924	3	rs838069
0,0331553	1,1454	621	0,668277	4559	0,63753	0,641216	4,5377	1	2	rs838066		

Значение p	RR	#пораж	Част пораж	#конт	Част конт	Част H0	X2	инфо	аллель	Маркер	аллель	Маркер
0,0420433	0,8752	632	0,312809	4627	0,342145	0,33827	4,1335	0,956363	4	rs2084154	2	rs895398
0,0490954	1,1279	632	0,413766	4627	0,384915	0,388382	3,8721	1	2	rs11680207		
0,0545878	1,1328	625	0,6944	4602	0,667319	0,670557	3,6946	1	2	rs4849765		
0,0565963	1,1614	632	0,216878	4627	0,19254	0,195451	3,6344	0,887609	3	rs1867750	2	rs11680207
0,1002333	1,1066	609	0,536125	3962	0,510853	0,51422	2,7018	1	4	rs2084154		
0,1098139	1,4252	630	0,983333	4049	0,976414	0,977346	2,5569	1	2	rs838087		
0,2444521	0,9299	632	0,450621	4627	0,46866	0,466267	1,3547	0,930417	3	rs7562894	3	rs2084154
0,2972591	1,0681	632	0,36608	4627	0,350928	0,352761	1,0865	0,9763	3	rs1439945	3	rs725346
0,324048	0,9409	632	0,600571	4626	0,615101	0,613351	0,9725	0,985757	2	rs6542504	2	rs12711923
0,3389162	1,0619	632	0,363391	4627	0,349623	0,351284	0,9145	0,992987	2	rs330763	3	rs10496556
0,3390482	0,9427	632	0,596933	4627	0,611053	0,609353	0,914	0,984363	2	rs6542504	2	rs4849760
0,3489783	1,0609	632	0,360144	4627	0,346646	0,348274	0,8772	0,98708	2	rs330763	3	rs725346
0,363347	1,0608	632	0,321891	4626	0,309141	0,310678	0,8263	0,984258	4	rs6542504	2	rs12711923
0,4496627	0,9444	632	0,212841	4627	0,222576	0,221424	0,5715	0,927154	1	rs10496555	3	rs745888
0,457771	1,047	631	0,385103	4610	0,374295	0,375596	0,5513	1	4	rs6542504		
0,4606554	0,9296	632	0,098529	4627	0,105201	0,104399	0,5443	1,01297	1	rs1867750	2	rs7564432
0,47764	1,0451	630	0,374603	4610	0,364317	0,365553	0,5042	1	3	rs725346		
0,4879223	1,1566	632	0,980221	4626	0,977194	0,977558	0,4811	1	3	rs1545360		
0,498098	1,045	632	0,343589	4627	0,333725	0,334913	0,459	0,948386	3	rs10496555	3	rs725346
0,5144968	1,0412	632	0,380538	4619	0,371076	0,372215	0,4249	1	3	rs10496556		
0,5466543	1,0398	632	0,348742	4627	0,339941	0,341001	0,3633	0,950967	3	rs10496555	3	rs10496556
0,5484699	0,8812	632	0,019407	4627	0,021965	0,021657	0,3601	1,017257	1	rs1545360	1	rs10165398
0,6700167	0,9696	632	0,216772	4627	0,222066	0,22143	0,1816	1	2	rs11680207	2	rs935361
0,7411272	0,9806	632	0,413979	4627	0,41875	0,418177	0,1091	1,048229	3	rs10496555	1	rs725346
0,7734425	0,971	632	0,090995	4627	0,093457	0,093161	0,0829	1,032733	4	rs330773	3	rs745888
0,8205628	1,0258	632	0,07753	4622	0,075725	0,075942	0,0514	1	1	rs12711923		
0,8541421	0,9688	632	0,968355	4627	0,969313	0,969196	0,0338	1	2	rs4849760		
0,9360615	0,9944	628	0,23965	4614	0,240681	0,240557	0,0064	1	2	rs838086		
0,9536622	0,9957	632	0,794304	4627	0,795007	0,794923	0,0034	1	2	rs7564432		
0,9754573	1,002	632	0,70174	4627	0,701318	0,701369	0,0009	1	3	rs1867750		

Таблица 9. Гаплотипы в правой части блока rs4848543:
(Приведены следующие коды аллелей: 1=A, 2=C, 3=G, 4=T)

Значение р	г	#пораж	Част пораж	#конт	Част конт	инфо	
2,65E-02	1,23	653	0,131	4480	0,109	0,96	2 rs895398 1 rs4848543 1 rs6759589 2 rs838066 3 rs838100 1 rs838086 1 rs12711924 4 rs3731603
1,51E-02	1,26	653	0,123	4480	0,100	0,94	4 rs895398 1 rs4848543 1 rs6759589 2 rs838066 3 rs838100 1 rs838086 1 rs12711924 4 rs3731603
1,70E-02	1,43	653	0,051	4480	0,036	0,90	4 rs895398 1 rs4848543 1 rs6759589 2 rs838066 3 rs838100 2 rs838086 1 rs12711924 4 rs3731603
2,97E-01	1,17	653	0,043	4480	0,037	0,99	2 rs895398 1 rs4848543 1 rs6759589 2 rs838066 3 rs838100 1 rs838086 1 rs12711924 4 rs3731603
1,19E-02	1,55	653	0,039	4480	0,026	0,89	4 rs895398 1 rs4848543 1 rs6759589 2 rs838066 3 rs838100 2 rs838086 3 rs12711924 4 rs3731603
3,10E-01	1,36	653	0,015	4480	0,011	0,76	2 rs895398 1 rs4848543 1 rs6759589 2 rs838066 3 rs838100 2 rs838086 1 rs12711924 4 rs3731603
2,08E-02	0,85	653	0,244	4480	0,274	0,97	2 rs895398 2 rs4848543 3 rs6759589 4 rs838066 1 rs838100 1 rs838086 3 rs12711924 4 rs3731603
6,56E-03	0,77	653	0,107	4480	0,134	0,96	2 rs895398 2 rs4848543 3 rs6759589 2 rs838066 1 rs838100 1 rs838086 3 rs12711924 4 rs3731603
5,95E-02	0,81	653	0,072	4480	0,088	0,96	2 rs895398 2 rs4848543 1 rs6759589 2 rs838066 3 rs838100 2 rs838086 3 rs12711924 4 rs3731603
3,91E-01	0,88	653	0,043	4480	0,049	0,90	2 rs895398 2 rs4848543 3 rs6759589 4 rs838066 3 rs838100 2 rs838086 3 rs12711924 4 rs3731603
3,00E-02	0,59	653	0,014	4480	0,024	0,83	2 rs895398 2 rs4848543 3 rs6759589 2 rs838066 3 rs838100 2 rs838086 3 rs12711924 4 rs3731603
8,84E-01	1,03	653	0,022	4480	0,021	0,91	2 rs895398 2 rs4848543 1 rs6759589 2 rs838066 3 rs838100 1 rs838086 1 rs12711924 4 rs3731603
8,33E-01	1,05	653	0,022	4480	0,021	0,91	2 rs895398 2 rs4848543 3 rs6759589 4 rs838066 3 rs838100 1 rs838086 1 rs12711924 4 rs3731603
7,42E-01	1,09	653	0,015	4480	0,013	0,88	2 rs895398 2 rs4848543 1 rs6759589 2 rs838066 3 rs838100 1 rs838086 3 rs12711924 4 rs3731603
6,26E-01	0,86	653	0,010	4480	0,012	0,91	2 rs895398 2 rs4848543 3 rs6759589 2 rs838066 1 rs838100 1 rs838086 1 rs12711924 4 rs3731603

Таблица 10. ОНП маркеры НарМар, сцепленные с rs4848543 с показателем r^2 выше 0,2

ОНП суррогат	Реф ОНП	D'	R ²	Значение р	Положение в Build 34	Положение в Build 36	Положение в SEQ ID 4
rs11680207	rs4848543	0,653174	0,309094	4,55E-09	120023583	119644908	301
rs7583824	rs4848543	0,655931	0,322684	3,29E-09	120027724	119649049	4442
rs4848541	rs4848543	0,656671	0,243502	8,45E-06	120032169	119653494	8887
rs1446128	rs4848543	0,628705	0,317935	8,08E-09	120032378	119653703	9096
rs4849766	rs4848543	0,851071	0,31689	7,16E-09	120034354	119655679	11072
rs6715243	rs4848543	0,923886	0,362632	3,83E-10	120034730	119656055	11448
rs4142753	rs4848543	0,934092	0,413776	2,96E-12	120051961	119673286	28679
rs10188745	rs4848543	0,934092	0,413776	2,96E-12	120057058	119678383	33776
rs895398	rs4848543	0,934092	0,413776	2,96E-12	120058180	119679505	34898
rs4848543	rs4848543	1	1	0	120061096	119682421	37814
rs11684731	rs4848543	1	0,319149	2,08E-12	120062319	119683644	39037
rs12464139	rs4848543	1	0,294118	1,69E-11	120062996	119684321	39714
rs6754664	rs4848543	1	0,963415	4,90E-33	120063070	119684395	39788
rs6754799	rs4848543	1	0,673648	6,09E-22	120063176	119684501	39894
rs6759589	rs4848543	1	0,534483	2,36E-19	120064976	119686301	41694
rs1562256	rs4848543	0,951061	0,50597	1,03E-15	120066318	119687643	43036
rs1446125	rs4848543	1	1	4,29E-31	120066620	119687945	43338
rs1446124	rs4848543	1	0,963415	4,90E-33	120066705	119688030	43423
rs11677262	rs4848543	1	0,30137	6,60E-12	120067290	119688615	44008
rs7565771	rs4848543	1	0,307692	2,81E-12	120070400	119691725	47118
rs7595954	rs4848543	1	0,289474	5,19E-12	120070666	119691991	47384
rs838066	rs4848543	1	0,3	2,40E-12	120071508	119692833	48226
rs895397	rs4848543	1	1	4,32E-35	120072400	119693725	49118

ОНП суррогат	Реф ОНП	D'	R ²	Значение p	Положение в Build 34	Положение в Build 36	Положение в SEQ ID 4
rs838059	rs4848543	0,841262	0,353861	2,97E-10	120075395	119696720	52113
rs838102	rs4848543	1	0,321918	4,92E-13	120079878	119701203	56596
rs3769659	rs4848543	0,849119	0,718225	2,57E-20	120081128	119702453	57846
rs865108	rs4848543	1	0,376923	4,62E-14	120081529	119702854	58247
rs708670	rs4848543	1	0,302817	4,27E-12	120082382	119703707	59100
rs708672	rs4848543	1	0,632075	3,65E-22	120082499	119703824	59217
rs708673	rs4848543	1	0,3125	1,13E-12	120082871	119704196	59589
rs838100	rs4848543	1	0,3	2,40E-12	120083710	119705035	60428
rs838098	rs4848543	1	0,3125	1,13E-12	120084226	119705551	60944
rs838096	rs4848543	1	0,310811	1,10E-12	120084741	119706066	61459
rs838092	rs4848543	1	0,321918	4,92E-13	120091298	119712623	68016
rs838090	rs4848543	0,780443	0,224773	5,57E-06	120092304	119713629	69022
rs12467944	rs4848543	0,865803	0,535809	9,44E-15	120094306	119715631	71024
rs10196144	rs4848543	0,84638	0,686857	4,46E-19	120095312	119716637	72030
rs6728660	rs4848543	0,872481	0,70845	3,68E-18	120095756	119717081	72474
rs10188946	rs4848543	0,74499	0,446479	2,07E-12	120098925	119720250	75643
rs895406	rs4848543	0,790753	0,304687	1,52E-09	120104365	119725690	81083
rs708675	rs4848543	0,707874	0,418005	1,81E-11	120104502	119725827	81220
rs1867856	rs4848543	0,77707	0,512514	2,87E-13	120104672	119725997	81390
rs3754847	rs4848543	0,717855	0,22469	4,70E-07	120105720	119727045	82438
rs12711924	rs4848543	0,792931	0,561965	3,43E-15	120107316	119728641	84034
rs1867749	rs4848543	0,75345	0,214069	6,37E-07	120109057	119730382	85775
rs838073	rs4848543	0,709981	0,247536	5,75E-07	120109754	119731079	86472
rs3731603	rs4848543	0,784597	0,392356	2,49E-10	120117062	119738387	93780

Таблица 11. Ассоциация аллеля А ОНП SG02S728 (rs13387042) с риском рака молочной железы

Когорта	Значение p - корр*	Значение p	r	95%CI	#пораж	Част поража	#конт	Част конт
«УмерПред» Illumina	4,4E-03	2,1E-03	1,20		652	0,496	4475	0,451
«УмерПред» Исландия дублир.	2,3E-03	1,8E-03	1,37		200	0,528	7966	0,449
«УмерПред» Исландия объедин.	8,5E-05	1,5E-05	1,24	(1,12,1,38)	852	0,504	12441	0,449
«любой РМЖ» Illumina	4,0E-05	2,1E-05	1,19		1598	0,494	4475	0,451
«любой РМЖ» Исландия дублир.	3,8E-03	3,3E-03	1,20		583	0,493	7966	0,449
«любой РМЖ» Исландия объедин.	2,0E-07	5,0E-08	1,20	(1,12,1,28)	2181	0,494	12441	0,449
Испанская когорта		1,8E-02	1,21	(1,03,1,42)	446	0,581	977	0,533
Шведская семейственная когорта		2,2E-01	1,11		346	0,516	1734	0,490
Шведская консеквативная когорта		2,0E-04	1,31		482	0,558	1734	0,490
Все шведские когорты		8,1E-04	1,22	(1,09,1,37)	828	0,540	1734	0,490
Все когорты	3,8E-11		1,20	(1,14,1,27)	3455	0,516	15152	0,460

*после коррекции на родство между индивидами

Таблица 12. Границы блока НС rs13387042 на хромосоме 2

NCBI Build	Начало	Окончание
34	218,062,001	218,141,002
36	217,567,902	217,646,679

Таблица 13. Разнообразие гаплотипов блока НС rs13387042

Приведены следующие коды аллелей: 1=A, 2=C, 3=G, 4=T.

p	r	#пораж	част поража	#конт	част конт	инфо	34	1	45	2	3	30	4	9	3	42	44	60
8,4E-03	1,12	1599	0,283	11122	0,261	0,99	3	4	3	1	4	1	1	2	3	1	1	3
2,7E-03	1,18	1599	0,144	11122	0,125	0,98	3	4	3	3	4	1	2	4	3	1	3	3
3,0E-01	1,13	1599	0,031	11122	0,027	0,96	3	2	3	3	4	1	2	4	3	1	3	3
2,9E-01	0,85	1599	0,014	11122	0,017	1,00	3	4	3	3	4	1	1	2	3	1	3	3
6,4E-02	0,65	1599	0,006	11122	0,009	1,00	3	4	1	3	2	1	1	2	3	1	1	3
3,0E-02	2,24	1599	0,005	11122	0,002	0,72	3	4	3	3	4	1	2	4	3	1	3	1
3,2E-02	2,79	1599	0,004	11122	0,001	0,57	3	4	3	1	4	1	1	2	3	1	1	1
7,6E-01	1,13	1599	0,003	11122	0,003	0,75	3	2	3	1	4	1	1	2	3	1	1	3
1,6E-01	1,70	1599	0,003	11122	0,002	1,00	3	4	3	1	4	1	1	2	3	1	3	1
6,7E-01	1,27	1599	0,001	11122	0,001	0,99	3	2	3	3	4	1	1	2	3	1	1	3
4,3E-02	0,91	1599	0,242	11122	0,258	0,99	3	2	3	3	2	3	1	2	3	3	1	3
8,1E-02	0,91	1599	0,128	11122	0,139	0,99	3	4	3	3	4	1	1	4	1	3	3	1
1,4E-01	0,88	1599	0,049	11122	0,055	0,98	1	4	3	3	4	1	1	4	1	3	3	1
1,1E-01	0,84	1599	0,031	11122	0,036	0,95	3	2	3	3	2	3	1	2	3	3	1	1
9,2E-01	0,99	1599	0,016	11122	0,016	1,00	1	2	1	3	2	1	1	2	3	3	3	3
6,3E-01	0,93	1599	0,015	11122	0,016	0,96	3	4	3	3	4	1	1	4	1	3	3	3

р	г	#пораж	част пора	част конт	инфо	84	7	45	2	9	30	4	9	3	42	44	60
6,3E-01	1,09	1599	0,012	11122	0,011	0,92	1	2	3	3	4	1	1	4	1	3	1
4,0E-01	0,80	1599	0,005	11122	0,006	0,98	3	4	3	3	4	1	1	2	3	1	3
6,4E-01	0,87	1599	0,004	11122	0,004	1,00	3	2	3	3	2	1	1	2	3	1	3
1,9E-01	0,49	1599	0,001	11122	0,002	0,86	1	2	3	3	2	3	1	2	3	1	3
2,9E-01	0,49	1599	0,001	11122	0,001	0,99	3	4	3	3	4	1	1	4	3	3	1

Таблица 14. Ассоциация ОНП маркеров и 2-маркерных гаплотипов в блоке НС rs13387042 с риском рака молочной железы (фенотип "все формы РМЖ")

Приведены следующие коды аллелей: 1=A, 2=C, 3=G, 4=T.

Р корр	г	#пораж	Част пора	#конт	Част Конт	Аллель	Маркер	Аллель	Маркер
7,58E-06	1,1913	1658	0,493969	11547	0,450377	1	rs13387042		
8,59E-04	1,1892	1660	0,178916	11556	0,154855	2	rs6735174		
8,04E-03	1,1216	1660	0,29247	11563	0,269307	1	rs6716542		
1,17E-02	1,1281	1660	0,796385	11561	0,776144	3	rs2372943		
1,24E-02	1,1214	1660	0,769879	11560	0,748962	3	rs13011060		
1,54E-02	1,1074	1659	0,6956	11560	0,673573	4	rs4491709		
2,65E-02	1,1007	1659	0,723026	11561	0,703399	1	rs12621130		
1,03E-01	1,0691	1660	0,655723	11563	0,640491	4	rs6435957		
2,37E-01	1,088	1659	0,920735	11560	0,91436	3	rs10191184		
7,63E-01	1,0122	1660	0,383133	11560	0,380277	4	rs6435959		
7,74E-01	1,0367	1660	0,975602	11555	0,974729	3	rs10171745		
8,45E-01	1,0078	1660	0,417169	11563	0,41529	3	rs10490444		
9,54E-06	0,8405	1660	0,489993	11563	0,533373	3	rs10171745	3	rs13387042
4,24E-03	0,8771	1659	0,22895	11562	0,25292	1	rs12621130	3	rs13387042
6,79E-03	1,1253	1660	0,288931	11563	0,265289	4	rs6435957	1	rs6716542
7,96E-03	1,1219	1660	0,29198	11563	0,268789	1	rs6716542	4	rs4491709
8,04E-03	1,1216	1660	0,29247	11563	0,269307	1	rs6716542	2	rs6435959
8,04E-03	1,1216	1660	0,29247	11563	0,269307	3	rs10171745	1	rs6716542
9,38E-03	0,8938	1660	0,286833	11563	0,310333	2	rs4491709	1	rs10490444
1,03E-02	1,118	1660	0,289694	11563	0,267286	1	rs6716542	1	rs10490444
1,29E-02	0,9003	1660	0,683133	11563	0,705426	3	rs10171745	3	rs6716542
1,80E-02	0,8986	1660	0,304278	11563	0,327368	4	rs2705593	3	rs13387042

Р корр	г	#пораж	Част пора	#конт	Част Конт	Аллель	Маркер	Аллель	Маркер
2,92E-02	1,0966	1660	0,315096	11562	0,295536	1	rs6735174	1	rs13387042
3,27E-02	1,0947	1660	0,312348	11563	0,293242	4	rs4491709	2	rs6435959
3,40E-02	0,9119	1660	0,275355	11563	0,294127	3	rs10191184	3	rs12621130
3,60E-02	0,9128	1660	0,274287	11563	0,29282	4	rs2705593	3	rs12621130
6,69E-02	1,0758	1660	0,405656	11563	0,388167	3	rs10171745	1	rs4674132
1,08E-01	0,9339	1660	0,315371	11563	0,33033	3	rs10191184	2	rs6435957
1,27E-01	1,0702	1660	0,293447	11563	0,279572	4	rs6435957	1	rs10490444
1,53E-01	1,0653	1660	0,270411	11563	0,258112	4	rs2705593	4	rs6435957
3,74E-01	1,0381	1660	0,335709	11563	0,327421	2	rs2705593	3	rs10191184
6,52E-01	0,9825	1660	0,544171	11563	0,548548	1	rs12621130	1	rs6735174
6,88E-01	0,984	1660	0,61747	11563	0,621271	1	rs6735174	3	rs2372943
7,59E-01	0,958	1660	0,028119	11563	0,029317	1	rs10191184	2	rs6435957
8,45E-01	0,9923	1660	0,582831	11563	0,58471	1	rs6735174	1	rs10490444
8,97E-01	0,9947	1660	0,355026	11563	0,356245	2	rs2705593	3	rs10490444
9,07E-01	0,9954	1660	0,403688	11563	0,404807	3	rs6716542	4	rs4491709
9,61E-01	1,0068	1660	0,026373	11563	0,026199	4	rs2705593	1	rs10191184

Таблица 15. ОНП маркеры НарМар, сцепленные с rs13387042 со значениями r^2 выше 0,2

ОНП суррогат	Реф ОНП	D'	R ²	p	Полож в Build 34	Полож в Build 36	Полож в SEQ ID 5
rs1882419	rs13387402	0,692153	0,331807	1,92E-09	218059508	217565211	301
rs2252488	rs13387402	1	0,258085	5,60E-11	218059992	217565695	785
rs2272525	rs13387402	0,849123	0,271771	1,07E-08	218061250	217566904	1994
rs4396687	rs13387402	1	0,268293	1,29E-10	218063010	217568713	3803
rs12614767	rs13387402	0,785371	0,245063	1,01E-07	218063508	217569211	4301
rs4571035	rs13387402	1	0,344501	1,91E-13	218064351	217570054	5144
rs10195963	rs13387402	1	0,324895	7,05E-13	218065238	217570941	6031
rs10207736	rs13387402	0,573221	0,236068	4,26E-07	218066096	217571799	6889
rs12464018	rs13387402	0,847812	0,263762	1,33E-08	218067143	217572846	7936
rs10932689	rs13387402	0,880377	0,389169	5,54E-12	218068532	217574235	9325
rs12613030	rs13387402	0,917661	0,281049	3,18E-09	218068906	217574609	9699
rs2888450	rs13387402	1	0,323308	4,19E-13	218069383	217575086	10176
rs4254482	rs13387402	1	0,304413	5,14E-12	218069959	217575662	10752
rs2372932	rs13387402	1	0,268293	2,98E-11	218069975	217575678	10768
rs4255939	rs13387402	1	0,300466	2,28E-12	218070431	217576134	11224
rs10192415	rs13387402	1	0,33518	1,75E-13	218073196	217578899	13989
rs7595393	rs13387402	1	0,319631	7,19E-13	218073796	217579499	14589
rs2372935	rs13387402	1	0,331897	4,68E-13	218076131	217581834	16924
rs6733648	rs13387402	1	0,318182	6,69E-13	218077438	217583141	18231
rs6435957	rs13387402	0,586033	0,244261	1,46E-07	218080751	217586454	21544
rs12614773	rs13387402	0,911373	0,292062	1,74E-08	218084485	217590188	25278
rs17833842	rs13387402	1	0,345455	5,50E-12	218087663	217593366	28456
rs10177578	rs13387402	1	0,331897	2,95E-13	218088201	217593904	28994
rs12623304	rs13387402	0,919746	0,300741	1,15E-09	218090342	217596045	31135
rs6716542	rs13387402	1	0,33518	1,75E-13	218093185	217598888	33978

rs17777330	rs13387402	1	0,33518	1,75E-13	218094234	217599937	35027
rs10211546	rs13387402	1	0,312977	1,07E-12	218094939	217600642	35732
rs2372937	rs13387402	0,818494	0,335949	6,72E-10	218095256	217600959	36049
rs4315498	rs13387402	1	0,345455	5,50E-12	218095412	217601115	36205
rs12998806	rs13387402	1	0,33395	1,54E-11	218096305	217602008	37098
rs12465515	rs13387402	1	0,347368	7,21E-14	218096945	217602648	37738
rs4491709	rs13387402	0,81627	0,334787	8,36E-10	218097298	217603001	38091
rs2372938	rs13387402	0,818494	0,335949	6,72E-10	218097651	217603354	38444
rs12621130	rs13387402	1	0,345455	5,50E-12	218098168	217603871	38961
rs13399995	rs13387402	1	0,321168	7,70E-13	218100083	217605786	40876
rs12052807	rs13387402	0,788298	0,255063	4,96E-08	218100783	217606486	41576
rs10199394	rs13387402	1	0,347368	7,21E-14	218101273	217606976	42066
rs2372940	rs13387402	1	0,354179	2,00E-13	218101566	217607269	42359
rs4583440	rs13387402	0,818494	0,335949	6,72E-10	218101794	217607497	42587
rs2372941	rs13387402	1	0,315673	1,25E-12	218101850	217607553	42643
rs6721811	rs13387402	1	0,347368	7,21E-14	218102838	217608541	43631
rs12615418	rs13387402	1	0,345455	5,50E-12	218104345	217610048	45138
rs12621884	rs13387402	1	0,340033	2,71E-11	218104452	217610155	45245
rs2372943	rs13387402	1	0,345455	5,50E-12	218106318	217612021	47111
rs4522583	rs13387402	1	0,431818	1,56E-14	218106548	217612251	47341
rs2888451	rs13387402	1	0,33518	1,75E-13	218107055	217612758	47848
rs2270398	rs13387402	1	0,33395	1,84E-11	218107518	217613221	48311
rs13412666	rs13387402	1	1	5,65E-36	218108321	217614024	49114
rs13387402	rs13387402	1	1	0	218108374	217614077	49167
rs13426489	rs13387402	1	1	1,42E-35	218108788	217614491	49581
rs6713249	rs13387402	1	0,345455	5,50E-12	218109630	217615333	50423
rs12620095	rs13387402	1	0,863636	3,14E-29	218110056	217615759	50849
rs4621152	rs13387402	0,782931	0,262305	4,22E-09	218111527	217617230	52320
rs6721996	rs13387402	1	0,964912	4,45E-34	218112005	217617708	52798
rs12622764	rs13387402	1	0,345455	5,50E-12	218112403	217618106	53196

rs6723013	rs13387402	1	0,331897	4,68E-13	218113452	217619155	54245
rs2372945	rs13387402	1	0,385965	4,45E-15	218116654	217622357	57447
rs12694403	rs13387402	0,934804	0,388184	4,75E-12	218117956	217623659	58749
rs12613955	rs13387402	1	0,863636	4,73E-29	218118594	217624297	59387
rs11894340	rs13387402	0,781203	0,257717	6,21E-09	218119709	217625412	60502
rs17778329	rs13387402	1	0,35514	5,16E-12	218122353	217628056	63146
rs4442975	rs13387402	1	0,964912	4,45E-34	218123311	217629014	64104
rs10932693	rs13387402	1	0,324895	7,05E-13	218123663	217629366	64456
rs17778427	rs13387402	1	0,345455	5,50E-12	218125555	217631258	66348
rs17835044	rs13387402	1	0,345455	5,50E-12	218126147	217631850	66940
rs7562029	rs13387402	1	0,33518	1,75E-13	218126803	217632506	67596
rs13000023	rs13387402	0,923607	0,31255	5,05E-10	218126936	217632639	67729
rs735361	rs13387402	1	0,345455	5,50E-12	218127665	217633368	68458
rs13409592	rs13387402	1	0,210526	2,27E-09	218128870	217634573	69663
rs2287289	rs13387402	0,856459	0,284427	2,96E-09	218134039	217639742	74832
rs12329133	rs13387402	0,722969	0,239375	1,69E-07	218137658	217643361	78451
rs13022815	rs13387402	0,722969	0,239375	1,69E-07	218138666	217644369	79459
rs13011060	rs13387402	0,849849	0,279401	1,41E-08	218140719	217646422	81512
rs13011326	rs13387402	0,853195	0,289251	9,17E-09	218140952	217646655	81745
rs4674132	rs13387402	0,725082	0,249037	1,45E-07	218141061	217646764	81854

Таблица 16. Ассоциация ОНП SG16S299 (rs3803662) алеля Т с риском молочной железы

Результаты исследования ассоциации аллеля Т ОНП rs3803662 на 16q12. Исландия 1 представляет собой когорту, на которой были получены исходные результаты, а Исландия 2 представляет собой независимую дублирующую группу. Все значения Р являются двусторонними и скорректированы на наличие родства и другую возможную стратификацию популяции исландских случаев и контролей. Для объединенной когорты жителей других стран и для всех объединенных когорт значения OR и Р подсчитаны с применением модели Мантеля-Хензеля, а приведенные частоты представляют собой простое (арифметическое) среднее значений частот в отдельных группах.

rs3803662 (T)		Частота					
Когорта							
(Случаи/Контроли)	Случаи	Контроли	OR	(95% CI)	P	PAR	
Исландия 1							
(1600/11563)	0,278	0,238	1,23	(1,13,1,34)	4,7E-06		
Исландия 2 (594/1433)	0,274	0,227	1,28	(1,08,1,52)	4,9E-03		
Исландия объедин							
(2194/12996)	0,276	0,237	1,23	(1,14,1,33)	2,8E-07	0,101	
Швеция (819/1723)							
	0,333	0,250	1,50	(1,32,1,71)	8,3E-10	0,209	
Испания (455/1002)	0,348	0,308	1,20	(1,02,1,42)	3,1E-02	0,114	
Голландия (558/1384)	0,306	0,255	1,29	(1,10,1,50)	1,4E-03	0,131	
Не-исландцы							
(1832/4109)	0,329	0,271	1,35	(1,24,1,47)	5,1E-12	0,166	
Все когорты							
(4026/17105)	0,316	0,262	1,28	(1,21,1,36)	2,7E-17	0,134	

Таблица 17. Определение блока НС rs3803662 на хромосоме 16

NCBI Build	Начало	Окончание
34	52,291,041	52,436,127
36	51,069,957	51,215,026

Таблица 18. Ассоциация ОНП маркеров и 2-маркерных гаплотипов в блоке НС rs3803662 с риском рака молочной железы (фенотип "все формы РМЖ")

Приведены следующие коды аллелей: 1=A, 2=C, 3=G, 4=T.

р	г	#пораж	Част пораж	#конт	Част конт	инфо	rs111480	rs187679	rs9302555	rs1420533	rs4783780	rs3803662	rs3104755	rs4784230	rs3112612
2,46E-06	1,23	1715	0,283	12971	0,242	1,00				1				4	
4,35E-06	1,23	1715	0,267	12970	0,228	1,00				1			1		
5,16E-06	1,25	1715	0,234	12972	0,196	0,93		4				4			
6,12E-06	1,23	1715	0,274	12972	0,235	0,99						4		4	
6,91E-06	1,22	1715	0,280	12971	0,241	0,99					2			4	
7,39E-06	1,23	1714	0,266	12967	0,228	0,98			4			4			
8,18E-06	1,23	1715	0,266	12970	0,228	0,99					2		1		
8,36E-06	1,23	1715	0,256	12973	0,218	0,95								4	4
9,13E-06	1,22	1713	0,273	12971	0,235	1,00				1		4			
9,86E-06	1,23	1713	0,244	12971	0,207	0,97						4			4
1,06E-05	1,24	1715	0,240	12967	0,203	0,93	1					4			
1,08E-05	0,82	1712	0,725	12953	0,763	1,00						2			
1,08E-05	1,22	1712	0,275	12953	0,237	1,00						4			
1,10E-05	1,22	1715	0,266	12972	0,229	0,98						4	1		
1,30E-05	1,22	1712	0,271	12965	0,234	1,00					2	4			
2,67E-05	1,22	1715	0,254	12973	0,218	0,94							1		4
3,31E-05	1,21	1715	0,324	12971	0,284	0,87		4		1					
6,62E-05	1,20	1715	0,322	12971	0,283	0,86		4			2				
р	г	#пораж	Част пораж	#конт	Част конт	инфо	rs111480	rs187679	rs9302555	rs1420533	rs4783780	rs3803662	rs3104755	rs4784230	rs3112612
1,04E-04	1,18	1715	0,358	12970	0,321	0,98	1			1					
1,05E-04	1,20	1715	0,292	12973	0,255	0,86		4							4
2,30E-04	1,17	1711	0,356	12952	0,321	0,98	1				2				
5,78E-04	0,87	1715	0,370	12972	0,404	0,94		2				2			
1,24E-03	1,15	1715	0,325	12973	0,295	0,94	1								4
1,39E-03	1,14	1714	0,458	12972	0,426	0,96			4	1					
1,89E-03	1,14	1714	0,390	12971	0,360	0,95			4						4
1,99E-03	1,13	1714	0,457	12962	0,426	0,96			4		2				
2,01E-03	1,13	1714	0,588	12958	0,558	0,99	1		4						
2,19E-03	1,13	1715	0,585	12971	0,555	1,00	1	4							
2,36E-03	0,89	1715	0,415	12960	0,445	1,00		2							
2,36E-03	1,13	1715	0,585	12960	0,555	1,00		4							
2,38E-03	1,13	1715	0,585	12970	0,555	1,00		4	4						
2,67E-03	1,13	1715	0,545	12972	0,515	0,96			4					4	
3,37E-03	0,87	1715	0,272	12973	0,301	0,86		2							2
3,43E-03	0,86	1715	0,238	12971	0,266	0,84		2			1				
3,44E-03	1,13	1715	0,496	12973	0,467	0,93		4						4	
3,45E-03	0,86	1715	0,239	12971	0,266	0,84		2		3					
3,64E-03	0,89	1713	0,541	12971	0,570	0,97						2			2
4,13E-03	1,12	1712	0,492	12956	0,463	1,00				1					
4,13E-03	0,89	1712	0,508	12956	0,537	1,00				3					
4,51E-03	0,89	1713	0,508	12970	0,536	1,00				3	1				
4,58E-03	0,89	1713	0,506	12971	0,534	1,00				3		2			

р	г	#пораж	Част пораж	#конт	Част конт	инфо	rs111480	rs187679	rs9302555	rs1420533	rs4783780	rs3803662	rs3104755	rs4784230	rs3112612
4,71E-03	0,89	1713	0,492	12971	0,520	0,99				3					2
5,00E-03	0,89	1712	0,506	12965	0,533	0,99					1	2			
5,18E-03	0,90	1715	0,507	12970	0,535	1,00				3			1		
5,33E-03	0,90	1713	0,493	12972	0,520	0,99					1				2
5,63E-03	1,12	1713	0,490	12970	0,463	1,00				1	2				
5,70E-03	1,12	1713	0,411	12971	0,384	0,99				1					4
6,05E-03	0,90	1715	0,507	12971	0,534	1,00				3				4	
6,60E-03	0,90	1705	0,510	12922	0,537	1,00					1				
6,60E-03	1,11	1705	0,490	12922	0,463	1,00					2				
6,69E-03	0,89	1715	0,614	12967	0,641	0,94	1					2			
6,77E-03	1,12	1715	0,536	12971	0,509	0,95			4				1		
6,90E-03	0,90	1715	0,496	12970	0,523	0,98	1			3					
7,04E-03	0,90	1713	0,574	12967	0,600	1,00									2
7,04E-03	1,11	1713	0,426	12967	0,400	1,00									4
7,10E-03	1,12	1715	0,488	12972	0,461	0,92		4					1		
7,50E-03	0,90	1715	0,509	12970	0,535	0,99					1		1		
9,23E-03	1,11	1713	0,409	12972	0,384	0,99					2				4
9,69E-03	0,90	1715	0,509	12972	0,535	0,98						2	1		
1,04E-02	0,84	1715	0,117	12973	0,136	0,83		2						2	
1,07E-02	0,85	1715	0,124	12972	0,143	0,83		2					3		
1,09E-02	0,90	1715	0,509	12971	0,535	0,99					1			4	
1,13E-02	0,90	1715	0,516	12972	0,541	0,99						2		4	
1,20E-02	0,90	1711	0,498	12952	0,523	0,98	1				1				
р	г	#пораж	Част пораж	#конт	Част конт	инфо	rs111480	rs187679	rs9302555	rs1420533	rs4783780	rs3803662	rs3104755	rs4784230	rs3112612
1,47E-02	0,90	1714	0,257	12967	0,279	0,98			3						
1,75E-02	0,90	1715	0,267	12970	0,288	1,00		2	3						
1,80E-02	0,90	1713	0,267	12919	0,288	1,00			3						
1,80E-02	1,11	1713	0,733	12919	0,712	1,00			4						
1,85E-02	0,90	1714	0,266	12958	0,287	1,00	1		3						
1,87E-02	0,91	1715	0,535	12973	0,558	0,97								4	2
1,92E-02	0,90	1715	0,268	12971	0,289	1,00	1	2							
1,96E-02	0,91	1715	0,522	12973	0,545	0,96							1		2
2,25E-02	0,90	1714	0,231	12972	0,251	0,95			3	3					
2,47E-02	0,90	1714	0,231	12962	0,251	0,95			3		1				
2,59E-02	0,90	1714	0,228	12971	0,248	0,93			3						2
2,89E-02	0,81	1715	0,027	12971	0,033	1,40			3				3		
3,33E-02	0,79	1715	0,022	12972	0,027	1,40			3					2	
3,38E-02	3,50	1713	0,002	12970	0,001	0,99				1	1				
3,41E-02	0,87	1715	0,108	12967	0,122	0,87	2					2			
4,73E-02	0,92	1715	0,529	12973	0,549	0,95	1								2
5,76E-02	0,85	1715	0,043	12973	0,050	1,26	2								2
6,49E-02	0,92	1715	0,244	12972	0,261	0,95			3					4	
6,73E-02	0,88	1715	0,098	12967	0,109	0,94	2						3		
7,04E-02	0,89	1715	0,097	12971	0,108	0,94	2							2	

Таблица 19. ОНП маркеры НарМар, сцепленные с rs3803662 со значениями r^2 выше 0,2

ОНП суррогат	Реф ОНП	D'	R2	Значение p	Полож в Build 34	Полож в Build 36	Полож в SEQ ID 6
rs4784220	rs3803662	0,675773	0,314847	8,65E-09	52314403	51093311	301
rs12598982	rs3803662	0,939522	0,331013	2,84E-11	52315478	51094386	1376
rs4784222	rs3803662	0,93875	0,331747	4,06E-11	52316179	51095087	2077
rs17271951	rs3803662	0,959839	0,88572	9,27E-27	52316633	51095541	2531
rs9933638	rs3803662	0,935921	0,321096	2,07E-10	52318190	51097098	4088
rs9302556	rs3803662	0,936406	0,334336	1,07E-10	52322238	51101146	8136
rs7190749	rs3803662	0,939522	0,331013	2,84E-11	52322671	51101579	8569
rs12443621	rs3803662	0,939123	0,326882	4,01E-11	52326630	51105538	12528
rs9933556	rs3803662	0,937324	0,325388	9,50E-11	52326895	51105803	12793
rs1362546	rs3803662	0,935967	0,329172	9,67E-11	52328528	51107436	14426
rs1075367	rs3803662	0,939522	0,331013	2,84E-11	52333595	51112503	19493
rs8046979	rs3803662	0,932586	0,342705	3,56E-10	52334335	51113243	20233
rs1420529	rs3803662	0,931004	0,310285	1,29E-09	52337266	51116174	23164
rs11642645	rs3803662	0,936189	0,329297	5,09E-11	52337500	51116408	23398
rs1420533	rs3803662	1	0,362637	3,83E-14	52342219	51121127	28117
rs1362548	rs3803662	1	0,924812	8,14E-30	52342544	51121452	28442
rs2193094	rs3803662	1	0,362637	3,83E-14	52345047	51123955	30945
rs4783780	rs3803662	1	0,362637	3,83E-14	52350025	51128937	35927
rs3112581	rs3803662	1	0,358407	5,51E-14	52350122	51129034	36024
rs3112580	rs3803662	1	0,370787	2,40E-14	52350194	51129106	36096
rs9931232	rs3803662	1	0,362637	3,83E-14	52351417	51130333	37323
rs1123428	rs3803662	1	0,366516	3,46E-14	52356420	51135336	42326
rs3095604	rs3803662	1	1	9,56E-34	52360562	51139480	46470
rs3803662	rs3803662	1	1	0	52364943	51143842	50832
rs4784227	rs3803662	1	0,812734	3,45E-25	52377790	51156689	63679
rs12922061	rs3803662	0,813366	0,393164	1,28E-08	52413602	51192501	99491

Таблица 20. Перечень всех маркеров dbSNP125 в блоке неравновесия по сцеплению rs4848543
(приведены позиции в соответствии с NCBI Build 34)

Полож в Build 34	Название	Полож в Build 34	Название	Полож в Build 34	Название
		119996150	rs10864979	120010592	rs3980308
		119996950	rs13433007	120010654	rs1317848
		119996967	rs6748754	120010949	rs4643545
		119997458	rs28574898	120011133	rs6705476
		119998119	rs4848537	120011458	rs2121217
		119998463	rs13409086	120011699	rs5833777
		119998504	rs13411872	120011761	rs7600045
		119998551	rs13420759	120012134	rs7563952
119987193	rs11883887	119999028	rs11895163	120012410	rs4849761
119987218	rs7601379	119999114	rs7597287	120012531	rs1446130
119987236	rs10189928	120000098	rs13020121	120012783	rs1446129
119987488	rs17012230	120000215	rs9308770	120013821	rs4849762
119987701	rs17798163	120000292	rs13425098	120014233	rs6542508
119988214	rs6542503	120000559	rs895404	120014416	rs11299793
119988417	rs7606204	120000798	rs10207133	120014424	rs6724034
119988662	rs17012243	120000837	rs10207198	120014425	rs6706020
119988746	rs10174027	120000969	rs11898182	120014553	rs6542509
119988777	rs17012263	120001134	rs9308771	120014683	rs6542510
119988845	rs12623644	120001139	rs9308772	120015600	rs1867750
119988905	rs10184937	120001461	rs6146899	120015668	rs5019100
119989602	rs7570704	120001581	rs17012317	120015928	rs1867751
119989690	rs7573686	120001748	rs9308773	120016277	rs7420091
119990027	rs6542504	120001768	rs9308774	120016397	rs4390775
119990110	rs7557802	120002407	rs10202747	120016788	rs6732650
119990364	rs12987676	120002993	rs17190912	120016904	rs6742613
119990863	rs12993619	120003189	rs11884636	120017026	rs7579840
119991237	rs6737956	120003199	rs4849760	120017542	rs2197539
119991499	rs10529569	120003758	rs7562612	120017573	rs10205385
119991618	rs6741215	120004530	rs745884	120017765	rs6542511
119991739	rs11677153	120004532	rs745885	120018695	rs6716071
119991908	rs6542505	120004821	rs745886	120018841	rs10597874
119991993	rs6728751	120005202	rs745887	120019534	rs13407241
119992039	rs6542506	120005459	rs745888	120019716	rs4849763
119992209	rs11386094	120005770	rs13395242	120019836	rs4848539
119993577	rs6721529	120005783	rs7560869	120019839	rs4848540
119993825	rs6739574	120005868	rs11885255	120019852	rs1867752
119993914	rs12711922	120005995	rs17012332	120020267	rs10531839
119993983	rs13397501	120006051	rs11900137	120020619	rs2197540
119994264	rs12711923	120006067	rs11884243	120020703	rs6727359
119994321	rs13017443	120006186	rs2197538	120020806	rs2197541
119994574	rs6740231	120006239	rs11889299	120020814	rs6737324
119994784	rs10716622	120006306	rs970013	120020831	rs2197542
119994803	rs1374121	120007130	rs7562248	120021103	rs7577922
119994954	rs1530522	120007189	rs12468740	120021157	rs7575097
119995078	rs10205049	120007959	rs6542507	120021526	rs6713238
119995182	rs1530523	120008874	rs4848538	120021548	rs7564432
119995238	rs5018771	120009246	rs6758051	120021672	rs11682799
119995420	rs13417283	120010134	rs7593041	120021695	rs11687834
119995638	rs2860753	120010278	rs11308896	120022550	rs2121218
119995975	rs4447611	120010394	rs7556873	120022556	rs2121219
119996042	rs4485581	120010590	rs3980309	120022644	rs13418186

Полож в Build 34	Название	Полож в Build 34	Название	Полож в Build 34	Название
120022871	rs2166408	120037242	rs7562256	120049031	rs4849767
120022882	rs2166409	120037469	rs7557444	120049111	rs4849768
120022986	rs6732687	120037538	rs6735202	120049158	rs4849769
120023118	rs12474874	120037631	rs7584874	120049468	rs13017938
120023404	rs6748897	120037714	rs7584975	120049808	rs11680905
120023422	rs4322838	120037998	rs7598853	120049895	rs2028852
120023516	rs2901763	120038031	rs11123504	120050448	rs2028851
120023583	rs11680207	120038363	rs11384356	120050503	rs13410792
120024083	rs6722032	120038373	rs11464659	120050761	rs13402151
120024822	rs7592528	120038374	rs11393319	120050907	rs7586792
120024900	rs6542512	120038698	rs1446126	120051111	rs11675684
120025087	rs7371346	120038699	rs10168744	120051961	rs4142753
120025150	rs10864980	120039011	rs4848542	120052171	rs4849770
120026424	rs10188216	120039313	rs4277535	120052802	rs2084155
120026569	rs2044428	120039785	rs13416713	120053731	rs11691024
120026637	rs11123503	120040375	rs13417356	120054452	rs11681612
120026789	rs2044427	120040703	rs12614389	120054477	rs11692475
120027283	rs10203658	120040851	rs2084154	120054484	rs13400379
120027724	rs7583824	120040967	rs12621053	120054490	rs11692479
120028226	rs10174654	120041245	rs10167806	120055083	rs4142752
120028314	rs10186930	120041777	rs10641092	120056573	rs11695796
120028820	rs6542513	120041779	rs895400	120057058	rs10188745
120028843	rs6542514	120041797	rs7599963	120057116	rs11685085
120029239	rs6542515	120041798	rs10658512	120057332	rs12475875
120029586	rs10532262	120041840	rs7560645	120057333	rs12463432
120029885	rs4849764	120041953	rs10551390	120057493	rs5833780
120030487	rs17804511	120041983	rs7600179	120057623	rs746945
120030834	rs935362	120042110	rs13017664	120057815	rs5833781
120030947	rs935361	120042545	rs10206284	120057816	rs3053192
120031161	rs2860752	120042980	rs3053193	120057829	rs5833782
120031486	rs11893808	120042985	rs5833778	120058071	rs876062
120031828	rs12623976	120043240	rs13024716	120058160	rs895399
120032169	rs4848541	120044181	rs7607939	120058174	rs10186657
120032367	rs28653619	120044182	rs13405317	120058180	rs895398
120032378	rs1446128	120044735	rs13020459	120058601	rs5833783
120033081	rs1838763	120045023	rs13015999	120060148	rs10181114
120033129	rs1470474	120045027	rs12052957	120060448	rs2860751
120033135	rs4528761	120045116	rs12468083	120060852	rs11693876
120033150	rs1838762	120045623	rs963662	120061012	rs11683050
120033268	rs2121216	120046099	rs12472931	120061029	rs10184332
120033353	rs3980311	120046932	rs7579602	120061096	rs4848543
120033359	rs1446127	120047249	rs5833779	120061139	rs10184438
120033749	rs13024710	120047386	rs13030441	120061708	rs11123505
120034173	rs4849765	120047590	rs991844	120062319	rs11684731
120034354	rs4849766	120048218	rs13000062	120062891	rs6754436
120034730	rs6715243	120048226	rs13022674	120062996	rs12464139
120036029	rs10166010	120048234	rs28579959	120063070	rs6754664
120036225	rs7561309	120048244	rs28719630	120063176	rs6754799
120036417	rs11886602	120048434	rs13027917	120063963	rs12997936
120036954	rs6542516	120048457	rs13027932	120064284	rs10206700
120037241	rs7601541	120048632	rs12466475	120064289	rs10194572

Полож в Build 34	Название	Полож в Build 34	Название	Полож в Build 34	Название
120064976	rs6759589	120077769	rs838110	120090847	rs10182241
120065127	rs10221610	120077991	rs838109	120091298	rs838092
120065310	rs11689491	120078174	rs838108	120091323	rs13401854
120066060	rs7560972	120078320	rs12373823	120091399	rs13386059
120066128	rs7558272	120078344	rs838107	120091931	rs838091
120066241	rs12463977	120078395	rs865688	120092184	rs864505
120066318	rs1562256	120078483	rs12373677	120092304	rs838090
120066453	rs1562255	120078615	rs7578870	120092610	rs7567193
120066620	rs1446125	120078754	rs838106	120092737	rs838089
120066705	rs1446124	120079222	rs838105	120092937	rs838088
120066972	rs1838761	120079745	rs838104	120093034	rs10174436
120066974	rs7601127	120079819	rs838103	120093360	rs838087
120067290	rs11677262	120079878	rs838102	120094306	rs12467944
120067736	rs10204391	120080519	rs10182677	120095038	rs17804991
120067927	rs6753314	120080542	rs10625730	120095312	rs10196144
120068051	rs6735649	120081128	rs3769659	120095426	rs13027072
120068477	rs28455610	120081529	rs865108	120095756	rs6728660
120068761	rs6709777	120082110	rs838101	120096263	rs13027293
120069635	rs10208455	120082382	rs708670	120096647	rs838086
120069642	rs838067	120082432	rs708671	120096791	rs838085
120069925	rs4848544	120082499	rs708672	120096949	rs10600943
120070154	rs12466050	120082585	rs6146900	120096982	rs10526493
120070400	rs7565771	120082871	rs708673	120097896	rs10196222
120070666	rs7595954	120083300	rs1106771	120097915	rs838084
120070908	rs708666	120083710	rs838100	120097952	rs3731606
120071508	rs838066	120084104	rs3223038	120098787	rs838083
120071996	rs708667	120084105	rs6146901	120098925	rs10188946
120072367	rs708668	120084215	rs838099	120099084	rs838082
120072376	rs708669	120084226	rs838098	120099515	rs838081
120072400	rs895397	120084235	rs13410686	120099779	rs10202732
120072476	rs838065	120084331	rs3769656	120099996	rs6738309
120072631	rs6755420	120084565	rs838097	120100457	rs17013371
120072650	rs838064	120084736	rs13400593	120100570	rs1665025
120072686	rs838063	120084741	rs838096	120100571	rs1624263
120072737	rs838062	120084781	rs838095	120100714	rs17805141
120073503	rs838061	120085360	rs13425946	120100985	rs5833785
120073744	rs11683528	120085486	rs838094	120100993	rs1530524
120073904	rs10209847	120085502	rs13010169	120101012	rs4592854
120074221	rs1562310	120085910	rs12470404	120101809	rs2592075
120074355	rs838060	120086734	rs895396	120102502	rs838080
120074581	rs7557910	120086898	rs5833784	120102508	rs6735209
120075395	rs838059	120086917	rs13432674	120102857	rs838079
120075418	rs838058	120086997	rs13408034	120103067	rs17013379
120076234	rs11676144	120087374	rs708674	120103130	rs12990907
120076668	rs2289894	120087538	rs2121358	120103529	rs7561146
120076919	rs13003725	120087544	rs1963247	120103701	rs734846
120076920	rs13034799	120087718	rs12465926	120103880	rs734845
120077150	rs3795927	120088448	rs7577784	120104365	rs895406
120077389	rs3795926	120088805	rs13391380	120104472	rs895405
120077721	rs838112	120089829	rs838093	120104502	rs708675
120077754	rs838111	120090797	rs10190429	120104672	rs1867856

Полож в Build 34	Название	Полож в Build 34	Название
120105219	rs838078	120118405	rs11694139
120105720	rs3754847	120118612	rs11682650
120105721	rs12467089	120119587	rs1530561
120105810	rs11888609	120119648	rs6721852
120105981	rs838077	120119899	rs13016861
120106044	rs7596511	120120260	rs13034803
120106433	rs838076	120120268	rs13034810
120107316	rs12711924	120120378	rs13012570
120108033	rs838075	120120617	rs6707849
120108409	rs838074	120120725	rs12165162
120109057	rs1867749	120121081	rs6723312
120109148	rs12104548	120121286	rs3054836
120109236	rs1045623	120122044	rs1545360
120109381	rs3731604	120122260	rs6740573
120109485	rs12104639	120122331	rs6730921
120109754	rs838073	120122555	rs13004570
120109782	rs12619779	120122566	rs5028235
120109902	rs6742425	120122626	rs11395093
120110130	rs838072	120123526	rs2860792
120110342	rs6542517	120123689	rs1867855
120110521	rs838071	120123862	rs5833787
120111317	rs12991476	120124074	rs28497293
120111319	rs12991477	120124123	rs12990281
120111544	rs12105674	120124221	rs12995866
120111606	rs838070	120125144	rs11692352
120111771	rs838069	120125454	rs13402922
120111887	rs838068	120126304	rs6729196
120112630	rs13400106	120126843	rs13034935
120112767	rs13388672	120127261	rs7599282
120112793	rs7575387	120127470	rs7582316
120113041	rs1610469	120127897	rs11123506
120113297	rs11891374	120128212	rs2084201
120114676	rs6720040	120128319	rs7592522
120114755	rs6723169	120128879	rs10864981
120115056	rs3054837	120128983	rs11123507
120115077	rs5833786		
120115093	rs13014052		
120115134	rs1374314		
120115780	rs11885203		
120116464	rs3769653		
120116484	rs10171408		
120116695	rs7590206		
120116930	rs7581339		
120117062	rs3731603		
120117527	rs6753005		
120117530	rs6753006		
120117532	rs12476967		
120117536	rs10571222		
120117768	rs10204553		
120118122	rs3731602		
120118333	rs4849771		

Таблица 21. Перечень всех маркеров dbSNP125 в блоке неравновесия по сцеплению rs13387042 (приведены позиции в соответствии с NCBI Build 34)

Полож в Build 34	Название	Полож в Build 34	Название	Полож в Build 34	Название
		218070106	rs6435956	218087900	rs13406843
		218070227	rs6713022	218088071	rs10198696
		218070431	rs4255939	218088201	rs10177578
		218070433	rs2372933	218088354	rs10177962
		218070638	rs10200456	218089512	rs4435428
		218071353	rs7426087	218089563	rs4276003
		218071353	rs4542823	218089745	rs13425828
		218071539	rs4456677	218089919	rs3084560
		218071539	rs7422519	218089928	rs13414201
		218071631	rs4264545	218090342	rs12623304
		218071948	rs10169454	218091057	rs2372936
		218073196	rs10192415	218092002	rs7425648
		218073796	rs7595393	218092308	rs13410769
		218073891	rs10169372	218092732	rs10191536
		218074854	rs13430080	218092782	rs10203987
		218075719	rs10175911	218093185	rs6716542
		218075836	rs12165215	218093638	rs10207757
		218076028	rs17833745	218093795	rs10173250
		218076072	rs2372934	218094186	rs28525300
		218076131	rs2372935	218094234	rs17777330
		218076869	rs12612444	218094822	rs10211354
		218076958	rs13423850	218094838	rs10187690
		218077438	rs6733648	218094913	rs10211289
		218077473	rs13389571	218094939	rs10211546
		218077681	rs12165173	218095256	rs2372937
		218077715	rs13389783	218095412	rs4315498
		218078368	rs6737508	218096305	rs12998806
		218078384	rs10209168	218096485	rs12999224
		218079285	rs13383189	218096566	rs17833945
		218079764	rs11884129	218096592	rs12465396
		218080177	rs6714036	218096776	rs12465506
		218080751	rs6435957	218096945	rs12465515
		218081124	rs10192982	218097298	rs4491709
		218082238	rs10932691	218097359	rs4674130
		218082405	rs10932692	218097448	rs3084573
		218082754	rs10188365	218097540	rs13417588
		218083373	rs6723019	218097624	rs13392238
		218083726	rs7349199	218097651	rs2372938
		218083877	rs13404341	218098168	rs12621130
		218084046	rs16856821	218098897	rs11387578
		218084485	rs12614773	218099133	rs4471865
		218086136	rs6732152	218099154	rs6435958
		218086701	rs11677001	218099578	rs6722681
		218086856	rs10171745	218100083	rs13399995
		218087663	rs17833842	218100630	rs7564605
		218087886	rs13393574	218100783	rs12052807

Полож в Build 34	Название	Полож в Build 34	Название	Полож в Build 34	Название
218101273	rs10199394	218107985	rs6734131	218113820	rs10192175
218101451	rs10210524	218108214	rs13386831	218113854	rs10192253
218101477	rs2372939	218108244	rs13412479	218113962	rs10192501
218101566	rs2372940	218108254	rs13412482	218114027	rs10204872
218101794	rs4583440	218108321	rs13412666	218114030	rs10192589
218101850	rs2372941	218108374	rs13387042	218114046	rs6708579
218101894	rs10187530	218108601	rs13426268	218114076	rs10192446
218101902	rs6760041	218108615	rs13387273	218114103	rs10192609
218102145	rs10190181	218108788	rs13426489	218114176	rs4672818
218102201	rs10190349	218108844	rs13413328	218114436	rs10195111
218102294	rs17834221	218108906	rs13413357	218114560	rs10171550
218102294	rs10202715	218108957	rs13387717	218114629	rs10195322
218102377	rs6705818	218108958	rs13426688	218114644	rs10195505
218102717	rs6706491	218109040	rs13413523	218114737	rs10195441
218102822	rs6735174	218109091	rs17777938	218114786	rs10207917
218102838	rs6721811	218109320	rs13416240	218115389	rs10198604
218102872	rs6735190	218109630	rs6713249	218115494	rs10198699
218102939	rs6735298	218109728	rs13391022	218115540	rs10198621
218103034	rs17777767	218109746	rs13416773	218115617	rs10174948
218103042	rs12464375	218109779	rs13416605	218115747	rs10175133
218103061	rs12476702	218109902	rs13430215	218115929	rs10198996
218103151	rs6435959	218109976	rs13416831	218116467	rs10167002
218103234	rs6435960	218109979	rs13416834	218116654	rs2372945
218103592	rs2372942	218110056	rs12620095	218116683	rs10201966
218104122	rs4594417	218110145	rs13433000	218116715	rs10201982
218104281	rs10709272	218110379	rs12327908	218116795	rs10178467
218104345	rs12615418	218110607	rs2372944	218117052	rs16856849
218104452	rs12621884	218110643	rs2888452	218117184	rs10204819
218104650	rs7572103	218110863	rs4618027	218117267	rs5838620
218104906	rs7421376	218111014	rs10490444	218117273	rs10566949
218105047	rs11681010	218111297	rs13397674	218117273	rs3084597
218105186	rs5838617	218111343	rs13423473	218117443	rs11901909
218105193	rs3084579	218111527	rs4621152	218117759	rs11901972
218105260	rs5838618	218111631	rs13386262	218117956	rs12694403
218105268	rs6711784	218111703	rs12329160	218118356	rs11892687
218105513	rs9973728	218111876	rs12329163	218118594	rs12613955
218105657	rs9973894	218112005	rs6721996	218118809	rs10208746
218105711	rs9973730	218112094	rs12329172	218119095	rs10185401
218105920	rs13402323	218112403	rs12622764	218119709	rs11894340
218106201	rs5838619	218112445	rs10206095	218119744	rs6435961
218106318	rs2372943	218112816	rs10206577	218120136	rs10191172
218106548	rs4522583	218113130	rs17778091	218120424	rs2372946
218106783	rs13431146	218113325	rs13411606	218120624	rs10168238
218107055	rs2888451	218113452	rs6723013	218120863	rs7586158
218107332	rs13383392	218113636	rs13386291	218121113	rs7600412
218107518	rs2270398	218113745	rs10685531	218121293	rs10183545

Полож в Build 34	Название	Полож в Build 34	Название
218121786	rs7589722	218134761	rs7598926
218122009	rs6740850	218134805	rs13018411
218122107	rs17778299	218134901	rs7599066
218122353	rs17778329	218135359	rs13415112
218123052	rs10198784	218135444	rs6754177
218123311	rs4442975	218135688	rs4361105
218123663	rs10932693	218135854	rs13025492
218124158	rs11291245	218136008	rs13025714
218124274	rs2888456	218137327	rs12328709
218124419	rs2372956	218137594	rs16825209
218124426	rs2888457	218137658	rs12329133
218125505	rs11674902	218137683	rs12329135
218125555	rs17778427	218137973	rs6734010
218126147	rs17835044	218138557	rs12622536
218126358	rs7588345	218138666	rs13022815
218126446	rs7561690	218138690	rs12996888
218126695	rs7561918	218138928	rs12997141
218126803	rs7562029	218139101	rs12328842
218126936	rs13000023	218139274	rs12328926
218127398	rs17835187	218139288	rs13002378
218127601	rs3085312	218139595	rs16856893
218127603	rs3085315	218140109	rs12478570
218127665	rs735361	218140238	rs12467063
218128212	rs17835248	218140453	rs11690147
218128870	rs13409592	218140482	rs13010040
218129396	rs16856867	218140537	rs12467163
218129599	rs2372957	218140719	rs13011060
218129616	rs2372958	218140839	rs4672819
218129746	rs5838621	218140888	rs4672820
218129748	rs3085335	218140952	rs13011326
218129820	rs16856873	218140976	rs11690464
218129963	rs17778798		
218130041	rs16856877		
218130142	rs7600279		
218130798	rs11682933		
218130803	rs4608489		
218131549	rs13034362		
218132117	rs10211044		
218133211	rs16856888		
218133717	rs5838622		
218133723	rs5838623		
218133948	rs2287288		
218134039	rs2287289		
218134273	rs16856890		
218134618	rs7598618		
218134620	rs7598805		

Таблица 22. Перечень всех маркеров dbSNP125 в блоке неравновесия по сцеплению rs3803662 (позиции приведены в соответствии с NCBI Build 34)

Полож в Build 34	Название	Полож в Build 34	Название	Полож в Build 34	Название
		52305075	rs9931647	52323870	rs7500427
		52305085	rs28484983	52325826	rs16951223
		52305086	rs10637121	52326630	rs12443621
		52305875	rs7188610	52326776	rs9923558
		52305996	rs7186907	52326895	rs9933556
		52306637	rs9923960	52327297	rs1362545
		52306726	rs2193092	52327389	rs4538010
		52306777	rs2216188	52327503	rs12922895
		52306900	rs9939506	52327601	rs12923628
		52307320	rs9926539	52327690	rs11415785
		52307440	rs9302555	52327938	rs7198754
		52308124	rs12597838	52328239	rs9936081
		52308541	rs28562017	52328264	rs8051286
		52309356	rs8064219	52328472	rs10647102
		52309429	rs11287146	52328482	rs5816851
		52309445	rs8044985	52328483	rs10645216
		52310401	rs9928129	52328493	rs3086690
		52310846	rs3086687	52328528	rs1362546
		52310948	rs1111481	52328629	rs2193093
		52311095	rs2077653	52328900	rs2160474
		52312052	rs8061860	52328944	rs12932707
		52312760	rs8051542	52329434	rs9939775
		52314189	rs11427470	52329968	rs8048043
		52314348	rs16951213	52333201	rs3086693
		52314403	rs4784220	52333595	rs1075367
		52314941	rs8046985	52334335	rs8046979
		52315230	rs28665779	52334886	rs1345388
		52315478	rs12598982	52335310	rs9941100
		52316033	rs10451108	52337266	rs1420529
		52316119	rs4784221	52337322	rs1420530
		52316179	rs4784222	52337427	rs1420531
		52316601	rs7195542	52337486	rs10708737
		52316633	rs17271951	52337500	rs11642645
		52316669	rs12934513	52338806	rs12918816
		52316689	rs12934308	52340832	rs4280232
		52316930	rs7184835	52341080	rs7197481
		52317345	rs10545412	52341404	rs12929984
		52317493	rs12600239	52342090	rs1420532
		52318190	rs9933638	52342219	rs1420533
		52318304	rs28530752	52342476	rs1362547
		52318333	rs7186498	52342544	rs1362548
		52318957	rs13337151	52342856	rs10538952
		52321003	rs8063421	52343467	rs28555277
		52321733	rs11326759	52343743	rs7193453
		52321900	rs4541074	52344234	rs9926163
		52321923	rs4517793	52345047	rs2193094
		52322238	rs9302556	52345209	rs28580623
		52322671	rs7190749	52345455	rs3095598
		52322726	rs16951220	52346438	rs12929605
		52323043	rs1362544	52346440	rs12929606
		52323132	rs1819860	52347923	rs28750243
52291041	rs1076479				
52292193	rs7188075				
52292669	rs28720857				
52293212	rs11642868				
52293398	rs5816849				
52293430	rs11454608				
52293504	rs9937283				
52294560	rs12921126				
52295118	rs189851				
52295192	rs28558853				
52295650	rs7198218				
52295682	rs11337972				
52295683	rs5816850				
52296050	rs7203597				
52296074	rs11304815				
52296100	rs7203778				
52296956	rs4784218				
52297566	rs12444429				
52297574	rs12447254				
52297595	rs12447113				
52297649	rs28689711				
52297698	rs3095662				
52297718	rs28493012				
52298081	rs4493036				
52298807	rs7205259				
52299817	rs11464652				
52299822	rs10641032				
52299825	rs11427655				
52299952	rs1111480				
52300224	rs12447430				
52301045	rs43143				
52301313	rs187679				
52302076	rs42542				
52302193	rs4784219				
52302954	rs11649553				
52303160	rs3086679				
52303163	rs9923098				
52303361	rs8061425				
52303886	rs1978343				
52304653	rs3095595				
52304699	rs12102837				
52305016	rs10637120				

Полож в Build 34	Название	Полож в Build 34	Название	Полож в Build 34	Название
52349175	rs3112583	52366478	rs28409238	52379702	rs3104746
52349177	rs3112582	52367399	rs13335398	52379713	rs3112571
52349414	rs28435355	52367692	rs7189773	52380052	rs11372822
52349558	rs11648477	52367823	rs7191347	52380131	rs11646414
52350025	rs4783780	52367830	rs13332085	52380217	rs8053784
52350122	rs3112581	52368150	rs3104771	52380227	rs11342984
52350194	rs3112580	52369037	rs28478483	52380243	rs3112570
52350390	rs11431881	52369057	rs28418561	52380252	rs9934281
52350614	rs9921569	52369268	rs8057406	52380383	rs3112569
52351417	rs9931232	52369377	rs7199494	52380473	rs3112568
52352109	rs11075535	52369741	rs28636082	52380526	rs3104747
52353864	rs12597887	52370476	rs11437687	52380961	rs3104748
52354487	rs3112579	52370482	rs3112577	52380977	rs3104749
52354492	rs4784223	52370737	rs28719661	52381045	rs3112567
52355066	rs9933351	52370774	rs28700577	52381091	rs3104750
52355242	rs1548910	52371185	rs3104808	52381630	rs3104751
52355727	rs8052175	52371458	rs28595571	52381735	rs3104752
52355897	rs28523905	52372064	rs3104812	52381752	rs3112566
52356089	rs1420534	52372504	rs16951246	52381753	rs3112565
52356352	rs7195388	52372582	rs7201350	52381770	rs3112564
52356420	rs1123428	52372830	rs3104818	52381945	rs5011736
52356800	rs1420535	52373327	rs16951251	52382513	rs3104753
52357133	rs1477074	52373331	rs3104827	52382735	rs11864809
52358132	rs3891563	52373794	rs28715219	52382879	rs16951272
52358441	rs13336638	52373795	rs28759054	52382932	rs11859735
52358491	rs8044585	52373894	rs28491753	52383045	rs16951279
52358571	rs8045285	52374821	rs28397852	52383698	rs3481
52358832	rs4784224	52375020	rs28561462	52383902	rs16951281
52358870	rs3095601	52375173	rs28366779	52384112	rs3104754
52359085	rs3095602	52375215	rs28690015	52384265	rs3112563
52359653	rs3095603	52375335	rs28472996	52384839	rs3104755
52360007	rs12930156	52375668	rs28437002	52385009	rs3104756
52360562	rs3095604	52375894	rs3104743	52386865	rs3112562
52360995	rs9921056	52376140	rs9922732	52387168	rs3104757
52361357	rs1362549	52376155	rs3112576	52387274	rs10521267
52361398	rs1362550	52376446	rs4594251	52387635	rs3104758
52361636	rs28463809	52376504	rs28685217	52387700	rs9921358
52361725	rs4784226	52376699	rs3112575	52388131	rs3104759
52362126	rs13335680	52377560	rs3112574	52388189	rs3112561
52362449	rs3095605	52377659	rs9937101	52388416	rs3104760
52362775	rs3095606	52377665	rs16951253	52388424	rs3104761
52362897	rs3095607	52377790	rs4784227	52388458	rs3112560
52363302	rs12926842	52377970	rs16951254	52388606	rs3112559
52364042	rs3112578	52377996	rs3104744	52389436	rs3104762
52364943	rs3803662	52378688	rs3104745	52389606	rs8055502
52365079	rs3803661	52378902	rs3112573	52390581	rs5816853
52365170	rs3803660	52378941	rs13334992	52390586	rs5816854
52365183	rs3803659	52379032	rs28503301	52390646	rs1109951
52365510	rs8053294	52379049	rs3112572	52390740	rs8061255
52365977	rs12446056	52379220	rs4784228	52391109	rs28636558
52366465	rs4322659	52379400	rs12919508	52391155	rs5816855

Полож в Build 34	Название	Полож в Build 34	Название	Полож в Build 34	Название
52391221	rs3112638	52405970	rs3104770	52413719	rs28709052
52391267	rs9708611	52406122	rs11645834	52413766	rs3112612
52391639	rs12935019	52406128	rs3112623	52413932	rs9925692
52392183	rs3104763	52406232	rs8046543	52414779	rs3104782
52393309	rs5816856	52406857	rs12930738	52414844	rs3104783
52394968	rs5816857	52406859	rs12926392	52414967	rs3104784
52395219	rs5816858	52406869	rs12926394	52415402	rs3104785
52395420	rs11284113	52406870	rs12930747	52415965	rs3104786
52395423	rs11326465	52406881	rs12102546	52416037	rs4578643
52395448	rs6498958	52406905	rs28642024	52416145	rs3104787
52396694	rs3112637	52406977	rs7198019	52416382	rs12597685
52396715	rs4784230	52407020	rs11860998	52416795	rs11374841
52396799	rs13330592	52407095	rs7192828	52417105	rs3104788
52397331	rs5816859	52407116	rs7193515	52417150	rs3104789
52397555	rs11645620	52407139	rs11860769	52417212	rs12599893
52398233	rs3112636	52407166	rs11861036	52417265	rs3104790
52398720	rs3112635	52407928	rs3104772	52417361	rs4290467
52398860	rs3112634	52408038	rs3104773	52417477	rs4467067
52399117	rs16951316	52408105	rs3112622	52417564	rs3104791
52399179	rs3112633	52408239	rs10692294	52417836	rs4238749
52399652	rs28567699	52408304	rs3104774	52417838	rs4238750
52399807	rs3112632	52408429	rs3104775	52418217	rs8048809
52399850	rs3112631	52408504	rs7200735	52418357	rs4459541
52399948	rs5816861	52408747	rs7203671	52418357	rs1112135
52400091	rs13335960	52408951	rs3112621	52419612	rs3104792
52400138	rs3112630	52409006	rs7206289	52419788	rs3104793
52400341	rs3104764	52409053	rs7206313	52420562	rs3104794
52400714	rs11640537	52409255	rs7204612	52420628	rs3104795
52401021	rs3112629	52409498	rs3112620	52420979	rs3104796
52401314	rs3112628	52409610	rs3112619	52422526	rs8046994
52402239	rs6498960	52409628	rs8062962	52422537	rs7193357
52402254	rs5816863	52409779	rs12597100	52422543	rs7186461
52402438	rs4262942	52409788	rs3112618	52423052	rs3104797
52402568	rs3104765	52410064	rs3104776	52423497	rs3104798
52403048	rs2335	52410319	rs3112617	52424706	rs3104799
52403100	rs10653756	52410349	rs3112616	52424978	rs3104800
52403137	rs3104766	52410566	rs11075551	52425772	rs3112611
52403194	rs7201410	52410952	rs3112615	52427035	rs12921031
52403223	rs12934093	52411233	rs3104777	52427036	rs12921032
52403340	rs3104767	52411546	rs12929797	52427156	rs3112610
52403795	rs28400414	52411635	rs3112614	52427333	rs3112609
52403802	rs3112626	52412049	rs7202142	52427444	rs8045104
52404000	rs7190208	52412254	rs3104778	52427468	rs3112608
52404011	rs28379686	52412259	rs3104779	52427539	rs12921801
52404038	rs17292403	52412298	rs8044824	52427910	rs6498966
52404154	rs3112625	52412310	rs10665605	52428293	rs4488440
52404215	rs12920540	52412516	rs3104780	52428649	rs3112607
52404998	rs3104768	52412602	rs3104781	52429238	rs5816865
52405153	rs3112624	52412765	rs3112613	52429511	rs3112606
52405181	rs5816864	52412855	rs8044760	52429906	rs5002422
52405833	rs3104769	52413602	rs12922061	52429965	rs3112605
Полож в Build 34	Название				
52430198	rs3112604				
52430316	rs11075561				
52430713	rs12599215				
52431082	rs4640177				
52431277	rs5816867				
52431322	rs4638564				
52431382	rs11647405				
52431582	rs11075562				
52431715	rs3112603				
52431739	rs4638563				
52431878	rs4426347				
52431941	rs3104801				
52432313	rs3112602				
52432373	rs11434012				
52432387	rs6498968				
52432454	rs3104802				
52432459	rs3104803				
52432718	rs7198792				
52432830	rs13338987				
52433165	rs11643841				
52433198	rs3104804				
52433468	rs3104805				
52433888	rs3112601				
52433897	rs12924424				
52433898	rs12924425				
52434496	rs28680336				
52434497	rs5816868				
52434960	rs7202932				
52435190	rs3112600				
52435864	rs3104806				
52436127	rs3104807				

Таблица 23. ОНП маркеры НарМар, использованные при исследовании мультиэтнической когорты

ОНП	Американцы европейского происхождения		Йоруба (африканцы)	
	r ² *	D'	r ²	D'
rs12922061	0,39	0,81	0,03	1
rs4784227	0,81	1	0,03	1
rs17271951	0,88	0,96	0,03	1

*значения r² и D' по отношению к rs3803662

Таблица 24. Данные по ассоциации с раком молочной железы, полученные на мультиэтнической когорте

Американцы европейского происхождения						
ОНП	p	r	#пораж	Част пораж	#конт	Част конт
T-rs3803662	0,0214	1,23	534	0,3408	572	0,2951
T-rs12922061	0,0724	1,19	534	0,2799	572	0,2462
T-rs4784227	0,0036	1,32	534	0,3034	572	0,2480
C-rs17271951	0,0154	1,26	534	0,3096	572	0,2630

Латиноамериканцы						
ОНП	p	r	пораж	Част пораж	конт	Част конт
T-rs3803662	0,0498	1,22	425	0,4190	426	0,3723
T-rs12922061	0,2487	1,13	425	0,3231	426	0,2972
T-rs4784227	0,1373	1,17	425	0,3429	426	0,3091
C-rs17271951	0,4867	1,08	425	0,2905	426	0,2753

Американцы африканского происхождения						
ОНП	p	r	пораж	Част пораж	конт	Част конт
T-rs3803662	0,0069	0,77	432	0,4747	457	0,5395
T-rs12922061	0,0358	1,41	432	0,1065	457	0,0777
T-rs4784227	0,0853	1,34	432	0,0972	457	0,0744
C-rs17271951	0,2060	1,24	432	0,0922	457	0,0755

Таблица 25. Риск рака молочной железы, ассоциированный с аллелем А rs13387042 и аллелем Т rs3803662 при опухолях, позитивных по эстрогеновым рецепторам и прогестероновым рецепторам

rs13387042 (A)			rs3803662 (T)			
Случаи	OR	P	Случаи	OR	P	
Эстрогеновые рецепторы (ЭР)						
ЭР-позит	2124	1,22	4,3E-09	2128	1,32	5,8E-13
ЭР-негат	589	1,06	0,37	589	1,07	0,30
ЭР-позит в сравн с ЭР-негат	1,15	0,036		1,21	0,0098	
Прогестероновые рецепторы (ПР)						
ПР-позит	1802	1,23	2,8E-08	1805	1,30	9,3E-11
ПР-негат	816	1,12	0,036	817	1,16	0,015
ПР-позит в сравн с ПР-негат	1,09	0,150		1,13	0,057	

Таблица 26. Последовательности, фланкирующие rs4848543 и сцепленные ОНП маркеры НарМар с r² > 0,2

```

>rs4848543
GCACTAGAAAGGTTACATACAATGGGAAACTGAGGCCAGAAAGGAGGAAA
GATGAAGTGCCCAAAGTTAAGGAGATAAAGCAGAAAGTGCAGGTAAGAGA
AGGAGTATTCATTAATGTCAGGAATCCAAATCTACAAGAACATACGCT
TGTTCTTTAAAGGGAGCCCTCGCTTTAAGCATATTTGCCGATCAATACC
[A/C]
ATTAATAATTACAGGACATGGGCAAAATATTCTAGCAAAGGTCCTGAGGT
CTTCCTAGATGTGCCCAAATAACAAGGACTCCCCCGCAGGTGGCACCA
AAAGTCGGAGGATGTGTTGGAAGCTGACCTTCCCACTGTCCCTGCCTC
ACCCACAGTCACCCACCCCTCTAGGATGTAATGGTGAATCATTCACACAT

>rs11680207
GTGACACTAGGGACGTGGGGGAATGCAAGTGGGTCCAGACCTGGGGGCC
ACCACTGCAGGAGCAAGGGCAGGCCGGCCATCAGAGCTGGGCGGGAGGG
GGGCGCTGAGCTTGGCGTGGGAGTTTCCCAACGAGCTGTCTTCCAGG
GACCTCGCCCTCATGCCCTCTGCGTCTGAAGGGAGGAGGCACACC
[C/T]
CTCTGTGGCATCTGCTTTTAGGGCTGACGATGGAACAAGCTCATGGAC
AGGAGTGAAGCCCCTGGCACATAGCAGGTGCCACAGAAGGGTGGTGAAGT
CTGCCCTAGATCAGAGATCATTCAGGGGTCTGTTAACTTGGATGAAGAA
GAAATACACTCTTATCACTAACATCAGGAAGGTAAGCAAGGAACACACAT

>rs7583824
AAAGGAAGGGATTTCTGGGCCTCAGAAAACCATGCCAACGCATGGGGG
TATTAAGGATAGGGGCTCTTGAACGCCAGGCTGGTAAATCTAGAACGG
CAGAGGGAGGAGGGTAATTTTCTGGTTGGTAGAAATAAAGTGGGGATG
AGGAAGAGCACGAAAGTCTGTCGGCAACCTGAAAGTGGGAGCGAGCTTT
[C/T]
GTCAAATCCAGGGTCAACCTGAGCTCTCATTAATTGGGAACTGGAGAG
TCAGATCCTCTAAAGAAAGTGGAAATATTTGTTCTAAAGAACCCCTGCCA
GGTCCCGTTTACGTGACTCATAAAGAGGCCTTCAGGCAGAAAGGATCCC
AGTGCACCACTAACAGAAGGGCTCAGGACCAGAGGAATGGACAGGGAAC

```

>rs4848541
TGGCCTATAGTCGAGTAGATAGAGGGAGGGCTGATAACTAAGTGACTACA
ACAGGACCTCCCTGTGTAAAGGTTCTAGCTTATTTCCACACTCTTCTGGC
ACAGACAGGTTGGCAGGGCAGCTGGTAGCAGGCACAGCAGATGGAATGTC
AGGAATTGAAATCCACAAGACTCCAAGCCCCCTCCAGGGTGAGCTTGCCAC
[A/G]
GGCCACACCAGGAAGCTGAGCTCCTCACCCAAAGCCTGCCAGTGCCTAC
GTGGCAGTCATTCCCTTGTCAACAGACTTCCCACTGGCTTCTCTTGCTCC
ATTATTAAGCAAGAACCAAACTTCACTGAGGGGATAGAATCCCACT
AATAAGCATGTCCACGCTTTGGGTCATCTTAGGGGCACAGAAATCTGAGT

>rs1446128
GAGACACTTTTCTGCCTGCCCAAAGTCACATAGCTCATTAGAGGCCAGGT
TGGGCCTGGAGTCTCTGGCAGGAATAATAGCTGAACCTTCATTATTCCC
TAACCCACCTGTGTTGTGAGGGCCCTGAGGCCAGCATGTTTCATGGCTGCA
GCCTCCAACCAACTGAGCTAATTGACAAGAACTCTGCTCTGCCTGATGCT
[A/G]
TGGGGGATACTCAGATTTTCTGTGCCCTAAGATGACCCAAAGCGTGGA
CATGCTTATTAGTGGGATTCTATCCCCTCAGTGAAGTTTTGGTTCTTGC
TTTTAATAATGGAGCAAGAGAAGCCAGTGGGAAGTCTGTTGACAAGGGAA
TGACTGCCACGTAGGCACTGGCAGGCTTTGGGTGAGGAGCTCAGCTTCCTG

>rs4849766
CCAGGCTCAAGAGAGAAAGCCAGACCCTCTCTCCTTCCACCTACCTGAGA
GAGGCTGTCTCAGGTGTCTGGCTCACAGATGGGGACCCAACTGGGAG
TCACTCCCTTCTGGAAGCCCCAGGCCCTATAAACAGGTCCCCGGCTCAC
TGCTGGGCCTCGGCCTGGAGACAGCGGCTCAGTGCAACACAACACGCCA
[C/G]
AGCTGATCAGGGGAAACCCACGCAGCAGTTGAGGGGGCTGGGGCTTTGC
GGGTGCACTGAGGGGACTGCTTCCCTTCAGTGGAGTTGCTGCCGCAGGTG
AGGGGAGGAAGATGAATCTGAAGGTTCTAACGTAGCCACAGGCTGAAG
CGAAGGCTGGCCTGAAGTGTTCACCTGGGCCCCACCTGAGCGGCTGAAG

>rs6715243
TGGGCCCCCACCTGAGCGGCTGAAGAGTTCTGTTTCTCAAGGCCACAG
AGCTCCTCAGAGGGAAGCTTTAGTAGCTATTTCTTAGGAAGGCAAGTAT
GTGTGCAACATTTCTTAATAAAATGCAAAGGCTGTTGGAGGCACTCCAGC
CCCCTCATTTAGCTCTGAGGTGTAATTAATACATAAAATTAAGCTCCTGCA
[G/T]
ACTCACAATGTCCAAGAAAAAATGAAAAATATAATTAGTTTGCCCA
TCTGACTCCATCTGTCAGCTCTGAGCTCAATGCTCAACGTAGCAAATCCT
GAGGGCTTTGGTCTCGCATTGGCTGGGCCCCGGCGCTGCTCCTGCCCTGCT
CCAGGGAAGTGGGCCCTCCGCAGCAGCTGCCAGAGCCCCGGGCTCTGGACT

>rs895398
TTACCAGCCCTCTCTCTCCCTTCTCCCTTTGCCTTCCACACACAGCCTC
TTGTCCAGCCACTCCCTCTTCTGGAAGGGCCTTACTGACAGAGTGCTGG
GGAGCCCTGTGAGCTGGCCACAGGTCTTAGCAAGGGCCACCCCTCCTTC
AGGCCTTACCGTGTCTTGGAACTCCCACTGCCACAGCCTCAGCGACAC
[C/T]
GTCCTTAGATCCAGGCAGGTGGGCTGTGACCCGGGACTGCACTTTGGAA
GGCCTTGCTATCCTCTCTGGCCATCTCCTTCTCCAGGGCCCCGAGGGATC
TGCAGAACCGGAGGAACATGCCAACCTGAAACCTGGGCTGTCTTTATCCG
GCTGTACTCCACCTGGGATTTTCTGCCCGGATGGCATCAAGATCGCTTCCA

>rs11684731
GCCTAAGCCAGAGGAGGCAGGAGGTTTAGGGCACCAGCCAAGAAGTGCT

TTCCAGATGCTGAGTGGGATCCCGAATAAGGGAAAGGGCATCTAACATAG
AATGAGGAGGAGATAACACAAGGGTAGAAATAACCAAGAAAGGGGATTTC
TCTCAGATGGCTTTAGAGGTGCTCCAGCCATTGCTGGGCTCACATGGCT
[C/T]

ACATGTCATAATTTTTGCCAAGGCGGTTTCACTGCCACCACCACCCCGC
TCAGCCTCCAGTGGCTTGTCCACCCATGGCTGTGGCCTCTAACTGGGCAG
GATTTTATAGGCTAGGAAGAAAAGTAGGGAATGTTAAGGGAGACGGAAGT
CAAGTTCAACTTGGAACCATCCAGAATGTCCAGCAGGAAAGGCCCATGGA

>rs12464139

GACACCTCACCTTGACACGCAGCCACCTGCCAGTGTTCACAAATGTTGGT
TTTCAGCTTTATGCGAGGGTAGGATCCAACAAGCTCATTCTTTAATGAAA
AAGAATGAAATTGTAAAATATCAGATCTTATTACATGTCCTAAGGGTAAC
TTTTATCTCATAACACTTTTGCTTCAATCATGAGTGTATAATTTCTATTT
[G/T]

ATTTTGTGTCTATGGGGTTAGGTGTGCAATGTAGAAAAGCTTGAAAACC
ATTGCCTTATACCCATTTGAAAAATGAACCACCAACAAAGGCTTAAAAA
CATGAATATTTGTCAACCAAAAGGAGAACTCCTGGGCTGGAAGAAACAT
CACGATTCAAGTGCTGCAGCACGAAAGAAGTTGAGGCTGGATTGTATTCA

>rs6754664

TCCAACAAGCTCATTCTTTAATGAAAAAGAATGAAATTGTAAAATATCAG
ATCTTATTACATGTCCTAAGGGTAACCTTTTATCTCATAACACTTTTGCTT
CAATCATGAGTGTATAATTTCTATTTGATTTTGTGTCTATGGGGTTAGGT
GTGCAATGTAGAAAAGCTTGAAAACCATTGCCTTATACCCATTTGAAAAA
[G/T]

GAACCACCAAAACAAAGGCTTAAAAACATGAATATTTGTCAACCAAAAGG
AGAAGTCTGGGCTGGAAGAAACATCACGATTCAAGTGCTGCAGCACGA
AAGAAGTTGAGGCTGGATTGTATTACGCTTGACTGTTTGTAAAGCACCA
TGCACCATGCAACTGCCTTCAGTTCTTTTTTTTTTTCTTTCTTTTTTTTT

>rs6754799

TGAGTGTATAATTTCTATTTGATTTTGTGTCTATGGGGTTAGGTGTGCAA
TGTAAGAAAAGCTTGAAAACCATTCCTTATACCCATTTGAAAAATGAACC
ACCAACAAAGGCTTAAAAACATGAATATTTGTCAACCAAAAGGAGAACT
CCTGGGCTGGAAGAAACATCACGATTCAAGTGCTGCAGCACGAAAGAAG
[C/T]

TGAGGCTGGATTGTATTACGCTTGACTGTTTGTAAAGCACCATGCACC
ATGCAACTGCCTTCAGTTCTTTTTTTTTTTCTTTCTTTTTTTTTTTTT
TTTTTTGAGATGGAGTTTCGTTCTTGTGCCCCAGGCTGGAGTGCAATGGC
ACAATCTCGGCTCACTGCAATCTCCGCTCCAGGTTCAAGAGATTCTGCT

>rs6759589

GTGGGGGCAGCTCCTTGTTTTGTTTCTCACTGTCAATTTCCAAAAGGCAG
ACAAGCCCTGGCCGCTTATGCCAGGGCTGAGTCCACACCATGCTACAC
GAGCACCTGAGGGGAGGCTCCTCTCAAAGTGCATTACCAGGGCCTGTGC
TAAATTTTTCAGCAGCCTCTTGCTCACTAATGGTCCCGGGAATAGACT
[A/G]

AGCCTCCATCCCCAACTAAGTGCTCAGGCCTCCCCACCTCTGTCCTGC
CTGGGCAAGAAGCCAGCGCCGCCACAGCCAGCTCTCCATGCAGCCT
CAGCTCCAGTGGCCCTCTGCCAAGGACTTGCGCCTCTGCAAGGGATGC
AGTGCCTCATTTTATCCATAATTGTACGTGGTCCCTGTGTGATGTCCGGGG

>rs1562256

GGCCAGACCACACTCTGTGGGGTGTGGGAGTTGGGTTTTCTTGCTGCAA
GGAATATATGTTCTTTAAGGCAGTCATCTGCCTTAAAGATCCTAAGGTGA
GGAATTAGGGGAGCCTTATACTTGAAAGGGAGGTCCCCTTTGGGCCTGGG
GAGCATTGGTCTGAAGGCTCAGGTATGGAGGATGCAGGAATATCTGCTGG

[C/T]

GAGGATTGAGCTGGAACCTCCCTCCTTCCCACTCCCATCTTAGTCCCTC
CGGCCTTTAAAAAATTTATTTTATTTTATTTTATTTTAAAGATAGAG
TTTCACTCTTGTACCTAGGCTGGAGCGCAATGGCTCGATCTCGGCTCAC
TGCAATTTCCACCTCCTGGGTTCAAGTATTCTCCTCCCTTAGCCTCCAA

>rs1446125

GCTGGTCTCGAGCTGCAGACCTCAGGTGATCTGCCTGCCTCGGCCTCCCT
AAGTGCTGGGATTACAGGTGAGAGCCACCGCGCCAGCAGACCAGACCCT
TTCTAATACCTTCTGACCCAGCAGTGGGGCCAGGCCACCTAACATGG
CAGCCTGGATGCACCTCTCTGTTCTAGGCCTGGTGCCTGGCTCAGTAA
[C/T]

GTTTTTTTTTTAATGATTGAAAGCTCATCTTTATGGTACTTAAAGTCT
GTTTCAATTACTAAGGGTCTTTGTGCGCTTTTTCTTAGGGCAGACTTGT
CAGGCCAGACCACACTCTGTGGGTGTTGGGAGTTGGGTTTTCTTGCTGC
AAGGAATATATGTTCTTTAAGGCAGTCATCTGCCTTAAAGATCCTAAGGTG

>rs1446124

GCTGGCTGGGATTACAGGTGCCATCACTATGCCAACTAATTTTTGTAT
TTTTAGTAGAGACGGGGTTTACCATGTTGGCCAGGCTGGTCTCGAGCTG
CAGACCTCAGGTGATCTGCCTGCCTCGGCCTCCCTAAGTGCTGGGATTAC
AGGTGAGAGCCACCGCGCCAGCAGACCAGACCCTTTCTAATACCTTCTG
[A/T]

CCCCAGCAGTGGGGCCAGGCCACCTAACATGGCAGCCTGGATGCACC
TCTCTGTTCTAGGCCTGGTGCCTGGCTCAGTAAACGTTTTTTTTTTAA
TGATTGAAAGCTCATCTTTATGGTACTTAAAGTCTGTTTCAATTACTAAG
GGCTTTTGTGCGTTTTTCTTAGGGCAGACTTGTCCAGGCCAGACCACACT

>rs11677262

TGATACATTATTTATGGATCAGGTGGCACAGCATTGCAGGCAGCAGCAGG
GCAAAGACAGTGTGGAATCAGAGTCGGGCAGCGTTGGATGGGATTCCGGT
TCTGCCACCACTGGCTATGTGATGGAATCTACCACATCTCTCCAAGCCT
TGCTGTCCTGTTTTATCAAATGAGGACATTGACTGAATCAAGGTCTG
[C/T]

ACAATGCCCCTGCATGTGCCTGGAATCTTCCCTCTGGTGTCCCTCAA
TGCCCATGGTGCTTCTGCACAGGGTGTACCATAGTTGCTTGTCTCTCC
TGCTGAAGCCTGCTGTGAAGAGGACTGTGTGGAGGTAAGGCAGTGGGCC
TCTCCTGTGAGGTCTCAGGGGGCTGAAGGCAAAGTCAGAAGTAAAAAAA

>rs7565771

TCTTCTTCAGTATTTCCCTTCTCTCTGAACACCGTTTCTTTTCTTTCA
GATTATCTGACAGCCTTTTAAATGAACACCTACTAAGTGCCAGCACTTTC
CTGGGAACTTAATTCTTACAACCTGTTAAGAGGTCTATACACTCATTTTA
CAGTGGAGGAAATGGAGACTTAGCGGAGATGTGAAATTTGAACCAGGTCC
[A/G]

TGTGGTTTCAAAGCCTGCGCATGTTCTTTACCCTATGGAAGCCTCCAG
AAGCCAAGGTCTCATCTCCCTCTTCCCTGGAGGCCCAAACCTGGGTAG
CTGTGCCCCTAGACCTTTGAGGGAGGGTCTCGAGCCATTATCTCCCTTAC
CTTTTCTTGCACTGACTTGGCTTCTCCAGGCTGCCGTAACCAGCAAACGC

>rs838066

TCACCTTTCGTGGTCTGGCACGTCTCTGCTGGATTTTTGAGCCAAGGA
CAGAGGGTTTGGGAGAAGGAGGCCTCCTTGGACTTCTGCCAGGAGTTACT
GAGAAAAGAAAACCAACCAGCTGGGGGGCCAGTTGGGGGGAAATCTGCA
CAGGATGTTCACTGTGCCCCCTCCTGCTTCTTCTGACCCCAAGAC
[C/T]

CACCTCAACACATCTGGAGTGTAGGGTGTGCTGCCCTGTCAGACCAC
TCTCCACACACAGCCTGGGGCAGGAACAATGGGACAGGTGGCTGGTGTCT

ACTCAGCACAGCAGGGGTGGTGGAAAGACACATGTCAGCTGAGGCTTGGC
TTCAGGGCAGGCAAGTCACAGCCAAGTCACAGCCCAGCCCAGCATAGGCAA

>rs895397

TCAACGGTCTTGGCTTGTGACTAGACCAAGATCGCCAACCAGAATGGAGG
GAGGCTGACCCAGCACATCCCGGAAGCATTTACAGAGCCAAGGAAGAGAT
TCTTCTTGCCCGTCTATGATAACATTGCATTGCCTTTTCAGCTGGGGGCGA
TTAGAAAGGCACTCACTACTAGATCATGGGAGATGCCTGGGTTTTAACCC
[C/T]
TTGCTGCTCGGAAGGTGGTGATCATGATGGCCTGGGAGGGTTTTAGAAT
GGCAGAGTGTGAGGCCACCCCAAAGCTCCCAAATTAGAATGGCATTTTAA
CAAGACCCCTAGGAGGTTGTATACGTGTTTAAGTTTGAGAAGCGCTGGT
CTACATTACAACCTGAGGAACTCACTACTCCCTGAACACCTGCATCAAGGT

>rs838059

ATAAGGAGCTTCAGGTTATTTTTAGGAAGTCTAAGAATCCTGTTTTCTGG
GGAGCCAGCCTCTGTTGGGGAGATAAGCCTCCGGAGGCCAGAGGCAGGAG
CTGGGCCTCAAGGAAGGCCCACTGGACCCCAAAGGGCACGGTGGGG
AGGAAGAGGCAGAGCTCTGGTCTCCCTGTGGAGGGAGTGCTGCTGTAAAG
[C/G]
GCCTTTCCCTCCACCCCACTCCCGCCGAGAACCCCCCACCCCGCA
GGCAGTTTGCTTTCCCTCCACAGCTTTCCTCTACCCGAATACCCACC
TCCCAGCTCTGCTCAGCCAGGGCCCAAGGCCAGAGCCTCTTCTCTT
GCAAGAACTTTCCTGTCCACCTTGCCCCAGGGCAACACCTCTTATGGCAA

>rs838102

CTAACTAACTGCCACATTCTCTGTCCAAATTCTGCCTCATCGGTTATTC
ATTCAACAAACACTTACCGAGCACTAATTTGAGCCAGGCACTCTAAGAC
CAGCTGAGTCCCACTTCTGGCAGGATGAGTGTTGTTGTCATCAACCACCC
CACAAACACCCGTGCCCTCTGCGTACTTGCGAATCTGAGTGAGGAGACAC
[A/G]
GTGGCAGTAAGGCTCAGGACCTATAGCTCATGTGAGTGAGCAAGGATGC
AAAACCAAAGACACCCGCTCCCTCACCCACTGAGCTGGAGGGATCAACC
ACTGTGCTGCCATCAGGGCTGATGGGCGACTGCTTGATGATCAGAACT
CTTGCGTTGTCACTTCAACCCAGGGGCAGAGTGTTCTCTTTCAGGTGTG

>rs3769659

GGGCCCAGCACTGTCTTGCTTCTTCCCTTCCAGTTCCAGGCACCACTT
CTTTCTGGCCAGTCCAGAGACATCTCCAAGGGCCAGAGGGTCTATCTCCAG
AGGGCCTGGTGTCACTTCCCGCAGCTGCTCTGGCCTTTACCCAGGTTGG
GGCGATTTGATGGACAGTGGTGGTCCCCTCAAACCTCACTGCATTACAG
[G/T]
TGCTACTGACTCATGTGCAGGCCTGGGCAGTCACCCAGAAAGCCCCCT
GGATGCCAGCAAGCCCACTGTGGCAATGGGCAGGCCAGACCTCACCTGGA
GGGCCAGCTCACAGCCACTGTATCCTCTAGCTGTGCACCCAGAGTTTAC
CTTCAGAGAGACACTTGCCCTGACTGGGGGTATACTGTTACTCTCTCAAGC

>rs865108

AGCAGCCAGAAATCAGGGTGATGGGTGGGTGGAGGGGCAGGCATCTGGGA
AAGCCCTGCCCTCCTGGTGTTCCAGATCTGAAGGCTGGAGGCTGGAAAC
CTTGTGTACCTGTGCTGGTGTGAGTTCCATTGGGGCTGCCGCTTACCTT
CCCATGACCTTTAACATCTGGCATGGCTTGTGGGGGTGCTGAGAGGAGG
[A/G]
GAAGTCTCAAGGCTGAGGGAGCTCAGGGTAGAGCCAAGGTCTCAGTGAG
AACCCAGGGGGCAGCTGTACAGCTGACCTGGAGCCCTCCAGGAACGCG
TACTCTCAGCTCCTGTACTCTCAAAGCAACAGAAAGGTTTGCTTTCTT
CAGGTGCGGTTTCTCACAGTGCAGCTTCTCAAAGAGATCTGGAGACCCC

>rs708670

GCAGGGACTGACTTCCACTACCCCTCCTTCCAGTCCCTGCCCACTCCACC
 CCCCACACAGCCAATGGCATTCTTCCATCTCCAATCTTGACAACCTCC
 TGGCCCATCCCAATCCCCCTTCTTAAAGATCCAGCCCAAGGCCTTGGGG
 ACTTAAGGCACCTGCCATCTCCCCAACGCCACCCCTGACAATCTCTCC
 [C/T]
 GGCTTTGAATTGAGACAGGCAGCTGTGTTAGTTGGGGGACCTAGCTGGA
 CAGAGATGTGTGTGAATCCGGTCCAGCCATGTGCTAGCTGTGTGACGTT
 GGGTAGGTTGCTTTCCCTCTCTGCTGATGATTAAACAGTGAATATGAA
 ACCCTAAACACATGTCTTCTTAGGACTGTTCTGTGTGCACCTTTCCTT

>rs708672

CCTCATTGTTCTCTGACTGCTGTGGCCCCCTGGGTCTCCTCTCTTCTACT
 GGAGTGAAAAGCCCCTGTGGGCAGGGCTAGATTCTGAATCTCCGCGGAGT
 CTAACACAGGCCTGTAGGCAGGGACTGACTTCCACTACCCCTCCCTCCAG
 TCCCTGCCCACTCCACCCCCACACAGCCAATGGCATTCTTCCATCTC
 [A/C]
 AATCTTGACAACCTCCTGGCCCATCCCAATCCCCCTTCTTAAAGATCC
 AGCCCAAGGCCTTGGGCACCTTAAGGCACCTGCCATCTCCCCAACGCCCA
 CCCCTGACAATCTCTCCGGCTTTGAATTGAGACAGGCAGCTGTGTTAGT
 TGGGGGACCTAGCTGGACAGAGATGTGTGTGAATCCGGTCCAGCCATGTG

>rs708673

AAACAGATGACAGCACTTGTCTCCAGGGCAAAGCATGGACGGGGGCTGAA
 GAAGACACAGGCATGTTTCTGCTGGCTGGGGATGGCTTTTCAGTGGACCT
 TGGAGGCCACACCGGCAGCCCAAGGAGGCTGGAGGTTTGGAAAGCTTTTG
 CATTGTTGAAGTGAATTATGTGCAAGCCCATCTTAGGGATAAGTGATT
 [C/G]
 TATTGTTTTCAACTTGTGTTGGAACCTAGCGATCTTGGAAAGGGCACAT
 CTAAGTATCCCGGGCATCTTTATCTCAAGCCACATGCAGAACCTGTTTT
 TCAGATCCCTGAGCATCTGATAAACAGGTTCTGGGCAGGTTTCCCTATC
 ATCACCACCAAGCCAAGGCCTCATTGTTCTCTGACTGCTGTGGCCCC

>rs838100

ATCAGTCACTACCCGGTCCAGCCCCTGTGGCCAAGAGCTGGCGTGCAG
 GCTGCGGGAGGCAGCTGGCTGTGCAAGACCCTGGCAGGGCCCTCGCCTCC
 TGAGAAACCGAGAGTCAGAACCAAGCCAGGCTGTCTGGTTGGAGACTG
 AGCCAGAAAGGGTGGCTCACCTCACGGTGAGGCTGTGAGTGACCTGAGA
 [A/G]
 CCTCAGACCCTCACGTGAGCCGGATGTCGACACAGCCTGCTGTTGGTAA
 GTCTGGCTAGGACGCAGATCCAAGGGGGCATGGGTGCTGTGCAAAACCCCT
 CCCACAGCTCCCTCACACCTGACCCCTGCCCTGCTGGAGGGTGCCCT
 TGAATCTCCAGGAGGTCCGAGGAACCTCGACACCTCCAGAGATCTGTACTG

>rs838098

AAATTTGAGGGGTTCAAGGATTTATCACACAGTTGCGTGTGTGTGTGT
 GTGTGTGTGTGTGTGTGTGTGTGTGTGTTAAGGATGTTAATGTTTAT
 GATGCATTTCTGTGGAATGCAGGCCATTGGGAATTGAAATTTTCTTA
 TTGGTAACCATTTCTGCATTGGTTAGCACGGAAGTGTGCTGCAAGTAA
 [A/T]
 TGAACCAGATTTTTCTGGAATGATTGAAAATCTAGCTGTTTGTAAGAAA
 GCTCTCTAAATTATTCATTCACTTTTATTTTGCACAGTTGTGTGACGGTA
 AACTCTTTTCCATCAGACGAAACTTGCGCGCACGTGATTTCTGCGGGGA
 TTGGCTCGAGTGTGCTTGAGTTTTCTGTATTCTTTTGTAGGAACCTCTC

>rs838096

ATGCCACAGAGAAAAATAGAAATCTACCAATTAGATGTGTGTGGCTATA
 AGTGAGCTTTTAAACCATGATTTGAGACTGGCATCCTCTGGCCTCTCTGC

CCTCAGGTCAGCTGTGGGAGGCTAGGGGTGGGATAGCAAAAGGGACCTGG
 GATCAGGTGGGGCTTCCTGGACCCAGGCCAGTAAGAATGAAGCCTAGG
 [A/G]
 GTTCTAGGATAGGGGCGAGCCCTGGATTGCGAGAAACAGCCTGAGAATG
 AGGGGCTGGAAAGTATGAGTTGGAGCAGCTCAGTGAGGAGACAGAAACT
 GAAGCAGAAGAAAGCAGAAGACACCAGAGTCTGGTGCTGACTGGGTGACT
 GGCTGTTACTCACATCCTGAGCAATGAGCCACAGACCCTCCCCACCCAGA

>rs838092

CATTTCCCTGAGGGCTAATGATGCTGAACATCTTTTCATGTGCATTTGTA
 TATCTTTTTTGGAGAAAGTGTCTTTCCAGTCTTTTGCCCACTTTATGTAG
 GTTGTTTGTCTTTTAAATACCGAGGTGTAAGAGTTCTTCAGATATTTAGA
 AACAGGTCCCCTTATCAGTTATATGATTTCTGAACCTTTCTCCATTCTG
 [A/T]
 GTTGGCTTTGCACTTTCTCGATGCTGTGCTCTTTGATTTCTGTCAGCAG
 CGAATTAGCAGGACCTGGGGGGCTTTAAGAGGTGAGCAGGCAGACAGACA
 GCGCCACAAAGTCCTGAGGATGAGAGAAGAGGCAGAGCTTACACAAAGG
 CTCTGTGAAAACGAAGATGGACCCAGGCAGGTTGGGGTGATCCCTGGGG

>rs838090

TCAGCCTCCGGAGTAGCTGGGATTACAGGCACGTGCCAACACACCCAGGT
 AATTTTTGTATTTTAGTAGAGATGGGGTATCACCATGTTGGCCAGGCTG
 TGTGCAAGAATTTATGTTTTCACTTGCGCTTGCTGAGCCCTGGTGCTGG
 CTGTGGGCACATACTCGTTACCCAGGGACATGCACTGTTGATGGGTCTT
 [C/T]
 GCACTCTGGGCTATCTTGGGCCCCGTTTGCTCAATGCAGGCCCTCTTT
 CTGAAGCGCTTTTCTTGGGCTCACTCTCAATTGCTCAAAGTTGCGGCTCAG
 GCCCACCTCCTGCAGGAATCCTCCAGTCCCCGCTGTCCGAAGTCAGTAC
 CCTGTGAGAGACCTTAAACACTTTAGTCCACTGCCTGGCCCTCCCATGG

>rs12467944

AAGAGGCCCGCTAGGGAGCAGTACAAGGCAGCATTACAACCTGGCATTTT
 CCTGTTGCAGACTGCATTTTGGTGGCAAGAGAAGCTTTGGAGGTGCTTAA
 CGTTTAGTTTGCATTTATCCTACCGCTGAAAAGCAGGGACCCTTAGGACC
 ATAAGTCTAGGACTTTTATAGCACCCCTAGACTGCTGGGGATGAAAGGGG
 [C/T]
 TTGGAGCTGAGCTGGTGCAGCAGCTCCTGTTAGAGAAGCCACCCAGAGA
 GGCCAGGGGACCTCCTATGTCCAGCGCCAGTTAGGGGCAGACGTGGGAC
 CAGCTCCCCCTCGGCTGCAGCGGCCTCCATATCAGCTCCCTGGAAGGGAGC
 AGATGCCAAGTCAGGTTCTGAGCAACACCCTCAAGGCCAGCCCTGGCAAT

>rs10196144

GACTTGGCCCTCCCGACCACAGGAGAAGAGGGATTGAATTCACCTCGGC
 CCAAATTGCAGATGCTGGCCTGCGAGCCCTGAGGCCATCTGCAAGGCTCC
 TGGGCTGACATAACTGTCACCCAGCTCTTCATCCTCCTTCCCCGGGGCTG
 AAGGGGTGCCTGCAGGGTCTTCTGCCCTCCCTCCACCTGGGCCCTACTGC
 [C/T]
 GGCACGGAGCTTCATTTCTGTTGGCAGCCAGTGTGTTGCTTAGTGATTAC
 TGAATGCCGTTTGTGGGTGCTTATGATGTGCTCAGCACTGTTCTAAGTG
 CTTTACACAATTGACTTATTTACTCCTACAACCTGCCATAAGGGAAGTG
 CTGCCATGGTCTCCCTTTATAGCTGAGAAAACCTGAGGCACAGATAGGTCC

>rs6728660

AGGCGGACTGGCTCCAGAGCCCGGCTCCTGGACCGAGCAGCCTTGAGTAG
 GGACCTGGTGCAGACCCGCCTGCCGTGACTCACTATAGGATGGGTGAGG
 AGGCCCGTTGTACCTACAAAATGTAGCCCAACCCAGGAAACGCTCCC
 AGAACCCAGGCAGAAAATACTGTCTCCTATTACAGGGTGGCTGTGCCTGG
 [C/T]

GTGAGTGTCTTGGCGGGTGCCTGCATGTCTGTGCAGGAGTGCCACCCCC
 ATCTGTGGACTCGGAAAGCATGATCCCCACAAAAGCAACCCCTCAGAGTC
 TCAGCTCAGCCAATCCCTGTTATCCTGTGACAGGCAGGGCCATGTTTCAA
 AACCATGTTAAAGATGAGCAAGAGGCTAAGGAAGTTTCCCATGCTGGCAGA

>rs10188946

ATAATACCCATCGTAACAATAAATATAAATGGCTAATTTTCATCGAGTGC
 TGCCAGCATGCTCTTTTGATATGTTATCATAACTAATCCTGCATCACTG
 CAAAATTCATTTTACAGAACAGGAAATGTGGCTCAAAGACATTAATGA
 TCTTACTGAAAATTCATAGCTGGTAAGTTGGGATTTGAACCCAGTTCCT
 [A/G]
 TGGATCTGAAACCCATGTTCTTTCTGCTATGGTAACAGGGTTTCCTAGA
 GGGCAAGGAGGAAAAGTTGATGGCTTTTGTCTATTTGGAGCCAAGAAGG
 AGAGAGACTTGGCTTTTGAATTTGGACTCAGAAGGAGGAAAGTAACACAC
 TTATGTACCCGGCCCTGAGGAAGGCAGGCATGTCTCAAATTTTATCGTATT

>rs895406

AGGAGGAATAGGAGTTTGCAAGCTGTCAAGGTGTCCTGGGGTGGGCGTTC
 CAAGCCGAGAGGAATGAGCTCTGCCCATTTAGAGGACAGTATGGGACCCA
 AGTGTGACACGTACAGGAGAAGGGGAGATGAGGCTGCGGAGGGAGTCT
 AGCTGCATAGTGAGGGTTCTACCATTATCCAAGGACACTGGGGTCACTG
 [A/G]
 ATCTTTTCAGCAGTGACATGGTCAGATGAGAGTTTCAGACCATCCTGTG
 CCAGCCATGTGCAGATGTAGCTTGGGGTGAAGGTAATAGGAGAACGGGGT
 GAAGGTAACAGGAGAACAGTGTGCAGGTTAAACATTTCTTTGATCTTGAT
 CTTTGAGTTTGATGTTGAAATGTGTATTTCCAGGGAAATACACAAGCAAAT

>rs708675

TGGACCTCTGAGTCGGGTATGAAGTAGTGTGTAAGAAAAATAATAAAAA
 GTATTTACAAAATGCAACAGATTACAGAGGCAGGTATGGGGAAGTAAGGG
 GATAGTCTGGGGGAGCTTAACAGAGGTAGGGTTTTGAAGGAGGAATAGGA
 GTTTGCAAGCTGTCAAGGTGTCTGGGGTGGGCGTTCCAAGCCGAGAGGA
 [A/G]
 TGAGCTCTGCCCATTTAGAGGACAGTATGGGACCCAAGTGTGACACGT
 CACAGGAGAAGGGGAGATGAGGCTGCGGAGGGAGTCTAGCTGCATAGTGA
 GGGTTCTACCATTATCCAAGGACACTGGGGTCACTGGATCTTTTCAGCA
 GTGACATGGTCAGATGAGAGTTTCAGACCATCCTGTGCCAGCCATGTGCAG

>rs1867856

CCATACTGTCCTCTAAATGGGCAGAGCTCATTCTCTCGGCTTGGAACGC
 CCACCCAGGACACCTTGACAGCTTGCAAACTCCTATTCCTCCTTCAAAA
 CCCTACCTCTGTTAAGCTCCCCAGACTATCCCCTTACTTCCCCATACCT
 GCCTCTGTAATCTGTTGCATTTTGAAATACTTTTTTATTATTTTCTTA
 [C/T]
 ACACTACTTCATACCCGACTCAGAGGTCCATGACTGGGTCTTCTACACC
 TTGGGGTCCATGGTTCCTAACACAGGTTACAGTGCATGCTCACTGAGTG
 ACTGATTTTAAATTGGCCACCACAGAGCCAATATCAGGCATCTTAAGAG
 GAGTGCACGTAGAGCTTGAGCAAGAACTAGTCTAACTGCCTTGCTGCCAC

>rs12711924

CCTGATCCCAGCTCTGCATCTGTTAGCTACATGTGACCTTGGGAAAGTGA
 TTTGATCTTTGTGCCTCATTTCTCCATCTCTAAAAATGGGCTAATAGTGA
 TCTCTACTTTGTGGGGTGGTTTTGAGGATTCGATGAGTTAATGCCTGAGA
 AACGCTTATCACAGAGCCAGGTAGCGTCAGGTGCACTGTGAGCGCCCGCC
 [A/G]
 TCTTGTCTTTCTTGTGTCCTTCATCATTGCCCTCCCGGGGCACCAACCT
 TTGCCACAGTGTTCATGGTTTTCTTCCATCCATGTCAGGTCTTGCCAA
 CAAGAGCCACCTCTGGGTGGAGGAGGAGTCTGGCGGATGGAGATCTACC

TCTCCCTGGGAGTGCTGGCCCTCGGCACGTTGTCCTGCTGGCCGTGACCT

>rs1867749

GGACTCAGGGACAGTGACAGCGTAATGCTCTTACGTAATCCAAATTAATT
GCTCCTGACCAGGGATGCTTCACATCCATTGGAAACAACAGAGTTGTAG
CCGTGTGTTTTTAAAGTGTGTCTGCCTCTGAATCTCCCTGGCTCTGCTC
ATCTCAGAAGTCCCTGGAGGAGGGCTCATGGCTGACACATGTTAGCATA
[C/G]
GAGAAGGGGAAGAATTTTATTTATCTGCCACTAATCCCATGTGTCTCA
TACCAAATCTAGGATTTTCTCTTCCACCTGGAGTATCCGGCGTGCAT
GGTTTATCCACAATTAAGCAGACTGCAGGCTGCTCCCTCACCAGAAAG
CCGCTCCCTTAAAGCAGCCATAGATTACATGCCCTGGCTCTGCCTGCCAGAG

>rs838073

TGCTCCAAGAGCCTCAGTTTACACCCCTGAAAAATGGAGTTAATGCCTTA
CAGGGATTTTATGAGCCACAGTAAGAGATTATGTGTGAAAGAAAAATAAA
AAGCATTTTGTATTATTCTTAGGGTATCTGGCCACATTTCCATTATCTC
TGACTAAGGTACCAACCACCGATTTCTGAGTTCACTAGGAGACAGTGAGA
[C/T]
AGAGGCCAAGGAACCCCAACCCTGGGCCTAAAAAGTGGTCCAGTCAGTGT
TTGCAAACTGGCGGACACTCACTAGTGGGTCTGCAATTGATTCCATGGA
TCACAAGCAGCATTTTAAAAATGAAATAAAATAGAATGACAAAGAAAACA
GTGGAGTGTTTTGCATGGTGTGCGGATACTTTTTCAGTTATATGGAATTGC

>rs838069

CTGACCTCATGATCCGCCTGCCTCGGCCTTCCAAAGTGCTGGGATTATAG
GCGTGAGCCACCGCACCCAGCCTTCTTTGGTATTTTTTTGTTATTGTTT
TGGTTTGGTTTGGGTTGTGCTAAGTGTTCATTCAGCCAGTTTAACTG
GCTGTTAGATATTAGACTGCAAAATACTCCAGTCAGAGCAGCTCACCTGC
[A/G]
CAGAAAAATGAGGGTTTCAGCTTTTAGAACCTAAGTTCTTTCCAGAGTT
AGAGGAACTGCCTGTGACCTTGCCCACTCCAGCCACCTGCCTGAAATAAT
GACAAGACCAGCCTGACCCAGCTTCATGAGCAATTGTCCAAAGCCCCTT
TGTGCCGTAGTAATTGAGCACCAGCAGGAGCTACAACGACCTGGCTGTTTC

>rs3731603

TCCTGTGGCCCCGCGGTGCTCTTCCCTCTTGCTTTGGGGCAGCAGCA
TTTTAGGTTTTGTAGCTTTGGTGGAAAAGTTGCTCCCCACTGTTTGAAGT
GTAATGCCCATGAGGAGGAAGTTGGAGCCAGTGTGGTTTTCTGGGTCTT
CTGCCAGAGGACATTTGGAGTCCCAATATTTTCTCAGGACTGACTTC
[G/T]
TGAAGACAGTCCTTCAGCCCAGGCAGGGCTCAACTCTGACGCCCCGGGA
GCTTTACTCCCTAGAGTGTTTTCTTGTCTCCCCCTTGGGCTGATTTT
TAAAGGGTCTGCTCTAGGCCTTCTGGGAGTCATTCAAACTGAAGATGGA
AAGGGTGAGAAGGCACAGTGTTTAAAGGACGCTCATCGCCAGGTTAAAGG

Таблица 27. Последовательности, фланкирующие rs13387042 и сцепленные ОНП маркеры НарМар
с $r^2 > 0,2$

>rs13387042

CTGGCATCCGAAGGTATAGGCTCAAATCCCAGCTCTCATGATTGCTAGCT
TTGAGGTTTCAGGTTCAATTTCTTAAGTGCTCTGAGCTTTGGTGCCCTCCT
CTTCAAAATGGGAATGAGGAGCATGTCTAGCTCCCTGTTTTGTTGCACTG
AATAGCTACATGAAGAACAGCTAAACCAGAACAGAAAGAAGGCAAATGGA
[A/G]
GCTACAGAAACCAAGGATTTCTTGTGAATCGAATCTTCTTCAATCT

TCCTTCACCACACTAGTGGATCTCCCTGTGGGAGGGATGTTGAGAGTGCT
CCGTGTTTTTTTTGTTTTTTTTTTTTTTTTTTGTGATGGAGTCTCACT
CTGTCACCCAGACTGGAATGCAGTGGTGCGATCTCGGCTCACTGCAACCTC

>rs1882419

AGCCATGGGGGTGACACAACTCCAACAAGCACCAGGGTCAAGTAGAAGC
TAACTGACCCAGGAGAGAAAATAATTCTGTCAAGTTAAGTAATTTGCTGG
TGTTCTCTGCTAACTCACATCTGCGTCCTGGGTCTCCTTAGCTCTTCCTT
GAAAAAGCAATGAATGTTTTCTAACTGTAGCTACCCCTAGTTCTTCAGCA
[C/T]
GTCTGTGGCTGGCTGCTTGCCACCCAGGCATGAGAGGATGCTGTCCTCT
GGGAGTCCCAGCCTTGTTGTAAGAATGCAGGCTGAGAAAGACTGTAAGA
TTCAATTCAGGATCTTCCACTTGGTGGGCAACCCTAAGAAACTCATTTA
ATCTTTATTAGCCTCCGTTTCCCCTTCTGTAATAGTACCTGCCTCACAGGG

>rs2252488

AGACATGGTCAGGGCCTCAGACCCTGTGTGTGGCCATTAGCTTTGCCCT
GTTCTGTGGCCACAGTCCATGCTCTAAGTACTGACCAGATGTCTTACAAT
GCACACGGCACACAGTTATGGTCTCTGTGGGGATGGGGGAATAAGGTGTG
GACAAATCTGTTAGCTCTAAATCCTTGACCAGAGGAGCAGCTCAGAGCA
[C/T]
GAATCCACAGGGCAGTGGAGCCATGTGCAGGATCAAGCACAGCGGCATC
ACTGACTTGCAGGACACTGGAGTGTGGGGCGTGGTTCTAGAGAAATGCC
CTCAAGTCTCGATAGATCCTAAATGTGGTTAGTTCTCTTATTACAAAG
ACAGTGGAAGCTGAGGTCCATGCAGATTAAGGCTAAGCTGTACAGTAAT

>rs2272525

GGACCCTGGTCAGCCTTAACACAGGCTGTGATCATCTCAAACCTTTGAG
ACCTTTATCCATAAGAAAAATAAGAGGAAATTTATTTCAAAGTAGCATGA
AGTGATATGCCAACTCACTTGAACAGCATGACTATTTTCTGGTTCTGTTT
TGAGTTACTTTTACGTGGAGGCCCATCTGCAAAATAGGACAAAGGAA
[A/G]
ACTCACTGAGACCAAGTCTGGGAATTCTGAACCCAGACCACTTTCTTCC
TCCTTCCCTCAAACACACTCACTCTCTCACTCTCTTCTCTTCTCTCT
CTCTTTTTCTCTTTCTTCTCTCTCTCTCTCAGGTCCTGGGGCACTC
TGGACACAGTCTGAAAACATTGAGGTACAGCAAGTACCACAGAATGGAAT

>rs4396687

TTTCCAAGGACTTAAGCAAACAGAGTCAACTCCTCTGTTTTCTTGTG
GGTATATTTTGAAGATGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT
CGTTGTTTATCCACCTTTTCTCTTCTTCTTCAAAGCTCTAGACCTGCCTCC
CTGCCGTGTTCTTAGCCCACTTTTGAAGAAAGGACTTTCCAACCCCATCCTC
[C/G]
AAAGGGCAGGGTCTCTCTCCAGGGCTCCAACAGAGTTCTGAAAGAACAC
GGCTTCCACCATTTGTGTCTTCTGTTTTGTCAACTGCTGAAAGACTGAC
ACAGTAGACAATGCTGCAGACCCCTTGGGAGATAGTCCAAGTTTAGGGTG
GGGAGGGAATGGTGCATTCTGGGTGTCCACATATTTGTTGTTGCTTAGGA

>rs12614767

ACAAGAAAAGGGAGTGAAGTGTGTGGGGTTTTTAAAGTGGTGGTGGGGG
GCATTAGGAGATACCCAGAAGCTCTCAAGATTCAAACAGTCTGCTTTTC
TTTTAGACATTACACTCCAAGAAAAGAACTCTAGTGTATAACAGAGGG
GCTCTGTCAAAATTTAACAACATAAATATTGGGCAAAATGAAGTTGTTTCT
[C/T]
TGGTGTATCTCAACCTGCCAATGGAGAAATATCTCAAAAAGTCAAGAA
TTGAAGCCATTACTCAGAGACAGCATCCTATGGCATGGTGATCCAAAACA
AGTACACCTCTACCTTTTCCCTTTTCAAGCCCAACCCCAAAACACTGGAG
AATATTGGGGCAGGCAAGGCCCGTGAATAGGTGAGCAGAGCTGGAAGAAGA

>rs4571035

AAAGGCTTTGCCAGTAGTGCAGCCCTCTCCTCCCAAAGTCTAGAGGCCAA
 TTAAGCTGAAATATGGAAGAAAGAGGCACATCTGAAAGCCCAGTGCTGTC
 TTCTTGTCTCACATCCCTCAGGATCTCAAGGTCAGCCTAGTAGAGGCAT
 TGCTTTACCATAGTAGCAACTCAGGGAATATCTGTGGAAGTGGATTCTC
 [C/T]
 GAAGTTGGACAGTTGCTTTATCCCTGCTAGAGGCTCTCACTACTTAGAC
 CCACACAGTCCCTATAGATAGAAGAAACCTAGAGAGAACCACCTCTCACAC
 TTTTATTTTATACATGGAGAACCAAGGCCTCAAATGGTGGAGTGATTG
 CTGGGTGCTACAGTCATTGTCACAACTTACATTGTGACATGGGCTTCATA

>rs10195963

GCTGCAAGTCAAGTTGACATTCAACTATTTCTCGGGGTCCTTTTTTGG
 GAGAAGCAAGGCTTTGAGTTCACTCTTTGACTTCAAGTGAAGTCCATTTT
 CCTCTTAGACACTACCCAGTGGTTCTCACACTCAAGGAAGCACCAAAGT
 CCCTGGAGGACCCGTCTAAATGAAGGATGCTAAGCCCCACCTCACAGAGA
 [G/T]
 TGATGCAGTACGTCCCAGGAATCTGCATGTTAAAGGAGCTCTGCAGGTG
 ATTCTGATGGAGGTGGCCCTTGCAGGATGCTTTGAGAACTATATGTTCT
 ATCCCTCGAACTCTGCCTGCAGAACAGCCTGAGATGGCTCTTCCATGCTT
 TGACCATTGGCCCTCCACAGTAGTCCCTGAGCTGCGTCTTCTCGGACCA

>rs10207736

TGTCTAACCCCTTTTCTGTTTATAATGAATGCCATTGCTTGTGTAT
 AGTTGCCTACATTAATAGACAACAATGCAATATTCTCCTTATGGAACCC
 TACCTAGATTGGGAGGCACCATGAAGCCACTTTTTCTTCTTCTCCATCT
 CTACCCGATTAGCAAACATCTCTAATTATTGTGACTCCCAACCCAGTC
 [A/T]
 GATCATTTTATTCTTTTCCAGTATCTATGCAGGGAAGCCTGGCAAAGG
 CTTAAGGGATGGGGTGGGGTGTAGCATGTGAGCAGGATTTGGAAGTATGG
 GCCCCATACTTTTGGCACCTGCTGCATTTTCTCCCCACCTGCTGCACCAT
 AATTCATCTGCTTGTATGAGTTTAGCAGAAAGAGCCCCCAGTCCCTGTAA

>rs12464018

GGGAAATTTTCAGTATACATAGAAGGCTTAATTCAAATGATATAATGTCA
 ACATTTGGTATAAGAAATTATTTATCTGGCAGAATATATACTCACAGATG
 TCAAAATGGGTCTAGAGAGTAGAGAATGGAGGCAATTATCCAGACACTGA
 GGCCCAAGGCACAATACTTTTAAGTAATCATGCCCTCAACTGGCATTTC
 [A/G]
 TAGTCTTAAGCCCAAAGCATTGGGTGGAACCATCTATAAGCAAAGCC
 TCAGCTATGAAAACAACCTTTTAAATTTGACATTCAAAAAGATAAGATCAAC
 AACAGGTGCCCTTTGCTTTCCGGCAGTAGGTGTGTTCTTGAAAAGTTGGT
 GTCAATATTTCTTTTCCCATTAATCTAATCACATTCTAAATATAACAGGG

>rs10932689

GGCTCGAATTATGCTTGCAGTGAACAAGGCTAGTCTAAGGTATGCAATTG
 CCCTGCTTCTTCACTACTTTCAAGTGTAACAAATATGTCAGCATTATCC
 CTAGAAGGAGGGTTTAGGGGTTTTAAGTTGAATACTCTTCAAGTGAAGAA
 AATCCATTAATCTCCAGTTTTATTCTTGAGTTAGCTTTCAAATCTACATT
 [G/T]
 CGACAATAATTACTTTGGGATCTTTGAATTTCTTTTATGCTTTCTGT
 TCTATGAAGCCACATGCGTAGTTTATGCATTTTCCAATTTAGGCTAGAA
 GTCCAATGTGCTGTCCATTGCACCACAGAGCCTGTTTGTCTATGCATTTT
 AAGAGTAAATTTTCAAGAAATACTTAAATTAAGAAAGATAGATTTGGAAAG

>rs12613030

TTAAATTAAGAAAGATAGATTTGGAAAGCATATCCAAATGTTGCTTTGAA

CTTCTTATTCTGGACAATTCCTGGACTGGGGAGAAGGCATGCCCAAGTAC
 ACATTGCTTGCAGTGGCACCTCCAAACCAATGCTGGTGGGACTCTTCCT
 ATTACCTCGCAACAGCAATGACTGGGCCCCATAATACAGCCACGGTCATT
 [A/G]
 TAAAAGATCATTACAAAGCATCAAATTAAGTCTATTAATGAAAGCCAG
 AAACACGTCATTAGGAGGTGCAATGGCTGGGATGGAACATACAGGGTTG
 AATAATGGTAGGAGCCTGGGTTTGAATTAGAAAAGTGGGACAGGAATCC
 TACTTCTGCTATGAATCCATTGCAAGTGAAGTGGGCAAGTTGTTCTCACTC

>rs2888450

CAGATCTCAATCCGTCTTATCCTACCCTCTTCCTGACCTCTAGCTACAAG
 AACAGCAAGGCACAGCCCCCTGGTTTTGAGAATGCTAATGCTTGCCAGCA
 TTCTTGTCTGCAGGGCATGAAGTATAGTCTGCACAACCAGCAGGCTCT
 GTGACTGCCTAACCTCATCCTTGGTCACTGACCTCTGTCTTAGTAGCAGT
 [A/T]
 CGTGCCATCCTACCTAGCTTTGCTGGGTCAAGTGC GTTCACATCTTAAA
 ACCGTTTGAAGCGTTTGTCTAATTAACATGCAGAAAGCATATACTGATTT
 TTATTTTACCACCATGTATGTGAGGGCCTATTTTATGGCATGCAAAGCA
 TCTCATTTTTATTTCAATTCTCTGTCTTGAACTTTAAATGAGGCTGAAAT

>rs4254482

TTTTCCGGTTTGTGACTTACCTTTTAATTTTGCTCATAAATTACTTTTAA
 CTTACAGATAGTAATATTTTTGAAAGGTTACATTCATCAGTCTTTGTATT
 GTTAAAAAGAGTCTTTATTGTTATACTTTTAAAGTAGCTTTTCCACTCCT
 AGTCAGATACTTTACATATAATAGTTGTCTACTTTCTCCTGAATTAAT
 [A/G]
 GGGATTTATTTTGGCATATGGTGTGAAGTAAGAAGCTAAATCTTTTTTC
 TTTTCTCTGCATCATCAATGATCTCTTACACTGTTAAATAAACTAGTCTT
 TTCCTTCTGCTTTCCACACCCATTTAACCTAGCCCACTGTTGACATCGTC
 AGAGAAAACCTTTTATCAGAAAAGTCTGAACAGCACAAATTACTGTATCAA

>rs2372932

TTACCTTTTAATTTTGCTCATAAATTACTTTTAACTTACAGATAGTAATA
 TTTTGAAGGTTACATTCATCAGTCTTTGTATTGTTAAAAAGAGTCTTT
 ATTGTTATACTTTTAAAGTAGCTTTTCCACTCCTAGTCAGATACTTTTCA
 ATATAATAGTTGTCTACTTTCTCCTGAATTAATGGGGATTTATTTGGC
 [A/G]
 TATGGTGTGAAGTAAGAAGCTAAATCTTTTTCTTTCTCTGCATCATC
 AATGATCTCTTACACTGTTAAATAAACTAGTCTTTTCTTCTGCTTTCCA
 CACCCATTTAACCTAGCCCACTGTTGACATCGTCAGAGAAAACCTTTTCA
 CAGAAAGTCTGAACAGCACAAATTACTGTATCAATCCAGCTTAAATTACA

>rs4255939

TGATCATCGATATACATGATGTCAAAAATTTGGGTATAACTGAAGAATTT
 TAGTAATTTAAAGATTTTCAAGATTTAAATAATGAAATATCTGATACATTC
 AATTGGAAATTTGGTTGAGATCAGCTGACTTGTTAATTGGACTTCACTGA
 TAAACATTTGCCTTAAATCTGTAGAATCTAAAGCAGGGTATACCTAGTA
 [G/T]
 GCGAGTAAGACATCATTTCTGGAATCAAGGGGGCTTTTCAGTGTTGACC
 AGATGTTTATCACTGCATGGGCAAATGGCCAAGGAGTATGGTTGTGGCT
 TGGCTTATTTCTTTAACGCAGATGCAGTTGCTTTCTTTCTGGGTGGCA
 GTGTGGAGGATAAAGTGGGGTGGGTGGGACATAAAGTCTCACAAGGGGAGG

>rs10192415

TTCTTTATTGTATCCAGTATAACATACTCGAAGATGTATATATACACATA
 TATAGATATACAAAGCATCTATACATCTAGACATGTATATATACATATAT
 ATATACACACACGTATATACTGTACTTCAATCATCTATATTAAATTGTGA
 TAATATCACATTACCACATGTTGATGGCATTCTTGAATCTGCCTGAAAT

[C/T]

CATGAAAACACATGCAGTGGTCTGGCTCATGTTGTCTCAGGGAGTGCAT
TCTGGCTCCCAGTGATCATCAATTACTTTTATAAAAGTTCACCAATCACT
TGTGGTTTCCGAAAAACAAATTTAAAGCAAACATTATTAATGTCAAAC
TACACATTTTGTCTTCTGGAATCAACTTTTTTATTTAAAAATTGAGAATA

>rs7595393

CAGGGTCCACGGTGTCTTCTGTCCCTCCCTGCCAGCCTCAGCAGGAG
CCCCCTGCTCTAGACCACTCCCTTACTGAGTCTGTAGATGTTCCAGCA
CACACAGTGCCTCCCTTCTCAAGGCCCCAGCCCCAGCCCCCTCTCAGGTC
ATTTGGGGGCTTACATCACCTGCTACCTGTGTGGAGATGGAGGTCTGTT
[C/G]
CCAGCACTGCCTTGCTTTTCTCTGCCCCACTGCATTTAATCACTGTG
ATAAGAGGAAACAAAGAGGGAACGTCAATTGGTCCCATTCCTCTGAATCTT
GACCACGAAATATCTTGAATTGCCTCCGTGTCCCACCCAACATCGCCT
CTTGCACTGCTACCTCTGCCTGAGTGCACATCAGCCAAGAGCCAGGAACCA

>rs2372935

CCCAGCCTAATGTGTACTTTTAAAAGCACTTTTCAAATACTTCATCTTGT
TTAATCTTCTTGGCAGTTCTGTAAAGTGTGAAGCATTGTACAAAGTTAAC
TCTCATTATTATATTGGAACCCAGGGTTTTAGAGGAGAATCCTCAAGGA
GATAGGGCTTCTGCCTGTGCCATATAAGGTTGTGGATCCCAAGGGA
[C/G]
GAAAGCATTTTCATCAGACTGGGAGAGGCAAGAAAAGGTCACTGGCAGT
ATAGATTCTTCTGGAGTTCACTGGCAACTTGTCCACTTCAGTTCTCCCT
TAATCTTCCAGATGTGCAGATGTATAGCCTGGGTGGATGAGCCAAAATG
AATTCAAGGCTAAACAGGTCAGCTGGCTCAGCCCATGGGAAGCATGCTCAT

>rs6733648

TGGAGAAACCAAGAGGTTGATTTCTCAGCCTTCAAGAGAGCTCATCCTC
AGCTCTCTAAAGCTGACCAGAAAGAGGAGAGTGGGACCTATGAAAACAGC
AGGAGATGGGAGGGCTTGAATTCCTAATGCTATACCCATAGGTCAATTGTG
CCATGGTCTCTGCACCCCCAACCCCCAACACTGATCCTGAAGGTCTC
[A/C]
ACAATCTTCATCTGCCTCTCTCCTTTCACCTTCTTTACACTCTCATTGA
TAGCCTTGGTTTTTCTATACCTCCAGTATTCCTATCCCATGATAGCCA
TTTAAGGACAGGGCACTTCTGGTCACCATAGGAATATTGTGGGAATGT
TCTGCTCATGGATATTGCATCAGAATAGGTCTGGAAATGAAACACACCAAC

>rs6435957

GAGACAGATGGAGAGTGAGAGAAGGGAGGCAGGTAGGTAGGCAGAAGAAT
GAGGAAGAAAGCAAAGGAGTTGAGACTGGGAATGGACAGAAGGAAACCTG
TACATGCTGTGATCTGGGCTTCTATGTTCCCTGATCCAGGGTCTTGAGA
TCTTGGGGTACTCGCCACTGCATTGTAAGTTTCTGTGGATGAAGCTGGCA
[C/T]
ATAAATACATCTGTAAAGCTGTAAGTCTGCGGCAATTGTCCCTGAAATTT
TACTGATGTCTAGATTCTAGGTATGGGATTAAATGGGTCCACATTTTGTG
TTTTTCTGAGTACCAATTACAACAAAACCTTGCTCATATACTTATGCCTA
TCTTAAAAATAATTTATTTTAAATAAATAATAAAGGCAAGAAAAAAG

>rs6723019

TCATTATCTCCAGAAAGCTTTTCCAGCTTCCAGTCACTGCTTCACTCCAG
TTTTCTTTCTTTTCTTTTCTATCTTGGCTTATTTTAGTTGGCTTCTATGCC
CATGACTATCACCTCAACCAAGACAGAGAGGCTCCAATGCAGGAGGCA
GACAGATTGGGCTTGAACCAATGCCGCTTCTTGCTGTGTGAGCTTTG
[G/T]
CAAGTTATTCTATTGCCTAATTTTGGGGATTGTATTATGTTATGTTAC
ATGTTATGTTTTGTTACATGATATTGTGTTATATTACTATCTTCTTCATG

GTGTTATTATATTATTAAGATTGATGAGATAATGTGTGTCAGATTTCTAG
CAGGTTTTCTTGGCATATTGCAACCACCAATACACTGAACATATACTGTCTA

>rs12614773

CCACATGTTTCTCATCCTCCTCCCAGGATTAGTGGGCAGCCCAAGCATGT
CTTTCTTCAAAATGATGCTGGATACATAAGAGAGCAAGCAAAAACATACA
TGGCTTCTTGAAGCCCAGGCTTGAATTGGTATGCCATTGCTTTTACTGG
CCAAGCAAGTCACATGGGTGGCATCAGGGATGGGCTTATGCATCATCTAC
[C/T]
GTGGGAGGAGTTTTCTAAGTATGATGAGGGCACAGATACAGGGAGCGGT
GACCATTACGAACATTAATGCGAATCCTCCACACTGTGTTTCTTGATCTG
CCAGGACAGGGTTCTCTGAAGGCCAGTAGAGGTTCAAGTCTTCCAAAAA
GGATGCTGGGCTCAGGTAGGTAGACAGAGTTCATCACCTCTATCATGACTG

>rs17833842

TAACCTATAAATTGACTACAATCCTAATAATCACATTAATAATTTTTTATA
GTGAAATTGATACTGAAGTTATATGGAAAAAATAAGAATAACTAGAGAAA
TATTAAGAAATTAAGAGAGCTACAATGAAAGATTAAGTCTGCCTAAAT
TAAAAGGCATTCTAAAGTTTCTGTGATTGAAGTAGTAAGAACCTTGTGC
[A/G]
TAAGTGGAAAGACCAACTAATGAAATAAATATGATCTGAAAGAGACCC
ATATGCATGTGAAATTTAGTATATAATAAATGCAAGTTCTAATCACT
TGATAGTCATTCTTTAAAGAGATACAATGAAATCCATACCTTAAACCAT
ATAAAGAATAAACCTCAAATGTATGAAAGATAGTATATATAGAAAAATAA

>rs10177578

TGACTAGGTAAAGTTTAGAAAACAACAAATTTGAATGTTCAAAAATATTAT
CAGCAGTAAAGTCAAAAGCCAAAGTGACGAACTGGTAGAAAATGTTTCTAT
ATATATCAGAGATAAAAGTTAATGTCCCAAGTATAAAGAATTTTTGA
AATTGAGAAATAAGCGGAAAACCAAAACCAATATAAATGGGCAAAAGATA
[C/T]
GTTTCATATAATTTAAGTGGTCCCTAAATATATAAATATCTTCAACTT
TCAATATCCTTAAACATATAAACATACATTCAAATATGTTTCTTGTAAAT
AAGAGGAATGAAATTAATATAATGAGATACCATTTTACCTATCAGA
ATGGTAAACAATTTTTTTAACTCTTAACACACTGTTGACAAAGCTGTG

>rs12623304

AAAAAAAACATGAAAAATGATGGCGATATATCAAAAAGACTCAGGAT
CCAGTGTGAGCAGCTCTTATTGGGCAAGTCTTGAAATTTGAACAACAAA
ATCAACAACGGTTAGTTACACATTAACATGTTTGACAAAATAAGAATCCA
TGAGTCCAGACTGATATAAATTAATGAGTGAATAAATGAATAGGTGATTG
[A/G]
GAAGAAAGGAAAGCTTCTCTTCACAATAGAATGCCAATTAACAAATACA
AAAGGAACAATGGAGTTAGAAAAGAGTTAGAACACTACCATTTTGCAGCA
ATTACAATAATAATTAGTTTAGGTAAGAATCATGAATGGGTGATACATTT
ACGGAGACTGTTTGATGAGGAACTGGATATTACATAGCCTCACAGTGTTTC

>rs6716542

AGAACACCTTCCGCACAAGCTCTTCAAATATCTCAAGTTCTCTTTTCCAG
TATGGTGTGAGGGGTGTAGTGAAAAATGCATGAGTTTGGAGTCAAACAG
ATTCCCGTTCCATTTCTGGTTGTATCCTATATTACCACTTGGAGTTGTTG
GCAGTCTTAGGTGAACACAGGAGGCAATCTTTTTTTACAGTGACTGACA
[A/G]
AGAGGTGAGCAAGAAATGTGAGTCTCTTATTTATCCAGCCCCCTCTAT
CTTAACCTGTTGAAGGTAAAGACCAAGTGGATCTGAATTTAGATTATTTAC
CCTCCCATGTGTGGGGAACACACACACGTGCACACACACACCAGGCTGGA
ATTTCTCAGTTTGTCTCCGGGTCTACTCTATACTTCATGGAAGCTGAT

>rs17777330

AGTCAGCAACATGTCTCCAATTAACATGACAGATTTTTGGACCACAGAT
 CTCTTTGGGAATCTTATGTAAGCTGTGAACTTCTCCCTAAAAAACCTA
 CATATTTAATTCTGTCTGTAATTTAGGACTTCACATATCCCCAAGACC
 ATTCATGATCCACCAGGGTTAAAGACTTGTTTAAGTAAAGAAAATATTC
 [A/G]
 CAGCCTCTGGAGTAGTAGACATCTCCAGAGAAAGTCTGGCATTCTTTGG
 TAACTCATGGCATTATTACATTAGCTTTGGATGGAACAACCCCAACCC
 CCAAGTCTGAGCATGTGGTTTGATGCTGCTTCTTTCCATTTTCACTCTG
 CTCTCTGGATTCTATGTAGCAAAAAGGTACCTTTCTAGGGTCACATGTG

>rs10211546

ATATTTCTAGGGTAGTTAGGTAACATAGAAGCTTTGTAGCTCAGCAGTCT
 TGGAACTCACTAGACTCTAAGCCCCATGTAGTCAGAGACTGCCTGCTTTTA
 TAATTCATCCCCAACAGCAGGCTATTGCATAACACATTGTTACCACCCAC
 ACCATTTTGTGAGTAAATGAATGCATGCATGAACGAATCAGAGGATTCT
 [A/G]
 TTGTCTAAATAAACTTTGGAGGTGACTTAGTCTTCTCTTCCCTCTGATG
 CTTAGATCTTCTTTATCTCTCAGTTCAACATATATTGCTCAAATGGTTCC
 TAGGTGCAAGGGATCTGGAAGTTATTAATGAAATAGATTTTGGACCAT
 GACCTTGACAAATTATGATCTACTAGGAGAGATAATATGGAAATCAGTATA

>rs2372937

GAAGTTATTAATGAAATAGATTTTGGACCATGACCTTGACAAATTATG
 ATCTACTAGGAGAGATAATATGGAAATCAGTATAATACTACAGAGACAGT
 AAAAAAGGAAAAAGAAATGTTTGCACTTAATAATTATCCACCATTTGTTCC
 AAAAGTTCAATTATAGGGAATAAGTTCCACTACTAATAAAAGTCAATA
 [C/T]
 GCATTGTATGCTTACTGCATGCTAAATATATTATGTCATTATAATATTT
 ATGTTCAAGGTGTATACTATGATTGTATGTTTACGACAAGGAAATTGAG
 GGCTCAGAGAGGTTAAAGAACTTGCTCAAGGTGACAGAGCCAGTAGGTGG
 TAGAACAAGACTCCTACATTCTGACCTAGTCCATATGTAAGCAGATTTAT

>rs4315498

TTCAATTATAGGGAACTAAGTTCCACTACTAATAAAAGTCAATACGCATT
 GTATGCTTACTGCATGCTAAATATATTATGTCATTATAATATTTATGTTT
 AGGTGTATACTATGATTGTATGTTTACGACAAGGAAATTGAGGCTCA
 GAGAGGTTAAAGAACTTGCTCAAGGTGACAGAGCCAGTAGGTGGTAGAAC
 [A/G]
 AAGACTCCTACATTCTGACCTAGTCCATATGTAAGCAGATTTATTGCTA
 GAAATATCTTGATGTTGCATTGAAATCTGCCTCTCCTTATTTTCCATTTG
 GAACTAATCCACTCTCCCAGCCAGAAAAACCAAGCCTCTCATTACTTG
 AAAAACCAATTAATGAGTTAAATTACATCAAGTTAGTTTAAACAGTCAGCAA

>rs12998806

TAGGTAAGTCTTCACAGAGAACTGGTATTAGGGCTAAGATCTAGCGGAG
 GACTAGGGGCTTTCTAGACAGACAATGGAGGCAGGACACTTCTGAGAGAG
 GAATGAAGCCTTGAGGTGTGAGACAGCATGTGGTAGAGCATGGAGGGTTC
 AGGCACTTGGAGGAGTTGGGTATGGGGAGAAATACAGAGCACCTGTGAGA
 [A/G]
 CCTGGGGCATGGCTGGAGAGGTAGGCAGGGGCCGGCTTGCAAAGGGCTT
 TACCTGTCATGCTGAAGGGCTAACTGCTGTTCTGGAGAATTTGCAGAGCC
 ACTGAAGGTTTTGAAGCATCCTAATGTTAGGATTCATTCAAGTTTTAGA
 AAGTTTAGTCTTTTGGCTCAAGGGTGGGGAGGGAATAAAGAAGCGAGAGT

>rs12465515

CACCTACTCAGGAGGCTGAAGCGGGAGAATCGCTTGAACCTGGGAGGCGG
 AGGTTGCAGTGAGCTGAGTTGGCACCCTGCACCCAGCCTGGGCAAAAA

GAGTGAAACTCTGTCTCATAATAAATAAATAGAGTGGACAAGGAAATACT
 GATGGGGGTGAGAAAAACAATTTGAGAGACATTTGAGAAGTAGTGAGATAA
 [C/G]
 ATTTGGTTACTGATTGAATATGGAGTGTTAGGATAACCCTGGGATCCTC
 ATTTTGGGCAATCACAAAGCTACTGGCTACAGCAAAAGGAGGGGAAGAAA
 AGTGGTGGTGGGAGGGAGGCATGAGAAATTTGCTGCTGGCCATGTTAAGC
 TCCACCTCCTGTCACCAAGTGGTGATGCTCAGGAGATGTTTGTGTCCACAG

>rs4491709

ACCTCCTGTCACCAAGTGGTGATGCTCAGGAGATGTTTGTGTCCACAGGA
 GGAGAGTGATCTGGGTGGGAGGTTTACATGTCCCTCCTTCCATCTCCCGG
 TCCITCTTTCCCTCCTCCCTCTTTCTTTGTTTGGTACTTCCTTTCTTCTC
 TCTTTCTTTTTCTGCCTTCTTTGAGAATGAAACCTAAGAGTGTTTATA
 [C/T]
 GTAGAGGGAAAAAGTGGAGCTAGAAGTTTAAGACACAGGAAAGGGAGGA
 AAAATGTCTAGTGTGAGATTCCAAGGAGGAAGAAGGGCCAGGATAAGGT
 GGGGAGAGGGATGCCTCTTTGCTCATTGCCTTTATCTGTCTTTGTCTGT
 CCTGGATTGATTGTCCTATCCTTTCTTCCACTTCGTGTCTTCTCCCGTAT

>rs2372938

GGATTGATTGTCTATCCTTTCTTCCACTTCGTGTCTTCTCCCGTATTT
 CTACACAAGGCACAGTCAAATAAGGGCTGGCACGTGCGCATCTTTGTGCA
 CACATGTGTGTGCACAAGCATTGAGGGATGCACAGGTGCACACACCTTGT
 AGAGCCTCTAGTAGCTCAAGTCACATTTTTTCGATTTCCAGATTGTAAG
 [C/G]
 CTTACTGAGGGGCAGATGGGACGTTCTCACAGGAAAAAGGCTGACCTCA
 GGGAGAAAGATGTGGCCCCACCTACAAACCAATGATGTCTACTCCCAAC
 AGGGGACGTCTTTCCCTGCGTGAAGCAACAGGACAGATTTGGTTAGGTG
 TGGCATCGATGTCCTAATCTCCTATGAGCACACAATGAATGAAGGTGGGGG

>rs12621130

GATGGGGCTCTGACCCAGTTTCCATTTGGCCAGCCGACTGGAAACAGCTA
 CTCCTGGGCACTGGGTTCTTTCAAAGAACAGTGAAGAACAGTGGTTT
 GGGGTAGGTGGGGGCAGGGGCTGGTGGGGCAAGAATGTCAGAAGCAAATG
 CAGGCGTGGCTGGTCCAGTAGTTCCAGTACCAAAGAAGCCGCTATCTGGA
 [A/G]
 GAGTGGGACAACCCATCCTAAACATGGGAAGAAGAAACAGGAAATTTCA
 AAAGATACAACATAATTAATGAAAGCACTCCTTAACACCCCTTTGA
 TAATTAGATCATCTTCTCCTAGAATGAGAAGGAATATTAGTCTGACTTTG
 CCTCCATATGGAGAAACATGGCAGGCAGGCTCTCTGTGCCTGGAGGGTCT

>rs13399995

GAAGGACACTGCAGCCAAGCTGAGCCTGTGTCCAGCTGAACACGGACATT
 TACTTCTGCTTCTGAGGGGAAATATTGTCATGTCTTATGGGTGTACCCAG
 CAAAAGATGTTTCCACTAAAAGACTATAGAGAATTGCTTATTTCTAAGGA
 CTGAATTCGAATGTTTTTGGTCAGGTTCTCTCTGAACAGAGAACCATTG
 [C/T]
 CCCTGAGCAGAAGCGTCTTACACAACACTAGCAACCATGTAGTGTTTTCT
 TCATTCAAATGGCGTTTGGTATTACCTGCCTGTCCTGGAACAAATGCAA
 AGTCAGGGTTGTTTCTAGCTCTCCCCAATATTGAAAAATGCTGATGAA
 CTGAGCCTGGTTCCATTTTACACCTGACAAAAGGAACTCAGATGAGAGTGT

>rs12052807

TGCTCTCATGTGTCTCCAGGAGAGTGTGGCCTGGGTTAGCTTAAGACTAG
 GAATGTGCTGACAGCTGGGAAGAAAGTCAAAGAAGGCATGGCATGTGGT
 TATGGGGCCGGAGGGACCGCTTGGGGGAGAGGTGGAAGCGCTGACATGA
 GCTTGCTCTGCATCGCTGGTCCCAGAGGTGGCCTGATAGGAGAGTGCCA
 [A/G]

CTGGTGGTATGCCGTGGGCTCTGGGAAGGGGCTAGAATGTTGTTGCAG
 CCCAGAGGGGCTGGGCTAGTGAGGAGCAAAAGAAAACATAAGCTGGGGGG
 AAAGTGCCTAATAGACCAATTAGGAGACAGGGTTGTGCCCTAAGGGATGG
 GTGGGAGCAGGGGGTCTAGAAGTAGACAGAGAGTAACTCTCAGCTCTGAA

>rs10199394

GCAAACTCATCTCATTTGCTACTTGTCTATAGGTAAGGATGGACACCAA
 ACCCAAAGAGACTTTTGGGATGTGGATAAAAATTGAAAACCTTGCCTCTC
 ATTAATCCCTTGGAGAATATAGCATCTCCTTCAAAGTGGACTGCTCTCCT
 TGATCCCTTGTCTTTACCTGAAAATATCTTTCAAGGCCTAGCTCATGCTA
 [C/T]
 CTTGGCCATAGGAACCTTCTTATCACCTCCATCAATGCTGTTTGTCCC
 TCCACTAGAATCTTCTGTTTGGTACTCCTCTTACTCAGCCTCAAGTTAA
 ATTTATGTGTACTGGTTTTAACCCAGTACCCAGTAAAGGTTCTGAATAC
 ACGTGATGAAATAAGTCTCTCCTATCCCATGGCACCTAGTATAGCATTTAG

>rs2372940

AAGTTAAATTTATGTGTACTGGTTTTAACCCAGTACCCAGTAAAGGTTT
 TGAATACACGTGATGAAATAAGTCTCTCCTATCCCATGGCACCTAGTATA
 GCATTTAGTTGCTATACTAATTATTATAATTAGCATTTATTACGGGCTTA
 CTATGTGCAAGTGCCATTATTCCATTATTATTACAAAAATATTTAGT
 [C/T]
 TCTACTATTCTCTAATTTATTAGATAACCTAATATAAATCTTTTTTGT
 TTGATTGTTGCTAGCTGATTGAAGGAAATAGAATTGTGGTGTGGTAGTAG
 GGGATGTTGAGTGAGTCTTCTCATTCTTAATATTTCCTTCCACTAGGAT
 AACGCCCTGGGCTCTTTCATCAGGGGAGGAGGAGTGACTTCATCTGATA

>rs4583440

CCTAATATAAATCTTTTTGTTTTGATTGTTGCTAGCTGATTGAAGGAAA
 TAGAATTGTGGTGTGGTAGTAGGGGATGTTGAGTGAGTCTTCTCATTCT
 TAATATTTCTTCCACTAGGATAACGCCCTGGGCTCTCTTCATCAGGGG
 AGGAGGAGTGACTTCATCTGATAGATTGGAAGGGGAAATGTTACCCAGT
 [C/G]
 GCTTCTCAGTCTCTTCTCTTGGAGCGAGCCCATCTCTGGAGCACATA
 CACTCTCCCCTGCTCCTGTAGGTCTCGAGTTCTCAGGGGAAGGGTTGCC
 CTGTAAGAGAGCTTGTCTTCTCCAGTGAGGCCTGTCACTGGGAAGATAA
 TTCTGTACAGTGCTGGGATTCAAATTTAGAAAGCCCACTTGCCCCATAATT

>rs2372941

TGTGGTGTGGTAGTAGGGGATGTTGAGTGAGTCTTCTCATTCTTAATAT
 TTCTTCCACTAGGATAACGCCCTGGGCTCTCTTCATCAGGGGAGGAGG
 AGTGACTTCATCTGATAGATTGGAAGGGGAAATGTTACCCAGTCGCTTC
 TCAGTCTCTTTCTCTTGGAGCGAGCCCATCTCTGGAGCACATACACTCT
 [C/G]
 CCCTGCTCCTGTAGGTCTCGAGTTCTCAGGGGAAGGGTTTGCCTGTAA
 GAGAGCTTGTCTTCTCCAGTGAGGCCTGTCACTGGGAAGATAATTCTGT
 ACAGTGCTGGGATTCAAATTTAGAAAGCCCACTTGCCCCATAATTAAGCC
 ACATTCACCAGCTTTACAATGACAGAGGTGATATTGAGTTCTCCAGTGAT

>rs6721811

TACCAGGTCCTTCATTATCATATTCTACACATTGAAAATGAAGGTAAGAG
 AGCAAGCAACTTGCCCTAGGCTACCCAGGGAATAGGGAGGAAAGTCTGGA
 TTTGAACCTACTGCTGCCTGACTCCGAAGCCCTGTTCTTTCTCTCACCCT
 GGAGGATGGATGAGGGTGGGATGGGGCAGTTAAAAGTTGAGGGGTATCTC
 [C/T]
 AGGAGAAAGATGAGACACATGACCATTGGGTCTAAGGAGCCAGAGAGAG
 GCCATTGTGTCTGAGCCTGGACAGCCAAGCTTCTAGAGACAGAGGATGG
 GATGAGATGACTTCTGGCTGACTCTGCCTCCTTGAGCTCTTAGGATCCC

CTGCAGAATCTCCCAAATGATGCTCATTGGAAACAGAGAGAGCACATTCA

>rs12615418

TAGAGATGGGGTTTCTCCATGTTGGTCAGCCTGTTCTCAAACCTCCCGACC
TCAGATGATCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGTGT
GAGCCACCACACCCGGCCAAATTTTTTTTTTTTAAAAATATGGCATTAAAA
TGGTATTTATCTTGACTACTGAGTTTTTGGTTAACCCCTTACATTTTGTG
[C/T]
CCAAGACATCTCACTCATGTTATCCTAGGCCCAGCATGGCTGGCCTTGT
CTCTAAATTCAAATTTTTAAATTTTTTCATTCTCAGGCTTCTCACTTGT
AAAATAGAAATAATAATAGCATCTCCCCACAGGCCTGTGCTGAGAGTGA
AATGTGTTACTACACGTGCCTCACCTTCATTTTCAACAGGACCTATTGCCA

>rs12621884

CACACCCGGCCAAATTTTTTTTTTTTAAAAATATGGCATTAAAAATGGTATT
TATCTTGACTACTGAGTTTTTGGTTAACCCCTTACATTTTGTGCCAAAGA
CATCTCACTCATGTTATCCTAGGCCCAGCATGGCTGGCCTTGTCTCTTAA
ATTCAAATTTTTAAATTTTTTCATTCTCAGGCTTCTCACTGTAAATAG
[A/G]
AATAATAATAGCATCTCCCCACAGGCCTGTGCTGAGAGTGAAATGTGT
TACTACCGTGCCTCACCTTCATTTTCAACAGGACCTATTGCCAGTAAGT
GCTCGTTAATGTTAGCCATTATTCTTGGCATTGCTGTTGTTATTAGTGT
AGCCATCATTCTCTCAGCCAAGAGATCCTCCAGAAGAGAAGTTCTCCATG

>rs2372943

CTTGAGTCTCAATTTCTTCAAGTTGGGAACATTACATGACAGGACAGGAC
TGTGAGGATTAAATGTAAACAACATTTTTTTTTTTCTATAGAACAATACT
GTGGGGCCAGCACCATGCAAGGTGCTTTTATCTACAATTTATTTCTTTCC
CTTCCCATGCCTTCTGCCTCTTTGAATAACTGCTGCAATGGCTGTAGAGA
[A/G]
ATAAATGCAGTAGACCAAGACTTTTAGTGATTTTTAAACAATGTTCTAG
GGCAAATACATGGGTATGTTTACATACTCCACAAAGTGAAAAAATTATGC
TAGTCAGTGTGCAAAGTGTGTTTTCCAGACCTGCAAAAAGTGAATCAA
CTTCATATTTCCAACCTTCATATTCCCTTCATCCCTTAAGCTACCAGGCA

>rs4522583

ATTTTTAAACAATGTTCTAGGGCAAATACATGGGTATGTTTACATACTCC
ACAAAGTGAAAAAATTATGCTAGTCAGTGTGCAAAGTGTGTTTTCCAGA
CCTGCAAAAAGTGAATCAACTTCCATATTTCCAACCTTCATATTTCCCTT
CATCCCTTAAGCTACCAGGCAGAGGATGGGAGATAAAATTTGGCCATAAA
[C/T]
TTGTCATGGGTCCATCTTGGTCCTAATTAGAATTATCTACTTCCTAATC
CATTAACTGCATGGATAGATAGTCTGAGGGCCTCAAGGGCAAGAGATGAC
TTCTACCCCTCCACCATCCAGCCCTGACCCACCCTTCTCATATACATTG
GAGAAAGCTCTGCTCTGTGTAAAACACATCAAGTGCAGGACATGGTTCTTT

>rs2888451

GTAGAGTTTGGGAAATGAGTAACATGTAGATCTGTGAATAGATATGTCTG
AGGAAGCACCAGGCAGCGAAATAGCTGGTTGCAGGAGCAGTTTGTCTG
CCACCTGAATCTTCAGGTTCTTCAGCTAAGCCTTAGGAGAGGAGGCTG
GAGAGCTGCTTGTGGATCATCTCCTTTGGCCCCCTCAGACTAGAAT
[C/G]
CACCTCCTCTCTACCTAGCTCTGTGCTCCACAGGCTGCCTCTGGGCTGC
ACTGATGGGCTCCCTTGACTGCTGGCTCCTGTTTCAAGGTTGGAGGAGGAG
GATGAAGTCATGGAATAGTGTCCCATGCTGGGCTGCAGGTACAGTTGCT
GCTTCTCTACTGAAGACCCCTGACAGCAACCCTGTCTTACAGCCAGGGTGT

>rs2270398

TCAGAGCAGGAGACATCATTTCCCTTAATTCTATTCTGCCTTGTAACCC
 CTTTCTTAGACTCTCTGCTGAGTCTCTGCTTATGTGTTGGGATTGCATTG
 ATGCAGATGCTTTAGGGCCACGTTGTGGAGGATACAAGATGATGAGACAG
 GAGGCAGCGTATGGTGAATCACTTCTCACAGCTCCTAACTAAGTTCTGGG
 [C/T]
 ACATGAATGGGTACTCAGCACATACCTACTGACCAAGCCTATGGAGAGC
 TGAAAACCTCATAGGCTGACTTTGAAGCGAGGCCTTTGAAACCATCGTCCC
 AGGGGAAATAATCTAGGTCAAAGAGACGGCTGTGGCAAGGGTGGGAGCTG
 GCAAGCATGAGGGCTGACAAAGCAGGCAGCACTCCTCCTGAGCAGATGACC

>rs13412666

TATTCCTCAGGCTGGGCTGGAGCACCTGAGGCTACTATGGTTACAGAAAGG
 AACCTGGCATCCGAAGGTATAGGCTCAAATCCCAGCTCTCATGATTGCTA
 GCTTTGAGGTTTCAGGTTCACTTTCTTAAGTGCTCTGAGCTTTGGTGCCCT
 CCTCTTCAAATGGGAATGAGGAGCATGTCTAGCTCCCTGTTTTGTGCA
 [A/G]
 TGAATAGCTACATGAAGAACAGCTAAACCAGAACAGAAAGAAGGCAAAT
 GGAAGCTACAGAAACCAAGGATTTCTTGTGAATCGAATCTTCCTTCAA
 TCTTCTTCAACCACTAGTGGATCTCCCTGTGGGAGGGATGTTGAGAGT
 GCTCCGTGTTTTTTGTTTTTTTTTTTTTTTTTTGTGATGGAGTCTCA

>rs13426489

CAAGTGATTCTCCTGCCTCAGCTTCCCAAGTAGCCAGGACTACAGGCATA
 TGCCATCACACCCAGCTAATTTTTGTATTTTAGTAGAGACGGGGTTTCA
 CCATGTTGGCCAGGATGGTCTCGATTTCTTGACCTTGATCTGCGCACC
 TCTGCCTCCCAAAGTGCTGGGATTACAGGTGTTAGCCACCACACCAAGCT
 [G/T]
 CTCTGTGTGTTTTAATCAATAGCAATATGTACCATATTTAGTAGGGAC
 TATAAGAAATTTGTGCTAGAAACACCTGGTTATTTTACAGAACTTTGCCA
 GTGAGTACAAAAATATGCGTAGCCACAACTCTCTCCCTAGAAAAATAC
 CTCCTGTGGGTAGAATGAAATGGGTAGGTGTTCTTACCATTATTAGCATT

>rs6713249

ATTAACCCCTTCTTATACATTACCCAGTCTCAGCTAGTTCTTTATAGC
 AGTGTGAAAAACAGACTAACACACTCTTAAAGTAAATGTTATTACCCACC
 TTTTACAGATTAAGAACTTAATATTACATCAAGGAATTAGTTTATAGCA
 CTGGATGATCTGGCCTCAAATCTCCAGTGATGATCAGTATCAAGCTTTTCT
 [C/G]
 ACTGTATCAACAGAGTTTTCTAACAATCATAAGACACAAGTAGTTTC
 TCCAACCTTAATCTATTATACCCATAATTCACCTTGCCTCCAGGGAAAAAA
 TTAGATGAAACAGACTGTAGTGGCTTTGTGTAACAGCAGCTCCATTCC
 TTTAGCCCTATTCTAAACAATAGACCATGTGCTGATAGCTAGCTAACCTC

>rs12620095

GCAATTTGTTCCAAACACATTTTACATGGATTATCCCATTTAATCTTTACA
 CTAAAGTATAACTCTTATTACCCCATTTTAAAGAAGAGAAAACTGAGG
 CACAGTGAGATTAAGGAATTCATCCAGTTTCTGCTGCAAGTAGGGTGGCT
 AGCATTGTAATCTGGGCATTTGCCATGAAGCTGTGCCCTTAACCACTCCA
 [C/G]
 TATGTGCTATAAGGTGGAGCTGACTTCATGGGGGCTAATTCAACCAGGG
 GAGTGTGCTCATATTAACAGATCTGTGAACCTCCCTGGTTTGTGTTGTGAGA
 CAATTATCTATTTGCTTACCTGAATCACAGCAATAATAGTTATACCTCCC
 ACTTACTGCATGCCTACTTGGTAGCATGTACCTGGACAGAAGGTTTTTCAAT

>rs4621152

GGCGTGAGCCACCCCGGCTGGTCGCATTCTCTATTTAGTGTTAATACCC
 TAAGACTCCCCAGCTGACCACCAAATGTAGGCAGGTGGAGGGACTGACCT
 CCAGGGGGGCACATGTCTCAGTTGACAAACACATTTTCTTGTATCCTTT

TTTTCAGTATGATGAAGTATTCTACAAGACTCTCATTGAACCTTCTGATA
[C/T]
CACATTCTAACCATTATATATAAAAGTCAAAAGCTCCTTGTCTTCAGGAT
TCCAAAGTTCTTAGCACAAAGGTCCAGGAAGTGTCTGAATTCTTTGGTGT
CTGGTTCAAGTATGGATGTTCCGGTCCCTCTTAGACTTGAGATAGAACAGG
AAAGAGATGAATAGACAATGACAGAGGTAGAGAAAGTACCAACCTTCCCA

>rs12622764
CCTCTCATCTCCTGAATGGCTGACCAGTGCTTGCTTCAAGTTAGATTAGA
GTTTGAGCTTTGTAACATAAGTCCCCTTCTGTCTGATCAATGAAATGAGG
GATAGGTACTTGCCCGAGAGTTATGGGACCTGGCATTTAACGTGAGAAACA
TGATCTTCAGAAGCACTGGTGAGACTTCAGAATTCCTGAAAGCACTATT
[C/G]
TTGCAAGGCTTGAATGTACCTTCTCCACTCCACCATGTCCTCTCCCAAC
CCTACTAGCATGATATCAAGCTTGGGTGGAGAGCTCGCTTGCAACTTTGT
CAGATACCCATGAGCCACTCTTGAATTTCTGACAATGAACAAGTATGTG
GGAGGCCAACCCCATCCAACCTGACCCATCTGAACCCAATCACATGTTTTC

>rs6723013
AGAAGAGTCCCTTGTGACTCTGATTGGCTTTGTGACTTCCTTGGGGCTGG
TGTTGGTCTCAGGCTTGTGTCTCGCAGTCATCCAAAATGTGGGTGGGCCA
TGTCACCTTAACGTGTTGTGAACGCAGAGGATTCACTCAACTTTCA
GCATGCAACAATAATCAAATGCAGTTTTATGGTTCACAGAGGGCTCCACC
[G/T]
TCATGATTTCAATTTGATCCTCCTAACCAACCCTGGGAGCTCAGAGAGACT
AAGGAACCTGTCTCATATCACATACCCAGGAGTTGACTGCACGGCATCTG
ATGCCAGCTTCTGTACCACTGTCTCTGGTGATACTAGAGTAAATAAA
ACATAACTAGACAAGTAAATATTTGTACAAAGTAATCTGCTAAAAATGT

>rs2372945
GAGCATTTTCTCTTGGTAAGTTCTCATTTTCCCTTAATCATTAGAAGACA
CGGCACAGTAAACGCGAGCTCATGAGCAGAAGGCTACATAGTTGTATACA
TGTGTGTGCTTTTTTACACCAGAGTTTGATGAAGCAAAATACCCAAAG
GTCTTTTAAAGTAAAGAGGTTTTAAGAGTAAAGAAATTTAAGAAAAAA
[A/C]
CAAAAAGAAGTAAAGATGGTTAGGTCTCCTGTTGCTCTCCTTTTGCAA
TCAGTCTGTCTCGCTCTGTGAGCTGACTATGCTGCAGACTTTTCTTTTAC
CTCCAGCCTTTGCTTGTAGGTTTGAGGCTAGGGGCGGGGAGAGAGAGT
TCAGCCAGCAGGCAGCAAAACCCTATGTGCAAACACGATGTGTGAACATT

>rs12694403
TCACTCTCTACCCCTGCTGCCCTCCCAATGCTCAGGTGGGCCCTAGATT
ATCCAGAAAGCCAGGGTGTGTGCTTCATGAAAGGGCCAATACCCCTAATG
TGGGACCTATACAATAAAAGGCAAGGGCAGAAAGAGAGAAGAAATAAAAG
GAAAAGGGGTGTCTTGAGGGGAGTACATGATAGACTCCTGTGCCTGGAAG
[A/G]
GACTAGGAAACTTATACTGACTGGTTCAGTGCATTTGCTCAAAGAGTAT
TCATCTTCTTGAGTTTGTGAGGTGACTGTAAGAGCTGGGTCAAGTTCTT
TGGTTCATATTATATGCCTAGCTGAGCACAATGCCTGGGACTTAGAGGG
AAGGTCTGGGACTTCACCAGGATTCCTCAAAATAGAACAAAGCATTTGTTA

>rs12613955
TGAAATATGCCAGCCAGCTTAAATTAGCAAAGCTGGAAGTGTGGCAGAAT
GAATAGCCCGTGTCTGAGTGAGAGGGAAGGAAAGGGTCTACTGTGTGATA
GGCAATGGTTCAAAAGTAGGAGTAGGGGCCTAGTGTGGTCATGCCTTCTG
ATATGTCAAGAAAAGGCAAAAGTTGATATTTATTTGAAATATTCTGATT
[G/T]
TATAGGTTGGTGATCAATTAAAAACTAACAAATACAGTTAACAAACAT

GTTTTAATGCTTGCAGACAGGATCCTTCCTGCCGATGTTTGGCCTTTGTC
TTTAATGGAATTTGGGCTGCCCATAAACAGCCACCCTAAGCAAACTTC
TTTTCTGCCTTCTTTTATATTGTTTTTATGGAATGTAGAGGGATTTT

>rs11894340

CTGCTTCTATGCTGTATTCTAGTCATTAAAGGAGACCATCTCACACTCAA
GGGGAGAGACTACACAAGCCATAGTGCTGCCATGGCAACACAACCATTAG
GATGATGCCCTTGAGCCTTCCTTCTCATATTTCTTAGACTTCCTTTCTGA
CCTAGAGGAGCCAGAAGGACTTGGGACTTGTCATACCGTCCAATCCCAAG
[A/G]
AATGACAGGAATGAAGGAATTTTTAAAAATTCAACTCATTTCAACACAC
ACACAAAAATGAGAGAAGTGAGTCAATTTGAGAACTGAAATTCTGGTAG
ACTTTTGGGGAGACACTAGAAGGGCCATGTTTCATGAGTAGGAACCATAGA
TTCTGAGTCAATTGGGCATGAAAGTTTTCTCCTGGTGTCCCGATCTTGG

>rs17778329

ATTGTGTTGGGGGAAGGGAGGCTTCTGAGACAGAGCTGAGAGTCCCCAGC
AATGCCATCCCTGGCTTCTGGCCATCTGAATGCCTTACTCTTTCATGTA
AGCAAGACATGAGGTCCTGCTGTTTGAGTTGGCTTGGCCTGCTCCGTGA
TTTCAGTTCTTGTCTCTCCAGCGTGTGGGCTCATTTCCGTCCTCTGAGC
[A/G]
CTATTGTTTTCTGACTTGAACAGTCATCCAGAAAACCTCAGTAATAAG
GCTTTTGTATCATTATTTACTTTGCTAATTATATATGCTTTCCCATCCC
CAGTGCTTGAAGACATCAAATTGCAGTTGAAGTAACGGCAGAGGGAAAAAG
CAAGGGCAACAACAACAACCACTTCACTCCTCTGGGTGGAACATGCAA

>rs4442975

AATTTTGGTGATGGGAGGTGTAGAGGGTAACTCTTGATGATATCATTTG
TGTACCTGGGTCCAGCCATGCCTGAACCTTTCAATTACTTGACCCAGTAA
ATTTCTTATTCAATTAATGTTAGAGTTGAGTTTTGTCTCTTATAACTG
AAAGAGTGACATGTGTTTATATTAAGTGCAGGTCACTAATATAGTCTATT
[G/T]
GTTTCCCCAGGCTTGTCTCCTTTTCTCAAATAGGACTGATCAGATAGT
AAGTGCTCCATACCTGTAGGATACATAAAAGAGCAAATGAACAAGGAGCA
GGAGATATGGTGGCTGACGGTTGGACTATAATTTTTCTTGATTTTTCTG
ACTTTATTTTGGCTGTGATCTAGGAAGTGAATCCAGGTCCTTCTAAACC

>rs10932693

TTTATTTTGGCTGTGATCTAGGAAGTGAATCCAGGTCCTTCTAAACCCC
AATTTAACTGTGGCTCCTGAGGGAGATTTTCAGATCTCAAATGATCTGTA
AGTTTAAAGGAGGATGATTGGTCCCTGTACACTCTCTGTGAGAGTTCCTT
TGGAATACTGGTTTCAGAATGTACAGCCACACAGGTAAGTGGATTAGGGA
[C/G]
GGCCAGTGGTTTTTATGAATGCTATTGAAGACACTGAATGGAAGCAAT
TCTCACCAGAACTCTTTTGTCTCATCTTTCAAGGGGCTAAGTTGACACT
AGAACCAATCAATGATCCCTGAACTGGGAGAAATGGGTGGCCTTGGCTGT
GATTTTTGGTTGTCTGCACATTTCAAGATGACAATATTGTTTTAAATTA

>rs17778427

TCATCCTGTTATCCATAATGAAGGAAGATTTTAGGATATTCCAGCTTAGA
TCCAGGTGTCCCAGCTTCTATACTCACTCTGAAGAAGTGTACTTTATCT
ACTCTTTTACACCATTTATCCATTTTGGCCATTTCTGAAAGGAAGATC
CTAGAAGCCAGGCCTCTTGAAGAAGTTGGAGCTACAGCATTGATAGGAA
[A/T]
ACCTGACACCTAGAAAGGCAAGTCAGTTCTTCCCATGCAGTGGCTTTGG
GGACTGTAATTTAACAAGCTCTTTTTACTTGTGTAAAGTGATCTCCAT
TCCACTTTTTCTTGAAGTATTCTAATTAAGTGAAGCAATTGTTATCA
TCATTTAGGATTATAGTTAAAATACCCCTCAAATTGTTTCCATTTGGAAT

>rs17835044

AAATATTTCTCTTGGTCACTTGAGTCTAGCATCTAAGTCATATAATGCTA
 TTATAGCCTTGTTTTGCTTTTGATTTAGGTCATATCTCTTTTTCTGATC
 TTTCTTAAAGTGAAGGGACCATTTGCTATGCTCACAAAAATTCTCAAGGC
 TGGAACTACAGCATGCTCTGTTTCTAAGCCATACACATCCAGGTAGAAG
 [G/T]
 GTCAATTTAATTGCCAATGATCTCAGCCTTTTCTAACACTGTTCTTAA
 GTTTCTTTGGCACTGCATGGCAAGCAACTGAGTCATGCATGAAAAATTTCT
 CTTCTAGACAACCCATACCTACAGAGGGAGGATTTGGTTTACTGCAGATG
 CTGAGTGCTTCTGATGGCACACATTCCTTCTTATTAATATCCCCCTGGCC

>rs7562029

GAACAGTCAGCATTGTCTCCGCTTCCGCTCCACCCTGCAGTTCTGAGCTG
 AACAGATAATTTATCTGTGGCACAACAAAGTATATGACTTGGAAGAATGT
 TAAAGATCAAGTCCACCTTCTCCTTGATTATATGGCTCTGAGAGCCAGA
 GAGGGGAGGTGACCTTCCCAGAATGTGAATGTGGTAGAGCCTAGCTTACA
 [A/G]
 TTCATGTCCAACCTTAGGTGTGAAGTTCTATGAGACTATGCTGCCCTCTG
 CCTCACAATGCTTTCTCCCATGGGAGAGGAGGCTCAGAGGTACAAAGCAG
 CTGCAGAAAGACTCTCTCAAGCTATTGACTCTTGGATGACAACAATTAAT
 GTTCAACTTGGCTTTACCCCCAACATGCACTGCACAGATGCCTGCAGGTC

>rs13000023

TGGCTCTGAGAGCCAGAGAGGGGAGGTGACCTTCCCAGAATGTGAATGTG
 GTAGAGCCTAGCTTACAATTCATGTCCAACCTTAGGTGTGAAGTTCTATGA
 GACTATGCTGCCCTCTGCCTCACAATGCTTTCTCCCATGGGAGAGGAGGC
 TCAGAGGTACAAAGCAGCTGCAGAAAGACTCTCTCAAGCTATTGACTCTT
 [A/G]
 GATGACAACAATTAATGTTCAACTTGGCTTTACCCCCAACATGCACTGC
 ACAGATGCCTGCAGGTCTACAATGCCGAGACCAAGTTTGGGGGGCATTTT
 CGCAGCTCTGTCTATTGTCTCCTTTCTCTTCTGTGGTCATTCTTCCCA
 AGGGATAGGCCTAGTCCAGGGTGGAGAAAAACAATTCTTGGTGGTAGGGG

>rs735361

GGTTTATAAGGCAACTCTTAGCTTGACCCACTCTACTCCTCCATCAAAAT
 AAAGAAATAAATAAAGTTTTCAAATAAGAGCATTAAAGAATGATAACT
 TTTCTTTTTCTTGGTACTATCTATTGACCTATACCTGGTTTTCACTGTAA
 CACATTTGTGGTATGAAAAATGACAATATGATGTCAGAGAAAATGAGAGA
 [C/G]
 ACATTTTAGAATTTATCAGAATGTGACTATGGTCACTTAAAGTCCAACCT
 TTCTCATGTGGGTTTCAGCCTGTGCTGGAAACAAATAGCCTCTCTTACCT
 CTAAGTACTTGGTTCTGTGCTCAAAGTGGCATATTAAGTTCTGCTGATT
 ACATCAGAATTTGGATGGTGTCTCAGGATCTCAGTTTCTATCAGTGGAA

>rs13409592

TTAGAATAAGACTTCCCTAAGCAAATCAATGTGATGATTGATACAAACAG
 CCATCACTGGGGCCTCCTCTATTTCCAATCTCTTCCTCAGTTTCCAGTT
 CTTTCAGAACCTCACTGGGAAGTTCAGGGATATATGCCTGCAGGCTGGCC
 TGGCGCAGATGGATGATGAAGGGTCTGCCTTGAGTGAGGGTAGAGTGA
 [C/T]
 TCTGAGTGCTTTCCAACAGACTGTAAATTAAAGACACTCCAAAAAGCCC
 CTGGTGGGGATAGATGGGGAGAAGGGGTATCACTAAGTCTTCATCTTTTC
 TTCTAGGGCCAGGGTCAGTGGTTCAAATAAACATGGTGCCCCACAGAGAT
 GGCTGTGGCCAAGTAGCTCTCAGTGGGTTGACTCAAAGTCACCTGGACA

>rs2287289

GATCCGGGATCTGGTGATCCTCGGCAGGACAACCTCACCATTCTGTTCGGT

TGCCCCGCTTGCTCCTCCCGCTGGCTCCAGCTTCTGTGCCTACAATGAAG
 GCTGTGCCATTCCCTCACGTGCCCGCCGTGGCCTGACACCTTCTGGGC
 CCACCCTAACACATTCTCCGTTGCTTTTAACGACTGTTTACAGCTCGCT
 [A/T]
 CCTCTAGGACGTGGCTTCTGATGAAAACTGTCCAGACGCTCTGCCTCCG
 GTGTAATTCAGAGCCCAGCTCTCCTCCTTAAATTTCTTTTACGCTATTTT
 CCACAAGCATGTGCTTCTGTTGGCAGTGCCATAAGTGTGATTGATCTTTA
 CCATAAGTTCCTTGAACCTTGCTGTGACCTGGATGCTAAGGTGGCGGGAG

>rs12329133

AAAGTCCACGGGAGACACAGCTTCAGCCACCACCGGGGCTGCGACCTGTC
 CTCAGGATCCTCCTCAGCCCCCTCCAGGAGTCTCAGCTTTCTGCTTTTA
 GCCCCAAACCCAGATTCACTTTCCAGGCCATTACCCGTCTGAGGCAGCC
 CAGTCCGCTTTGTCTTCTCCTTCCCGAAGGCTCTCAGCCCCCTGGGCAG
 [C/T]
 ATCAGAATTTCACTTTGATCTCCTTTTCCCTGGGAGCCAGAATAGCTTA
 GTCTCTTCAAAAGGAGAGTTAGAAAGTGCCGTTTCATCTAATAAAAAAT
 AACTCCTGAGAGTTATCCAATCCAATACAGGTCAATTCGCTGGAGGCTA
 CCCCAGCTAGGAAACACATTTTAGGACATCAGCAGAGGTAAGATGAGAG

>rs13011060

TAAATTTAGCACATGCTATATAGTTTTAGGATTAGTCCCTCTCTAGGT
 ATTTCATCTGGGGGCTGGCGCCTCCTTGGCTGGCAGGGGCTGGTGATCGGGA
 CTCCCAGCGAGACAGATTGCACCCAGAGCAGTGAGGCCCCGAGCGCTTC
 CTGGAGCCTGGGGAAGATGGCATTCAATTTTAAAGTCCATTTATTGTGCA
 [A/G]
 TTGTTGAAAGTGGGACACAGAAATTTGGACTGATATGCCTAGGCCACAT
 GACTACTGATTGGGCCAGGGTAAATGCAGTCTTTTTGTTGATGGAGAAGG
 TCTTTTCTGGGCCAGCACAAAGCAATGGATCTTCAGTCTTAAGAAATGC
 ATTGAAACCATCCTTGTGCAGTTTGGGTCCCTGGTAGGCCAAGAGCTG

>rs13011326

ATATGCCTAGGCCACATGACTACTGATTGGGCCAGGGTAAATGCAGTCTT
 TTTGTTGATGGAGAAGGTCTTTTCTGGGCCAGCACAAAGCAATGGATCT
 TCAGTCTTAAGAAATGCATTGAAACCATCCTTGTGCAGTTTGGGTCCC
 TGGTAGGCCAAGAGCTGGAAGGCCCTTAGACTCTGAGTGCTGTAGCCCC
 [C/T]
 CTTCTAGGTGATGGTCAAACTTGAAAACGGCTCTAGACTCTGGACAGA
 AATGCATGAGAATTGAGTGTCAAAGCTTGGTGCTTTGACTTCTTTTAAAC
 TTTCTGTAGACTGAAATATGAATAAGAGGAAGGCATGAGGCCCTCTCAG
 TTTATGGAGCCATAGCTCCAAACCTCCCACTTTCCAACCCAGCAACCCTG

>rs4674132

AGAAATGCATTGAAAACCATCCTTGTGCAGTTTGGGGTCCCTGGTAGGCC
 AAGAGCTGGAAAGGCCCTTAGACTCTGAGTGCTGTAGCCCCCTTCTAGG
 TGATGGTCAAACTTGAAAACGGCTCTAGACTCTGGACAGAAATGCATGA
 GAATTGAGTGTTCAAAGCTTGGTGCTTTGACTTCTTTTAACTTTCTGTAG
 [A/G]
 CTGAAATATGAATAAGAGGAAGGCATGAGGCCCTCTCAGTTTATGGAG
 CCATAGCTCCAAACCTCCCACTTTCCAACCCAGCAACCTGTGTGCCTA
 GGAATTGTAGCGGGTGTCCAGAATGCAAAATCACTCTGCTAGAAGTGG
 AAAAGGAGACGCCTGCCTTTTCTAGCACGATCTTAGTAAATAGATTAG

Таблица 28. Последовательности, фланкирующие rs3803662 и сцепленные ОНП маркеры НарМар с $r^2 > 0,2$

>rs3803662

TTTTATTGTTCTATGGTTATTAATAAATAACATGTCATATAAATTAAGCC
TAGAAACATGGATGTTTATCAAAGAAAGGATTGTCATCCAAAGCACCAC
TATGAGAGATATCTATGTCAATGGTATATAGATCTGTCATAGAAGGGTT
TAATTATATCTGCCTAATGATTTTCTCTCCTTAATGCCTCTATAGCTGTC
[C/T]
CTTAGCGAAGAATAAACTGTGGACTGACCCCCACCCATTTGCGAAGAA
AGTACTGGGTCTTCAGCTTTTATTGTTTACGCCGGTGGTCTTTGTGGACAA
CACCAGGGGGCCCTCCTGCTGAAGACTGTCTGGACTTCCTGATTTTCACT
CCTGTATAAACCCCTCCCCATTACTACCTTTGTCTCCTTGGGGAAAAA

>rs4784220

AGAGAATTACTAATATTTTCTTCTCCAACCTTAATGCTCACTCCCTGAGC
TATGTGGATTATACTAGTGAAAGTACTGCCTTAACCCACTAATGTAAACA
GATTCCTTTTAAGAAGCCTCCCTATCGAGAATGGTCTGTAAGAGACAACC
ATGTCAACAAGTAAAGTTCTTCTGCCTTCCCTCCTGGGGTGGGCTAAGA
[C/T]
GGACCCCTTCTCACCATTCCCAATTACCTGCAGGAAGGCAGCAACAGAAT
AGAGAAGGAGGGTAACACAGTTTCTGTACTTTAAATACCTCCCTGCCTC
CGCAAGTAAACTTGAGCCATGCCAATGAGGGAAAAAGTGCCGCGATT
CCAGGTGGCTTAGTAATACACAGTTTCTGACACACTGCAACTCCATTC

>rs12598982

CACTACCAAATAATATTGCTTCTTGCATTAAGTATAGCATTGGTTTGCTA
GAATAATCCTAACAATGACAATGCAATGTACTTTAAATTGGGGTCAAAAA
ACTTTTTTCTTAAAGTACCAGATAGTCAAGCTTTTAAAGCTCTATGGGTCA
TCCACAGACACACTGCAACTATACAACTATATCACCCCTGTTGCAACTA
[C/T]
ACGAAAAGCAGGAAAGCAACCACAGAAGACAGACAATACATAAACAAT
GAGTATGACTGTGTTCCAATAACATTTTATTAATAAAAAACAGGCAACTGG
CCCATGGGTGGTAATTTGCTGAACTTGGTTTAAATTATTATTAGTATCT
ACCTTTTTTAAAGATGATGTTTTTGGGTCTTTCCATCATGAAAAAT

>rs4784222

TAATGATACTATAAAACACCCAGCATAACGCCAGCATAAGGATGTGAAT
GAATGAGCAAATGAAAGGAGGGAGGGCAATGGCAAAATTTTACCTTGGTT
TCCAAGCTAAAGTATAACCTGAGTTTTTAAATGCTTTACAAAAAGGAAAG
GCAATCAGAACCATAATTAATTTAAGCTGTACTGACAGGTTTTCCAGG
[C/G]
AAAATTCTATAGAAGTTGTTATTCTCCAACCTCCAAAAATAAGCAATAA
ATTTTTTTTAAAGAGCTGTTATTTTAAAGACTTTATTTGAACCAAGAATG
TGTTTTATGGTAACAACATACATGTTGTAAATTAATCAGTTTCAATGTTAT
AAAATAATACTTCCAAGAATAAATGTATCATATAATCTATGTACCTCATG

>rs17271951

TCATTGTTTGTAATAATCCAATTGTCATACTCTGGTATTTGTTCTCCAAAA
GAATCATCACTTTTATATTCACTACAGCATGAGATGAGATATTTAATAAA
TAATAGGACATCATAATCCAGGGATTAAGTAGTTTCTTTAGCAAAATTT
AATATGAGGGTAGCATATGCCAGCCGTTCTAGTTGAGCAACAATTTCTG
[C/T]
GTTGGGAGAGGATAACAAAGCCAAGTATTTTCCATAAGAAGTTCACAC
TCTAAAAGGCCAGGCGGGTGGCTCACACCTATAATCCAGCACTTTGGG

TGGCTGAGGCAGGCGGATCACTAGGTCAGGAGATCGAGACCATCCTGGCT
AAGACGGTGAAACCCCGTCTCTACTAAAAATACAAAAAATTAGCCAGGC

>rs9933638

GGAATAATTACAGATGATGCTAGGCAATATCACTAGATTGGGTTTTTCATG
AAGCTCGTGACTGGATTTGCCACAAAAATCACTAACTGCTTTGGCTTCC
AGAGAATGGCCTCATGGATAAGCCCCAAGTTCCAAAACTACAAAATGTC
AAAATCACATTCACTAATTAATAATTTTATGTCACAATAAGAATTAGGA
[A/G]
TCTCTAGAAGGCTGGGTACAGTAGCTCATGTCTGTGATCCCAGCACTTT
GGGAGATCGAGGCAGGCAGATCACTTAAGCCCAGCAGTTCAAGACCAGCC
TGGGCAACATAGGGAGACCCCGTCTACAAAAATATAAAAATCAGCTGG
GCATGGTGGCACATGCCTGTAGTCCCAGCTACCCAGGAGGCTGAGGTGGG

>rs9302556

AACACATAAGCCATTTTTTCCAAAGCTAGACAGCTTTTCTCTATCCACGT
TCCTTCAAATATTTTACCTAAGACAGAGAATCTCATATTAAATTCAAATTT
CACACTTGCCAATCACACACGCCCACTCACAAATTCCTCCAGATGATGCAA
TGTGGCCTAATGTGTATGCTTATGTATGTACATGTGCACACACAAACAGG
[A/C]
ATAATGTTTTAAATACATATAAACACATTCATTGCTAGGTATTGGTTAAA
ACAAACAGGCTTACAGCATACCAACTTATTAAACATAATGATTGATATTA
GTAGTTTCATTAATGTAATTAACCTTTAGATTTGAATTATCAAGTTCAA
GTCTAAGACCGAGAGCAGTGGTTCTCAGGCCGGTACAATAATGGCAAGTC

>rs7190749

CAGTCCTGCCCTACCAGAAACCAGAGATATTTGCTGCTGGTAATAATGA
GGAGGGGGTTTCTAAGTACAAGATCAGCTCTATTCTTTCAATTCCTGGAA
ATCTGCCCTCGCCTCTGATACCTCAAAAAATATCTGTTGCATTTCCAAGG
GAAAGAACATGATATAAGCCTTGCTGTTAAAAAAGGCAAGGCTGCTGCTG
[A/G]
CTACTTACAAGATAAAAGGCAAGGTGTTTTGTTTTTTCCCTTTCTCT
TAGACTAGGCATGAATAATCTCAAAATATTTGTTGTCAATAAAGTGATT
AAAAATAAGTGATAGACAAGACATAATGATTAGAGAAATGAATCAAAGTA
ATATTTGGGACCAGGGTCTTACAAATCTAGAGTAAGTATGTTGAGCAGAA

>rs12443621

AGATCAAGATGTCTAGCATAATTTGAAAGTATGTTTCTATTTGACAGAAA
CCTTGGCTTGAAAAAGTTAACTCTAAATATTTGTACAGAAAGAAGTATT
CTGGGTTTCATTCAATTAAGTCAATTGACCACTGCAGAAAAGGGAGAGCA
AGAGATTTACATATATTTTTCTTGACGTTTTATATGCATTAGGCCTGGCA
[A/G]
TGAACCTTGAGGTAGGTATTACTATCTCCTTATTTCTAAGGGAATCAGAA
TACTGTAAACAAAGACATGGAGCTAAATAATGGAGTCAAATCCTGGCTTA
CTAGGCTCCACATCCATGTTGCTCCACTTATCAGACGGCCACCCAAACAT
ACAAGTTTAACAAATTCAAATATTGCGGCTAAATAGACACTTTTTTTG

>rs9933556

CATGGAGCTAAATAATGGAGTCAAATCCTGGCTTACTAGGCTCCACATCC
ATGTTGCTCCACTTATCAGACGGCCACCCAAACATACAAGTTTAACAAAT
TCCAAATTAATGGGGCTAAATAGACACTTTTTTGTTCATATTGCCCT
TCTTCCCTTGATTCTGAATATCATGTCAACCATGAATCATAAGCCAATG
[C/T]
TTACAAATTTAACTTCACATTATGAGTTTTGAGCCCACTGATGCCTAA
CCACTAAACCATTCATATAACACACTGAAAAGAAGCCATTTAAGAAATGA
GAAAAATAGGTATTGTTTTCTGTATGGCTGTAAAAGCTGAGATAAGAA
CATTTTGGCAAGTTCAATAGAGCCGAAAGGCTTGTGAAGGCTATGGGAGG

>rs1362546

TCCACTATTTGCACTAGCATTAAATCAACATCCAAAAAGTGACAAATTAT
 TTTATCTTTATAATCAATCTTCTGAAATGCTAACTTGAAAAATATAAACAG
 CACAATTACTTTTTCTCTTTCTTCCCATTTGAAATGTGTGCTTGTGT
 GTGTGTGTGTGTGTATCTGAATCTGTACATATAAATCAAATACACACA
 [C/T]
 ATATACTTGTTTTAAAAAATCCTCAAAGAATCAATCACATATTACAGA
 ACTGAAGGAATATTGGAGTTCATCTGGTCCAAAGTGTTGGTATTGTAAAA
 TCACTTTCAAATTTACAGTTATCCAAATGGTAATGAATATTAAATACCTA
 AGAAATCAATCAAAGATTAACATTAAGACATTAGATTTTAAAAATCATT

>rs1075367

TTTTAGCTACAAGGATATGTTTTAAACCATATTGAATTACATAGAAAG
 CATTTCCATTTTCCTTTAGGCTTTCTTATTACATCAGTGAGAAAAATCTC
 AGTTTGGTGCGAATATGTCTTTAACACCTAATACTTGCTAATCTCTTTT
 TAGCAGAGAGATTGAGACACATTATAGGGGAAGAGAGAGAGAGGGGAGAGA
 [C/T]
 GGAGGAGATCATACAGGACCCAGTTACAACCTTCTGTAGCTTTGGTTT
 GCTAGAATTAATAGATTATTTGTATTTATTGTATTTATTTGATAGTT
 ATGTGCTATTTATAGAAAGGAATACCAGTTTCCATGTATGAAATAAAGT
 GTCCTTTGTAAACAAAAAGCATATTAATTTTATGCAAATGCATGAAT

>rs8046979

CAGCTTCAGTGGATGCAAGTATGATCTCTCCGATGATCCTGGGTCCCTTC
 AGAGGCCTTTGGATGGTTCTGAGCAGTAGAACTCAAACCTGTATCCAGG
 AAGGCAGCATCAGCATCACCTAGCAACTTGTAGAAATGCAAATCTTGG
 GCCCTACCCAGACCTACTGAATCAGAACTCTGGGTGTGGAGCGGCAAC
 [A/G]
 TGGATTCTAACAAGGCGTTTCAAGTGACTCTCATGCTCAGTGAAGTTTGA
 AAAGCCTGTGGGTAACGAGCAGAGTCTCTGGAACCTGCCTGTTGATTGA
 AGGAATCCCAGATCAATCACCTCCTCACCTGAGACTTTGAACAAGTTATT
 TAATTTAGCCATGCTTCCATTTCCGTTTCTGTTAACTGGGACATTAATA

>rs1420529

ATGAACTCTTCAGTGCCCTCCCCTTCCCACAGAGGGCTGAAAGGGACTCA
 CTCCTTGTTACAAGAACACGGCCACTCCTCCCTCTCTTCTCTTCCCAGTC
 CATCCCTCACCCAGTTTCAGCCACCATCCAAAAAGAAATCGCAGCATGG
 CCCTTGAAGCACATTGAGCTTAAAGTGCCTAGAAGATAAAGCCTAAACT
 [G/T]
 CAGCACGACAGGCAAGACCTTTCAAATCTGGCCTCGACCATCCTCTCT
 GGGTCTCCTTTCTGAGAACCCCTATTTTCTGCCTTTTCACACTCCACAC
 TTATGCATGATGCTGTTTTCCCTCCCTGGCTCACCCACCCAGTGTTCTT
 CTTCTGTGATGGTAACCTCCCAGAAAGTCTCTTGAGCTAGAGTGATCAT

>rs11642645

CTCGACCATCCTCTCTGGGTCTCCTTTCTGAGAACCCCTATTTTCTGCCT
 TTTCACAACTCCACACTTATGCATGATGCTGTTTTCCCTCCCTGGCTCAC
 CCACCCAGTGTTCTTCTTCTGTGATGGTAACCTCCCAGAAAGTCTCTT
 GAGCTAGAGTGATCATTATTAATTAATAAAAAAAAAAAATCCTACTTTGTT
 [A/C]
 TATTCTGATTTCAAAAACTGAAAGGCAAGATATTTGAGGGAAAAAAGA
 GAATATTGTTTTGTTGCTTTTTAAAAAATAAATTTGCAGGTCCTGACAAT
 CCCGTGCTTTGTGTGTGCTGATCACACTGCAATCTAATCAGGTACAGTAA
 ATAAATGACTTGAGAACCTTTCTGACGAATGCAAAGGCTGATGATATT

>rs1420533

TGTGATTTTTAAAGGTAAGTAACAGGTGAATCAAAGCAAGGAAAAAGTA
 AATTTGAGAAAAATAATAAGCCTAGATGGCCAGCACACAGGAATGCATAT

CCTACTGCCCTAGACAGCTACTAAATGGGTTTCCTGTTTGAAGCTCAAGCT
TCTTGGCCATCAATGCAAAGAAAGAAACAAGAATGGTAAGATTAAATGCCCA
[A/G]
TAAGATACAGCGAAGTGGATTTTTTTCAGTTCTTCTTGATATAGTAACC
TAAGAAGATTCTCTGTAAATCATCATTTAAAAAGACACCATGAGGTGTAA
TGAGCTATACATTGTAAATAGAACTTTTTAAAAAAAATGATGTTACAGAA
ACTATTACTCATAATCCTCTCAATACAAAGTGAAGCCTTAATCAGTGTGC

>rs1362548

TTTTAAAAAATGATGTTACGAAACTATTACTCATAATCCTCTCAATA
CAAAGTGAAGCCTTAATCAGTGTGCAATTCAATAAGCAGAGAGGGCCTG
TGAAGGCAACTGTGAAGTGGGGTGCCACATAATATGCCATCGTGAGCATT
CAGCTCTCTCACAAGTCTGGCAGAAATCTTATCAAATTAGGTTATCTG
[C/G]
ATGAACGGAAGGAAGCAAGGCATCTTCTTCATGCCATTCTCCTTAGATA
TATTATTTCTTCCAACAGATTTTATTAACACAACTGGGCAGCAGGGACT
TCAAGAAGGACAAAAATCCTGTGTCAAACCATACATATGTTGGCAACA
TTTATCATCAGGAAAAGCGTCCCAGGAAGTTGAATTCTAATGTATGTGTG

>rs2193094

CTGGGCCTGGACTCAAGGGTCTTGTGTACTTTCTGCTGTCCGAAATAC
TGCCATCATGAACAAGCCAGGCTGGCGTGCTGGAGGTGAAAGATACATG
CCAGAGAACTGAGGGACCCCGGTGGACAGCCAGCCAACCTCCTGAAGCAA
ACCTCCCTGCCAATCTGCAGTTGACTGCAAACACATAAGAGACCCAGAA
[G/T]
AGGATAAGCACTGCCCTTCTGAGCCCAACCTAAATTGCTGACATGCAGA
ATCAGAAACAAAATATGTGTTTTGAAGTCACAAAGTTAGAGGCGGTTT
ATCACTCGCGATAGATAACTGGTACAGACAACCGGGTCTGCCTAGAAAAG
TGAAAAACATACGAAGGCAAAACAGTAAGAGGTCAAGCTTTAGGAGGATA

>rs4783780

TCTCATATCCATAGAATATCCAGGTGACCTATAGGTTTCTTGAAGGCAGA
CCACATCCAATCTCAGCCTCAACAAGAAGGGCGTTCCTAATATAAACAAT
ATTTATGATATTTACATTACATAAGCACAGGCTCATTTTGGTTTCCACAA
AGAATAATAATAACATAGACAATGGAAATCCTCAGAAATCTTCAAAATC
[A/C]
TCAAGCTTCACTCCAGTTAGCCGTTTCAATGGACTTTGCTTTCTATCC
TACACCCACATTCTCTCATTGTGACATCAGCACTTTGTTTTCTCCGAA
GAACCACTGCTCCATACTCTTAGCTCCTGTGGCTTCCAAGAAGATGACAG
TGTCAGCTCCCCCACTGGCATGCCTACGTGGCTACGAGCCTGGGCATGG

>rs3112581

TTACATTCTCACAGCTCAGCCTCTTCAAGAGGGAGAGATTCTTTCTTGCA
TGTTTCAGCAGCTTTTCTGGATCAAGTCTCATTGGATAGACTTGGGTCCC
ATGCCCAGGCTCGTAGCCACGTAGGCATGCCAGTGGGGGAGCTGGACACT
GTCATCTTCTTGAAGCCACAGGAGCTAAGAGTATGGAGCAGTGTTCTT
[C/T]
GGAGGAAAAACAAAGTCTGTATGTCACAATGAGAGAATGTGGGTGTAGG
ATAGGAAAGCAAAGTCCATTGAAACGGCTAACTGGAGTGAAGCTTGATGA
TTTTGAAGATTTCTGAGGATTTCCATTGTCTATGTTTATTATTCTTT
GTGGAAACCAAATGAGCCTGTGCTTATGTAATGTAAATATCATAAATAT

>rs3112580

TTTTTTTTTTTTTTATCATATCAGCCAGGCTGTACTCAAGTAGTGGGAA
GTTGTCCAGCAGCAGCTCCAAGTTACATTCTCACAGCTCAGCCTCTTCAA
GAGGGAGAGATTCTTTCTTGCATGTTTCAGCAGCTTTTCTGGATCAAGTC
TCATTGGATAGACTTGGGTCCCATGCCAGGCTCGTAGCCACGTAGGCAT

[A/G]

CCAGTGGGGGAGCTGGACACTGTCATCTTCTTGAAGCCACAGGAGCTA
 AGAGTATGGAGCAGTGGTTCTTCGGAGGAAAAACAAAGTGCTGATGTCAC
 AATGAGAGAATGTGGGTGTAGGATAGGAAAGCAAAGTCCATTGAAACGGC
 TAACTGGAGTGAAGCTTGATGATTTGAAGATTCTGAGGATTTCCATTG

>rs9931232

TTGAATTAAGATTAAGATAGAATGTTAATCATACAAGAAGGGCCATTTT
 CCAGGGCCTGGCATGTTTTACACAAGCATGAAACATATTTAAGAGAACAT
 GGCCTCCCTGAGGAAGAGATTAAGGATACAAACCTAGGAAACAAAGAAAA
 TAAACGATCTACCTCAAAGGGATAAGATTAACCTGTGAGAAGCTCATGC

[A/G]

GACCTTAAATAAAAGCCTCTACTCCTTATATTTTTAAACAGGATTTTT
 TTTTTTTTGTAGAGCACTAATTCACCTTCAAAATATATTTGAGTTTGTA
 AGTATACTCAACAAAAATCCTGTCTTTCTTATTAACCATACTCAAAACA
 GATTCTCTAAATAGTTATGACACCATTAAACATTTTCATGAAATCCAAAAAC

>rs1123428

CCTGGCTATATTTTATTTGAAGTTATTCTCCTACTTGCAACGACTTTTGA
 ACCCACAAAACCTTTTCATCCATTCTCCACAGACACAGTACACTTTTCATG
 AGCACTGACATTTGAAAAGGATCCATATACTCACAGAATGTAAGATAAT
 ACACATGCATATAATTGATACAAAGGACTCTGGTCATTTGTTAAAGAGAG

[A/T]

AATGAATGAAACCTATTATTACGGTTTGTAAATAAATTTAAAGATAA
 GCGCTAACTTTCACTTGTTAAGCCTAGAGCCAAGAGTTCCGGTGTTCCTC
 TAAAGTTTCCCAGATGGTCATTAAAAACAAGAGAGGGAAGAAACATACTTT
 CTTTGAATCCATCATTTCCCATCCAGAAAGATAATACATGTCATCAAAA

>rs3095604

TCATTCACTGCTCTGGGAGAGAAAAATGTCCCTATTTTCTAAAAGGTCCT
 AGCTGATTAACTTCCCCACTCTACACTCACCCACCCCTCTTATTTCTTT
 ATCTTTCTTAAGCTCCATCTTTTTATTTAATTTTTTTCTCATCGGCTTT
 ATTGCTTTGAGTTAATCTGAGAGTTTGGCCTTCAGCCTTGCGGAAGGGAA

[C/G]

GGGCTTTTACTAAATGGGAACGTTGTGGAGAGGCCGTGAGATCTGCGC
 CCTCCCCAGTTTAGTTATCTGCCAGCTACTTTAACCAGAGTGGCGACATT
 CTGAAATATTTCTATGACACTCGCTAACAGTTTCATCCCCTGGTCACAC
 CTTGCTGGCCCAGGGCCTGAGCTTTTTCAGACAAGCATGACATGGGTTTT

>rs4784227

GAGTTGATCAGTAAATATTTGCTGAATGAAAGAATACATGAATGAAAAGT
 CAGAGCCCTATAGGTGAGCATGGACGGCGGTAAAGGAACCTGGCTGAGCC
 TGAAAGAGAATGTGATCTAAGATTAATCCAGGATATGCTGGTAAATGTT
 TAACAGCCAACCTCTTTGGGGAGGAAAAAAGTCCCAATTTGTAGTGTTTGC

[C/T]

GATTATTGTGATGTAATACTCCCATCATGACCAATTTCAAGCTACCAA
 CATGCTGACACTGAACTTGGAGTTGGAAGGAGATGAACAGGCATAATCAG
 GTCTCGTGAGATGGCCCAAGCCGCCCCAGCACTCCACTGTTATATATGA
 GGCTAGAATTACTACATAACTGGAATAGCACTTTCTGGACCATATGCCT

>rs12922061

ACTTCATTTTGGGAAAAAGAACAGTTACTACCATTCAATTCATTCATCAT
 CACACCAATATTACTGAATCCTACCCTGGACCAGCATTGTTATGGTAGAT
 CTTGATAATATAGTGACAGACAATACCATCCTTCCCACCATTTACATAGT
 AGACTGCCTGGGGCTCATAGTCTACTCACATCTCAGCACATGAGATAGT

[C/T]

CATTATCACATTCTCTTCTATCTCTCTCACAGGCACTCACTCACCAGCC
 ATGGAGAACTTAATGGGCATTAGGTTCTTAATCCTGCTGGGTATAATTGG

AGATAACCACCCTGGCTGACTGAGCACTATACCAGGCTACTCATGGATGA
 TTTGGTTATTAGACGTAATCAGAGTGGCCCTTGATGCTAAAAAGGAGAC

Таблица 29. Мультиэтнические когорты

```

>rs4784227
GTAAATATTTGCTGAATGAAAGAATACATGAATGAAAAGTCAGAGCCCTA
TAGGTCAGCATGGACGGCGGTAAAGGAACCTGGCTGAGCCTGAAAGAGAA
TGTGATCTAAGATTAAATCCAGGATATGCTGGTAAATGTTTAAACAGCCAA
CTCTTTGGGGAGGAAAAAAGTCCCAATTTGTAGTGTTC
[C/T]
GATTATTGTGATGTAAATACTCCCATCATGACCAATTTCAAGCTACCAA
CATGCTGACACTGAACCTGGAGTTGGAAGGAGATGAACAGGCATAATCAG
GTCTCGTGAGATGGCCCAAGCCGGCCCCAGCACTCCACTGTTATATATGA
GGCTAGAATTACTACATAACTGGAATAGCAACTTTCTGGAC

>rs17271951
TAAATCCAATTGTCATACTCTGGTATTTGTTCTCCAAAAGAATCATCAC
TTTCATATTCACACAGCATGAGATGAGATATTTAATAAATAATAGGACA
TCATAATCCAGGGATTAAAGTAGTTTCTTTAGCAAAATTTAATATGAGGG
TAGCATATGCCAGCCGTTCTAGTTGAGCAACAATTTCTG
[C/T]
GTTGGGAGAGGATAACAAAGCCAAGTATTTCCATAAGAAGTTCACAC
TCTAAAGGCCAGGCGCGGTGGCTCACACCTATAATCCCAGCACTTTGGG
TGGCTGAGGCAGGCGGATCACTAGGTCAGGAGATCGAGACCATCTGGCT
AAGACGGTGAAACCCCGTCTCTACTAAAAATACAAAAAAT

```

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ определения подверженности раку молочной железы у человеческого индивидуума, включающий определение наличия или отсутствия по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера в образце нуклеиновой кислоты, взятом у индивидуума, или в массиве данных генотипа, полученном у индивида, где маркер выбран из полиморфных маркеров rs13387042, и маркеров в неравновесии по сцеплению с ними, и где наличие по меньшей мере одного аллеля указывает на подверженность индивидуума раку молочной железы.

2. Способ по п.1, где по меньшей мере один полиморфный маркер расположен в составе сегмента генома с последовательностью, приведенной в SEQ ID NO: 5.

3. Способ по п.1 или 2, где по меньшей мере один полиморфный маркер выбран из маркеров, приведенных в табл. 15.

4. Способ по любому из предшествующих пунктов, где подверженность, обусловленная наличием по меньшей мере одного аллеля или гаплотипа, является повышенной подверженностью.

5. Способ по п.4, где наличие аллеля А маркера rs13387042 указывает на повышенную подверженность раку молочной железы.

6. Способ по любому из предшествующих пунктов, также включающий этап определения наличия или отсутствия по меньшей мере одного генетического фактора с высокой пенетрантностью в отношении рака молочной железы в образце нуклеиновой кислоты, взятом у индивидуума, или в массиве данных генотипа, полученном у индивидуума.

7. Способ определения риска развития по меньшей мере второй первичной опухоли у индивидуума с диагностированным ранее раком молочной железы, включающий определение наличия или отсутствия по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера в образце нуклеиновой кислоты, взятом у индивидуума, или в массиве данных генотипа, полученном у индивидуума, где маркер выбран из полиморфных маркеров rs13387042 и маркеров в неравновесии по сцеплению с ними, и где наличие по меньшей мере одного аллеля указывает на риск развития по меньшей мере второй первичной опухоли.

8. Устройство для выявления генетического индикатора рака молочной железы у человеческого индивидуума, включающее

компьютерную память и

программы, хранящиеся в компьютерной памяти,

где программы адаптированы к выполнению на процессоре с целью анализа информации по маркеру и/или гаплотипу по меньшей мере для одного человеческого индивидуума в отношении по меньшей мере одного полиморфного маркера, выбранного из маркеров rs13387042 и маркеров в неравновесии по сцеплению с ними, и генерирования выходных данных на основании информации по маркеру или гаплотипу, где выходные данные включают показатель индивидуального риска по меньшей мере по одному маркеру или гаплотипу в качестве генетического индикатора рака молочной железы у человеческого индивидуума.

9. Способ идентификации маркера, пригодного для применения в оценке подверженности раку молочной железы, включающий:

а) идентификацию по меньшей мере одного полиморфного маркера, находящегося в неравновесии по сцеплению с rs13387042;

б) определение генотипического статуса в выборке индивидуумов с диагностированным раком мо-

лочной железы или подверженных ему и

в) определение генотипического статуса в выборке контрольных индивидуумов;

где достоверное различие в частотах по меньшей мере одного аллеля по меньшей мере одного полиморфизма у индивидуумов с диагностированным раком молочной железы или подверженных ему, в сравнении с частотой по меньшей мере одного аллеля в контрольной выборке указывает на то, что полиморфизм может применяться для оценки подверженности раку молочной железы.

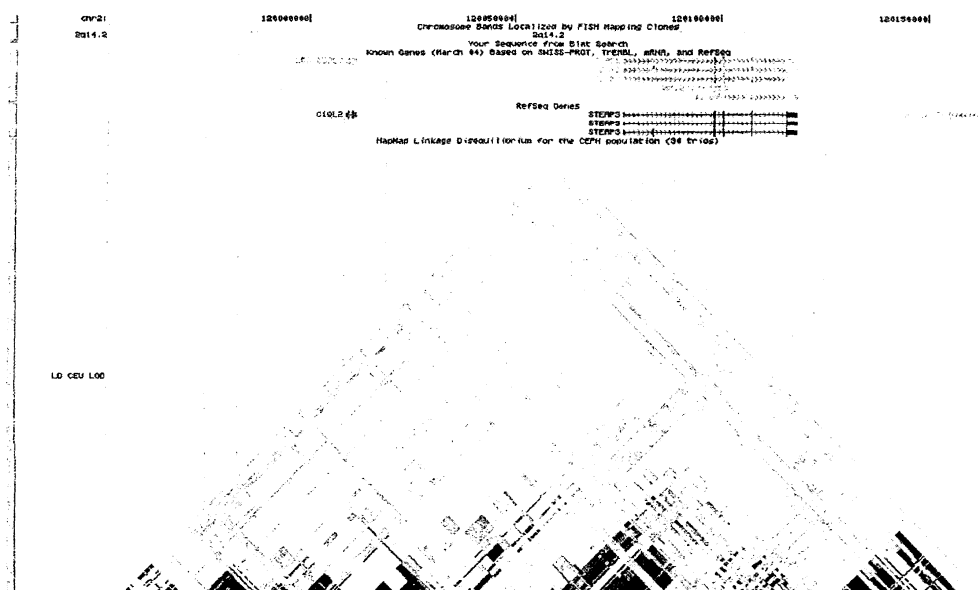
10. Способ по п.9, где увеличение частоты по меньшей мере одного аллеля по меньшей мере одного полиморфизма у индивидуумов с диагностированным раком молочной железы или подверженных ему, в сравнении с частотой по меньшей мере одного аллеля в контрольной выборке указывает на то, что полиморфизм может применяться для оценки повышенной подверженности раку молочной железы и где понижение частоты по меньшей мере одного аллеля по меньшей мере одного полиморфизма у индивидуумов с диагностированным раком молочной железы или подверженных ему, в сравнении с частотой по меньшей мере одного аллеля в контрольной выборке указывает на то, что полиморфизм может применяться для оценки пониженной подверженности, или устойчивости к раку молочной железы.

11. Способ определения прогноза у индивидуума с диагностированным раком молочной железы, включающий определение наличия или отсутствия по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера в образце нуклеиновой кислоты, взятом у индивидуума, где полиморфный маркер выбран из группы, состоящей из полиморфных маркеров rs13387042 и маркеров в неравновесии по сцеплению с ними, где наличие по меньшей мере одного аллеля указывает на менее благоприятный прогноз рака молочной железы у индивидуума.

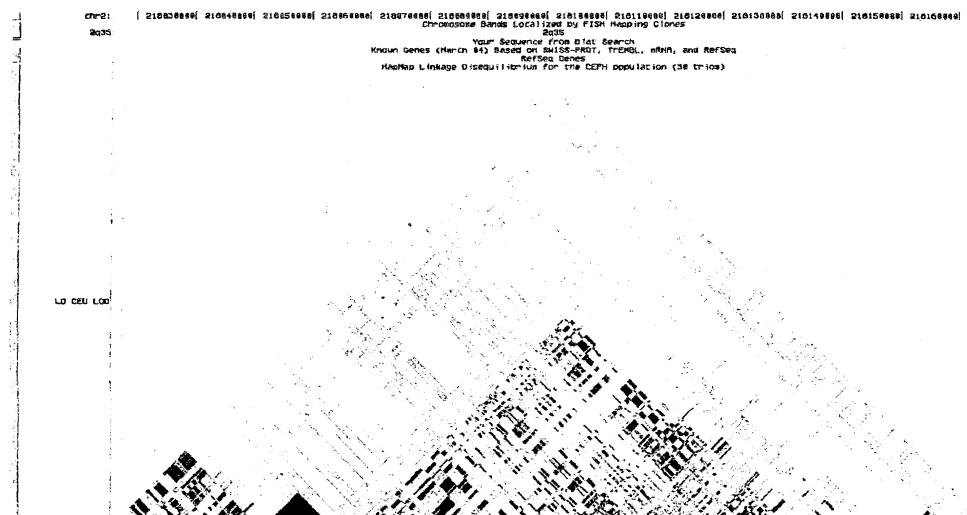
12. Способ мониторинга хода лечения индивида, проходящего лечение рака молочной железы, включающий определение наличия или отсутствия по меньшей мере одного аллеля по меньшей мере одного полиморфного маркера в образце нуклеиновой кислоты, взятом у индивидуума, где маркер выбран из маркеров rs13387042 и маркеров в неравновесии по сцеплению с ними, где наличие по меньшей мере одного аллеля указывает на исход лечения у индивидуума.

13. Способ по любому из пп.1-7 или 9-12, где неравновесие по сцеплению характеризуется значениями r^2 по меньшей мере 0,2.

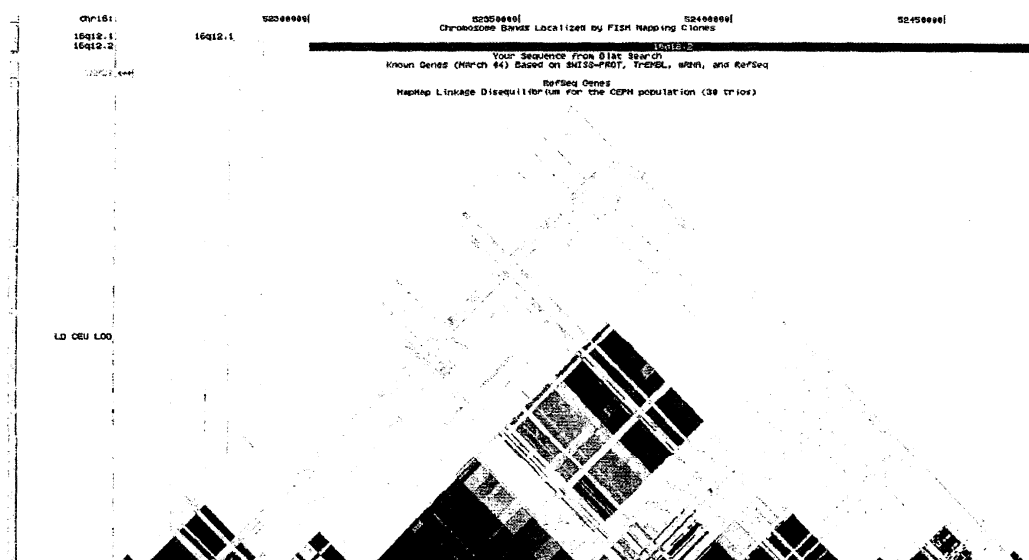
14. Устройство по п.8, где неравновесие по сцеплению характеризуется значениями r^2 по меньшей мере 0,2.



Фиг. 1



Фиг. 2



Фиг. 3

Перечень последовательностей

<110> deCODE Genetics ehf

<120> ГЕНЕТИЧЕСКИЕ ВАРИАНТЫ CHR2 И CHR16 В КАЧЕСТВЕ
МАРКЕРОВ ДЛЯ ПРИМЕНЕНИЯ ПРИ ОЦЕНКЕ РИСКА, ДИАГНОСТИРОВАНИИ,
ПРОГНОЗИРОВАНИИ И ЛЕЧЕНИИ РАКА МОЛОЧНОЙ ЖЕЛЕЗЫ

<130> 2506-PC00

<160> 6

<170> PatentIn version 3.5

<210> 1
<211> 600
<212> DNA
<213> Homo sapiens

<400> 1
caggataaga ttatttctga ggctgcttcc atcagtggtt ctgtgagctg gcccatgtat
60

tcagccagtc agtaggccat caagcatgca tctacgcaag gactagaag ggttacatac
120

aatgggaaac tgaggcccag aaggaggaaa gatgaagtgc ccaaagttaa ggagataaaa
180

gcagaagtgc aggtaagaga aggagtattc atattactgc aggaatccaa attctacaag
240

aacatacgct tgttctttaa agggagccct cgctttaagc attatttgcc gatcaatacc
300

mattaataat tacaggacat gggcaaatat tctagcaaag gtcctgaggt cttcctagat
360

gtgccccaaa taacaaggac tcccccgca ggtggcacca aaagtcggag gatgtgttgg
420

aagctgacct tccccacctg tccctgcctc acccacagtc acccaccct ctaggatgta
480

atgggtgaatc attcaacaca tgggccacct tgggaagtcg cttaacctct ctcccttttc
540

acgttataaa atgggcatag tcattacaca aatgcataaa tcaacctcaa ttttctcatc
600

<210> 2
<211> 600
<212> DNA
<213> Homo sapiens

<400> 2
ctgacatagc cgcctgtgca ggtttttggg accaggtctg ccatctgtat tcccaggctg
60

ggctggagca cctgaggcta ctatggttac agaaaggaac ctggcatccg aaggatatagg
 120
 ctcaaatccc agctctcatg attgctagct ttgaggtttc aggttcattt cttaagtgtt
 180
 ctgagctttg gtgccctcct cttcaaaatg ggaatgagga gcatgtctag ctccctgttt
 240
 tggtgcagtg aatagctaca tgaagaacag ctaaaccaga acagaaagaa ggcaaatgga
 300
 rgctacagaa accaaggatt tccttggtga atcgaatctt ccttcaatct tccttcacca
 360
 cactagtggg tctccctgtg ggagggatgt tgagagtgtc ccgtgttttt tttgtttttt
 420
 tttttttttt ttttgtgatg gagtctcact ctgtcaccca gactggaatg cagtgggtgcg
 480
 atctcggctc actgcaacct ctgcctccca ggttcaagtg attctcctgc ctcagcttcc
 540
 caagtagcca ggactacagg catatgccat cacaccagc taatttttgt attttttagta
 600

 <210> 3
 <211> 600
 <212> DNA
 <213> Homo sapiens

 <400> 3
 gctagtcctt ggctgttctg tgatcaaatg caattgttcc ccatttctta agaaacagaa
 60
 tctttattag actcagttct tttcttggtt atttttacaa ttttattgtt ctatggttat
 120
 taaaaaataa catgtcatat aaattaagcc tagaaacatg gatgtttatc aaagaaagga
 180
 ttgtcatcca aagcaccaac tatgagagat atctatgtgc aatggatatat agatctgtca
 240
 tagaagggtt taattatata tgcctaataga ttttctctcc ttaatgcctc tatagctgtc
 300
 ycttagcgaa gaataaaact gtggactgac cccacccat ttgcgaagaa agtactgggt
 360
 cttcagcttt cattgttcag ccggtgggtc ttgtggacaa caccaggggc ccctcctgct
 420
 gaagactgtc tggacttcct gattttcact cctgtataaa cccctcccc cattactacc
 480
 tttgtctcct tggggaaaaa aaatgttttt cacttaagca aaataataac taccagtat
 540

ttgtttctgg agtgaaattg taaacttcat tgtacattct tcttggttt gatgcttata
600

<210> 4
<211> 94080
<212> DNA
<213> Homo sapiens

<400> 4
ggttgacaca aacacgggaa gattctgcag gaaacacaag gcagccacgg cctctgtccc
60

tgagaggagc agtgaggag ccgcagggt gggctgcggc gtgacactag ggacgtgggg
120

ggaatgcaag tgggtccaga cctgggggcc accagtgcag gagcaagggc aggccggcca
180

tcagagctgg gcggggaggg gggcgctgag cttggcgtgg gagtttcccc aacggactgt
240

gtcttcagg gaccctcgcc cctcatgccc ctctgcgtcc tgaaggagg aggcaccacc
300

cctctgtggc atctgctttt tagggctgac gatggaacaa gctcatggac aggagtgaag
360

ccccggcac atagcagggt ccacagaagg gtggtgaagt ctgccctaga tcagagatca
420

ttcagggggc ctgttaactt ggatgaagaa gaaatacact cttatcacta acatcaggaa
480

ggtaagcaag gaaccacaca tagaattgac aataagtgtg actttgccac caagaagaat
540

cacaaagttt tattttattt tattttattt tattttactt taagttctgg gatacatgtg
600

cagaatgtgc aggtttgtta cataggtata cttgtgccat ggtggtttgc tgcacccatc
660

aaccggttat ctaggtttta agccccgcac gcattaggta tttgtcctaa tgctctcctt
720

ccctcgtct cccaccccca aacaggcccc ggtgtgtgat gttcccctcc ctgtgtccat
780

gtgttctcat tgttcaactc ccacttatga gtgagaacat gtggtgttta gttttctgtt
840

cctgagaatc atagagtttt catagcccag tacaataatc acagatacat gacatagagt
900

tgtgtgacat gtagaccca atagtccgt taatgggatg ctttagttaa aaagtacata
960

gattactgta ttataagttt atgttttagg atattttgat gattactttt tagaaatagt
 1020
 ttcctttgta acctttattt tacttaatgc atatgaaaga atcattcaga ggaggggtcc
 1080
 atggcttcac caagatgaaa caggtcattg gcatgaaaaa ggtaaaaacg ctgtgccttg
 1140
 gaaccaggag tttatgatcc aggcctgcag catgctgtat gaccttgcc acataacttc
 1200
 attgtgcctc agttgcctca actataaaat ggggataata aacctgtgac atatgatgct
 1260
 tgggggggta aatgtgacaa aagtcttagc actctgtgtg ccatacagtg agggccccag
 1320
 catgggcatt gatgcaccgt cttgattcct cctgccttcc ccaggtagag ggagcttccc
 1380
 cagcctgcct gcatctgaga agattccacc agagaaacag aaccaatagg agacatatat
 1440
 taagagttgt gttgcggcca ggctgggtgg cccatgcctg taatcccttt gggattacac
 1500
 tttgggatta cagcactttg ggaggccaaa gcgggcctat cacctgagtt caggagttcg
 1560
 agaccagcct gaccaacgtg gagaaacccc gtcctacta aaaatacaaa attagccagg
 1620
 cgtggtggca ggcatctgta atcccagcta ctcgaggagg tgaggcagga gaatcacttg
 1680
 aacttgggag gcggaggttg cagtgaactg agatgggtgc attgcactcc agtctgggca
 1740
 acaagagcaa aactccatct caaaaaaaaa aaagagagag atgtgttgca agaaatcggc
 1800
 ttacgtgacc gtgggggctg gatccatagc gcaggctctc aggaggggca gtccagaact
 1860
 cttaggtgta tgctgcagga cacagggtgg atttcttctc aagaaagcct cagctctgct
 1920
 cctaaggcct ttcaactgat tgactcatgc cactcagat tattgagggc tatctccttt
 1980
 atttaaggtt aaccggccca gtgcagtggc tcacgcctat aatcccagta ttttggcagg
 2040
 aagattgctt gagggcagga gtttgagaca ggactggaca acagagcaag agcctcatct
 2100
 ctacaaaaa taaaaaaatt agcccggtgt ggtggcacac tcctatactc ctagctactc
 2160

tggaggccaa gccgggagga tcacttgagc ccaaaagttc aaggctatac tcctagctac
2220

tctggaggcc aagcggggag gatcacttga gcccaaaagt tcaaggctgc agtcagctat
2280

gatcaagcca cttatactcc agcccagaaa acagagtgc accctgtctc taaaacaaca
2340

acaacaaaaa accccaacca attgtagata tcaactacat tgctgtggct tgaatgtctc
2400

ctccaaaact caatctgaaa cttaatccct aatttaatag tattggcagg tgacagggcc
2460

tttcagaggt agttgggtca tgtgggcatt catggattaa tgggttaata gattaatggg
2520

ttatcatggg agtagattag ctatcacaag agtgagtctg ttacaaaaat ccagcttggc
2580

tctctcttgt gagccccctc ccatacgatg ccctgcactg accattagca agaaagcctt
2640

catcagatgg accccctcaa ccttgacat ctcaatcttc agaactctaa gaaataaatt
2700

tcttttcttt ataaattgcc cagtctctgg tattcagtta tagcaacaga aaatggacta
2760

agatacacat ctacgagata cctccacagc aactcctaga ttagtggtga tggaataact
2820

ggggattgta gcctagccaa gtcgacacct aagaccggcc ctcatgtcac ccaaggacaa
2880

gacagcgagg tcctgaagag tttcatgaag aggaggggaa ggagagagtg gggagctcaa
2940

tgtggacccc tcaattaatt tccaagtgc tcaatgagge ctgccccacc tctcatccag
3000

ctccttggcc aggcctgccc cccactcacc tctcttatgt gggacacgtg agcccccca
3060

gtgtgccagg agcacatggg tacaggcaca catagagaat gccgtgtcca caggaaaccc
3120

agtgagcctg aagggccttg gcgtgccatt tacaggtcac tcacttgggc caggcctcac
3180

tagggatgcc aggtgaaca agacgtggcc agtgctcttg aagggtccc acctggtgga
3240

ggtagcagac aagtaagggg acaaaggccc ctgagaagta cacatctggt gggccgggcg
3300

cggtggctca catctgtaat ccagcactt tgggaggccg aagcaggcat cacttgaggt
3360

caggagttcg tgaccagcct ggccaacagg tgaaaccag tctctactaa aaatacaaaa
 3420
 attagccaga cgtggtggcg cgtggctgta atcccagcta ctcggaagc tgaggcagga
 3480
 gaatcacttg aaccgggag gcggaatttg cagtgagccg agatcacgcc attgcactcc
 3540
 agcctggaca cagagcgaga ctccatctca aaaaactaaa aaaataaata ataaaaagta
 3600
 caaattaggt gtggttatcg cctggagagc caggagcatc ctggggaggg gacttgggga
 3660
 ggggtccgct gggctgctct agcccactct gggccagaat cgtgattcta attgcaggta
 3720
 ggctttaccc tatcaaagcc ccctgggctg gggactctgt gaagaagcaa gatgtgcttg
 3780
 aaacctccct ctctgctgaa aaatctagac aagctcctga ccggctgcat tgctgcttca
 3840
 gccctcctc aggaagcgg gaggccctgg gtctctacgg gtgctctgtg aaagactgca
 3900
 gaggtggtg gggctgcatt cctgtctgtc tgcctaggac agtgctact gacttctgac
 3960
 tgctgcccc tggtcaggg ctccaggtac accactgcaa tgtcactttg tgcttggggt
 4020
 tccaactaga caaaagcccc acgttgaca gcccttgagc cgtcccggt agcatgcccc
 4080
 ctggtggccg tgggcagtta ttaccgttaa tcctttcttg ccagatgcc tctagcgtag
 4140
 gccacagcag ccggccgggt tcgttgtgtc tggaaagacc agagaagggt tcccggagaa
 4200
 agaggccgtc tgatcggagt tttatgaacc agaggataaa gaaaggaagg gatttctggg
 4260
 cctcagaaaa ccacatgcca acgcatggg gtattaagga taggggctct tgaacgcccc
 4320
 ggctggtaaa atctagaacg gcagaggag gaggggtaat tttctggtt ggtagaaata
 4380
 aagtgggat gaggaagagc acgaaagtct gtcggcaaac ctgaaagtgg gagcgagctt
 4440
 ttgtcaaatc cagggtaac ctgagctctc attaatggg aaactggaga gtcagatcct
 4500
 ctaaagaaag tggaaatatt tgttctaaag aaccctgcc aggtccccgt ttacgtgact
 4560

cataaagagg ccttcaggca gaaaggatcc cagtgcacca gtaacagaag ggctccagga
 4620
 ccagagggaa tggacaggga accaccaccc acaccaagca caggccatt gtcggcatga
 4680
 ggagcctcgc ccctgcagat cacctcccag aggggagcgg atcagagatt aaattccaca
 4740
 cctttctaca aaatcatagc agtttgaaag tgccaagggg agtaggtgag gaggtttcga
 4800
 ggcatgtatg cagccgtttg gagacaccag ccatgtcgcc tgagttttct gatgtaatcc
 4860
 tcataacaac cattccagat ttaagtgatt agtcctgttt tacagatggg gaaactgaag
 4920
 ctcagagagg ttaagcaact tccccaaggt catacagtag ccaagtagca taagtggat
 4980
 tcgaaccacag ttctatttgg agccttcaca ggctttttgc tgcattccac aagagtctgt
 5040
 atattcagaa gcaagtcttc cacagccatc ataagtcac ctcacagccg gctggcaggg
 5100
 aggggaacatt gccactgtag ctgggatcat agatatcgtg gtctccattt gacggataag
 5160
 actgaggtct ggaagagagt cagtcagtct gtgtgataag cctgaaccac cctctccttc
 5220
 agttcctccc ctaacccaac atcctccggg tctggaagga cagtgtcttg agcaacttct
 5280
 tgcagctata cttctccagc taatagtcta gaaaaaccgag cctgggtccc agtctcctcc
 5340
 tcgggggtacg agccgagggc actgatgggg gtcaggaggt gctggctaaa agacgggtgg
 5400
 gcaggtgccc ctggctttaa gcgcgccag gttataaaaa gcttggttcc catgctggcc
 5460
 ctaattctcc atgcactgag taccttaggt cccttctctg ggtctgttgc tcccctgtaa
 5520
 aacaagggg tgggagaaat gaccttcaag ctcaaaccac atatgtccac ctgggttggt
 5580
 atagcagatt gcattagttt gtaggactg tttaacaaa tgccacagac tgagtggctt
 5640
 aaacaacaga ctttattct ctcacagctc tggaggctgg aagtctgaga tcaaggcagt
 5700
 ggctggcagg gttgtttctc ctgaggctc tctccttgta gccaccctc ttctcctgt
 5760

gtcttcacat ggttgccct gtgtgtgtgt gtctgtgtcc taatcacctc ttcttataag
 5820
 gacaccagtc acagaggatt accggccacc cataggactt cattttacct taattatctc
 5880
 tttaaagtct ccatctccaa attagtcaca ttctgaggta ctgggagtga ggacttaaac
 5940
 atacgaattt ggaggatagg acacaattca gcctataaca cagatagaag aactgctca
 6000
 ttttaaaaaa gaagtagcag gctgggcgca gtggcccata cctgtaatcg cagcactttg
 6060
 ggaggccgag gcgggaggat cacctgaggt caggagtctg agaccagcct ggccaacatg
 6120
 gtgaaacccc gcgtctacta aaaatacaaa aattagccga gcatgggtggc atgctcctgt
 6180
 aatcccagct actcgggagg ctgaggcagg agaactcactt gaacctggga ggtggaggtt
 6240
 gcagtgaact gagatcgac tgctgcactc cagcctgggt gacagagcaa gactccgtct
 6300
 caaaaaaaaa aaaaaaaaaa gaagtagcag tcccagccct ttaacagcga acttctgata
 6360
 catttaaaac agaactgat tctcactttt tcttccaggc ccacttcaat gccactaatg
 6420
 aggcctatct gtctgcaatc tgatattaga ggccaaaaat aatctttcta tttgccactt
 6480
 actgggcacc aatgtcctac ctgacacatg ctggctcagg tgggggagga ggtacgtcc
 6540
 ctggccctag gaacgctcca tcttagaaag agacagaatg tgctggggtc agagtaaaaa
 6600
 gtgcatgggc ttcagcagca gatcccagca tggctaccac cagctggggg accttgggca
 6660
 agagccccct gtgagcctca gttttctcac ctggtcaatg ggactcacac tactgcggca
 6720
 cacattgtgc gccacattg tctcgggtgca caggggccac ccttcaaact gccaggtct
 6780
 gcacctctgt gcccaaggtg cagaaagggc tggctgccag aacctactct gcattttgcc
 6840
 tctttgcacc cgaggctgga agagccctgg aatcactgac tcgctggagc atccattcac
 6900
 cagtggccat gggagctgat gcataaatat tccagcgctc tggctcctcg agaggacac
 6960

tctgtgacat gcgttgctca cctgtgacat gcccctocca gcgtctctg acctcccctc
7020

cctaatgcac tacttgcaact tgctctcatt cccattctgc ttgctcacag gctggatcac
7080

acaagaccac acatgcgtgt aactcgtgc ctgcccattg caagtattca aaaactgctt
7140

gctattttgt ggggtggtgt gaagggaata gtatatttat aatacttcag gccaggaata
7200

tacatctcct ctgttttctg acccttccag gagtaggggtg gattccaccc gcttttttgt
7260

aagccctcag aagtgtcca ggctctctgc acacccccag atctcccctg agcttcagcc
7320

ctcaccacag tccccagtc atttggtccc tgtgacgatg ttggtggggg gggcggggg
7380

aggggaaaca gctggccctg gctcctgaca ggcaactctg tcccatctca gctgcctggc
7440

cctgcctgtg gcacccctag agaccagtg gtgagtactt ccttggccct ttaggccct
7500

ttccccaatc cagtagtgag acacggggag gggctggcag agacacaagt ctgccaccct
7560

ttcctggaca gccttctcct ctctggcagc aggaaccac cccctactcc cactccagac
7620

ctacgtcac actccccgtc aacccaagt gccactccag tctgggggtc ctccctgagc
7680

tttcataagc attgcaaatg cccaattcct tcattggctg gggagggcaa agcctaaatg
7740

ggcagtgggc tatagatggt ggccttccag gcactggcag caggacactt cagggcacat
7800

ttcaaatgca gaggggagag cattctgccc aagggcaggc ctatttgggg tgggcagggc
7860

agacttataa tggagactac catacttacg tggctctctg caggccacgt gacccttcc
7920

ctctgctcca tgtgccagg attcaaaactc aggttagtct gattcccaa actgtggg
7980

tcagttttg gaaaactgag aaaattgggc ctcaaatag ggatagggat ggagatggag
8040

atggaagaaa ataaaccagg ccctctgcct gatccctcc ctactttggc tctggggtg
8100

ggagatgctg cagaggcttc atccactctc tgccccgacc cctaactcc ctctcctgct
8160

ctgtccactg ggctggaccc tgtctgctgg gcctatcagg gctccccgtg tgtcctccac
8220

cagccagcaa tggccagaga gccagctgg cccagccta tccctgccc acctgggcc
8280

ctgctcagag gacaaggctg aactgattct gtcaggaagg gttagtcttg tctcaggcct
8340

ggctgaatgc aaaaggacga gacaaatgca tgtggacagc ctcgatccgg agtgctctgt
8400

caaacatgct ggccacctct ccggaagat cactcagact ccttgcattc catcttggga
8460

aggtggcct gaggtccag agagatgttg ggagatgcca ggtgctgca agccaagggc
8520

cctccgggga aggtgtcag gaccgctgc ttctgcctc ttgcagagag ggcaacagct
8580

ggctccttct cctgccccag gtgaagactg atcagatact gattccagga tgaatccagc
8640

aacattaggc cctgctggt cccatggtag catctggctg ggcttttggc ctatagtcga
8700

gtagatagag ggagggctga taactaagtg actacaacag gacctcctg tgtaagggtt
8760

ctagcttatt tccacactct tctggcacag acaggttggc agggcagctg gtagcaggca
8820

cagcagatgg aatgtcagga atgaaatcca caagactcca agccccctcc agggtagct
8880

tgccacaggc cacaccagga agctgagctc ctcacccaaa gcctgccagt gcctacgtgg
8940

cagtcatcc cttgtcaaca gacttccac tggcttctct tgcctcatta ttaaaagcaa
9000

gaacccaaaa cttactgag gggatagaat ccactaata agcatgtcca cgctttgggt
9060

catcttaggg gcacagaaaa tctgagtatc ccccatagca tcaggcagag cagatttctt
9120

gtcaattagc tcagttggtt ggaggtgca gccatgaaca tgctggcctc agggccctca
9180

caacacaggt gggttaggga atgaatgaaa gttcagctat tattcctgcc agagactcca
9240

ggcccaacct ggctctaata gagctatgtg actttgggca ggcagaaaag tgtctctgct
9300

atggtttgaa ggcaccccc aaaggcctcc aatgcaacag tgttgagaag ggggaccttt
9360

aagaggtacg tgggtcacga gggctctgcc ctcacgaatg gatttgcct catcgcagga
9420
gcgggctcat tatcatgaga gtgggttggt tataaaattg aatgtggccc cattccttct
9480
ctccctttcc cccatgtaag gccccaccag acatggccct ccagtgggtg agttcccaac
9540
ctccagaacc acgagacaaa tgaactgtta ttttttcata aatgactcag cctgcagtag
9600
tctgttacag catcacaaaa tatgctaaga caccctgag actcatcggc cacctgtctg
9660
tcagctttgc cgcaaaaatg atttgggagc agctatagtc aactttcctc agactacctc
9720
ctgctggcac attacggcca cctgctgtgt gcctgggtgt tttctaggtt cttggaataa
9780
agtgaacagg ccgggtacgg tggctcacgc ctgtaatccc agcactttaa gaggctgagg
9840
tgggcagatc acctgaggtc aggagttcga gaccagcctg gccaacatgg tgaaaccccg
9900
tctctgctaa aaatacaaaa atgtgctggg catggtggca ggtgcctgta atcccagcta
9960
ctcgggaggc taaggcagga gaatcgcttg aatccgggag gcagagggtg cagtgaaccg
10020
aggtctcgtc agtatactcc agcctgggca ataagagcaa aactccaagt caaaaagaaa
10080
aaagaaaaaa gaaaaaaaaa acagaaagaa agaataaagt gaacaaagca ggcaaaaata
10140
cctgccccca gaaagatcac attcagtgga tgtgcccattg ttccttcctc ttcctaaagc
10200
actcaciaat atcaacatto cctgaactc accactgaga gtcacttctc agcgtgacaa
10260
tggtgaccag tcttcacctg ctgtgtaaag caggcctgcc ctatagtgac ccccaaaaca
10320
ggtgcagaag gcaggccac catcagccca tggccagcac accagttaga tacctgctgt
10380
tgccaggcag ctgtctgcca gcccgtaggc actagagagc cttggggcct ttcctgagcc
10440
cttttaaaga agtgcttcct ttagacatgg actctcaggc ctgctcagag cggggagggt
10500
gctgctccgc cccaaggcct tccctccatc ctttctctg gcaatcaccg tgatagcagc
10560

aacgcccagc ctgaggcctt gggtgctgct tccctcagga aaactggcca aagagtagac
10620

agcagatact ttctccatat tcccaccccc acaacacagg ccagcagcac agcagctgta
10680

gacaagattt cataaagaac tggaggacaa gaaaggggag attaaggagc cctaggctct
10740

agggcctggc tcacacaaga atgggtcttc tcttgccaac ttcttggcct ttacaaaggc
10800

cccacctgac ccagcccagc ctctgggca ctggccaagt tccacaccaa gactaccaga
10860

ccccgccaag tccaggctca agagagaaaag ccagaccctc tctccttcca cctacctgag
10920

agaggcctgt cctcagggtg ctggctcaca gatgggggac ccaactggga gtcactccct
10980

tctggaagcc ccaggcccta taaacaggtc cccggcctca ctgctgggcc tcggcctgga
11040

gacagcggct cagtgaaca caaccacgcc agagctgac aggggaaacc cacgcagcag
11100

ttgaggggct tggggctttg cgggtgcagt gaggggactg cttcccttca gtggagttgc
11160

tgccgcaggt gaggggagga agatgaatct gaaggttcta acgtcagccc acaggctgaa
11220

gcgaaggctg gcctgaagtg ttccacctgg gccccacct gagcggctga agagtctgt
11280

ttctcaagg cccacagagc tctcagagg gaagctttag tagctatttc cttaggaagg
11340

caagtatgtg tgcaacattt cctaataaaa tgcaaaggct gttggaggca ctccagcccc
11400

ctcatttagc tctgaggtgt aattaatata taaattaage tctgcagac tcacaaatgt
11460

ccaagaaaaa aaatgaaaaa tataattagt ttgccatct gactccatct gtcagctctg
11520

agctcaatgc tcaacgtagc aaatcctgag ggctttggc tcgcattggc tgggcccgcc
11580

gctgctctg ccctgctcca ggaagtggg ccctccgag cagctgcca gagccgggc
11640

tctggactca gcctccaggg cagcaaggac tctggagtac acagactgca ggctgggagc
11700

cggggcactc ggagagcag ccaggacctt ctctgtggcc ctgagcgctg gctaggggtg
11760

gcaagggcac cccggggaat ttgcatgtgc ctcttgctgc catgttccta tttgtgcaga
 11820
 cctggaagtg ggcggcagag ccctgaatgg gctaagactc gccctcctcc tccaacctgt
 11880
 cacctctgtc gcttactatc tgtaagaaga gaaaggctgc aaaggggaaa gggaactcga
 11940
 gaggaaggca ccgaggagaa gagaggagac aagagggcat tgggggagca aaggaggcca
 12000
 ctcacacagg aggattagaa ttaacaacta cccccacccc agctcccccc aacacacact
 12060
 tacacacgct tgtattacac cggggatctc acacacatcc ggtcactcag tattcagcac
 12120
 tctaggaaaa gaacaaacaa atggcttgctc tttcatcccc caaaattact tatgtcataa
 12180
 aattaatcca tatttgtttc agaaacattt aaataagaca taaggggtat ggcattaaac
 12240
 gtcccccaa atgcctcttc cccttacaag tggggatcat cattaaaaac ttttgtccca
 12300
 gaagtgggta ctcaggcctg taatcccagc acttagggag gctaagggtg gaggatcact
 12360
 tgaggccagg aggtcaagac cagcctgggc aacactgtga ggtcctgtct ctaacagaaa
 12420
 aaagaacagg aaaatttaaa aggaggaatg aaaacaattg tttaaataat aatttgggtg
 12480
 gtattattta agtacatta atatttggtg caatcgtttt tgaaaggag attgtactat
 12540
 acatgttggt tggaactta ggtgctcgct caggcaccag gccacacaca tcttgggcat
 12600
 tccacatctg ggttttgaaa gtaaaattat aataagaaaa taagcaaagg tcaaaatggg
 12660
 ttggaaaaaa attatttaca cagtgaagtt tctactttat cataccaaga agatgacttc
 12720
 tctaattatt attatttttt tttaagatgg agtttcactc tgtcaccag gctggagtgc
 12780
 agtggcgcaa tctcagctca ctgcaacctc cacctcctgg gttcaagcga ttctcctgcc
 12840
 tcagccttcc aagtagctgg gatgacaggc tcccaccacc acatcaggct aatttttgta
 12900
 ttttttagtag acatgggggt tcatcagggt gccaggggtg gttgcgaact cctgacttca
 12960

ggtgatccac ctgcctcggc ctcccagagt gctgggatta caggcgtgag ccaccgtgca
 13020
 cccggcccct aattattatt tacgtagaac aatatcaaag atatttttct tcctaactct
 13080
 atactccttg acaagaaaaa aaaagaaaag aaaggaagag cccgcgatat ataatggtac
 13140
 atttgaaaat gaagcatttg tatgtgcgct cacatgagta catatatatg tattacatcc
 13200
 atatgcatgt atgtgcatag actgtttcta gaagaattca caagaaactg gtaacatttg
 13260
 ttgcctctgg gaaacagatt ggcaactggg agaaggtagc tgacttcatt ataaacatat
 13320
 tttgtaatgt tttaattttc tgtcacatgc attattactt ttttccttag aaaggatttt
 13380
 ggttttgtct attttatcgg aagaggcaaa ttttaacaat aaaaacaaa tgactgacat
 13440
 tcagccatag tcccacctct ctgaattact cagatctact gatttacatg tagattaaat
 13500
 gcaaatgagg gtggcactat ggttgattat ttagcagctc agcaatatga ccaatgcctt
 13560
 gttcccttct acccttctgc tctgccccct cggttgtag actttgtcct tgggcttcct
 13620
 aacctcattg ctacaagagt caggcacatg atcaaaacaa gacactgcc caaggtagga
 13680
 aagaagccac ctcttccttc ctatttttag gagaaaaatc tttcccccac agccctcctg
 13740
 cagacttagc ccacatatcc cactggccat atttatttcc catgtctatg cctgagccca
 13800
 tctctggcaa gggaaaggga cccctttgga catcccaggg caagcaggcc tcctcagtgt
 13860
 tatcagacca actacccttc ttgaaaacaa tgactgtatt cttgtaacca cccaacctca
 13920
 atccagaaac aaaattcaca gatcacagaa tgtacctaca aaaacatatt tcaaaaaaga
 13980
 atcaaaaacga agccctagct gtcacgtaaa ggagaaagta aaagtaattc aagacaaaag
 14040
 agtatgtttt ataataagca aaggctcagg cagcagactc tacagaacat aatgaagtgg
 14100
 tgagatgctt gcagctgcac acggaaatgc caaggttgcc ccatcataca tatagacaga
 14160

tgcgggagat gactcaccca gcacgaggcc attgccactg gcggcatgat tttccaaaat
 14220
 ggcgacaac tgttggtaaa gcttcaaaca aaacacagtg caacctcccc ttgatttata
 14280
 cagtaggata cttcctagga aatttagggc atatttataa acatgcaaaa agatacttgg
 14340
 agctcataca tcaaatggaa tgaagtttta ggcttaaaaa ggactttcag ctacataaaa
 14400
 gtctcatgag acatcagaaa gtcacgtggg acatggagcc atttcctctt agaaaggact
 14460
 gtctcacaca ttgtaaggta tttagccttc ctagaccccg tttattgaaa gatgatagca
 14520
 agggccctcc ccacaaaatg tcaaagcccc caccaatttc cagaacacac ttagggaaatg
 14580
 gttctgtccc tttcgagaac tcttggtta ggccaacctg cctctcccag ccaggtgac
 14640
 ctcagagctg ggtggttgc tgaacaacat tcaggttctt gcaggaagca tagaaggcga
 14700
 ttagaggttg gatagatagt caaccgtctg ttcacacag aagcatttgc acatatttgc
 14760
 ctctagcttt gcctttgtat acatattttt aaatacaaat aaaatacgta tttaaaatat
 14820
 gccaggtgt gttgactcat atctgcaatc ccagcacttt gagaggctga ggtgggagga
 14880
 tcaacttgag cacatgagtt tgacaccagc tgggcaacat agtgagatcc catctctata
 14940
 aaaaattagc caggtgtggt ggcattgtgc tgtggtccca gctactcaag aggtgatgt
 15000
 gggaggatca cttgagccct ggaggttgag gctgcagggg gcaactccagc ctgggcaaca
 15060
 gtgaaactgt ttcaaaaaaa aaaaaaaaaa aagacaaaca actgatttac cgaatgggca
 15120
 gaacttcatt gcacatactt ggctttcttt ttctttttcc cccatcactg tttttcagt
 15180
 ctggctggtc agctctcttc tttcagccct tggatgacag ctggggctgt cccatagctt
 15240
 ttcagtcact ttgagctatg tcataactgt ggtcatattt tcaagcatga aaagtacaga
 15300
 tttgtctcaa tcaaaactca gtctcacaaa agccactttc ctcccagct cagtgccac
 15360

ccagtgggca ggaggagggg caggctctgtt ccttttctcc ctgctgccca ccttcagctc
 15420
 aagttcgggt ggggaagagt tgattagagg atggggcaaa gtcttagagg gccggacgtg
 15480
 ggtgatccat cacacatacc cattctatct tggcctctag ggtggatttt gaggagcccc
 15540
 ctccccatga tgatctctgc actccacaag gagttcctcg atgtgacggg gcaactctcta
 15600
 gctgggtcttc tgtggcatgc cccctctggg acccagtggt tcacccctcc caaccagcct
 15660
 ttccctgggg gacctctcct acccagcagg tgagtctcct gagtgggctt ctttcaagat
 15720
 acctgaaaaa ccaactcctg ggagggtcac ttgctccaag gaaacttgta catctttgcc
 15780
 ttgcctaaga gtggaggagc agcttcccca caactgacct tgtaatagaa aaaaacaagg
 15840
 tgttttgcct attctcacac acagaactcc caacaccaga tgtgtggagc atttcccaca
 15900
 taccaggcaa ttctcccttg gacaccagct ggggtgcctc taattcagtt cagttttgat
 15960
 cctatctacc tggagatagc atcagatccc acaggctgat ggctcagtc tgcgagactg
 16020
 cccctctaga gatgccaatc tcaagtgcag gttgtgacct gtgctttgaa ccattgacta
 16080
 tatattgggg ctccctcgac tccctcgttg ggtgagatta atctgctaga gcagctcata
 16140
 aaactcaggg aaatgctcta tttatgttta ctggtttatt aataaaagat gtaataaaga
 16200
 atacagataa gcaccagatg aggagataca tagggtaagg tctagagggc cagcagcttc
 16260
 tgtctctgtg gagttggagt gtaccaccct cccagcaagc agatgtgttc tccaacatgg
 16320
 aagttctctg aacctcatac ttagctgagg gatttttttt tttttttttt gagacagggt
 16380
 ctactccgt caccaggtg ggagtgcagt ggtgccatct cggtcactg caacctccac
 16440
 ctcccggtga caagtgattc tcgtgcctca gcctcttgag tcgctgggat tacaggcgct
 16500
 tgccgccaca cccaactaat ttttgtatth ttggtagaga cagagtttca ccgtgttggc
 16560

caggctggtc tcaaactcct gatgttaggt gatcccccca cctcggcctc ccaaagagct
16620

gggattacag atgtgtagct cagggatttt tatggagtct tcatcacata ggcattgattg
16680

attttcaagt ccatttccag cccctttctc ctctccagaa aatgggtcat ggggctgaaa
16740

gctccaagct tccaatcatg gcttggctctt tctgctgaca agccccatc cagaagccca
16800

ccaagagttg cctcaataga acaaaagatc ctctttttac ccaggaagtt ccagggactt
16860

taggagctct gtgtcaggag ctaggggtcaa aaaccagata ttagaacaaa agattctcct
16920

agcacctgta tctacaagga tattaggagt gctgtctcgg gaataaggac agagacccaa
16980

cattggaaca aaagattatc ctgacacttc tatcacttag aaaatcacag ctgacccttc
17040

aacaatgcaa gggtaggag tgctgacaac tgcccaactc ctgcctcgca cacggttgaa
17100

aatttgtgta taacttttga ctccccaaa gacttaacta caagcctact gttgactgga
17160

atccttatgg gtaacataaa cagtcaatta acagatatat cgtatatatt ttatatactg
17220

tattcttaca ataaagtaag ctagagaaaa gaaaacatta ttaagaaat cataagagaa
17280

aatatattta tgggtgctcta cggtatattat tgataccata ggtctgcacc atgtgtttac
17340

aacatgtatc atctgtctga aacgtcaggc aacctcagct gctgacctta gtctatggta
17400

catatgagc aattcaactt gttcctgtga ggtcatgacc ttacttctta ggagctcttc
17460

cagcatcact agtggcactt cgtatgggtc ccatgatgct attcaagggt tacagtattg
17520

cactaaccat gaaaataaat gagaactgag agagatcatg gttttcactg caatatgcaa
17580

tttactggag attgcaatag taaattgtaa attgctatat gcaatctact gctcacgtgg
17640

agatggtcag tatttcatag tgttttaagc agacactcgc aacacgagct caccgcaata
17700

gcaacaggag gtagctacaa aaatattaca gtagtacagt gcttactaca gtgaatttta
17760

tgtagttatg attgaatact tcatctttac atttgtttac atttccctca actacaaatg
 17820

gtgccatgta tgatctgtgt tttttgcatg ttttcataaa ttttaacttt ttataataga
 17880

tttgtgtaat tttatagtag taacttataa aatagactag tactcgcata tactttatgc
 17940

attcatgaca tacctaactt tttgttggtc ttttttgata tttctgggct acacaagggc
 18000

agtgacatga tcgtagcgca ctgtaacctg taactgtaag ttttttcaaa ctgtcacaaa
 18060

tctcaaaaaa aaaatcta atatttaatg aaaactatcc aagtataagt tgacccttgc
 18120

agttcaaacc tgtgttggtc aagggtcaac cttacaagga ttttaggaac tctgctggga
 18180

actggggaca gagagcaatg tgtatatatt gtgttatctc aggtgcgttt gctcatgcct
 18240

ataatctcag cactttggga ggctgaagg ggagaattgc ttgaggccag cagtttgaga
 18300

ccagcttggg caacacagag agatcatgtc tctacaaaa agatcccaaa aattcatatt
 18360

ctccctcctc actttacctc acacgatgct ccaaatcagt ctccacctt aagaattctc
 18420

cataagcaaa agcccatctt gaatcccaca ctctcaggga cccttgccac actatcccc
 18480

caattttctt ttctttcttt atttatattt ttatctttaa gacagagtct tgctctgtcg
 18540

cccaagctgg agtgcagagg tgcgatctcg gctcagtga acctccacct tctggattca
 18600

agcaattctc ctgtctcagc ctctgagta actgggacta caggtacatg ccaccacagc
 18660

tggctaattt atatatatat atatttagta gagatggagt ctcaccatat tggtcagtct
 18720

ggtcttgaac tcttgacctc agatgacca ccgcctcag cctcccaaaa tgctgagatt
 18780

acagggtgta gccccgtgc ccggcctccc aattttctca ttggccagtc actctggtgc
 18840

cctggggact agcatccagc taagggatag atgccagctg ctctctcagt gctatggttt
 18900

gaatgtttgt gcgcctcca caattcatgt tgaaaatccc cagcacaaca gtatcgaggg
 18960

gtggggctct taggaggtga tgaggacatg aggctctgc agtcatgaat ggattagtgc
19020

cctatagaag ggcttttagag gctgagtcac accttccatc cctccaccat gtgaggacac
19080

agcatttgct cctccagag gatacagcac aggtgccatc ttttaagcag tgagcagccc
19140

tcaccagaca ccaaacctgc cagtgccttg aacttggact tccgtcctct agaactgtga
19200

gaaataactt tccattatct ataaactacc caggctcaag tgttttgta cagtggcaca
19260

gacgtactaa gacttggtgg taccacaat tccataagcc attttctttg cttggcctgg
19320

gcatgacccc ctgtgctgct cccgtgaaga aaaaggataa accaacaac tctatgaatt
19380

aacttcagt aatctcgtat gtttctaata ttctgctttg gagctttggc gataaagaaa
19440

gattaggaaa ttagattaaa aaaaactaac attatcttaa atgaccata ttcccaatca
19500

atatctcttc ccttgaaact tgggtgtgat aaatccaata ctttttaatg tatttgata
19560

tacatatgta aacatatata aatatatata tcatatatat gtatgtatct gcctaaataa
19620

ccacctggag ttagcaggtt ccatacctgc aggatgcaag cagggatggg cagaggcagg
19680

ctggtcagct tccaggtata tatatatata tatactcgga aaatacttgc tattagtttg
19740

gtggttttat tattaacgta aatgatgtct tgcgcacat attattctgc aagttggttt
19800

tccacataa tgtgttgga atcttttagc ccatttccac actcttagcc tgccttttgc
19860

acaagagttg ctgtggcaac aagctgccc cagggtgtaac ctgacagctt cacttcagct
19920

gcaccaggta tgcctttctt tctgccccag gcatcacggt atacaggcaa atcttgaagt
19980

gccaggaggt caaccctccc tggagcaatc cttgccaat aggagaagga agccaatgaa
20040

taaatatgcc aggcctcgag gtcttaggcc ctcagtctaa ggcatagtct acatggtttc
20100

tctgggagtc ctcagttaaa ctgagccccg actgccaca gcaatggcca gtgttaaaag
20160

aacaacctta accagattaa atttaacaca gtttaactga gcaaaggaca attcgtgact
20220

agggcagtct gctgagccag agtaggttta gacagactct ggtggaagat ttatggacag
20280

aaaaaggaaa gtgaagcaca gaaaacagaa gtgaggttca gaaacagctg aatgggtcgc
20340

agctcgtcgt ttgccttatt tgaacacagt ttgaaaagtt ggccaccttt gattggccaa
20400

aactcagtga ttgacacaag agtacgttac agcctgttta catttccttg taggctatag
20460

tttactatgc acagagaaac ctttaggctg aacttaaaat atgtaaggag gaagctttaa
20520

gctaaatttg atttttgttg ttgttggtgt tttgtttgt tttgtttttt gttttttgag
20580

atggagtttc actcttggtg ccagggctgg aatgcagtgg catgatctcg gctcactgca
20640

acctccgcct ccaggttca agtgattctc ctgcctcagc ctcccagta gctgggatta
20700

caagtgtgca ccatcatgcc caattaattt ttgtatattt agtagagggg gtggtttcac
20760

catgttggcc aggctgggtc ctaactcctg acctcaggtg atccatctgc ctcagcctcc
20820

caaagtgtcg ggattacagg cgtgagccac cgcgccaggc cactaaacct cagcctccca
20880

aagtgtggg attacaggcg tgagccaccg caccaggcca ctaaacctca gcctccaaa
20940

gtgctgggat tacaggcgtg agccactgcg ccaggccact aaatttaatt taacagtagc
21000

ttaacacatg gctggatggg cttttcctac ttccttggtt cattcttccc agtcctccaa
21060

ctcttgactc ctagatttat ggacccaaat aaatgactca aatgcagtcc ctttgtctca
21120

ggtttttggt ggaatccaag ctaaaacact aaataaaaaa tgttggttaa aataagttaa
21180

ttttggtcat tattgatcca cttggagtgg tcctcaaaat gggatctagg gaatcgggcc
21240

ctgaggctgg tttctgggaa tccacaaatg caaaactatt tccatagtca cactaaggag
21300

gtgtggcctt ttcaagaaat tacattaccc gttaaagtga atagacggaa aacagaagcc
21360

catatgagaa tccagctatc ttctagcaag ccagaagtaa aagatgccag gtgcagtgga
 21420
 tcacatctat aatcccagca ctgaggattg cttgagcctg ggaatttgtg accagcctgg
 21480
 gcaccatagt gagaccccca tctctacaga taatttttaa aattagccag atgtggtggt
 21540
 gtgtgcccac gatcccaact actcgggagg ctaggaagga ggatcacctg agtcaggagg
 21600
 tcaaggctgc agtgagctgt gattgcacta ccacactcca gcctgaatga tagagtgaga
 21660
 ctctgtcaag aaagaaagaa agaaagaaag aaagaaagaa agaaagaaag aaagaaagaa
 21720
 agaagaaag aaagaaagaa agaaggaagg aaggaaggaa ggaaggaagg aaggaaggaa
 21780
 ggaaggaagg aaggaaggaa agaaagaaaa agagatctgc aaaaaaatgt aaaacaatac
 21840
 cactcttctc actaaatttt tgtacttggg aaacatagtc atttttcaca aaaatgctat
 21900
 ttatattaac ataagtattt tgatgtgcaa tggacttatt tttgttattt tcacataaag
 21960
 attgtttaaa tttcttagtt ttaatttcta atactgtaaa tatcagtaga tataacccat
 22020
 gtaaacaaaa gctttttggg gtcctcaata attttcaaga gagaaaaaga ttgccaagac
 22080
 taaaaagttt gatagctatt aatagagttc attgctttta acttctgagt cccatgcata
 22140
 cttcattatt attattatta ttatttgtaa agacagggtc ttgctgtgtt gcccaggctg
 22200
 gtctcaaaact cccaggctca agtgattctc ccatcttggg ttcccaaagt gctgggatta
 22260
 caggtatgag ccacagtgcc ctgcctcttc ttttaaaatc taattatgta tttgttatct
 22320
 cttacatttt tcctatcttt cttttttctg attatggccc aaaaggggaa gatttctgcc
 22380
 catgctcatg ttactatctt gaaattattt cttttcagac tctctagttc aggaaaaagg
 22440
 tcccctgcc catcaagtgg ccacggagga gcttgggcct gtgatgtccc ctcccttctg
 22500
 cctaaatgac cacctggagg cagcaggtta catacgtgga ggatgcaggc ggggatgggc
 22560

agaggcagac tggtcagttt ccagggcaca cttcgggggc atacaagga caaaaatgtg
 22620
 acggtggctc aggagggcac caaagatggc caaaagcttc tgcttttaag gtggcagtag
 22680
 gaggcagtta atggaaaaat agaaaccaat caagctacgt ctgccaccaa tgaaaatgaa
 22740
 ttattattta ttcaacaaaa tataagcttt ttaagaaaa aaaaatagca aacgctcaac
 22800
 ccacttgcta agttcgttg ggggtataatg gcatcagctc tgagtgccaa cttaccttg
 22860
 caagggtggt gcacgccctg aggatatgta caccttctcc gctgctggcc agctagatga
 22920
 ccttgagcaa gccaccacc cttctgaact tcgtctattt ggaaaataca gacattaata
 22980
 cctcatagtt tgtaatgagg ccttagagtc cacatattaa tcgttcagaa cacagctcaa
 23040
 aaaaatctaa aacttctggt ggggcacagt ggctcacacc tatgatccca gcactttggg
 23100
 agcctgaggt gggaggatag cttgagccta ggagttagg accagcctgg acaagataga
 23160
 aaggcccat ctctatacaa aatataaaaa ttagccaggg gtggtggcac atgcctgtag
 23220
 tcccagctac tcaagagact aaggtaggag gatcgattga gccaggagt ttgaggctgc
 23280
 agtgagctat gatcacacca ctgcactcca gcctgagtga cagagcaaga ccctgtcacc
 23340
 aagaaaaaaa agagagagag agagagagag aggtggcttc agtgggccga atagcagaga
 23400
 tccactggga gcattagcaa catcagggcc attgcttcag cagcctggaa ctaaagagga
 23460
 agaaagcaag aagcttcagg agattgaaca gtaggaagga aggaaggaag gaaggaagga
 23520
 aggaaggaag gaaggaagga aggaaggaag gaaggaaaga aggaaggaag ggagggaggg
 23580
 agggagggag ggagggaggg aggggagag ggaagcatat gagaaagcga ggctaggagt
 23640
 cttggagtca tccttgactc atctcttatt gtagtcattg ccagtactca cagctctcag
 23700
 acaacagtga ggaaacaagt tattttcaac atttcaggaa gctaaatgga aaccatgcac
 23760

acaacaaggt ttcaaagtta attgaggatc aagattttct gctggtgtct ggtaagttat
 23820
 caattctgat ttttttaaac ctgattacag gcatccacaa aaaaatatgc aactaacatc
 23880
 atgtttaatg ttgaaagact aaatgctatc ctctgggatt gggagcaagg aaagggtatc
 23940
 cactctcatc actcctgtgt acttaaaaaa aaaaaagaaa taaaacatat gtccacacaa
 24000
 aagtctgtat gcaaagtgtt atagtggctt tatttataat tcacccaac tggaacaac
 24060
 ctaacatgac cttcagctgg tgaatgaata agcaaaccat ggtacatcca tacaatgaat
 24120
 tactagacag caaaaaaga aactggtgat ataggcaaca acctgatgaa tctcaaagtc
 24180
 actgtgctta gtgaaagaag ccaggcccaa aggactacat actgtatgat tcaatgttgt
 24240
 gacatactgg aaaaggtaaa attataggaa caaatcagta gttgtcaggg tctagtggct
 24300
 agggaaataa ttgactacaa aaggacagca agggggaatt ttggggggaa ttggaactgt
 24360
 tctgtatctc aaagatgggc atgattgatt acatgattat gccaaaattc attcatcaaa
 24420
 aatgtaaatt tttatgtatg tagatagatt ttattttaag tgaataaaag gaacaggcaa
 24480
 cttggaagca aaagagcagg agccgggcca ggccctgcggc tagattcaga aatagggtta
 24540
 gagggaaatga tcgaggggaa tgaccagggc agccacagct ctggagagac ggctgcaggt
 24600
 ggacaaaaag caggtggaag cgcccagcag gaccttgacc agacaggcaa tgggaacata
 24660
 aaagacagca acagattgca acatcatgtc catgggctcc caaacctgat atctcagaat
 24720
 cacttcaga gctgattgga aagacagatt ccattgctaa gcatagtggc ctgcacctgt
 24780
 agttccagct actccagagg gcttgaaccc aggaggttga ggctgcagtg agctatgatc
 24840
 gcaccacagc gctccagact gggtgacaga atgagacaag tcagaatgaa agaaaggaag
 24900
 aaacaaagaa agagagaggg ggaaagaaag aaagaaagag agagagaaag aaagagagag
 24960

agagaaagaa atagaggggg gaaagaaaga aagagagaaa gaaagagaga gaaagaaaga
 25020

aaagaaagaa agaaagaaag agagagagag agaaagaaag aaagagagaa agaaagaaag
 25080

agagagagaa agaaagaaag agagagagaa agaaaaggaa ggaagggaag agaaagagca
 25140

aaagagacag agggaaggaa ggaagggaag aaggaggga gggaggagg gaagaggga
 25200

agaaggtgag aaggaggag agaggagg aagaggaga gaaggtgaga aggaaggag
 25260

gaggaggaa agaagggaag agggaggga ggaagggaag aaggaggga gggagggaag
 25320

aaggaaggaa agaagggaag aaggaaggaa ggaggaggagg agggaggga gaagggaag
 25380

aaagaaagga aggcagattt ctactgaggc tgcctcttta tccccctctt tattccctc
 25440

agagaggtca tttaacctct ctgatctctg tgggtgctgt agggactaaa ggacacaatc
 25500

cttgggaaag tgcttgacac agaacttttc agcttgctat ggggagtgc agccaccgc
 25560

aacaggtgtg actgcagggg aaagccccgt gcagggtgaga gaaaggggccc agaagagaa
 25620

gggctttgta aggctgggac cactatgttg ggtctgctat ggagggggga gctcaggagt
 25680

aggcaggga gatttgaaca tgggaagggg aaagctgctg ctggaagcag tggcccatgc
 25740

ccagggtgcat ggagaagaga gcctggaagc ctctcttggt ccctctaggc tggcaggagc
 25800

agggcccccac attggcagtg gtgtgaacaa tttccccca tcatggggct gctcatggct
 25860

gtgaaacatg ggatggcact agatttggtc attaaagtaa taggataaag aaataattgg
 25920

aaaaatggatc cctgcctgga aggagtggaa ttgttcctgg ggctattgtt gtccggcatc
 25980

ctgctttctg agtaatctcc cgagccaggc aaggaaggat gtcaggctta gggctgatga
 26040

gaaacttaca cttgaaaaga gctcaaaatt acccatgaca atggaaaagc ctttgtatcc
 26100

tgaacagaga gaggaaaaag actctgaatc ttctctgggg cgggaggaga ggggaagatg
 26160

gagggcagtc tgcagtcaca gaatgtcaaa cccaggaccc ctgtgtcttt ccaagttggg
 26220
 ggcaactggcc aggaaggaat cctggttctt cgggtgtctg ggatgatgag gctggtcctc
 26280
 aggcagtggg aggctggagg agggggaatg tgtgaagacc acatgctttc gctgcaaga
 26340
 gcatccatcc atgtcttggc cactgcaatt tacatcctcc cattcaaccc tcagcacacc
 26400
 caccaaacca agctcctcct tgcaccttga tgggcagaaa gttggagtga tgattatcta
 26460
 agacagagcc ctgaggtgaa gaactggttt tctggtggag ggggtgaggg gaggatggag
 26520
 ggagctatcc ctccctgtc cccgcaaata cttcagcaaa aaatccatag tgttgagtca
 26580
 aaatgccaat cagatgtgcc tctcagcagc tttaggcctg gactttgacc tcagccca
 26640
 cccccacccc cacaggcact cagcctaagc tcccattctg caccctcccc accctggttt
 26700
 gcaaaatgca tcaaagcaga agctgtgctt ttgcctttcc ttgtaacttc ttgtctgcaa
 26760
 gctgtattag atgcttgctg gtgtttaaaa cattttctat ttctctaaaa aaacaagttg
 26820
 gggcacccca aaatgaagcc tccaagttc ggtcaaacat gcaagattgc cactcaattt
 26880
 ctccattcta gatgagaata gcgagtcca aaggctcgtg ggtgtctgag gcgttgcaat
 26940
 cacttaactc acttggtcaa aatacagggtg ccagggcctc cccccgagc cacagaacct
 27000
 aaatccccag aatcaataac tttctttaa taactagatg cttgatttaa agggaaaact
 27060
 ctatttaaag tgctgtatcc ctctttaaag acagattgtc ctacacgaag cagctacatg
 27120
 caaaatgaat tataggagtg tccttggttt tacaaaaaat gaaaagcatt gcaattctcc
 27180
 aaacatgata agaaaggaat ttttttgta aactctgaga gcaaaacaat ttactagaag
 27240
 atacagatgc aagttgagta agcaagccaa taatggagaa aggaggaatt ttacagaag
 27300
 cagtttttcc tcaccccatc tgcatttcat gtctattaaa acacatgtca tgcaacgaaa
 27360

aaatcaaagc agtatgggtg ggtacacaca ttagtggtt tctgtatggt tagtgagatg
 27420
 aaaggacaga gaaagctatg ctggagtcac caggacatgg tagagccctc ttgtgagctg
 27480
 ctggaacaca tcagttgcac gtacgttctc ttgctggtgc agcccatccg gtcagaatct
 27540
 gtcgctcaca gtagctgtta ctgaatactg tctgcttctt agctgaaacc ttaccgtagg
 27600
 accatccggc cctgctcctt tcccgtaaac atgtaagttg ttctcagacc ctaccttttc
 27660
 ccttggtttc tcatggaccg tggccaggac ttgggtgtcc acagccacac ccacaaaaga
 27720
 agtgcctggt ccaccagga cctctgggtt aaatgacctc tctgaggcca cacaggatct
 27780
 acagcacaaa ctcagctcag attataaaac ttgcactctg aacgactgcc ctagccagcc
 27840
 ccaccaaggc catgaggggc ctcccagagc caagtgcgac ggcttctctt ggctctgcc
 27900
 tccttgacct gcaagcaata ttgacccca cttcccttcc cgtctcacac cctcctgacc
 27960
 tcctcagctc tcctttctgg tctctcttcc cctctcctca cacctccaat tggatcttca
 28020
 gccttctctc cctctcactc cacatttttc agctggattc ctgtgggttt cactcttctc
 28080
 tctccctatg cctgccaaat ctgtgtgttc agctgtccga acttctcagc tctacatgtg
 28140
 tgtttgcaaa ggcttcacca tcctcccccac acttctctca ccagaggctg caccatgggt
 28200
 cactgggcct caatcttctt gccactcag gtagaggctc tagggtcagc ttgctctcct
 28260
 gccggatgct gcactgctgg aatcagctgc ccacacttct ctcttttgt atgggtgtct
 28320
 cactgaaact tgggcggtgt cccttgaggt gccttgagag ctgatgattg ccaccctctc
 28380
 taagagccta cagacccctg atggatatgt gtcctggagc cctcacaca agcccagccc
 28440
 tctgtgttag ctagacagtc aagtgggtgc acacagggtc agcaccaccg cggctcagaa
 28500
 gccaccttcc cctggaagcc actccttata tctgggcact ttctcctctg gactaccctt
 28560

gcattgtggt gttctatagg aggtgctgga gacagtgata agaaccatg gccctgccc
28620

tcaagaagtt ctccagtttt attggtggga gcaaacaat aaacagacag acatggagaa
28680

aatgcatcag gtcgggtttg catgcaggaa gttgtcactt gcgagggaat gacagaagca
28740

aaagtgggca gagggagaag ttaaaccaag gtgcagtccc accagaggcc tcagaagaac
28800

ccatgagaag ctctgaaact ggatgggctt tcagagttgt accaaatgga ggcaagaggg
28860

ctggcctcta cccctgtctc atccactaga atgtgggctg ctcccagaga ggggtgtgagt
28920

gagttggggc aatacatctg ccttcagtgg aaagcaatta attcccagag agggttgcag
28980

ccattgctcc tccagcagcc agggatatga gtacctgtgt gcagatgggg acctgggcaa
29040

catgccacag catccatgac acggcacaga cagggatgcc caaagtgggg gctggccaag
29100

tcaaagatgg gggagtaat gaggtgatca tcagatcttc ccatttcacc tgggcattga
29160

cctacatgta ggtgtttttc aagtggacaa tctttagtct cctctgctcc tcatcataca
29220

agtataccag caatggctaa tgttatcttt gcacttgcta catgcccagc tgtgtcttag
29280

ctcttcacat gcatagctta ttaatatctc acaagaacct tacgaattaa gtaatactgg
29340

ctaggcacag tgggtcctgc ctgtaatccc agcactttag aaggtaagt ttggagaatt
29400

gctagagccc aggagttaa gaccagcctg ggcaacatgg tgaaacctca tctctacaaa
29460

aaatacaaac aatttgctga gcatgatggc agacacctat agtacctgct actcaggacc
29520

ctgaggtggg aagatcactt aagcctggga ggcagaggct gcagtgcagc aggatcgcac
29580

cactgcactc cagcctggat aacagaatga gaccctgtct caaaataata aataaataag
29640

tgggccaggc acggtggctc ttgcctataa tcccagcact ttgggaggcc aaggcgggtg
29700

gatcacttga agtcgggagt tcaagaccag cctggccaac atagtgaaac cccatctcaa
29760

ctaaaaagac aaaaattagc cgggcatggt ggcacactcc tggaaatccca actacttggg
 29820

aggctgagac aggagaatca cttgaacctg ggaggcggag gttgcagtga gctgagatca
 29880

tgccactgca ctccagcctg ggcaacagag caagactcca tctaaataaa taagtaaata
 29940

aataaataaa tacataataa aactattgcc cctgtcttaa tgtaaattga acatgaggct
 30000

tagagaggta aaatgactca aggtcacaca gtaagtcaaa ggctgagcta tctctgccct
 30060

ttcttcttcc tgggtgtacac ctttcaccct gcaaagctca gtgcctacct tacttgcctt
 30120

atcaaacttt ccccatcttc ctcagaaaga ccaagtcccc ctctgtgtct ctagagaaca
 30180

agctgccctg gaccacacac cttggcttcc agccccttca ccatcctaga cgaggtcacc
 30240

tagaaggttc catccagact gtgaatgtcc ctctttaact acagtctcca gtgaacccac
 30300

aacctcaggg acatatgacc ttggaggggt aagtgggatt tcatttcctt gactctggac
 30360

ctgggctttg ctagcacacc tctcccacta tgatagcatt cgcttaaggc agctgagccc
 30420

cgctcagggg tccttgagct tgtggtcagc tcaagccatc acatctttct catctgtgct
 30480

ctgggaagct gcctctgctt ggtctacaca gctgagacaa tggagcctgg gtgggctcag
 30540

gggccccttc tgctgggaac ctgctttgaa acaacatatc cattgcattc actggaggag
 30600

catcagaagt agacggtaat ttgtgttttt tcttactagg atgccctagg gtgggttccc
 30660

ccgaagcaga ccctgagatg agcatttatg caactgatcg aggatgtgcc caggagaacc
 30720

ggataagggg ctgagagttg gaggggaagat actaagaaaa tcatgctcca ggcaaagtcc
 30780

cagcctcggc ttgatataaa ttacactgac ttgaggcaag ggagctgggc tctcacacc
 30840

cggctccagt cagccactgt ttcagggtcg cccagaggag acaccagctc ccagcactct
 30900

tgccgtctgt gcatgcattc agcacactgg caccacctca aaggccagcc atttaagagg
 30960

ggcacagggtg ctggggccac aaagcaccca gagcccagtg atgggtgcac aggcactgta
 31020
 aaagtgatat gagggccggg tgcagtggct catgcctgta atcctagcac tttgggaggg
 31080
 tgagggtgggc ggatcaccag gtcaggagat tgagaccatc ctggttaaca cggtgaaacc
 31140
 ctgtctctac taaaaatata ataaattagc tgggtgtggt ggcacacgcc ttagtccca
 31200
 gctactcggg aggctgaggc aggagaatca cttgaaccca ggaggcagag gttgcaatga
 31260
 gctgagattg caccaccgca ctccagcctg ggcgacagag tgagactcac ctcaaaacaa
 31320
 aacaaaacaa aacaaaacaa aaaaagggat atgagaggat cagagcagcg tccccttgcc
 31380
 tctgtttacc aagaaggtag caagccgttg ttccctcagg ttgttcctca ttccatgttg
 31440
 ctcagccagc caaaatcaga agctgccaaag ccacctgggg aaacaaggca cccaccgaag
 31500
 ctcttgtttt caaaagtttg agctaattaa agttgatgga gctggatttc ccctgaagga
 31560
 aaggatacca atgtttggga aggttaggat tcagagaaag aagcatttga agtcttcccc
 31620
 agggcctaga aattaaagtc tagagatttg ggagggtctc tagaagggca ggaacgcagg
 31680
 ccttcacctg aaccacttcc taagcagata gccaacctca tatgggagat tcctaaaggg
 31740
 gctgacatca aatcctgaga ctccatagtg aagaccccaa gctgcagcct cgctggaag
 31800
 tataggaact taccagagta tccaatgaca ggagaagtcc agagggtctc accctgaatg
 31860
 ttcaccatgg ataagtctg ccccaaacct gacagatata gatttctcat ctatcttgac
 31920
 aaatgggaac ttgcagccag tttaatttgc ctcaaagaaa ataaaataat gtgacattcc
 31980
 ttcaccccca gcaagaacta atccctttct ccagtgtctc tacagcaggg aggacatttc
 32040
 cctgatattg atttttcatg actgtccttc ctccctagga gaccctgac tccttaggga
 32100
 agagggtgtg ctttttaaac catttctgca gccaggtgc ctatcaaggc cccagcatcc
 32160

agtgggtctc aataaatgag acctatcata attaacattc ccctacctca gcacatgact
 32220
 catcaaagta ttacatgttg tttgcacttc tctacccctt ccccttccc atcacctcaa
 32280
 agctctccaa gcacacatga gccactgctc ataaatgttg catgaaagaa gggatggatg
 32340
 gatggatgga tggatggatg gatggatgga tggatggatg aatggatgga tggatagaca
 32400
 gacagatggc ctgatggatg aggctggtga gcaagtgaag gagtgcttgt tttagaataa
 32460
 gctatttagg ccactgccat ttctgttggt ctgagcaatg agcatgccac agggccaatc
 32520
 agctctaaat gcaaggcggg gcatggccag aggaaagttc ctgctgtctc tcttcaccc
 32580
 tgacagccat ggccccaagg gctaacagag gcccttgaga ccaattctgc catcaggaag
 32640
 gaaaagccct gccaaaatat ttgcattcag aaagttagca tttataacac agtcagagaa
 32700
 ctatgaaata tcttcaggc caagcactgg gcaatgccat gttggtgggt tctgatgagc
 32760
 cctgggaaac ttacttgac cttttcagcc cctaacccta ttacacaggc ccctgcatgg
 32820
 aacaaatgcc ttcactctca catgttgctg ttgggagtgt actttgatac agcctctact
 32880
 caagaacaat ttgggaatac ctatagacat tttactaaca aatgtgcacg ccatttgacc
 32940
 caggaattct actcctagga attaactctg cacatatact cacgtatatg tacaaaatcc
 33000
 ccatatacaa tgtttgtaaa aagcacgggt tgaaagccat ctaactcttc atcagagggg
 33060
 ctgggtaaat aaaggatgtt ctgcagccat caaaaagaat gaaagttcta catgttctga
 33120
 tatggaacaa tccccaaaat atattgctaa gaaggaaaaa acaagggtgca gatcagtgtg
 33180
 catcatttac taccatttgt gatggtctta tatagcacag accagttctg gaaggatcca
 33240
 taagaatggg tccaagggc taggtgcggt ggctcatgcc tgtaatccca gcactttggg
 33300
 aggccgtggc gggcagatca cctggggctg ggagattgag accagcctga tcaacatgga
 33360

gatatctcat ctctactaaa aacacaaaat tagccaggcc tgggtggcaca tgcttatagt
 33420

cccagctact caggaggctg aggcaggaga attgcttgaa cctgggaggc ggagggtgca
 33480

gtgagccagg atcgagccat tgcaccccag cctgggcaac aagagtgaaa ctcggtcaca
 33540

aaaaaaaaaa aaagaagaat ggggtcccag agttgcctcc aaggaaggga ccatggggac
 33600

tggggactga ggactgggag acaagagtag agaaatgctt ctcatattgc acaattttga
 33660

atttttttgt atcatgtgta tgtattactt atgcttaatc aaataactcaa ttttaatat
 33720

aatttaaaaa aataataata tccccatcca cccagccagg tgccatagta tctgacattc
 33780

tgccaatagg gagtcaacaag gttgaagtct tagattgaac attgtttcaa catcatttct
 33840

cactggaagc tctttataat gcaagcctta ctagaactgg ttttaagtgtg atccaacct
 33900

tcagaccaac tttcctcttt cgggcataca aaaaacttca ctgggcattg actttgtcct
 33960

acacacttgc ccaggcactg gagagacagg gagaacaggg ctgacctcac agtcagccag
 34020

gtggacagct gcacgcacag agtgggtggca gtgcacaggg agtgggtggct ctgcacatgg
 34080

actgtgggcc ctgcatggga ctagtggcac tgtgtacaga atgatggccc tgtacaagga
 34140

gtggcagccc tgcataggga cctatggccc tgcacaggga gtcgtggcct tgcatgaggc
 34200

cgggtggcggc ctgcacggga ctgatggcac tgcacagaga gtgtggccct gcacaaggag
 34260

tgggtggccct gcatagggac ttatggccct gcacaggag ttgtggccct gcaaaaggag
 34320

tgggtggccct gcacagggac ggggcagata taccctgaca tttgaggaga ctgaggcaag
 34380

gatataagtg ggagagagtt agaatttga attctcagac tccttgggtt tcatgctgga
 34440

acacgctggc atgagagaac caggccttgg cctgtggccc agctccctcc tcttcccacc
 34500

tcctgccctg tcccaccca tcaggggctt gcacacacac acacacagac ccttcagccc
 34560

tcagagccaa gctccttgca cacagactct ttccacagtc agcgccctta gctcctgtca
34620

agcctccagg tacacacaaa cctgcgtctc tgctcgctg cacaaggaag gaccagcca
34680

agtgtcccaa tagggcctgg aagcgatctt gatgccatcc gggcagaaaa tcccaggtgg
34740

agtacagccg gataaagaca gccaggttt caggttggca tgttcctccg gttctgcaga
34800

tccctcgggc cctggagaag gagatggcca gaggaggata gcaaggcctt ccaaagtgc
34860

gtcccgggtc acagcccacc tgcctggatc taaggacggt gtcgctgagg ctgtggcagt
34920

gggaggttcc aagacacggt gaaggcctga aggagggggt ggcccttgct aagacctgtg
34980

gccagctgac agggctcccc agcactctgt cagtaaggcc cttccaggaa gaggagtg
35040

ctggacaaga ggctgtgtgt ggaaggcaaa ggggagaagg gagagagagg gctggtaacc
35100

caattcctct tcctgcctcc ctacacttca gcatcacctt gttcacgccc cttgcagatc
35160

tcccacagca ctccctgtgt acacctttgc tgctcaaagg gtctacacaa gccaaagcaca
35220

gtggctcttg caaaggtggg aagatgtctt gaggccagga gtttgagacc accctaggca
35280

acatagtaag actctgtctc tattaataat aataataatt tttttttttt ttggagacag
35340

agtcttgctg tgttaccag gctggagtgc agtggcgtga tctcggtca ctgcaacctc
35400

ccctcccca gttcaagcta ttcccgctgc tcagcctccc aagtagctgg gattaccggc
35460

acctcccatc acaccagct aatttttgta tttttagtgg agacagggtt tcacatggt
35520

gccaggtg gtctccaact cctgaccca agcaattcgc ccacctcggc ctcccaaagt
35580

gctgggatta gaggcgtgag ccaccacaac cagccaataa taaaaatttt ttaaagggtg
35640

atctaagatc agcagcatgg cctcggaat tggaaatgca gacctctgg cccacacctc
35700

gaaccactga attagaattt gtatttgat aaaagtctca ggtgagttct gcaaactttg
35760

taatttgaga agcactggtc tacaccacac tccagcacct taacgcagct ccatttagca
 35820
 gtggtggaca ttgcatcatg cagtgcctgg ttcttcttg agtgcagctt atggaaggag
 35880
 ggggagtgaac aggactgaac ttattgttca aatccccctg tggctctggg gttcttggtt
 35940
 gctcactgca gaggttatag gtgaaatgag cagctgtcct cacctctcca gcacccttg
 36000
 aatggaagag gctgcctga gctgggctgc catctcctgg gtcctcagtc tcaagccgct
 36060
 ttctgatgtc agtgagctct ttgaagtcag gtccgcatcc tacacagggt aggtacagg
 36120
 ggccgggaag gcatagaata cacattaggt tggatctagc atctgatcca cctccccatg
 36180
 tttcacgaat gggccagtgg gtgagcttta gggagggcta ggcctgcct cttagaggaa
 36240
 ggcttaaaag gcccacact actcttcttg gcctggctag ctggaggcag gcagatgcc
 36300
 caagttcaga cagtcagata ctcccaccag ggatttggt ctatggtgag tgacatctag
 36360
 aagcagggcc agccagaagt ccttcttggg gctgggcaca gtgactcgtg cctgtaatcc
 36420
 cagcactttg aggggctaag gcaggaggat tgcttgaagc taggagtttg agaccagcct
 36480
 ggacaacata gcaagacct catctccaca aaagtaaaaa taataataat aataaccagc
 36540
 catattagag cattcctata gttccagcaa ctcaagaggc tgaggtggga ggatcactgg
 36600
 agcccacgag gtcaaggctg cagtgaagta cgattgcct cctacagccc agcaattgca
 36660
 cacctaacc caccatacaca taatcaaata gagattatac attcacgatt gaagccaaag
 36720
 tgttatgaca caaaacttac ccttaggaca tgcataaga tctgatgttt tatattttta
 36780
 ttcttttttg ttaaagaaag agcttggtga tccaaccctc taataaagct gaaagccaac
 36840
 atttgcaaac actggcaggt ggctgcgtgt cagggtgagg tgcactgaa cagggtgata
 36900
 aagcaagatc ttgtctctaa aataaataat ataataata aaatataata taattaagta
 36960

gaataatgaa atatcctgct ggccaaggctg gcatccccgc cccagcagc aaagtcagca
37020

gggccagtgg tggcaagctg gtgtccacac cagtggccac agtgtcatgt ccaggcaatc
37080

tctgccctat gacctcagct gcctttccag ctgtctttgg tttctgcctg tttcccttca
37140

gcctccctag attctgggag gggccggtgt cctgctccag gtaggcagag tccaagtctg
37200

ttgtgggcag aactctgatg atcattgcag ggggtgggagg agggatggag ggagcatgat
37260

gctggcagaa cctccatggc atctggggaa ttgaaggctg ttttttgctc agttctcacc
37320

atccaaagca gcatgaacag ttttcaccac gcctctcttg tggatgcctg tgttcttggt
37380

acacacctct gctatgacac ttatcacctg gtattagtta agcatctgcc tccttcattt
37440

gctgtgggta tttagaaagc aatccatatt tcctctctgt atgcctcaat ttctccacct
37500

ttattagagg acccaggata agattatttc tgaggctgct tccatcagtg gttctgtgag
37560

ctggcccatg tattcagcca gtcagtaggc catcaagcat gcctctacgc aaggcactag
37620

aagggttaca tacaatggga aactgaggcc cagaaggagg aaagatgaag tgcccaaagt
37680

taaggagata aaagcagaag tgcaggtaag agaaggagta ttcattattac tgcaggaatc
37740

caaattctac aagaacatac gcttggttctt taaagggagc cctcgcttta agcattattt
37800

gccgatcaat acccattaat aattacagga catggggcaa tattctagca aaggctctga
37860

ggtcttccta gatgtgcca aaataacaag gactcccccc gcagggtggca ccaaaagtcg
37920

gaggatgtgt tggaagctga cttccccac ctgtccctgc ctcaccaca gtcaccacc
37980

cctctaggat gtaatggtga atcattcaac acatgggcca ccttgggaag tcgcttaacc
38040

tctctccctt ttcacgttat aaaatgggca tagtcattac acaaatgcat aaatcaacct
38100

caattttctc atctctaagt gagaagggtg gactggaatt aacaagagga tccttgaagt
38160

tctttcgagc tctgacatcc tcatccattg caaagcggct gtgaaactgg attatgtaag
 38220
 gtccctgtggc cccaccaatt gtttggcaca tccctgtctt ttccctagct agataatgat
 38280
 gtggctgagg gcagggactg tgcccacact caccagggtc atgtgagtca gcgtctggat
 38340
 tgtgtggccc cgttggggct gccctggctg tatcacatgc ctcttgtgcc catctcatac
 38400
 ccctcagcca ccctctgacc acagcattgc ccggcgtgcc cctgcaggc ttttattctc
 38460
 ctgcagcccc agggcctcct caacactatg atgtgagtca cctgagaagc ccagtccac
 38520
 aggcattgcac agctcaggag tgcaggactg tcaaggactc ttgcagcaac cctcagccat
 38580
 gtggtaggta tgtgggggtgc aggggtggga ggcagtggca tacacactcc ccagaagctt
 38640
 cagtgggatg gagatcagct caacaacaca cctttgactt gacttccct ccttctccac
 38700
 ttactcccc cagccctca ccctgctccc tgagctctcc tcccaaataa actgcatgca
 38760
 tgcaggcaca catgaccagg ctgaggctct gctcttagtg ggataccagg caaagacacc
 38820
 tgccctaggg attggagcct aagccagagg aggcaggagg gtttagggca ccagccaaga
 38880
 agtgctttcc agatgctgag tgggatcccg aataagggaa agggcatcta acatagaatg
 38940
 aggaggagat aacacaaggg tagaaataac caagaaagg gatttctctc agatggcttt
 39000
 agagggtgctc ccagccattg ctgggctcac atggctcaca tgtcataatt tttgccaagg
 39060
 cggtttctact gccaccacca ccccgctcag cctccagtgg cttgtccacc catggctgtg
 39120
 gcctctaact gggcaggatt ttataggcta ggaagaaaag tagggaatgt taaggagac
 39180
 ggaagtcaag ttcaacttgg aaccatocca gaatgtccag caggaaaggc ccatggatgt
 39240
 tataggctac aaccctctca gtgctcagcg gaggaactag gagcagagag ggtgagtaag
 39300
 tcaaaggcac ccagacctg tagggaatcc tgtgtccctg caggatgcag aatggatgct
 39360

tctgaaccct gccacaccct cagggtacat gtcagagcac gtgtgtggct gggacaatcc
 39420

tccatctatg ttctcactt acaccggctc cagttagga aacccaaaag gctacaacca
 39480

acccaccacc acagtgaacca caatgtactt ggtgacacct caccttgaca cgcagccacc
 39540

tgccagtgtt cacaaatgtt ggttttcagc tttatgcgag ggtaggatcc aacaagctca
 39600

ttctttaatg aaaaagaatg aaattgtaaa atatcagatc ttattacatg tcctaagggt
 39660

aacttttatc tcataacact tttgcttcaa tcatgagtgt ataatttcta tttgattttg
 39720

tgtctatggg gttaggtgtg caatgtagaa aagcttgaaa accattgcct tatacccatt
 39780

tgaaaaatga accaccaaac aaaggcttaa aaacatgaat atttgtcaac caaaaggaga
 39840

actcctgggc ctggaagaaa catcacgatt caagtgcctgc agcacgaaag aagttgaggc
 39900

tggatttgta ttacagctga ctgtttgtaa agcaccatgc accatgcaac tgccttcagt
 39960

tccttttttt tttctttctt tttttttttt tttttttttg agatggagtt tcgttcttgt
 40020

cgcccaggt ggagtgcaat ggcacaatct cggctcactg caatctccgc ctcccaggt
 40080

caagagattc tgctgcctta gcttctgag tagctgggat tgcaggtgcc tgccaccaag
 40140

cctggctaata ttttgatatt ttagtagaga tggggttttg ccatgttggc caggctggct
 40200

tcaaactcct gacctcaggt acttggcctc caaagtgcct ggattacagg catgagccac
 40260

catgcccagc ctgccttcag ttctaaccac taaaggattc gtgttctcaa ctgagaaaca
 40320

gccatatgga ggggcttctt gcaagggcac tgaagatgtc agagatgatg cctgtggagc
 40380

acctggcaca gaggagctgc cctgctgact gcagcacagc cagttcacag gggtagagcc
 40440

cggcatcagc ttagggagca tcccttcttg cagctcctcc ccctagcaca ctctcccacc
 40500

tctgctcgca caggacctct gtccaggagt gcttacggca aggtttctgc ctcaaaaaca
 40560

gggcttttgg ggctaactgg gagtgttaagg gtaaaaggaa agcttccact agggcccccc
40620

gattgttcac tggaaatgaa gggacaatag gagattaata ggagaagaag gcatacaaat
40680

gtattaacat acataaacac aggagccatg aaaatatgag gctcaaagaa gagccagatg
40740

actgaagttt aaatagcacc ctotcatag gagagaggga aatgggggta ggtaattttc
40800

aaggtttagta aataattttg agggaaatca aatggacca gaaggcagac attttctgt
40860

aaatgtctc ttaggaagct gaatgggact gcaatttaag ggaaagtga gggcaaaact
40920

gcattgtgaa caaaggttgt cttattacat aggtaaagtc tccctagtca tccctcaaag
40980

ctgccctcag aagaatagat gaaaagcctg ggcttagtga cagcttttac tcttttctcc
41040

ttgcagggtg gtggttaatc tttcctggtt atttgatgag atttctaggg agggggcctt
41100

aagacagttg catttctttt ggaaagaagt gtccatagtc agatatggaa attctagaca
41160

gcctcttttt cctgcacttg gggagggggc agaggagaaa caagagaaga ttggaaaatt
41220

cttggttctg aggcagcttc taaggacttg cattttcctt taattcaaag tgctcagcct
41280

gccaacactt cctacttttg ggtatcattc tctgtgacct cacaggggaa tggggctctg
41340

cctcagagaa tgggggagga cctcaaggca gcagggcaga actccagggt gagagtcaaa
41400

gatgggtttc tggacctgag actccagtgg ggttcattgct agagtcacct cagtcttcag
41460

gctcactgcc caaacacgca cttgaggaag tctgtggggg cagctccttg ttttgttct
41520

cactgtcatt ttccaaaagg cagacaagcc ctggccgcct tatgcccagg gctgagtcca
41580

caccatgcta cagagcacc tgaggggagg ctctctcca aagtgcatta ccagggcctg
41640

tgctaaattt ttcagcagcc tcttgtctca ctaatggtcc cgggaaatag actaagctc
41700

catccccaac taagtgtca ggctcccc acctctgtcc tgctgggca agaagcccag
41760

cgccgcccc cagcccagct ctccatgcag cctcagctcc cagtggccct ctgccaagga
41820

cttgcgcttc tgcagtggga tgcagtgcct cattttatcc ataattgtac gtggtccttg
41880

tgtgatgtcc ggggacgtga aaagttgtgt ctcatttcat ttgggatcca gttaaattcc
41940

caaagtctgc aactgagcag gaatctcagt gctcagaggc cagcagtagg aggagtctgg
42000

ctgctggggc cggacacctg gcctctccgg cctctgacca gctctggcca ttggtcacca
42060

ccttggtttt gttgggtatt aatccaccac actctgcttg gctgatccct gagtcagagc
42120

caggtggga tttgtgggtc tctgagccca ccagcctgtg actcccaaaa agagagcacc
42180

cttgctactt ggaaggcttg ccaccatgcc tagccagcac tcttgctttc taagctgaaa
42240

tcattataag atagtccct actaccttct acccagaaat gtatccctc acaaagtata
42300

tgttcccacc ctatcccaat cctaccatcg tgatttctga gtactttcca gaaaactgtc
42360

tccactggct cccaaggga gagcacttc tccagcctct ggaatgtcca ccatgttgat
42420

ggtgcatgga tgataatgac cccacggatg agctctctcc agtgggatca cgaagggaga
42480

atgggatccc tgactcccca cagaacagct cggcccagct gtgaggttga gcaaatacac
42540

tcaaagctca gcgttaaccc aagtccagcg ctctcagctt acacaggtgg cagcccaatc
42600

ttagtgctcc actcaatcca tgtctatgta tacttaattc tctggaacta aaatggggaa
42660

tagggccagg caccatgact caagcctgta atcccagcac tttgggaggc cgaggtaagt
42720

ggatcaccta aggtcaggag ttcgagacca gcctggccaa catggtgaaa ccccatcttt
42780

actaaagata caaaattagc caggtgtggt ggcgggcgc tgtaatcca gctacttggg
42840

aggctaaggg aggagaatca cttgaaccca ggaggtggaa attgcagtga gccgagatcg
42900

agccattgcg ctccagccta ggtgacaaga gtgaaactct atcttaaaaa ataaaaataa
42960

aataaaataa atttttttaa aggccggagg gactaagatg ggagtgggaa ggagggaggt
43020
tccagctcaa tcctcaccag cagatattcc tgcattcctcc atacctgagc cttcagacca
43080
atgctcccca ggcccaaagg ggacctccct ttcaagtata aggtccccc aattcctcac
43140
cttaggatct ttaaggcaga tgactgcctt aaagaacata tattccttgc agcaagaaaa
43200
cccaactccc aacaccccac agagtgtggt ctggcctgga caagtctgcc ctaagaaaaa
43260
agcgacaaag acccttagta attgaaacag actttaagta ccataaagat gagctttcaa
43320
tcattaaaaa aaaaaacgtt tactgagcca ggcaccaggc ctaggaacag agagggtcat
43380
ccaggctgcc atgttaggtg ggcttgggcc cactgctgg ggtcagaagg tattagaaag
43440
ggtctggtct gctgggcgcg gtggctctca cctgtaatcc cagcacttag ggaggccgag
43500
gcaggcagat cacctgaggt ctgcagctcg agaccagcct ggccaacatg gtgaaacccc
43560
gtctctacta aaaatacaaa aattagttgg gcatagtgat gggcacctgt aatcccagcc
43620
agctactcag gaggctgagg caagagaatt gcctgaaccc aggaggcgga gggtgcagtg
43680
agccaagaaa gcgccactgc actccagcct ggatgacaga gtaagagtcc atctcaaaaa
43740
aaaaataaaa taaagaaagg gtctggaagc tctaggtgac ctcatgctca catctatacg
43800
cacacctga tacattattt atggatcagg tggcacagca ttgcaggcag cagcagggca
43860
aagacagtgt ggaatcagag tcgggcagcg ttggatggga ttccggttct gccaccactg
43920
gctatgtgat ggaaatctac cacatctctc caagccttgc tgtcctgttt tatcaaatga
43980
ggacattgta ctgaatcaca aggtctgtac aatgcccgtg catgtgcctg gaatcttccc
44040
ctctgggtgc ccctcaatgc ccatggtgct tctgcacag ggtgtaccat agttgcttgc
44100
tgtctcctgc tgaagcctgc tgtgaagagg gactgtgtgg aggtgaaggca gtgggcctct
44160

ccctgtgagg tcctcagggg gctgaaggca aagtcagaag taaaaaaga gaggggggac
 44220
 ccatcccaa ccagctgcca tcccaggga ctcagcctaa ggccactgca tttgcatgct
 44280
 atttacataa atctgcatat aactgtatt ccactctat cttgtaggca attattgcat
 44340
 ttaggcagct ctgcattcca ggaaggggca caagcagaga tgtgatcccc tgaacttcat
 44400
 gacaagaagg taccgaagc tcctctcca cctccagcag caaatgtctc cctggggct
 44460
 gtccgtgtgt gtgtttgaac cctgctctgg ggctcctgcc cgccacctgt agtggctacc
 44520
 atcagaagtt cctgttaatt ggggatggca gggtcaccac atctctactc ttctccact
 44580
 tctcagcagt tctaatttaa accaggtag aatcttcca cttggtagtc tggtagggaa
 44640
 attatagtta acaataattt attaataagt taaaatagcc agaagagaag aactgtaatg
 44700
 ttcccaacac acacaaaaa gataaatgtt tgaagtgaag cattgaagtg atgctccaat
 44760
 tatcctgact tgatcattac atattgcatg caagtaccaa aagatcacat gtacccaaaa
 44820
 atatatacaa ctatgatata tcaattaatt ttttaattga tttttttct tattttctg
 44880
 tgtagtgcca gaatcacagt tctactgcagc ctcaacctcc tgggggtcaa acaatcctcc
 44940
 caactcaacc tcctgagtag ctgggactat aggcctcac aacctgccc agctaatttt
 45000
 ttgtgtgttt ttgtagagat ggggtctca tttgttacc cagctggcct gtaactccg
 45060
 ggctcaaggg attctcttgc ctcagccttc caaatgatg agattacaag tgtgagccac
 45120
 cgactctggc tgatatatca attttttaa aaaaagggtc ttctccaggc ttcaatgcga
 45180
 cttcacctat accagggtgtg agcgccctc atgtccaggt ggggtggggc aggcactctg
 45240
 agggacctca ggcttcttgg attgtttggg atggaggccc aaggaggcgg aaggctctgg
 45300
 gtctcaccca agctggctct cccaaggaa taagagtaat cacatgggcc tggagcagaa
 45360

aagggaagag aagggtcccc aggaatgatg atttctgggg cttcaggaaa gagtgtggaa
45420

ctgaatggca gagctaaggg aagactagga cttacaacca ccaaaatcag gccctccagc
45480

tgggttcggg gccaaaggag agtgcagcct ggcttctgcc attctccttt catccccctg
45540

gagcttctga aatccacaga attgtggtga aggacaagga gccaaaggct cacaggcgag
45600

gggaggcggg gaattcaatg tgttcaaggc ctacattgca catctgccct tggtttactc
45660

tatctaattgt tcccatttga ccggttaagaa aactgaggct caagattcaa acccaggtcc
45720

ttgttaacttc aaaaacctg ttctttctag gtgtcaaagt caatagatgg gccacggggc
45780

tagcagtcac ttctcttctt ggacttggca gtcaagggat aggacctgaa ctaaattgcc
45840

tttctacaga cttttcagcc ccaaccagag gcctgctgca ggcagttggt tgaactcctt
45900

gattaaatgt tcagcccctg gcctgggcgg cagactccca acacacaaga cttactgagt
45960

ccccaaata gacttcaaaa cacagtcttg agactctgcc tgccagttgc ccattaagct
46020

catcccagcc tggcatcttc ctttgttggg ccactctccc tatggctacc agggagaaaag
46080

tcttttcaga gatgaataag ggcgtccaga caacagagat gtagtagtag gaggtggaac
46140

caaaacagaa ccaggggtct gaatcctggc tctgccactt tctgagatat gtcacatcag
46200

acaggtgact taacctcctg tgtctcagtt tccaaattca aaacggaaa atgtattatg
46260

acgatcagat gagttaatac agtcaaaatg cttcttagaa cagcatctgg gatgtagcaa
46320

gcactcaaaa acatattagc tattatgatc acacagagat gcataacctc taaaatcttt
46380

ctgggctccc accgccagca tgataaagcc tcacttctca ccaaaggcaa atcaggcctc
46440

tcatcatctg cctcctccct cccagcccc atctcccgaa gcgcctctca gggcacctct
46500

gcccttgac aaggcagggc ctgttcttag caactctcag agtagcagtg gaagtttcca
46560

acaccaagat ctcagatacc tcaagctggc aggagtgcga ggtccaacaa gaagagattt
46620

tatgaatcag ccagccatgg gggtagggca tggccaggtt tcttgattcc cagaccaggg
46680

tgccaacaac acctactgtt catgtgagaa aaggagaaaag aggaggaggg gccactggcg
46740

tcttttgccc acctgggggt cctcgtgtg ggagaggggg aggccagagg acagcaagtt
46800

ccctgagagc ctctcagaga aaaggagggt ttagatcatc acacaggata aggaagaaga
46860

ccttttccaa gggggtagg cagagcctgc ctggaggtaa atttcctttg ggaactgtct
46920

tcttcagtat ttcccttcct ctctgaacac cgtttctttt cctttcagat tatctgacag
46980

ccttttaatg aacacctact aagtgccag cactttcctg ggaacttaac tcttacaacc
47040

tggttaagagg tctatacact cattttacag tggaggaaat ggagacttag cggagatgtg
47100

aaatttgaac caggtcctg tggtttcaaa gcctgcat gttcctttac cctatggaag
47160

cctccagaag ccaaggtctc atctccctc ttccctggag gccccaaacc tgggtagctg
47220

tgcccctaga cctttgaggg agggctctga gccattatct ccctacctt ttcttgagct
47280

gacttggtt ctcccaggct gccgtaacca gcaaacgct aaagatggct ctgagacgat
47340

ggccttctgt gtgtgtgtcc atgaaatcag gtagtacacc acacagaacc agactctcag
47400

tgacatgcct gatattccac agacctgcat ccgagctgtg cagatctcag gtagacacaa
47460

ccagtgattg gggatggctg gtaggatctt gcctaggcag gccagggagc tctgtgagga
47520

gtgtttggga gaaaatcccc caggttcac tcctgcatg gactgcccct gctgggccct
47580

cacactaagc tggctcctgt cctcagtact gaggatcac cacagggtgt taggccacag
47640

ctcacaggag ccctccacct agcgggtcaa aggtagtga agatctgaga gtcaggcaga
47700

cctggtccaa atcccggatc tgccatttcc tctgcatg atactaaaa agtcacttct
47760

tcattctgaaa aatggggctg atcccaccca ttctgcaggg tctcagtggg gattgaaggg
47820

ggctgacctc aatgtagcac gatgcctgcc ccagggtggga cttcagcaaa ctctctttcc
47880

tttcccttcc ctccccttga aggccaagac caaggaaaca gaattgggca tgccagaaca
47940

tatcccagcc ctgtggggtg tatttacctc cacttctgat cacagtgtgt ttgcgtatga
48000

ccttgaggcc agtgactttt acaccttgcc tatgctgggc tgggctgtga cttggtgtg
48060

acttgcttgc cctgaagcca agcctcagct gacatgtgtc tttccaccac cctgctgtg
48120

ctgagtgagc accagccacc tgtcccattg ttccctgccc aggtgtgtg tggagagtgg
48180

tctgacaggg cagacaacac cctacactcc agatgtgtt aggtgagctc tgggggtcag
48240

gaaagaagca gggaggggca cagtgtgaac atcctgtgca gatttcccc caactggccc
48300

cccagctggg tgggttttct tttctcagta actcctggca gaagtccaag gaggcctcct
48360

tctcccaaac cctctgtcct tggctcaaaa atccagcaga ggacgtgcca cgaccacgaa
48420

aggtgacagc agtcaagggg aaagaggact cccttactgg agtgacaga agtttccctat
48480

ttattttgct tacgcggttg actttttcca aaggctgctt gggacggctt acctcgaaaa
48540

gcttagcaat aaggctagta gtaaatcaga caaaaaaatc atggactggg ggaagaaag
48600

tagcatttaa atatatthaa taaggacgat atagtcgtga acatgtgcct gagtatcctg
48660

gctaccacag taaaagggga aatatgatca gttatgggaa aagagtaaaa atattaatca
48720

tgattttttt tttcttggag ctttgtgtaa taggcactga tgttgaaaag ggtagagtta
48780

tttttccata tttgtagaaa ttcagaagat tatcagacta tttgcctcac aaggactttt
48840

tagcaagcaa aaagtatgat attgtcatgc accctggtga ggggaagtgt catccagttc
48900

taggtttctg tgaatctacc ttgatgcagg tgttcaggga gtagtgagtt tccctcagttg
48960

taatgtagac cagcgcttct caaacttaaa cacgtataca acctcctagg gggctcttgt
 49020
 aaaatgccat tctaatttgg gagctttggg gtggcctgac actctgccat tctaaaaccc
 49080
 tcccaggcca tcatgatcac caccttccga gcagcaaggg gttaaaaccc aggcattctcc
 49140
 catgatctag tagtgagtgc ctttctaata gccccagct gaaaggcaat gcaatgttat
 49200
 catagacggg caagaagaat ctcttccttg gctctgtaaa tgcttccggg atgtgctggg
 49260
 tcagcctccc tccattctgg ttggcgatct tggctagtc acaagccaag accgttgatc
 49320
 atacacaaaa gttaagtac ctactacgcg ccaggaagct ccctcagaaa gatggttggt
 49380
 cctgctcccc caccaccct ccaccctggc atcttttatc cactcttcct tcagcaactg
 49440
 cctttatcaa tgctcccttg agccaaaata aacaaataaa aagcaacttc ctttgaagc
 49500
 tactccatac tcttgctg gcacagtta catctccctg cagatacaca gattctcatg
 49560
 acacagagct cagacagtgc ttcctctctg gagggctcat gtttcgagag aagactggaa
 49620
 ggaggacaac cttatgctcc agtagaattc caaattctgg ggtgcaggga catttcgtgt
 49680
 gtgtatgtgg aggggggggt gtccaatgc cctgtaagcc aactttata atgctgtaca
 49740
 atattgggtc accatcatgg gccatggaaa gcagaatctg tgagggtcag tcccaggctc
 49800
 cagcaccctt ttgtcagctg gcctccccgc tcagtgcctt ctcccactac ctgcctgccc
 49860
 tcagctctga ctgctgctga tgtgcctcca cccatcaggt cattgtgatg gcaaagacaa
 49920
 tgccctcatt tgccctgaat gcagccttac ataggaaaga gcttcctgtg ctccagaact
 49980
 tggaggagtc ctgcagcacc tggcttagtg gggagagcat cagtggggag cggcaccttt
 50040
 cctgcagcca cggccactg ggaggacca gatccaagtg gcctggcttc tggctagggt
 50100
 gtcagagtac attggggcat acagcttccc gtggccatct cgttttcatg gaggatttat
 50160

tttgtgaagt ggacctgctc tgtcccatcc tataaatgct ggattaatcc aagaagctga
50220

gggtgagcaa catccctctc cttcctgatc aaataagaaa taaatctaga ttttctcctg
50280

gatgctattc tcttcttagc caaacagcta accttgctct caaggaaaaa gggcaatccc
50340

aggggctctg taaattgaat attcagcaac tccaggggca gacatcccta agtgcataag
50400

tggagcacc ctctgaatca gcatcacacc agctgggggc ataaaaagga gagcagagta
50460

ggcaggagct gtggggacat ttctgagatg ctaaaagagg caaggaagtg tgaggacggg
50520

tgtgcattca tgcaacaggt atttactgag cacctactat gagccaggca gtgctctatc
50580

cattagagat acggaattga ataaaacata aaaatccctg cccaggagga gctcacattt
50640

ctatgggaga agacagttac tgaaaaacag tccaaaaaac tctattttat agtgtgttca
50700

gaggtaatac aggctatggg gaaaaaaga aagcaagcag gtacatggat gaggaatggg
50760

ggcgatcagc gtgggcctca cagaaatgat gacctttcag ccaagacttg gaggaggtag
50820

ggagggtagc tggggatgcc tagggtccag gcaaggggaa cggcaggggt caggatcctg
50880

tcagattcat ctgaggtggg actgaatgtg agcagagaga gggagacaga ggatgagggc
50940

agagagaagg ccaatcacac tgggccacgg ctattgcaag gactttactt ttgctggggg
51000

agggggtaga gttttgagca gaggagaaat gcgatctggc aaacatcttg caaggatggc
51060

tctgactgtt ctatccttgg cagaggacac cagagtctgg cccatagaag gtgctcaaca
51120

aagctcagca cattgcactc ctccctggag taggaaaggg ctgctctgct ttgggaaatt
51180

tggctctcaa gcttctggaa acacctgacc ccatgccagg tgagacaagc tcacactcaa
51240

ggccccctta gtccaacggt cagtggacag aacaagcacc aaaggtaaac ggcgtggggc
51300

gtggcctctc tgccaaagtg tgtggggcca agtagggat tcacctccct ccctgtctgt
51360

gagaaatttc agaccctct actttttcca gcaggagttc atgttccttt taagataatc
51420

ctctatagac ctctggcagc ccagccttca aaatgtctcc ccctgtctct ctactcatg
51480

gagagccccg cttccccgcc tcaaaaaaaaa gtcttctctt ctcttactca tgcattggaga
51540

gccccacaaa aagacactgg ttcccaggga ttcttccaaa gatgctggaa gacctctcag
51600

tgtttgcatt tcccaagctc ggtttctctg tgctgtcctt gggcattaaa ggaaggctat
51660

tcagccatcc ccatacagc gagcaggcat caaaagaaca gctactgcag ggggagcata
51720

ttaggggcac cagagggggc taggaactgg gtcttgtctt ctcaaagttt acaaccact
51780

gcaggggcca ctcaccagcc ctggcaccct aagggcagga acctccattg ggcacgtgtg
51840

tctccagcac tgggagcact gtctggcacc tggggcactg tgagtgttga atggataaag
51900

gacctgtaag aattgccata agaggtgttg ccctggggca aggtgggaca ggaaagtctt
51960

tgcaagagaa gagaggctct ggccttgggc cctggctgag cagagagctg ggagggtggg
52020

attcgggtag gaggaaagct gtggaggggg aaagcaaact gcctgcgggg gtgggggggt
52080

tctcggcggg agttgggggt ggagggaaag gcgctttaca gcagcactcc ctccacaggg
52140

agaccagagc tctgcctctt cctccccacc gtgccctttg gggccagtt gtggggcctt
52200

ccttgaggcc cagctcctgc ctctggcctc cggaggctta tctccccaac agaggctggc
52260

tcccagaaa acaggattct tagacttctt aaaaataacc tgaagctcct tatcagaggc
52320

taatagtgtg tgtctcccag gaatactcta gaaagagtac caggggtcag gaacccccac
52380

tcagcctcca cgaatctgag cggaaaagcc cgggtctcct cacctgtaaa gaaaataact
52440

gcagccggca aatgtgacca caggtgagaa ggggcttcgg accagggcgc acacctctag
52500

agagggtccc acagacagcc tccccttact tacgccctgc ttcgctgcag tgctgtgaag
52560

ccatcgcatc tggaaagtgcg gagggcagcg ggcagtgcc cagggggaag gggggacggg
52620

tagggtttca gagcccagaa ggcctagaca gggaggcagc tccgggagaa aggtacccgc
52680

catcccaccg gtcttggaac ctgaggactg ctgggaaggg tcaaaagtct gccctaaggg
52740

tggacaaggg acagagattg gggggggacc acaaaggcct gggcgcaacc ttgttgagag
52800

ggaagggatc cggacccaat ctggagccct ggtctcttcc cacctgcaac cgtggctgcc
52860

acttgccctc ggttataaaa tggacaagga ttcgagggtc tcctcagggt aagagagggg
52920

ggcctagagg tcaccctggc ctcggtgcc tccgtggccg ggttaagagg aggtcccga
52980

gttctgctca cttcagccgt gtgccgggca ctgcaaata ggaagtgttg gcgccggctg
53040

gcgacctccc gcctggggcc aggggaggag ggtggttga cgtgccacc gctgccggg
53100

ctgtgcaggg ctgggcgggg agcgaggacc cggcggtcc tgattgcggc cccgggggag
53160

gtggccgagc cggataagct gcggcgggct ggaggcggc cacctcccct gcaggctcgg
53220

ccctcccggg cgggtggggc gcgggggagg aggagcctcg ggccgagcca ccgccttcgc
53280

cgcggacctt cagctgccgc ggtcgctccg agcggcgggc cgcagagggt agtgtacct
53340

ccccgggtct ccgcggggct gcgtgctgcg cccggtcccc gagacgccc cccgggtgca
53400

ccctgcgccg tcgtgcgcg gacctcgggt gccgccacac gtctggaggc gacttctgtc
53460

ccctgggacc gagccacgtg cggccggcgg cagagaaacc gggttccggg gccccaccc
53520

cgtgtgcctt cttccctag gcgtggaagc cgcttgccc gcgcagggtta ggcagggccg
53580

gcggcgacag tggcggggag caggctccg agccccgggt gcagatgtgg gcgcccctcc
53640

ggatgacccc ggctgagtc acaggctccg tgtccccac ggctggggct ctggccagcg
53700

gtcccaaagg agggctggtg gctgtgggcg gggatcttcc aggtcttcc tcgccagacc
53760

tcccagtccc ttgcagccg aacagggcag ggacaggggc aggtgcaggg gctgggcagg
 53820

ggcccccagc cttggtcccc aggaggaggc tggcctcacc agtgcgccgc tctgccgcgc
 53880

gagcccttgc ctggcaaagt tccccctcct gcccgcaaa gctccccctc ctgccactt
 53940

ccccgaaatt tggggacagg ctgcccggag aggtgtggtt ggaactttgg tcactctggg
 54000

taaatgaagg ggagggtgtg cccagataat atgcggtttg gggaccagaa agccggatcc
 54060

gtggggagaa gcaggaaggt cccgctggag ccctggcaga gggcacagag aggcctccca
 54120

aatatccagg gccacctcag agatgggctc cagaagtggg gattctggaa atgccccag
 54180

ggaggatatt ctgtccatca gagtgcagcc ataataaat taacagcaga tattcccatt
 54240

tatgtgtgcc aggcattctg ctaagggtt tgagcacaat atcccagta atcctcacia
 54300

caggaaggca ccattatctc ttgttaaaaa tggggaaaca gcctcagaga gtccacgtat
 54360

tttgcaagga tacacagcta ggaagttgca gaagtagaat ttgaaccag agttatctga
 54420

cttggagccc agaggaaact accactgggg cgaaacaggc tgtgtcctca cggggattct
 54480

ggccagccca agaatacagt gctgagagcc agaaggcca caaaaccaa gggctcacc
 54540

tgggtgtgtc ctctcagtc cctgcaaggc ttgttaaaaa tatagatccc caagtccat
 54600

ccgaagctga ctgaatcagc taggggcacc tggaggacc agactatgtc ttactcatct
 54660

ctgtccctgg tggcctgcag aggggctggc aaggagcagg tgcttagtga atgtctaag
 54720

ggtaagtggg cggggtgggg tagatgggtg cactagtgtc ctgaaggggg cttaaagga
 54780

gcaggtggcc tcagccaggc aatccggcag caggatatga agtgacacag tctaagctg
 54840

gcctcagagc cagcaggag ctgaggctgg gccactgtc cacctggggg ttgaacagcc
 54900

tgccccaggt ttgctgaggt cccagggacc aggcaggtc aagctgggat cacacatgtt
 54960

ccctctccag cctagagatg tcaaacaggt agattcctct ccattcata tctctatcc
 55020
 ttggcccaca gcccttcctt tcttgactt atcagagacc aagggtgctg gcaggcctc
 55080
 aggtggttaa aaagtgaag ttcttgagt aagtccaaag gcgcacacct gagagctgag
 55140
 tgggcaaaag gtcgctggct gaggctggg gatagtctgg ctttgagtc agatggacga
 55200
 gtccaaatct cagctcctta ccccgtaaca tgaagcctc agctctctga acctctgtt
 55260
 atttgcaaaa ccttgccaag ggcttcaaac aggatatact cataaaacaa gtatgcctat
 55320
 ttgggggact aagtgtgtga agtgcctgag tttaggagtg cagtaaataat cggatccctc
 55380
 cttcttcacc tcaggtttgc gaggtaacca gtgaggggac ttccaggcct agtgccttg
 55440
 tagaccgaac cactagcttt cctagagaag gcaactgagg cctagagagt ggcctccagg
 55500
 ttccgtgctc tgtggcagag ctggccttg ctccaagtcc cagggttccc agctcagtct
 55560
 cagctgatgc ctccagcct ggcggtctac gctgcctgct acattccctg tccagctaca
 55620
 aatgaacaac tgagagggct ggggtgagat tcaggcaggc tggggcaagg gaaggaagca
 55680
 ggtcatttgg ggagtcgggc tctctggtct tccagctgag gctggatcca ctcttctctg
 55740
 tccaccacg gtgcgtagt gccaggctt cctgagctga gcagggtgag cactatgaaa
 55800
 ctgctttcta gcatctagac cagggagttg cctcttgatt taaaatgagg atgtttacca
 55860
 gtttaagtag ttattaaaac ctaaggccaa gttggtgctt ttctccctc taggagttgg
 55920
 aagggtgtgt gctttttct gctcctacc caccctgcc aaaccatccc aatttgagag
 55980
 gctgtggaga ttacagaact tggagaggag aggggcaggg atcggggttg gagactagat
 56040
 ggttttcaaa cattgggcta cagaagtccc tcttctccc atctgtctcc tgcagctggg
 56100
 tccaccacc tgtccaacac agcctccac aggtccggg cctctttgtg ctccaaccgt
 56160

cacatccttg gcatgccaca caccctcacc atcctgcctg aaatgccagt cctcatcact
 56220

gacattcagc ctcccatgtg ccagcatccg ggccaaacac tttatctaca ggagctcatt
 56280

tgtacctgga gactcttagg aggtcgataa tgttattctc ccaaccatgt gactttccaa
 56340

ctccatatgg accccaacct ctttttttcc acccctaacc tgagactcct agccactaac
 56400

taactgccac attctcctgt ccaaattctg cctcatcggg tattcattca acaaacactt
 56460

accgagcact aactttgagc caggcactct aagaccagct gagtcccact tctggcagga
 56520

tgagtgttgt tgtcatcaac cccccacaa acaccgtgc cctctgcgta cttgcgaatc
 56580

tgagtgagga gacacggtgg cagtaaggct caggacctat agctcatgtg agtgagcaag
 56640

gatgcaaac caaagacacc cgctcccctc acccactgag ctggagggat caaccactgt
 56700

gcctgccatc agggctgatg ggcgactgct tgtggatcac aaactcttgc gttgtcactt
 56760

caaccaggg gcagagtgtg ttcctcttca ggtgtggcct ggtctcgtga gcctgcact
 56820

gaatgggcca ggcgttcatt gcttcagcac cgtaagaaca ctgtcaccgg agaattgttc
 56880

gataactcaac tcttcccagt cagtctgtct tctgcatgcc cacagttacc gggcagcgtg
 56940

ggcaacaagt gagcactgat tgttcgcaa tcaggctgtg acttactcca tctctgcctt
 57000

ctgcttaatg agagccagggt gggagtccag caaagatcac atgtttaagc cttgttatta
 57060

aacacgttaa aaaaatactg cccaagatt ataaagccga cttcggtaca tataaagaga
 57120

tttggggtcg gtgaatgggc ccactaagct ttattttgtg gtgtcagctg tcttggttca
 57180

tgtgtagaca gcagggtcc caaagccagc atgctgggggt ctgtgagga acaagcttcc
 57240

cattctctgt atttatttat ttatttat ttatttat ttttggaga cagaatctcg
 57300

ctctgtcacc taggttgag tgcagtgtg ccactctggc tcaactgcaac ctctgectcc
 57360

cgatttcaag caattctcct gcttcagcct cctgagcagc tgggattaca ggtgtgctcc
57420

accacacca gctaattttt gtattattag tagagatggg gttttgccat gttggccagg
57480

ctggttttga actcctgacc tcaggtgatc cacctgcctc ggccctccaa agtgctggga
57540

ttacaggtgt gagtcaccac agccggctgc ttcccatttc tataagcatc cgactggtgg
57600

ggtggccagt aggctgtggg gagtggcg tgattttcct ggtgggcttg agagagtaac
57660

agtatacccc cagtcagggc aagtgtctct ctgaaggtga actctgggtg cacagctaga
57720

ggatacagt gctgtgagct gggccctcca ggtgaggtct ggctgcca ttgccacagt
57780

gggcttgctg gcatccaggg gggctttctg ggtgactgcc caggcctgca catgagtcag
57840

tagcacgctg aatgcagtgg agtttgaggg gaccaccac tgtccatcaa atcgcccaa
57900

cctgggtgaa aggccagagc agctgcggga agtgacacca ggccctctgg agatgacct
57960

ctggcccttg gagatgtctc tggactggcc agaaagaagg tggcgcttg aactggaagg
58020

gaaggaagca agacagtgtc gggcccagca gccagaaatc agggtgatgg gtgggtggag
58080

gggcaggcat ctgggaaagc cctgccctcc tgggtgtccc agatctgaag gctggaggct
58140

ggaacacctg tgtcacctgt gctggtgtga gttccattgg ggctgccgt taccctcca
58200

tgacctttaa catctggcat ggcttggtgg ggtgtctgag aggaggagaa gtctcaaggc
58260

tgagggagct cagggtagag ccaaggtctc agtgagaacc ccagggggca gctgtcacag
58320

ctgacctgga gccctccagg aacgcgtact ctccagctcc tgtactctca aaagcaacag
58380

aaaggtttgc tttcctcagg tgcggtttcc tcacagtga gcttctccaa agagatctgg
58440

agacccccca gcagctttga caggaaaata gagtcctatc caccacacac caacacactg
58500

aagtagagac tctgggctg gggctgcccc gaaaccacc cacttaaact gccggggcga
58560

ttctggtggt cacaaatatt tgaaagcact gcctgacacc tgctggctca acaggtgtcc
 58620

acctaggctg gaggacagtg gtgccatctc ggctcactgc aacctctgcc tcccaatttc
 58680

aagcaattct cgtgcctcag cctcccagag agctgggatt acaggtgtgc tccaccacac
 58740

ccagctaatt tttgtattat tagttcacct aagcagagag cgatgtgagg ggctgtgagg
 58800

atgtgcaggg gtggagggca cttagggtccc tccctcagct gaccatagag caccagaggt
 58860

ggagaggggg ccacgtgcag tgagtggcca cagcagagca aggaaagggt cacacaagaa
 58920

cagtcctaag gaagacatgt gtttaggggt ttcattatca ctgtttaatc atcaggacag
 58980

agagggaaag caacctaccc aacgtcacac agctagcaca tggctggacc ggaattcaca
 59040

cacatctctg tccagctagg tcccccaact aacacagctg cctgtctgaa ttcaaagccg
 59100

ggagagattg tcaggggtgg gcgttggggg agatggcagg tgccttaagt gccaaggcc
 59160

ttgggctgga tctttaagga agggggattg ggatgggcca ggaggtgtgc aagattggag
 59220

atggaaagaa tgccattggc tgtggtgggg ggtggagtgg gcagggactg gaaggagggg
 59280

tagtggaagt cagtcctctg ctacaggcct gtgttagact ccgcggagat tcaggaatct
 59340

agccctgccc acaggggctt tcactccagt agaagagagg agaccaggg gccacagca
 59400

gtcagagaac aatgaggcct tggcttggtg tttggtgatg ataggggaaa cctgcccaga
 59460

acctgtttat cagatgctca gggatctgaa aaacagggtc tgcattgtgg cttgagataa
 59520

agatgcccg gatacttaga tgtgccctat tccaagatcg ctagggtcca aacaagttga
 59580

aaacaataga atacatttat ccctaagaat gggcttgac ataattcact tcaacaatgc
 59640

aaaagcttcc aaacctccag cctcctttgg gctgccggtg tggcctcaa ggtccactga
 59700

aaagccatcc ccagccagca gaaacatgcc tgtgtcttct tcagcccccg tccatgcttt
 59760

gccctggaga caagtgctgt catctgtttg tgggtctgtc tgtctgtctg tctgtccgtc
 59820

tttctcccca ctgtgagcac tgacaagcaa gtaaagatga caatgaagcc tttgtgggca
 59880

ccaaagctg acagcattat cactaagcca gtggcagcca tggtgactgc aataaatttc
 59940

cccacagaa attgataaat tgccactttc tgcagccatt atgtttgcct ttgtatcccg
 60000

aagagtaaac gcctctctcc ctttctctct ctcgctcacg tgtgtgtgcg cgcgttagat
 60060

gacatttatt cattttatgc atcctggggt ctactggctg tcccacctca gttcctgtag
 60120

caaagagact tgagtctgag ccactaatta tcacccgtga ggtttctcc ccgagcagga
 60180

agcagcaggc cagagctgcg ctctctcagt gcactctcca accaagcatc agtcaccact
 60240

cccggtccag cccctgtggc caagagctgg cgtgcaggct gcgggaggca gctggctgtg
 60300

caagaccctg gcagggccct cgctctctga gaaaccgaga gtcagaacca aagccaggct
 60360

gtcttggttg gagactgagc cagaaagggg ggctcacctc acggtgaggc tgtcgagtga
 60420

cctgagagcc tcagaccctc acgtcagccg gatgtcgac cagcctgctg ttggttaagtc
 60480

tggctaggac gcagatccaa gggggcatgg gtcgtgtgca aaccctccc accagctccc
 60540

tcaccacctg accctgccc tgctggaggg tgcccctga atctccagga ggtccgagga
 60600

actcgacacc tcccagagat ctgtactgcc ttggaacagg ctgcaaaaag tgcagtatga
 60660

aatgggaaat atgtatcagc tgcacgtttt catggcccc tttggaggaa agggagtcca
 60720

gggtattaag tcaggaggct tttaaatttc aggggttcaa ggatttatca cacacagtgt
 60780

cgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtttaagga tgttaatggt
 60840

tatgatgcat ttctgtggaa tgcaggccca ttgggaattg taaatttttc ttattggtaa
 60900

ccatttctgc attggtttag cacggaactg ctgctgcaag taaatgaacc agatttttct
 60960

ggaatgattc gaaaatctag ctgtttgtaa aaagctctct aaattattca ttcactttta
 61020
 ttttgcacag ttgtgtgacg gtaaactctt ttccatcaga cgaaacttgc gcgcacgtgt
 61080
 atttctgcgg ggattggctc gagtgtgctt gagttttctg tattctcttt tgtaggaact
 61140
 tctcagcctt ccttttaaaa agataataat tttttaatga tgtctaaata atgattccag
 61200
 gcagtattac aagaaatata acagtaaaac gtttgtaaga gtccatcaga aagggaanaa
 61260
 gccacagag aaaaatagaa atctaccaat tagatgtgtg tggctataag tgagctttta
 61320
 aaccatgatt tgagactggc atcctctggc ctctctgcc tcaggtcagc tgtgggaggg
 61380
 taggggtggg atagcaaaag ggacctggga tcaggtgggg cttcctggac ccagggcca
 61440
 gtaagaatga agcctagggg ttctaggata gggggcagcc ctggatttgc agaaacagcc
 61500
 tgagaatgag gggctgaaa gtatgagttg gagcagctca gtgaggagac agaaaactga
 61560
 agcagaagaa agcagaagac accagagtct ggtgctgact gggtgactgg ctgttactca
 61620
 catcctgagc aatgagccac agaccctccc caccacagag gcagcttgaa tgcagggact
 61680
 ggatggggcc agggcacagt ggatcgggg caggcaagag ccaactcggg gccagagga
 61740
 ggaccttta cctgactggg cctcagtttt ctcttctgca aagtagaagg aattatagcc
 61800
 ttctgcaggt caaacctgt gccacatcc acaggcacga tccccgactt cgctcgctg
 61860
 gcacttctct gcatttgact cctggttatt gcagctgtgc ccatgccagg catttgtag
 61920
 gtgctgggat gtgcctgcc cgctcctctc tgttggggaa gacggaattc atgctgtagg
 61980
 ggaggcccg agtttctaca tggggcagtc acagggtggg ggacttcctt gaaaacttct
 62040
 tatttgcata aaatgtgcgc tgcttttgac ttatattgta aatttcagat ccataaaatg
 62100
 ggaaagactt tgcagaagct gttattcact ttacctgact gggagatgta gattgttcat
 62160

atccaagggg gttagtttta atactaacat tagcatgtgg ggaaggggga tcgagatggg
62220

ggggagtagg gccccccacc acctaccctt gggacctcta agggaaagta cacaccaccg
62280

tatgctagtt gcatcctgtt acctgtgctg cagagccaag gggccagagg cagggaaatt
62340

aattctgcct tgtaaaggca ctgggccttc ctgttgaatt ttgctacttg cggggaaggc
62400

acttcccagg caaaaagatt aagcctataa cctaagcaga aagcaagcaa gagctttggg
62460

tggaagtgg ccagagacat agatcaggca gaccaggtt tggaagggcc aaggagacag
62520

cacatctgcc aatggcagca gctgggcctc acctgccctg acacccccac ctaatgggga
62580

ctgggcagaa cctccaaat ggttgccatg ccctggccca ggagcctcca aaaaccctga
62640

agccattgac tcagcatccc tgaagcctcc ttccatgagt acccccatag caggctgcac
62700

ctcaatgcac cttagatctg gtgtttgatt ttcactttga ccagatctgc agcagtctct
62760

ttgttaatat ttatttccag tgtttacatt catctttgca taagtcctca gtgttactaa
62820

gtttatcaaa tagcgaaatg cttggccagg tctaacaatg tcaccaagct gcagaagggg
62880

ttagaaactc atagtattat tatttcacat ccactgagta caagagagcc ccagctaggt
62940

tctggggaga acagctatga ctacaaaata ctcccctccc aactcacaaa tactcccctt
63000

agaatctagt ccttaaatca ttcaaaccac aaaaatgtat tgagcactta ctgtgtgcta
63060

ggccccggta ggtgttagag gcacagaaat gagtttttta ttggccccc tccaatgagg
63120

tgacagacaa gagaaccagt gtgtgcagcc aatgttaatg tgacaaaaag tgcagagtta
63180

gtcccagagc gtgaagtcct ccctgtccta ggcccagggt gaggccagga gctgctcccc
63240

gggctgcctg gaccagactt ttgagctctt gggaccctgg cgaacatcaa agctgttgca
63300

aaagtcagag ctcttgcaaa agttggagct ggcccatggc aggtggttga gaagtggcta
63360

catgactctt cccatgcaga gcggttagag ggtggaggcc gatttccagt atccctggga
63420

atccaaagaa agaggccact gaggggcagg ggaggggccc tccgaggctt ccagaaaaag
63480

gcagtgcagc agctgggcaa accaggataa ggcatgcatc cctaagccca aaaatcaggc
63540

gatgtgaggc aggccgagct ccatggcaga catgctgagc tgcctctgtg tgtcttggt
63600

aactctgccc ttctctaacc ccacggctca ccctgcccag gagcagggat gaggcccatc
63660

cctgattcag caaatcctgt gccataacaa ggggaggggt gggggcagcc acagcccaca
63720

agaccaccca gtcccagggt ttgtgcagct ccaggctctc ctggaagca agaaaccacc
63780

cagctccaga cgctgaggct gcagattcag ctgccactgc ctgatggcct ccaacctggc
63840

tggattaact ggcccatagg gagctggctg cccccacaag ccagctccct ggctccctg
63900

ggcaaaactaa agccctctc cctggctcca tagcgcccg tcatccctc attccagcac
63960

tttgcttctt cctgcattgt gtagctggcc tttctcctaa gagccttaaa agggcaggct
64020

atctttggcc acctgcatgg ccccgccgcc ttgcttcatg ctgggtacac agcaggctact
64080

ctgggaataa caaataagag aatttgtcag aggaggtttg gcaggatgga ctattgcata
64140

gaactgcaca gactgtgcac tgcacgcctc cagagggcac tctccacagg gattactatg
64200

taaatggcac ctctaaaagt tgtgcagcat agggacccta gggcttgagc atgccccacc
64260

cgctccgct gcttcaagcc cctggccttg gtgctgactg ttgggctttc ctcttccct
64320

tcccaaccag cactgaccat catccttccc atacgtggc gatgggtagt aatgaactgt
64380

cccagtgagc agttctcca gactgggctg tcaaagcatg ctagtgagag attcctggat
64440

ggcaggcagg ccattctggat tcttgctcag ccaggaaact ctgtcctctg ggaatatact
64500

ccagtcagg gtgttaggg gctcctggat cccaagatc ttcattgtcc cctagaactg
64560

ctcctctccc cagcctgcag tgggcctaata gaccacatg tagccaggga gtggtcccca
64620

tggggtttgt tttattgact cagctaacaa tgccacaatg actctgattg ccttcttctc
64680

acctgagaga ccagctctg ggcagccaag gggtagctag ggacagagca tatgcatgtg
64740

acctcaaga tctagaggag ctgagagaca gcatggatac tcgaagcccc cgtgagcagg
64800

gaggactttc tggaccaagt gggctgggag ctgggccaac tgtaaagacc cagggaaaaa
64860

aaaagtcaaa actgagttgc tcttgcagtc aagtgagtgg ttagttcatt gactccgtcc
64920

acagcccttt cctgagcagc tactgcagtg gttttcaaag tgtgggtccc acaccagcag
64980

caccagcaat tctggatgtg gagcccagca atatgttttc atcagtcctc caggtgacct
65040

tgatgcagct caagtttgac ctgatgtgtc ctaagtgtg gggtttgcaa aagtgagttt
65100

tacttagttc tcaccttgag gagcttgag tgggggagaa gatagacagg aaagaaaaca
65160

gtactgtcca gtgagacgta ttgctctgga gggccatgga ggacagaggag cagtgggtat
65220

aggcttgatc cttgcaaggg taagtggcct ggagaccca gtggcatgac tgtcctcggc
65280

aggaaccag atgttgggac cagcatcccc tagagggtc aaggagcca gtgccgacct
65340

cagtgcacca gtgcccagcc accagggtg tggccctgag cagcaggtcc tcctactctt
65400

ccctccacc cagatcttct gctctgctct tcccccttct tcttcccact cccaagtccc
65460

aaatagtagc aggtcttttc ctccgattgt cttggggaga aatgggcatg tacctccagg
65520

ccattatctt agagcacatt tttagaagat tttaaagtaa atttttccat taggcagaga
65580

acaccgtgta ttgatttggg tgcacttatt tctgctttac ccagacaaaa agatttacia
65640

aatatcctag agaggaatag aattaaaaat gagtaaattt aacagggtca gagcgtagca
65700

ctttatttgt ctcatataat ccttgccaca cctcatgaag gtgaggaaca tcagcccata
65760

gtgcagatga ggaaatgggg ctcaggaagg tgaagtgact tggccggggc gccagctagc
 65820

aagtggctctg gtgggatatg aaattccatc cctctgctgg gcccttgctg gctgaccttg
 65880

cagtttacta aggtgactga tctgaataa atccagagtg ttctaacgtt ttgctgagcc
 65940

cgtgtcctgt tggctgcatt gtggggagtg gcctggcctg gctagggag ctctcagtga
 66000

ccaattgtgg cttcagctgt gttatttaga gaatcccagg ggctgaaatg atcctctgcc
 66060

tggttagggg agggatgact ttaggaggaa gagggaccta atgccccact tagcagatct
 66120

gcagctgtgt gcatgtttcc acctgcacat ccttctataa accctcagcc atgtgggtga
 66180

gcagcagacg gtgcctcccc cactccccct ggaacacatc tgccctgttc actgacctgg
 66240

gagctggtgg ggtggacctc agaccccagc acttggcaca cacacatcat cccctgggcc
 66300

ctggaatggc ttgcatggat ggcagagaaa tagctgtgtg agcaaaggcc aagagaacag
 66360

tttagcccaa ggcaccattt ttacaacaac aacaggtaat ggtactgatt atataagcct
 66420

cacttgtcca gaggtcacgg ggcttaggaa cctcaattat gtggaaaaca aataggacat
 66480

aacttcagaa ggttgacag ccactgtgct agtgtgtgtc ttgtgttgct ataaaggaat
 66540

acctgaagct gggaatttg taaagagaaa aggtgttttt caaaaggaag acaaacattt
 66600

gaatgctctg gatatttca ggatcagtca ccttagtaac ctgcaaggtc agccagcaag
 66660

ggcccagcag aggagtggat tcatatccca cctgaccact tgctagctgg tggcctgggc
 66720

aagtcacttc accttcctga gcccatttc ttcattgtga ccattgtgtc agagatcacc
 66780

tggggagaga ggaagtaaga gagagagaga gggtagggag gtggcaggct ctttttaaca
 66840

agctgctctc gaggaaacta atacagtga aactaacagg gaacttactt gctcctctgc
 66900

cctcccccca gccaggagg gaatccaccc ccatgaccca aacacctctt attaggtgca
 66960

cctccagtat tggagatcag aattcaacat gaggtttttg ggggacaaac ctccaaacca
 67020
 cagcagccac caacattggt gttacaaaat ccatccaatg acataggctc atggagcgcg
 67080
 cttaagctgt catccaaaac atgtttactc ttggtttaca cacactagta gcagtctgcc
 67140
 atacctcaat ttcttttttt attttaggta acaacttttt aaataacctt actgagctat
 67200
 gcctcacatc accacgcaat tcacccattc aaatggtgag ttgcatggtg taccgtgag
 67260
 gatcactccc cattttctcc aaatccccc ttctggccac cactaatcta ctttctgtct
 67320
 ctattgattt gtttattcta gacatttata taagtgtaat tatgtaatgt ctggcctttt
 67380
 gtaattggct tctttccctt aatgctttca aggttcaccc ctgttggatg catacttagt
 67440
 tctctttatg gctgaatatt aatccaccgg atggataggc cacattttgt ttaaccattc
 67500
 ctcaattgat ggacgtttga attgtttcca ctttttgggt attatgaata atgggtgaaa
 67560
 atgtttttgt gtggatatac attcttcctt ctcttgata tatagctagg agtggagtgt
 67620
 ctgggtcata tagaaactgt ttaacgtttt gaagaattcc caggctgttc tccacggtgc
 67680
 ctgcaccatt ttacaatccc agcagcagtg taggaggggt ctgatttctc cacatcctcc
 67740
 ctgtctttta aaaaaatata tcattgtagc catcctagtg ggtgtgaagt ggtatctcat
 67800
 tgtggttttc atttgcatth cctgagggc taatgatgct gaacatcttt tcatgtgcat
 67860
 ttgtatatct tttttggaga agtgtctttt ccagtctttt gccacttta tgtaggttgt
 67920
 ttgtcttttt aataccgagg tgtaagagtt cttcagatat ttagaaacag gtccccttat
 67980
 cagttatatg atttctgaa cttttctcca ttctgagttg gctttgcact ttctcgatgc
 68040
 tgtgctcttt cgatttcgtc agcagcgaat tagcaggacc tgggggggctt taagagggtga
 68100
 gcaggcagac agacagcgcc cacaaagtcc tgaggatgag agaagaggca gagcttacac
 68160

aaaggctctg tgaaaactga agatggaccc caggcaggtt ggggtgatcc ctggggtgct
68220

ggatgaggat ctgattttgt actgcatttt ggaaggattt ctgccccag aggtacata
68280

ggctccagtg gtgtgtgcag tcaggcagag agatgcacag aacctctaca gtgtgccagg
68340

catggtgtgg ggttctgggt gcatgtggaa caacacacag caatcttgcc ctcccaaagc
68400

tcagggaagt gggggcgggg ggtggacaac gagcacatga gaaaaagacc acaggggaag
68460

aactcccaga aacatcccag ccacagcgca atgccccagg tgtatatggc ccttggtac
68520

ttcatgccct tgagggaaaa tcggtatttt cattttacac ttgatgaaac aggttagaga
68580

aaggactggc gtgtccagta cctcccagtg agaagatact cagagtagac ccagatcca
68640

ggtccttcct ctttccctct ctacacgttt tcagtctaga aaagagttct gtgcgagaat
68700

ttttttgttt ttgtttgttt gtttgagacg gagtctcact ctgtcgccca ggctggagtg
68760

cagtggcgcg atctcggtc actgcaacct cgcctccca ggttcaagcg attctcctgc
68820

ttcagcctcc ggagtagctg ggattacagg cacgtgcaa cacaccagg taatttttgt
68880

atttttagta gagatgggt atcaccatgt tggccaggct gtgtgcaaga atttatgttt
68940

tcacttggtg cttgctgagc cctgggtgctg gctgtgggca catactcgtt caccagggga
69000

catgcactgt tgatgggtcc ttgcactctg ggctatcttg ggccccgttt gctcaatgca
69060

ggcccccttt tctgaagcgc ttttcttggg ctcactctca attcgtcaaa gttcggctca
69120

ggcccacctc ctgcaggaat cctcccagtc ccgctgtcc gaagtcagta cccttgctag
69180

gagaccttaa acactttagt ccactgcctg gccctccat ggggcaaagg tctccagggt
69240

gtgtgtgcac ttccttgggg cttcttaagg ctcattccct gggattaggg aggacataca
69300

gagcctcctt gcacctacat ttgtttacat attacttata ccacatacaa cgttataggg
69360

cactgctgca taagtgtaca ttttgaatgt attttactta taaagaaata gagcagtaat
69420

aggaacatgt gcccctaaat gtaattttac tgaccagaca aaaaaaattt tttttaattg
69480

aaaataagta gctggacgtg atagcatgcg cctgtagtcc cagctactca gactcagaag
69540

tctgagacaa gagaattgct tgaacctggg aggcagaggt tgcagtgagt cgagatcgcg
69600

ccactgcact ccaggctggc gacacagcga gactccgtct caataaataa ataactaata
69660

aataaacaaa aatttaaaaa aattgaagat aactgtgggc tgcttgagga caagaactgg
69720

gttattttct ctctctttgc tcagtcagtc cccacatggg aggcgtgcga ggcaccccg
69780

tcttggcctg caggccgccc cactccatc tggctgcagg gtgccacagc cagcctctga
69840

ttgccagtgg ccccagcacc tcaatccata ccgacctcac agcaaaccct ccgtggccac
69900

agcctgccct ggccactcat gtctgaaagg tgatcggtgt ccatgaagaa gaggccatgg
69960

ggctgtggat gtcattgtggc caactgttag tggcctgctc aagagcagac aggagctggc
70020

aggcaggccc atgggccagg cgtggctccc agccataaac acaggctgta gccctgccgg
70080

gaagaaggag gggacagaca agcctgggca gagaccaga ggggaccagg acaccgcatc
70140

ctccatggca gttctgtcca cagcttggga ggagaggcag gcaggtctgc aggcaggaga
70200

caggcctcgg cccacagagc tgcaatctga tcccatctgc tccctctcca gttctataaa
70260

tagaaccccc agggagccag ggcccagaca gggtagtca caccctcacc tcaggccgca
70320

gggacaggcc tctgggactt ccaagacatg taccacctgt cctgccccca acccctgcca
70380

gccttcagca ggaggggttc tgctgagcat caggaggggg cagaagaatt gaacgtggcc
70440

actgtccatg aagaagttag agaagagtac acacaacaca catgcacaca agaataggac
70500

acgtgcacac acacacacac atgcatacaa gattaggaaa cacaagtgcac cacacaacca
70560

gtcagaaatg agctctgagc ttttacacat attaactcat ttcacccca aacagtcttt
 70620
 agagagggtat aattagtgtt cccgttttaa aggtgaggga actaaggcac agaggagcta
 70680
 tgtaactcac ccaaaggcac ccagctactg gcatgtgagg ccatgatttc tccccagcag
 70740
 cccacccac agcccggtgc ttaaccacc tgcctgacct cccccagcca agtggctccg
 70800
 tccctacatc agaaaggta aggaagaggc cgcctaggga gcagtacaag gcagcattac
 70860
 aacctggcat tttcctgttg cagactgcat tttggtggca agagaagctt tggagggtgt
 70920
 taacgtttag tttgcattta tctaccgct gaaaagcagg gacccttagg accataagtc
 70980
 taggactttt catagcacc tagactgctg gggatgaaag gggcttgagg ctgagctggt
 71040
 gcagcagctc ctgttagaga agccaccag agaggccag ggacctcta tgtccagcg
 71100
 ccagttaggg gcagacgtgg gaccagctcc cctcggtgc agcggcctcc atatcagctc
 71160
 cctggaagg agcagatgcc aagtcagggt ctgagcaaca ccctcaaggc ccagccctgg
 71220
 caatcccttc ccacctgtga ttaggtctgg ggttcaacaa aaataggact caaaatctgg
 71280
 atcttctcgc cttgcttcac tttgggtgg ttcagtcttg tttcgaggta aatcattcga
 71340
 ggtttttcag ggatccctcag atgcatgtct gcgtgaagcc gtccctggctt ccccttcggc
 71400
 tgccccctgg gcaagagcct gggccatacc tgtgccattc ctgagctctg tctctagctg
 71460
 ccctatgtcc agtcattggc gatgggtctt attcaagccc cctttccccc aaccacact
 71520
 agattgtaag ccctcagggt ggggaccgct tctcactcat ccttatatgc ttaggagtgt
 71580
 tcagcacaac gtcactcagg agatggcagg cagcggggag agggcagaga cagaggagag
 71640
 gcttctgtcc agagggttcc tacttaggat catccaaatc atcagcaaat gttggaaaga
 71700
 tgggtctggg gcctcaggct ggctactctt cctccccatc agggcaaac tccccgggg
 71760

cttgtgtctc cctgtggttg gggacaggcc tgggtgtatc tgtcccagcc ttgactgacc
 71820
 acaaggggag acttggccct cccgaccaca ggagaagagg gatttgaatt cacctcggcc
 71880
 caaattgcag atgctggcct gcgagccctg aggccatctg caaggctcct gggctgacat
 71940
 aactgtcacc cagctcttca tcctccttcc ccggggctga aggggtgcct gcagggtctt
 72000
 ctgccctccc tccacctggg ccctactgct ggcacggagc ttcatttctg ttggcagcca
 72060
 gtgtttgctt agtgattact gaatgccgtt tgctgggtgc ttatgatgtg ctcagcactg
 72120
 ttctaagtgc ttacacaat tgacttatct actcctacaa cctgcccata agggaagtgc
 72180
 tgccatgggc tcccttttat agctgagaaa actgaggcac agataggctc atggtcacac
 72240
 agttaacata aggcagagtc aggaccagaa tccaggcgga ctggctccag agcccggctc
 72300
 ctggaccgag cagccttgag tagggacctg gtgcagacc gcctgccgtg gactcactat
 72360
 aggatgggtc aggaggcccg ttgtcaccta caaaatgtag cccaccacc aggaaacgct
 72420
 cccagaacca ggcagaaaat actgtctcct attcacaggg tggtgtgccc tggcgtgagt
 72480
 gtcttgccgg gtgcctgcat gtctgtgcag gagtgccacc cccatctgtg gactcggaaa
 72540
 gcatgatccc caaaaagca acccctcaga gtctcagctc agccaatccc tgttatcctg
 72600
 tgacaggcag ggccatgttt caaaaccatg ttaaagatga gcaagaggct aaggaagttt
 72660
 cccatgctgg cagaggcagg acaggaacac atgtcccctg actcttgaca ttgagaatgc
 72720
 accccttcc acccacctgg cctcttgtcc cagccagtgt ggcccagct atgcagctct
 72780
 ctgggacatc tctggaaagc tcattgttct ctgtctgggc agtaagacgg agtgaggctc
 72840
 tgactgccga ggcagacaga cctgcctcct gctcttctc ccctcaatta tttgtgctgc
 72900
 ctgaaacaag ttcttgagcc tccctgagcc tcagttccct cctctgtaaa gtgggaatga
 72960

taactgaatc tacttcctgg ggttcttggt aggaccagag gcccatcacc cgaggaaagg
 73020
 cctgacagca gcaatgggtg tatggtaggc atccattaat aggaactgct gctgctgctg
 73080
 ctgccttata aaaggaaaaa tcacagtttc aaaaacctct taagattgca tgatttgtca
 73140
 tctgagagca gagattgtga tctccattga caaagatgga agctggggcc caggaggga
 73200
 cacagacttg tctgaagaca cacagctgct tcatagcaga gccggggcca gcctcctggt
 73260
 tcccctgccc catctttatg tcagtccctg gatccactcc ctcacctggt cagggtgtcca
 73320
 gaggttccag gagagtgttg tggaaaagta ggaagaggca gacgcccatt gaggggtgtg
 73380
 aaagccaccc cagctgctct gctgcttcag ggcaggggat ctttaagccct gtggaactgt
 73440
 aggctggtga agcttggtga cttctgctca gaataatgtt cttacatgca ccagacaaaa
 73500
 tagcttaggg ttacaaatga aatatattat atggaaacac agcgaaaata ttttttaaaa
 73560
 aagactacag tagcattagg catgcttctt tattagcaca ttaaattgaca agctctagt
 73620
 gcagatctaa tgtctaccat aattttatct ttacttttat tattattatt attattatta
 73680
 ttattattat tattattatt ttgagatagt cttgctctgt caccctaaact ggagttcagt
 73740
 ggtgcatca cttcagtctt caactcctgg gctcaagcga tcctcctgcc tcaatgtcct
 73800
 gagtagctgg gacgacaggc gtgtgccact gtgcctggct aatttttttt ttattttttt
 73860
 agagatggga cctccctgtg tggcccaggc tgggtctgaa ctctgacct caagcgattc
 73920
 tcctgccttg acctcctgaa gtgctaagat tacaggtgtg agccactgca cctggctagc
 73980
 atggttttta atttgtgctc aatttaaagt gtattaggat aaattattta caggtaaatt
 74040
 tatgatatga tcttggtcag tattaatgaa aaattatgtt gcctaccttc atacttaaag
 74100
 gaaatagcaa attttagtta gaagttgggtg acaattaata tataataatt tttccatcca
 74160

agtttactgt agaccgatgg cccgagttca gggtttcctc cacatocact catatcttgg
74220

gctcagcaaa gaagatcaca gggcaccac tctatcttag gggcccatc acttgtctct
74280

cagtcctgcc ctgagagatc aaggtttggg gttcattcat tcattcattc agcaaact
74340

gagcactgcc tccgtgccag gctatgatac acagtggagt ccaaggtcag atgagatgag
74400

gtcctgccct caagctgttc accatgcttc agggagagca taagaagaac tgcctcaaga
74460

tgcaaaatgt agacaagcat agggacatgc tcaccttcc tggatcatca acaccagcg
74520

atgggtgagg tggggttaga tggcactcc tggctagtga cactgtccat tgcctaacc
74580

ttttgaaaag aagtatggca ctgcatcaca ttggatcatg aaattccaga gttcccttc
74640

ccagcactct cactctgccg gtttttccc aggaaatcat tcagcaggag ggacaaagg
74700

tggagggcag aagtcaccc tgccttccct gagtctgtgc tctcccagg gaaggggctg
74760

tgttgtgttc agcttgttcc cctgggtgtg acacgtggta gcactcggg aacacctgct
74820

gggtgaatgt ctgactgccg ttagttacg atgtataaga ggagggagg ggcagaagg
74880

gcagtcagg agccctccac tgaccagggt cctgacttct cttgcagcca ccaaatgcc
74940

agaagagatg gacaagccac tgatcagcct ccacctggtg gacagcgata gtagccttg
75000

caagggtccc gatgaggccc ccaaagtggg catcctgggt agcggggact ttgccgctc
75060

cctggccaca cgcctggtgg gctctggctt caaagtgggt gtggggagcc gcaacccaa
75120

acgcacagcc aggctgttc cctcagcggc ccaagtgact ttccaaggagg aggcagtgag
75180

ctccccggag gtcactcttg tggctgtgtt ccgggagcac tactcttcac tgtgcagtct
75240

cagtgaccag ctggcgggca agatcctggg ggatgtgagc aaccctacag agcaagagca
75300

ccttcagcat cgtgagtcca atgctgagta cctggcctcc ctcttccca cttgcacagt
75360

ggtcaaggcc ttcaatgtca tctctgctg gacctgcag gctggcccaa gggatggtaa
 75420
 caggcaggta ggttctgggg gaataatacc catcgtaaca ataaatataa atggctaatt
 75480
 ttcacgagt gctgccagca tgctctttt gatatgttat cataactaat cctgcatcac
 75540
 tgcaaaattc cttttacag aacaggaaat tgtggctcaa agacattaat gatcttactg
 75600
 aaaattccat agctggtaag ttgggatttg aaccagttc ctatggatct gaaacccatg
 75660
 ttctttctgc tatggtaaca gggtttcta gagggcaagg aggaaaagg tgatggcttt
 75720
 ttgctatttg gagccaagaa ggagagagac ttggctttta gaattggact cagaaggagg
 75780
 aaagtaacac acttatgtac ccggccctga ggaaggcagg catgtctcaa attttatcgt
 75840
 atttattcta agggcttcca agaactgatt tctcagaaaa gtcaacagca ggtttttctg
 75900
 gggctagata atggctccac ggaacagcta gagctgcata cgtttcttgc cacgaggagg
 75960
 tcttaccaag acaaggtggt ttttaatgcc accttggttt ttaatgcatt ttcttggcga
 76020
 aagaatgcca gactccatgt gtgagaactt ccaccagtga gtcgagccca ggtgccaggc
 76080
 tctgtaggac agtttaaagg tgcccagcca tacctgacta cagcacctta aatgcaggct
 76140
 gcatcatggg actcccagg atggaccagg aggaacagtg gagagtgtg gatgcagggc
 76200
 ctggagggaa cggagcaagc aagctggctg gaaaagactt ggaggagggc tttgacctgg
 76260
 gccacacctca tgctgtcagg ctcagggagc agggagtgc ttcagctctg agggaacaat
 76320
 ctgaaaaggc acagagataa gaaacaatgc tcctagacct cagttttctc acctgaaaaa
 76380
 tgaagactat agtacctacc tgcgggctca tggcaaggct gagagatcat atatgtaaag
 76440
 ctcagcatag agtaggtgtt cagtctgtgt ttggagacta aactgaccaa tgcagtgtca
 76500
 acaaggccca cggggtgctt ggggatctga ggactgcaat gtgcctgcag ggggacctcg
 76560

tgggtcagag cagcgggctt caaggcggag agggcaggct ggagctggct tgtagaaggc
76620

cttatgtgct ggctgagggc tctactgagc cgaggacat catcaagggg ggtgattctg
76680

aactgagggg tggagctgca gccacagctg cccattctgc accatacgca aaactctttt
76740

gctgagcggc tcccagggca ggccacggcc ttggagacag ccatgtcggg ggcccatggg
76800

gactacctgg atgagcattt tgcttggtggt tgctgatgtc cacgagctgg gcaagatgct
76860

tactgtttgc agcttccttc cttgagtagt tggcctggct tggcaggtgg gccagaggct
76920

gggaggcctg ggcagcagag aaaagacccc tagcctggga gctagaacct cattggctcc
76980

cacctgtacc attgactggc agctcacttg aacaagtccc tgcccctctg gccctcggtt
77040

tcctcagctg ctcaagtagg ccctgcccac tcctgtggcc tgtgacgccg tctgactgcc
77100

ccacttttct cccgcagggt cccatctgcy gtgaccagcc agaagccaag cgtgctgtct
77160

cggagatggc gctcgccatg ggcttcatgc ccgtggacat gggatccctg gcgtcagcct
77220

gggaggtgga ggccatgccc ctgcgcctcc tcccggcctg gaagggtccc accctgctgg
77280

ccctggggct cttcgtctgc ttctatgcct acaacttcgt ccgggacgtt ctgcagccct
77340

atgtgcagga aagccagaac aagttcttca agctgcccggt gtccgtgggtc aacaccacac
77400

tgccgtgcgt ggccctacgt ctgctgtcac tcgtgtactt gcccggcgtg ctggcggtg
77460

ccctgcagct gcggcgcggc accaagtacc agcgttccc cgactggctg gaccactggc
77520

tacagcaccg caagcagatc gggctgctca gcttcttctg cgccgccctg cagccctct
77580

acagcttctg cttgccgctg cgccgcgccc accgctacga cctgggtcaac ctggcagtca
77640

agcagggtacc caccceatgc ccttcctccc tctggcaact cagcacatgc ttgtccagca
77700

cctccccccc ccaccaacca ggtgcagccg ataccacgg gtgcagcctt accaggtgcc
77760

aactgcagat tctgttccaa tggggcagca cccacatcct cctgccaag attttgagta
77820

tcagtgaaca aagcagatga aaattcctgc ccccatgggt acagcccagt tggggcagac
77880

ttaaacaata tacagcagtg agtcatagtg ggcgttgaga ggatcaggag tgctaaggcg
77940

caaaggaaag caagcgtggg cgggtgggggt gcaggtttga agatggtcca ggaggacttg
78000

cagagtgagt gctatgtgag cagagctgga aggaggggag agggggccggt ggcctgctgg
78060

ggtgatgttg caggccacgg ggcatagtcc ttgggcagaa gtgtgcctgg tacgttgag
78120

gagcagccag gaggcccatg tggctaaagc tgaacaagga gggggtggca ggcgaggcct
78180

gagaggaaga gcttcagcct taggctccga gtgaaatagg gagctggtga aggtttgaat
78240

ttaggtgttg atcggcctga cattttaagt ggttctctct ggtttttggc tgagaacaga
78300

ctgtaggagg tcaaaggcag ctagctgctg gggccagcga ggaagccgtg catatgccca
78360

gatttgaggc agtggaggct ggaggcagtg ggagcagtgg aggagcaagg agatcctgaa
78420

ggcacatgtg aagctggagt gcatggaatt gccgctggaa tggccccagg gttttgtggg
78480

tgatcaggga tgactctggg gtttgggtct gagcacctga aagagagagt tgccatcagc
78540

tgagatgggg aaggctgcag gagtgagcat tccaggttta ggtttgttgt tgttgttgtt
78600

gttctttgtt tgttttactg ggggtggcac taatcattca tttagaaaat tcatatgggg
78660

cttggttgcc agcagggttc ctgggaagtg gaggaccttc aggccacacc aggtggggac
78720

ccaggcatca gctggcacct gtgccacag aaacaggccc ctgatgctcc ctctgggtcc
78780

acaagctcat ggggcttgat tottagtccc aagctcacat cttcagcaac actgtcttta
78840

ttttctcaa attaatcca cctccagtgc ccagccccgg ggctggttga tgcataggtc
78900

ctgcccactc caggacgcta gtcaccaac agggttcagc ccagggtctg gtgagacact
78960

tctcatccac ccctcattcc agcctggcag cccctgcagc tctgtctcct gtcagaatg
 79020
 cagggacctt ggtgaggctg ggcctcccc taaacaccta gcacagggcc atccctccg
 79080
 agcatggggg attggcagga ctgggaagca gggctgggca cctcctgaga accacagacc
 79140
 tcagccccta ctactgtca gggttgtgc ttctctgcca gcagctactc gactgccctc
 79200
 gtcctttgcc acccagat accggcatca ttgcatctgt gggctcagaa tttcctaaga
 79260
 caaaggaaga acccagact tcaggcagaa aagcaagatg cttgatggaa atgagaggaa
 79320
 ggaagaaat cagtcatgaa cccttaacaa gagctcgtg tgttccagga accattctaa
 79380
 gagcttcgta caccactgtg aattcattta acccatctc atccctgtag gccaggtgct
 79440
 attatctct tatttcatag atgagaaatc tgaggcacag aggggtcaag taacttggcc
 79500
 aaagtcacac agtcagtaag tagcacagct gggattcaga tccagctatt ttccagatcc
 79560
 atgcttatac ccacgtgct ttctacaggc aaaccaaatt cacaattgag taaatgattt
 79620
 cagaactagg tcattcattc actcattcac tgagcaccta ctctgccagg cctggggcc
 79680
 acagaaagga gtaagaaaga cacagaccat ctggctctc gcctcagctg gtgtggggcg
 79740
 gtgtggaggc ctgaaggcta gaggtggac tctgtctctc cagcatcgtg gactggaaca
 79800
 aagatcttga ccttgagcaa gtcacagccc ctcttttctc ctcgaggccc tgatgtgtag
 79860
 atcaagctct ctctagtgc taatgtttc aatgctctt gacgatattt tcttactgac
 79920
 ttctacaaag caccagcccc ttggcaggag gctctgcggg gctccctggg gactagtaca
 79980
 atgttgctgc ccaggacggg gcagatctgc atggaggagc cactccctc tcccctgagc
 80040
 tggcagggtat ttccagcagc agtgggaaag acggggacgg tgcccactac ccccagggt
 80100
 aaaagacaca cactgctcc ctccctcccc tgccaccctt caactgcctt tgtctaaatt
 80160

ctctgccctg ccaaggctct ccctgatatt ctgttaaccc aattctgcc tgatgcaaat
80220

accagggcac agaggggcgc atgcaaggta attttccccg tgtcccaaaa atctttaggg
80280

ctgcaggccc tggtagggtt gccaggcaac catccctccc gcaggacgag atcagtcaga
80340

gacaccgaag gcctctcatg caaaggggct taaacttggt ctgcagagtt cttaaggga
80400

aaaccagggc tgatgtgggt atgtcggggt gggtaggtcg agggaggcag gggcttgaca
80460

cacagcagcc cagatacggg gaggtgccc agcctgggca catccaagca gattctggat
80520

ggtccatggt aagccatgcc cgcggttctg ggctgggca cctgaggctg ttctgccttc
80580

agcccaaac agcagggact cctgggtggc tcaggcttga tgtgtggctt ctgctgccag
80640

cccatgggca cctcaccaca gggccccac atggccactc accaccccaa ctgccacta
80700

ggccttggg tcaaggagc acattagtct aggcacaaga aactcctcca agatttctt
80760

acagccctga gtctgcagat taggaggcaa aggtgtatgc cagctgtaga gaaaacctt
80820

gagaccatgt ttacagacag tattattacc aacaataacc gatatttga cagggcattt
80880

ggatttgctt gtgtatttcc ctggaaatac acatttcaac atcaaactca aagatcaaga
80940

tcaaagaaat gttaacctg cacactgttc tcctgttacc ttcacccgt tctoctatta
81000

ccttcacccc aagctacatc tgcacatggc tggcacagga tggctgaaa ctctcatctg
81060

accatgtcac tgctgaaaag atccagtgc cccagtgtcc ttgggataat ggtagaacc
81120

tcactatgca gctagactcc ctccgcagcc tcctctcccc ttctcctgtg acgtgtcaac
81180

acttgggtcc catactgtcc tctaaatggg cagagctcat tcctctcggc ttggaacgcc
81240

caccccagga caccttgaca gcttgcaaac tcctattcct ccttcaaac cctacctctg
81300

ttaagctccc ccagactatc cccttacttc ccataacctg cctctgtaat ctgttgcat
81360

ttgtaaatac ttttttatta tttttcttac acactacttc atacccgact cagagggtcca
 81420
 tgactgggtc ttctacacct tgggggtccat gggtcctaac acagggtcac agtgcattgt
 81480
 cactgagtga ctgattttta aattggccac cacagagcca atatcaggca tcttaagagg
 81540
 agtgcacgta gagcttgagg caagaactag tctaactgcc ttgctgccac taagaacaac
 81600
 agccatttat tacattaagt gccactgtg tgctgttcc atgtactca ttccacaaac
 81660
 gtaatgcttt tagtctgca gtaacacagc aagatacaat gatgctatga gccaggaaag
 81720
 ggactgccca ggggtctcatg actgcccacc agtggcagag tcagcgtcag aatccaggcc
 81780
 tatgggtctct atgccccaaag gttgcccctc cacaccctgc cccccctcc tctctgagt
 81840
 ctggcaggca gtgagttggt accactctct tcccatccat ccccactat gaagacatgg
 81900
 agaggggcta atgctgcagc ctactgggtc cttcagcgtc tgacctgcaa cctagaggga
 81960
 aatcccacac ctgaatgctg tctttccatc tcttgttgat gaacactctc cttggccttc
 82020
 tcaaggagaa ggtggtgaac ttgatcatct ctaaacaatag agacttactt actaggggag
 82080
 ctgttacatc tgcgttccag ttagccatgt gtgccactgg tcaccaggga gctctgagca
 82140
 caacactgtg gtttaatagg ctggctgaga atctcatgga ctctaaggc tggcttaggt
 82200
 tctgtctca tgcgggtctt tgccatccac tgctccctag cccacctcat cctcccttc
 82260
 ccagcgtcac tgcagactgg aaatgttaga ggaaaaaac aaggcatttt acaagccaat
 82320
 gtggtatacg tgtgtgtaat catccaaact ctgacttgtg acattaactg cctgatcttc
 82380
 ccgaggcgca gagaggaatg actaatgaag cagagttaaa gaacaaaca cttccagcga
 82440
 ggagatgact tactcaattc atgctgtgag gtgtgagtaa gacttctctg ggcagaggta
 82500
 cggaggagac ctctctgcc tctgcacgcc tttccagcta atggtgactt caaccacagc
 82560

aggcacataa gcaaaatcaa gagaatttca gacatgacct catccaaacc tcctaggtag
 82620
 agaaactgag gccagagac agtcttgcca gagtcacatg ccactgggtg tcaggacaag
 82680
 aggctgggtc cctagactcc tggtcctta gtggcttaa gtcaccaaag tgagccaggc
 82740
 ctgccctaaa atagaacaga gccaatcatt tttcacagat tgcaggaaaa gcactcaatg
 82800
 gatggcattt tctctgtttc tttaaacttt ctcttataag ctgtacaacc attttgctca
 82860
 gccagtgga ttcactttta atattttatg tggtaactga gctcgtgttt ttttttttct
 82920
 gctgaaattg tttctcaaaa aacctgtatg agtcattctg ttttgagagg ctgcctatta
 82980
 aagtgtagtc acatcttcta atggaaatac tgttccttgt tctgctgaca ttatagaata
 83040
 aaagtaccag catttcacag tgtttcacag agaacacctt gacttccagg aactctgcta
 83100
 cccggtctag ttgcacctg ataaaagttc ttcctcatga gggcactatg cgagtgagtg
 83160
 ttggctgggc taggaggggg ctgagtgga aggtggggag agaaggggtca gaaatctcca
 83220
 ggttcatagg ggctcccaa tttatatgtc ctggaacttg gtgggggggc ccagatctct
 83280
 gtgagggcag ggtaagggca ggtatcttta ttttctgtag ccagggcct ggcacacagt
 83340
 aggtgcttcg taagtggta ctgaactgct tggatgaatg ccgtcagtct ctcttgctg
 83400
 gcaaagtga gccaaaggctg tgtgcctcag gcagaagtag agaactgggt aggaagaggg
 83460
 gaggtgaga gctttgggtc ggggagtggt gggggctcag tgtctgtatg tgtctttgta
 83520
 tgtcctgagt gtgagccagt gatgggggca gtgtggggct ccaggaagac agagggctca
 83580
 gggtgagaca ggctgggaag ggaagcagcc agtaaggtgt gagtcattg gctgagggga
 83640
 ttcagagcta ggagagctcc tctctctgt cctgggatc tggggccgga aaaccgtgag
 83700
 acggagccat gcttaagggc tgtatttctg gttgactcct tactctgctg atgtgactgt
 83760

tctgattggc tgccccgggg cagatatgtg aaagccccag gagccccggt cattggcacc
83820

aggctgagtg ggtcctgac ccagctctgc atctgttagc tacatgtgac cttgggaaaag
83880

tgatttgatc tttgtgcctc atttcctcca tctctaaaat gggctaatag tgatctctac
83940

tttgtggggt ggttttgagg attcgatgag ttaatgcctg agaaacgctt atcacagagc
84000

caggtagcgt caggtagcgt gtgagcgccc gccgtcttgt ctttcttggtg tccttcatca
84060

ttgccctccc ggggcaccag cctttgccac agtgttcatg gtttctcttc catccatgtc
84120

aggctctggc caacaagagc cacctctggg tggaggagga ggtctggcgg atggagatct
84180

acctctccct gggagtgtg gccctcggca cgttgccct gctggccgtg acctcactgc
84240

cgtccattgc aaactcgctc aactggaggg agttcagctt cgttcaggta aagtagtctc
84300

tagtctgccca gccagcttca gcgtggccct ggcccacctc tcatgagtaa gtgctctgcc
84360

tttgagaact gcctgggctc tgatcatctg cacagcagga cactcgggtg agatgccagg
84420

ccttctgac ccagagggcc catccaagcc tcttgatagg ggctgcctgc tacaagctta
84480

acctcccaac acatgaggaa gccagacaca ggcatgtgaa aaaattacag accaggacac
84540

tgagccctgg gtccaccatt ctcaagcttc tagaattctt gagctttcca aacaaaattg
84600

aatgacatga agtcattgga cattttgctg gtctaggtag gagtcatcat ctacctagac
84660

atccatctag gtaggagtct agggagttag cgagcctagc ccaccaccct ccaccctca
84720

tctcaacctg gctctgttgt tccccttaac cccttgattc atgcagcgtg gatgagacac
84780

aagcattagg accccagttt ataggggatg cgacacagct ttttcccaag gatgccattc
84840

ctatgaaaaa atccagaaaa aggaggtctg tcatcagatg aatttgagaa aagatttact
84900

ttcccctaag agatcacaat gcactttggc acatcaaaga ttcgaataac atctacaaca
84960

aacaacctt tgtaatcag acttattgga gtgttagcac cctccccctc cttttttcca
85020
aataacagct atcagcactc cacaagctgg gcgctgtggc atgtggctgg gatcccagct
85080
actaacagag gctgaggtgg gagaatcact tgagcttgag cgattcttcc aggagttaga
85140
agttcaagtg tccagcctgg acatcatagc aagaccttgc ctcttaaaaa aaaaaaata
85200
ctccacagaa catactctgg gaaatgctgc tagtggcaaa agtttctttt tgcatttgtt
85260
cattcaacaa acatttattg agcacttact gcaggccagg catcctgcta ggctgtgat
85320
gcaatagaga acaagacagg cagagccatc tttccatgat aaaggaaacg gaggcaggga
85380
cattgtttac atttctttgg catctccac tgttctccac acagaatctg cggtcagtg
85440
tggcggatta actcctcccc tcccatgagt gggagctgac taatgctgta tcctgccatt
85500
gtaacagggc agagaacagc catcctatcc acccccacta tcaacagaaa tctgtgcagg
85560
gaggacgtgg tacaggactc agggacagtg acagcgtaat gctcttacgt aatccaaatt
85620
aattgctcct gaccagggat gcttcacatc cattggaaca acacagagtt gtagccgtgt
85680
gttttttaaa gtgtgtctgc ctctgaatct ccctggctct gctcatctca gaaggctcct
85740
ggaggagggc tcatggctga cacatgttag catacagaaa ggggaagaat tttatttatt
85800
tgccactaat ccccatgtgt ctcatacca aatctaggat tttctcttc ccacctggag
85860
tatccggcgt gcatggttta tccacaatta aagcagactg caggctgctc cctcaccag
85920
aaagccgctc ccttaagcag ccatagatta catgccctgg ctctgcctgc cagagaaagc
85980
cggtggattt gattttgtta gcaagagctg actgagacct tagcagctgc attctacttt
86040
acatattaaa gagtggaaat acaaatagga cagtcctttt gccactgcaa gagcaggggg
86100
ggtaggtgg gtagaaaggc aggattttac ttgtggaaat cagtgcctgc cttccatcca
86160

caaaaaccag atgtcaatat ggaggaggca catgtcagcg ggcctggga agcctgagtt
86220

ccagctttcc ctccccacaa cctgctctg taactgtggg cagagcatgt ctgctccaag
86280

agcctcagtt tacaccctg aaaaatggag ttaatgcctt acagggatth tatgagccac
86340

agtaagagat tatgtgtgaa agaaaaataa aaagcatttt gctattattc ttagggatc
86400

tggccacatt tccattatct ctgactaagg taccaaccac cgatttctga gttcagtagg
86460

agacagttag acagaggcca aggaacccca accctggggc taaaagtggc ccagtcagtg
86520

tttgcaaaact ggcggaact cactagtggg tcatgcaatt gattccatgg atcacaagca
86580

gcattttaaa aatgaaataa aatagaatga caaagaaaac agtggagtgt tttgcatggg
86640

gtcgggatac tttttcagtt atatggaatt gctagtthta tttgcaaatg tagcatacct
86700

cgttgtctac cggacagcag agtaaaacat ggttcttact gtgagtcctt gtcaaatagg
86760

ctgaagccag agcctcttg gagaaacaga gataaccag gctgcctcc catccgtgga
86820

ctcagatgct gagggatggc tctggcaatc taththtgca gctccctaag tgactgatag
86880

gtgcgcatgc cgcagtgtgt ggcaacttcc agtgcaggaa acacaggcag tggaaccga
86940

agacctgaat ctcagtccca agaccccccac ttacctgccc cgcacatca gacaagtttc
87000

ctaggccctc ggagcttctg ctgctcactt gtgcctgtgt ccaccataa tggatcaaa
87060

gaggatttga gctggacacg ttaaatgcag gatgcgtgca gccaacagtg gcatgctggc
87120

ttttgaggta agctgcaacc ttctccttha aagttagcgt ggctagttag aagttacctg
87180

agcccatca cctccctca cacacatgtc cacagcagac accaggctat ggggacagcc
87240

atggttcttg aggggccttg gctgaccctt agatctgtaa tgcttttggt accggaaaag
87300

ggtcccgatc cagaccccaa gagagggtta ttggacctca cgcgagaaag aattcaggac
87360

ctgtccatag agtaaagtga aagcaagttt attaagaaag caaaggaata aaagaatggc
87420

tactccatag tcagagcagc ggcattggctg ctcagctgct tatacttatt gttactttct
87480

gattatatgc tgaacgaggg gtggattatt catgagtttt ccaggaaagg gatggcaatt
87540

cccagaacta agggttcctt accttttttag accatatagg gtaacttcct gacattgcca
87600

tggcatttgt aaactgacat ggcgctagtg ggagtgtttt ttaacatgcg aatgcattat
87660

aattagtgtg taatgagcaa cgaggtcact ctcgttgcca tgttggtttt ggtgcttttg
87720

gctggcttct ttactgcaag ctgttttatc agcaaggat ttatgacctg tatcctgagc
87780

cgacctcta tgtcatcccg tgactagcaa cgtttaacct cctgggaatg cagcccagta
87840

ggtctcagcc tcattttacc cagccccca ttcaggtgaa gtctctctgg tttgaacacc
87900

tcagacactt cgttttagcta aaaaccattg aaacagtgtg ttgattccac aaagaggtga
87960

cttgtgcttg catttacct tttgttccca aatctgtttt caaagtttta catgttgttt
88020

tgtgttcctt tgggtatttt tttttttttt tttttttttt ttttgagacg gagtctcact
88080

ccgtcaccca ggctggagtg cagtggcaca atctcggctc actgcaacct ctgcctcccg
88140

ggttcatgcc attctcctgc ctcagcctcc caagtagctg ggactacagg cgcccggcac
88200

cacacctggc taattttttt tttttttttt ttttgtattt ttagtagaga tggggtttca
88260

ccgtgttagg caggatggtc tcaatctcct gacctcatga tccgcctgcc tcggccttcc
88320

aaagtgcctg gattataggc gtgagccacc gcaccagcc ttcctttggt atttttttgt
88380

tattgttttg gtttggttg ggttggtgcta agtgtttcaa ttcagccagt ttttaactggc
88440

tgttagatat tagactgcaa aatactccag tcagagcagc tcacctgctc agaaaaatga
88500

gggttcagct tttagaacct aagttctttc ccagagttag aggaactgcc tgtgaccttg
88560

cccactccag ccacctgcct gaaataatga caagaccagc ctgaccccag cttcatgagc
88620

aattgtccaa agcccctttg tgccgtagta attgagcacc agcaggagct acaacgacct
88680

ggctgtttct tttttgctaa caagcttacg tataataaaa atggctctag cttggggatc
88740

ttattaatcg cctcatgaca atccctctcc ttttgggctt ctttgggttc tgggattttt
88800

aaattgttcc caaggagagt ttttgggaaa caaactccct ggaatcatgc aggcttccca
88860

ctgggacagg aaagtcaaag gccctttgtt attggctggc ctttggctgt cagcccttgc
88920

tatgaaccac tggcctgagg ctggaacttc ctgctgtctg caccatggga ttttgaatgg
88980

cacaagaatg aggctatatg aacaggagga gctgactgcc actggtgggc ctgggagcag
89040

ctctccaca ctgtccctgt gcagagccca gagtgccag caaaattcat cactctccag
89100

atcaggggct ggggacagtg gccacacag accaaacact ggagctacag ggcattgcac
89160

ccgagggaga tgcaaagacc tctactgggc tactaaggct ggctttccag gcctggatgg
89220

gtccccattg ctgcacctcc tgggaaagcc cacgacacag aggtggggcg caaaaggctc
89280

tcagtggggc tttgtctaag gaacagatag gtccaggctt tcctcaggca agtgacctga
89340

gcagaggggc ccgtgtggac ttttcaccct gaagctagaa ggggtgtggtg acgacagtct
89400

cctaaatgag tggctttagt cctagctgcc catcagcctc atccgccgaa ccttaaaaaa
89460

caacagatgc ccgccccaa accaacttaa tctgaatttc tggaagtggg gtttgggcac
89520

tagtatttta cccaagctcc ccgggtgatt ctaatgtgta gcaaggtttg aaaaccacag
89580

tctaaaagcc acagactgtt aaactcttgg ggataaaggg catcaccacg ccatcttggc
89640

aacaccaata agtcacaaat gaggaggggt gggagggccc acagttagtt agggctgggc
89700

acctggctgg ggaacatcag ggtgattgac ttcaagactt actctttttt tttttttctt
89760

ttttttttga gacggaattt cactcttggt gccaggtg gagtgcaatg gcgtgatttt
 89820

ggctcactgc aacatccacc tcttgggttc aagcgagtag ctgggattac aggcattgcgc
 89880

caccacacct ggctaatttt gtatttttag tagagacggg gacggggttt atccatgtcg
 89940

gtcaggctgg tcttgaactc ccaacctcag gtgatctgcc cgcctcggcc tcccaaagt
 90000

ctgggattac aggcgtgagc caccgcacc agccaagact tattcttgat ggtcagtgtc
 90060

ggggactttt ttaaccacag tcttcacac taagagatta gaatttagaa aagttgggt
 90120

ggctcagaag ttgcaggga tttggaattc ttgacagaca cttctggaag tggcaaggat
 90180

ttgcttatgt gcaactgaca gtcaaggagt cttctcttg ggaggtgacc acattaaaat
 90240

gataggaatt cactgtgggt gacagcataa agccaatcgt ggcttcctca ctcttcgtgg
 90300

tgagccagaa gttgttaaag actgggcaa gaagattcca aaaatggaca caaagccaag
 90360

aagaaccaga gaggtgagg aaggatagag aaggtttat ttggttatt tacttgggcc
 90420

caagaatgtt gaacgtgaga acctgtctgt gaaattgtag tgtgcagcca ttggaaaatg
 90480

ttattatta tcttagcaa aagctttaca agcaattctc ccacaccatg ttctttcact
 90540

cactgacgtt cactgtctg gtgtctgtcc tggagcagga aatgaaggga cacacacaca
 90600

cacacaccag ggaggagcaa gaacggggcc caatgtggaa tcagaccac ttagaaacac
 90660

ttctgagggg agtgcaaatg agaagcgtct gtcagccgag gctctggatt gatgccaaca
 90720

agctgggcta gaagggaatc caggaggctc agtccatgca gaccaccag aggtctcagt
 90780

gtgccgccac acttaaatta caccctcgca atgtctcctt ggagaaaatg catatgtgtt
 90840

cattgcttga tgggtcatcg ctatttccat cctgtgttaa tagacgtggg aggtctctcc
 90900

caaagctctc agtggctcac cagtgttagc agaattcagt cagagcacc ctgtgccacg
 90960

actccctcgc tctcgggtgcc ctctgctgtg tcaactgcct gccctgtcta ggggactaga
91020

tggaggctac tgcaactact cagcgcttcc cctctgact ttagctttgg aaattactta
91080

tctccactg aagttgatta agtcgcctgc ccagggttca tatcagttga accagcacca
91140

acagtagaac ctgcatctct tccttcgtgg tccacgagaa ctcgtaacatg ccaagaggct
91200

ggagtcatag agcagcactt cccagccatt ttcacattgc atctcttaca gaaagtcag
91260

ctttgggtgg tggatcagcc actctcggt ggaagtgacc cacaaaagcc cacacattgt
91320

tcaaggggtg tgaagctcca ttgtaaaatg gccgggacca tttccattta gaggttggat
91380

ggtctcttct cccaacggtc actaatctcc tccagatgga agagtaggcc cagcgggggt
91440

tccacccggg ttgtttataa cccagtgtgg ggtgggaggt gaaggcttag catgcacaga
91500

gggcatctgt gtggtgttct ggggcttct atacggtgcc aggtcaccca gaccacaggc
91560

cccgtagtc cccttcgcta ctggtgccaa atccacgtgc catctgtaa ttatttactc
91620

tgcatttata ttcaagctaa tttaaaatca atccactttt tgtctgatct taacttaggc
91680

caaatgttat ctgtgaaatt atgattttgg tatgcaagca gacacttggt actctaatat
91740

acattaaaa acacacataa ctagtataat ttatacacac acacacacac acacatatac
91800

atgtatatgt atatatatat ttagagacag ggtcttgctc tgttaccag gctggagtgc
91860

agtagtgtag tggcacaatc ttgggtcact gcaaccttga cctcttgggc tcaagtgatc
91920

ctcccatctc agcctcccaa gtagctggga caacaagcac atgctgccac agcctggcta
91980

atTTTTTgt atTTTTTgta gaaacagggt ctactatgt tgccagggt ggtcttgaac
92040

ccctgggctc aagcagtcca cccacttagg cctcccaaag tggtgggatt acaggcgaga
92100

gccaccgcag caggccttaa atgtttttta aacttctttg aacttaatto catctgaaat
92160

catccagcag cctccagaga taacaaataa caaagttaag gaaaaagcat ggtcatatag
92220

ttattgtatg gaatccgtgt gaattggtgg gtccttagac accttcagca tagaagcatt
92280

attcccattt tcctggttggt gaagctgagg cacagagagt tgagagttat cataagatca
92340

cactaagaga gtgacaggac tgggggttcca acccatagcg gtgagtacta aagccaccct
92400

ccttcccctc cgccaggcca gcagatgagt cgttggcagg atcactgcat ctgtcatccc
92460

ctcgccctca ctccagcctt tttttccctc cacagtcctc actgggcttt gtggccctcg
92520

tgctgagcac actgcacacg ctcacctacg gctggaccgg cgccttcgag gagagccgct
92580

acaagttcta cctgcctccc accttcacgc tcacgtgctt ggtgccctgc gtcgtcatcc
92640

tggccaaagc cctgtttctc ctgccctgca tcagccgcag actcgccagg atccggagag
92700

gctgggagag ggagagcacc atcaagttca cgctgcccac agaccacgcc ctggccgaga
92760

agacgagcca cgtatgaggt gcctgccctg ggctctggac cccgggcaca cgagggacgg
92820

tgccctgagc ccgttagggt ttcttttctt ggtggtgcaa agtgggtataa ctgtgtgcaa
92880

ataggagggt tgagggtccaa attcctggga ctcaaagtga tgcagtacta ttcagaatga
92940

tatacacaca tatgtgtata tgtatttaca tatattccac atatataaca ggatttgcaa
93000

ttatacatag ctagctaaaa agttgggtct ctgagatttc aacttgtaga tttaaaaaca
93060

agtgccgtac gttaagagaa gagcagatca tgctattgtg acatttgtag agatatacac
93120

acactttttg tacagaagag gcttgtgctg tgggtgggttc gatttatccc tgcccacccc
93180

acccccacaa cttccctttt gctacttccc caaggctctt gcagagctag ggctctgaag
93240

gggagggaag gcaacggctc tgcccagagc catccctgga gcatgtgagc agcggctggt
93300

ctcttccctc cacctggggc agcagcagga ggcctgggga ggaggaaaat caggcagtcg
93360

gcctggagtc tgtgcctggt cctttgcccg gtggtgggag gatggaggga ttgggctgaa
93420

gctgctccac ctcatccttg ctgagtgggg gagacatttt ccctgaaagt cagaagtcac
93480

catagagcct gcaaatggat cctcctgtga gagtgaagtc acctcctttc cagagccatt
93540

agtgaagcctg gcttgggaac aagtgttaatt tccttccttc ctttaacctg gcgatgagcg
93600

tcctttaaac cactgtgcct tctcacctt tccatcttca gtttgaatga ctcccaggaa
93660

ggcctagagc agacccttta gaaatcagcc caagggggag agcaagagaa aacactctag
93720

ggagtaaagc tccccgggcg tcagagtga gccctgcctg ggctgaagga ctgtcttcac
93780

gaagtcagtc ctgaggaaaa atattgggga ctccaaatgt cctctggcag aggaccaga
93840

aaaccacact ggctccaact tcctcctcat ggggcattac acttcaaac agtggggagc
93900

aacttttcca ccaaagctac aaacctaaaa tgctgctgcc ccaaagcaca agagggaaga
93960

gcaccgccgg ggccacagga cgtctgtcct ccagtcacag gccatccttg ctgctccta
94020

ctgactctag cttacttccc ctgtgaagaa acaggtgttc tcggctgagc ccccaacctt
94080

<210> 5
<211> 82154
<212> DNA
<213> Homo sapiens

<400> 5
cattaatagt gataactatt aataatatct gagtggttaac tatgtatcaa aactgtgct
60

atgggctttc tttaactatc tcacttatct ctcaataa ccctgtgagg caggactat
120

tacagaaggg gaaacggagg ctaataaaga ttaaagagt ttcttagggt tgccaccaa
180

gtggaagatc ctgaatatga atcttacagt cttctcagc ctgcattctt taccacaagg
240

ctgggactcc cagaggacag catcctctca tgccctgggtg gcaagcagcc agccacagac
300

atgctgaaga actaggggta gctacagtta gaaaacattc attgcttttt caaggaagag
360

ctaaggagac ccaggacgca gatgtgagtt agcagagaac accagcaaatt tacttaactt
 420
 gacagaatta ttttctctcc tgggtcagtt agcttctact tgacctggt gcttggtgga
 480
 gtttgtgtca ccccatggc tctgcttct gctctgaagc acaaggttac cattagcata
 540
 catcaggctg tggtaattt ggctgcccc tctttaagct aaggagacat ggtcagggac
 600
 tcagaccctg tgtgtggcca ttcagctttg cctgttctg tggccacagt ccatgctcta
 660
 agtactgacc agatgtctta caatgcacac ggcacacagt tatggtctct gtggggatgg
 720
 gggaataagg tgtggacaaa tctgttcagc tctaaatcct tgaccagagg agcagctcag
 780
 agcatgaatc cacaggcgag tggagccatg tgcaggatca agcacagcg catcactgac
 840
 ttgcaggaca ctggagtgtt gggcggtgtt tctagagaaa tgccctcaag tcctcgatag
 900
 atcctaaatg tggtagttc tcttcattac aaagacagt gaagctgagg tccatgcaga
 960
 ttaaaaggct aagctgtaca gtaattatga ggtccttttg attagcagaa ctccatctat
 1020
 tccaaaacaa attcatagaa agacatctt tctatgccta attgccttct catctttgtt
 1080
 gaaaatcaac agaccataaa cgtggagtct gttttgggac ttgacattct gttctgctaa
 1140
 tctatgtgcc tctctttgcg ccaatacttt actctcttga ttactgtagc tttttagtaa
 1200
 atcttgacat cagatatttt aagtcctcca actttatttt tgcttctaag ttgttttgtg
 1260
 tttacataca aaatttagaa tcaatttgtt caatttctac caaaaaagtc tgcttgggtt
 1320
 ttggttggga ctgcattgaa tccatagatc aactattggg gattaatttt tttctctacg
 1380
 ttttgactgg tttgtatacc acagattaga ggattctggg atggagcttc agaattcata
 1440
 gcacaaacat gaagatatc agaacaatta acttttatgc tgatttaaag ctaataagaa
 1500
 taaatcttta catcaaatat aagccacctg gagccagtat tattgggact tcattgataa
 1560

taagattatt tacccaaagg ggaatgattg tcaaagtacc tactgtgagc catatttata
 1620

aaatgtaagc catatttgaa aaataacttaa ttttcacaag aagctgatga ggtaggcatt
 1680

acctgctgta ctttacaat ggggaaacta aagtacaaag agatgtagca gttaagctga
 1740

ggtcacacag gtaattaaat ggtagcattg tcacttaaac ccaggcttgt ctaattccat
 1800

tctgtggtac ttgctgtacc tcaatggttt tcagactgtg tccagagtgc cccaggacct
 1860

gagagagaga gagagaaaga aagagaaaaa agagagagag aaagagaaga gagtgagaga
 1920

gagtgagtgt gtttgagga aggaggaaga aagtggctctg gggttcagaa tcccagactt
 1980

ggtctcagt agttttcctt tgtcctatct tgcagatggg gcctccacgt gaaaagtaac
 2040

tacaaaacag aaccagaaaa tagtcatgct gttcaagtga gttggcatat cacttcatgc
 2100

tactttgaaa taaaatttcc tcttattttc ttatggataa aggtctcgaa aggtttgaga
 2160

tgatcacagc ctgtgttaag gctgaccagg gtccaggcca gtacagtga gctgcttttc
 2220

ttctggattc tgaaccaccc tcagccttca gggcctgttg tctggaggga tcatggagtc
 2280

accagtgaac gtgaaagcca ctgctcctct ggtgtccatg tctattctga caggaggatc
 2340

tcagccccag gttgactgat tgctcgcagg tgaggctact gcctccgcag tgctcaccaa
 2400

atgctccggc cgctccagtc tgcacgttcc tgagcttggg ccacctcctg cttttcacag
 2460

gcagcttctc ctgatctcc tctctctttt tttcttcatg agcctcacc cttggcagcca
 2520

ccccaagaaa ttcagccttc ttcatctcct ttctctctca gaaactggaa agatttttac
 2580

tgaattctg atatttttgt ctttgccaaa acccaaaggc cccaatctgc catggggccaa
 2640

gaagtgccca gagagggtgt ttggacaaat ccagaaatga ccttgaatct ttactcagca
 2700

cttagatgac aggttctctc cttctccacc acaccctagc acacgtgctt cttgtcggat
 2760

caggaaatgt tttattgctt tttttccttt tcactttctc taccacacct tctttgtgag
2820

cccttcccgg ttatgctgtc cgtattgcct ctcttctgga aaagtaatca gagagtgaaa
2880

atcccaggac aggaagggaac cttgggaagg cattaaccag accttttcct ccaggcagca
2940

gccacagtcg ccggggtaaa gaacctctt ttccatcttt aaactttcat cggtttctgc
3000

atcttctcat tgaagcatct ctgggttttg ctttttgttg tcttatttgt tactttcatt
3060

tttaaaagtt ttctttcatt cagaggtttc ttgagactgc ctggtgcaca tcatgcgtat
3120

ccgtgatag atgctgttgt gacttggtga aacatgttgc gtgctcttca ggcagatgta
3180

tgtcaacaaa cccttttcct ggctcagcca ggatttctgg gagtcagatt tgattatggg
3240

gtgtcatagg taaaaactgg aggctgtctg tggtggtcca gctgtgcagt tagctgactt
3300

ggaggaatct ttggctataa ttggatgcca agagtggagg tggcacgaga gatggtacaa
3360

gcctatgatg atgtctaccc aggtgaaaaa tgatctaggt gcttgaacc atcacctga
3420

gcattcagtt ctacatccag ctattgtgga aaaaaatgat ttatattgat taattgacta
3480

gttggttcat tttagcttga agagtactga tcagctgtgt gcattaggca gtttcttgac
3540

acactgctgt gctgaggaca gatttcccag agcctggatg tgtataattc tacatgctcc
3600

ttttccaag gacttaagca aacagagtca actcctctgt tttccttggt tgggtatatt
3660

tttgaaaatg tctcttcttc tttcctttca tttacatttt accgttggtt atccaccttt
3720

tcctcttcct tcaaagctct agacctgcct ccctgcctgt tcttagcca cttttgagaa
3780

aggactttcc aaccccatcc tcgaaaggc aggttctctc tccagggtc caacagagtt
3840

ctgaaagaac acggcttcca ccattgtgtc ttctctgttt tgtcaactgc tgaaagactg
3900

acacagtaga caatgctgca gaccttgg gagatagtcc aagtttaggg tggggagggg
3960

atggtgcatt ctgggtgtcc acatatttgt tgttgcttta ggaagaggag ctgggaggtt
4020

ggagtcgcca cacatgttga aagatgtggc aaaaaacaaa caaacaaaac acctcataga
4080

gagaaatcct gactttgtgc acaagaaaag ggagtgaagt gtgtgggggtt tttaaagtgg
4140

tggtgggggg gcattaggag ataccagaa gctctcaaga ttccaaacag tctgcttttc
4200

ttttcagaca ttacactcca agaaaagaaa ctctagtgtg taacagaggg gctctgtcaa
4260

aatttaacaa cataaatatt gggcaaatga agttgtttct ctggtgctat ctcaacctgc
4320

caatggagaa atatctcaaa aagtcaagaa ttgaagccat tactcagaga cagcatccta
4380

tggtcatggtg atccaaaaca agtacacctc taccttttcc ctttcagccc caaaccccaa
4440

aacactggag aatattgggg caggcaaggc ccgtgaatag gtgagcagag ctggaagaag
4500

agactcctgc ccctgggtgg cctcaggcta tttgtggaat gatgtttgtg agatgaacat
4560

gttcgctccc tcacccggat gttttggaca taggagcctg gagtaaagag tgtatatctt
4620

gggctgtcat tttcccgtgt gaaagggtt ttgttttctc cccgtttgca caacttgttt
4680

cccttaaact gtgctgtgca tggctaggaa gctcagcaat agactaaaca gtttgcgatc
4740

ccacagtttt gctttctgtt tctcaggcct gacccaacac tccatgggtt ccctagccct
4800

tcgtaggtgc ttcgaaacac tcatggggtc cttgacctga tcccagtggt gggttgtgtg
4860

tggaagttag ggaacctctc tttgagagtt ctggtgcccc taactctgta ggtctgact
4920

tcacttctgc cccaataaga atgaaaggct ttgccagtag tgcagccctc tcctcccaa
4980

gtcctagagg caattactct gaaatatgga agaaagaggc acatctgaaa gcccagtgct
5040

gtcttcttgt cctcacatcc ctcaggatct caaggtcagc ctagtagagg cattgcttta
5100

ccatagtagc aactcagga atatctgtgg aactggattc ctctgaagtt ggacagttgc
5160

tttatccctg ctagaggctc tcactactta gaccacaca gtcctataga tagaagaaac
5220

ctagagagaa ccacctctca cacttttatt ttatacatgg agaaccaaag gcctcaaattg
5280

gtggagtgat ttgctggggt gctacagtca ttgtcacaac ttacattgtg acatgggctt
5340

catatccagt taaaacaaaa gcctctattc actgccccaa ggtgtgaagc cttgtctcca
5400

ggagatggag agtcaatgac tccaaggaga taagacttca gactcttctt gctgcaattt
5460

ttttttttt ttaagtaagt gccaccaat aactcctggg gaaagtaaaa taatctgggg
5520

gaggaaaatt gttctctctt ggagatggga ctctctgat ttggaactcc tctgtctgct
5580

ttcacttctt ggataggatt ttaaattccag caccagcct gctgctgtga agttcacaat
5640

aactttatca cctgagttcc cttgggtgctg cagcagatat gaggtaacaca cctggctctc
5700

cgagagtggg ttaatccaca ccagagtcct gacttactag tttcttccta ttgcagagcc
5760

tatgcagact gcttctctga gcaccaacat cattatgcac tgcctttctt attccaaga
5820

cttcctagtt gctgcaagtc aagttcgaca ttcaactatt tcctcggggt ccttttttgg
5880

gagaagcaag gctttgagtt cactctttga cttcaagtga agtccatttt cctcttagac
5940

actaccccag tggttctcac actcaaggaa gcaccaaagt ccctggagga cccgtctaaa
6000

tgaaggatgc taagccccac ctcacagaga ttgatgcagt acgtcccagg aatctgcatg
6060

ttaaaggagc tctgcaggtg attctgatgg aggtggccct tgcaggatgc tttgagaaac
6120

tatatgttct atccctcgaa ctctgcctgc agaacagcct gagatggctc ttccatgctt
6180

tgaccatttg gccctccaca gtagtcctg agctgcgtct tcttcggacc atgagcagaa
6240

tgtgtgtccc ctgactgact gtccttttcc ttgagaacca ggcatagtgg tgcactggta
6300

agagcatgga ctctggagcc aactgggtt tgaattatgt ctctgccact tactagctga
6360

gtgaccctag gcaactttct aaatctctgt gcttttgttt cttcatctgt aaaatgggga
 6420
 taataatact acttgcccca taggggtgtt atgaagatga accaagttaa tatctctaata
 6480
 aaatgtctct aaaatcccat gattgcaaaa agtagaaaat atgtcatgta gcctattata
 6540
 tcattatatt ataagggtgag gctattgaag gcaagagggg aaatgacttg ctgaaggcaa
 6600
 cagagctggt ctgcagcatg gccaaagacca aaaacagctc ttttcgctcc cagataaatt
 6660
 gtctttttta tttttttccc aaactgcctg tctaaccccc ctttcatctg gttataatga
 6720
 atgccattgc ttgtgtatag ttgcctacat taatagacaa caatgcaata ttctccttat
 6780
 ggaaacccta cctagattgg gaggcaccat gaagccactt tttcttcttt ctccatctct
 6840
 accctgatta gcaaacatct ctaattattg tgactoccaa cccagtcag atcattttat
 6900
 tctttttcca gtatctatgc aggggaagcct ggcaaaggct taagggatgg ggtgggggtg
 6960
 agcatgtcag caggatttgg aactatgggc ccatacttt tggcacctgc tgcattttct
 7020
 cccacctgc tgcaccataa ttcatctgct tgtatgagtt tagcagaaag agccccagt
 7080
 tccctgtaag aaacacatac ggcttttcca gggtacactg tggggattga gaaaatgcac
 7140
 agtatccatg tctggaaatc aatggacaac tcaactgctg aatcgtttgt gtcttgaac
 7200
 ttgcctcaag caacatctgt ctgaaaaatc tgggtgcttg ggctaaatga ctggatattc
 7260
 tcctaaggaa tggcagttag tgggttaatt tcacctctca tcttacttcc tgctggccc
 7320
 ctctaccttg cccaccatgc ctggctttgc tgctgatctt cggctgacct gaaattgaac
 7380
 accttaactt aaggacaact aagggtgtag gtagattggg aagtatgtgt tcttcttgct
 7440
 ttgccaataa ttgtcttctg gattcattgg ctttcttctt gggttccttc ttcaacttct
 7500
 ttaacaaatg acttcagcca tgccaagcct tcttttaagc actttccctc tcaactctct
 7560

gtcctataaa gatatctttc ttgaagcccc tccgtgccag gcgtccaaag tgaggcacat
7620

caactcggtc cttgaacgtc tcatgtttat cccaatagat ggtctagatg tggatgaatg
7680

taaaggaaat gtgcaaataa gatactttta aaattcagac agaattatac aggggtgggaa
7740

atttcagtat acatagaagg cttaattcaa aatgatataa tgtcaacatt tggataaga
7800

aattatttat ctggcagaat atatactcac agatgtcaaa atgggtctag agagtagaga
7860

atggaggcaa ttatccagac actgaggccc aaggcacaat acttttaagt aatcatgccc
7920

tcaactggca tttccatagt cttaagccca aaagcattgg gtggaacct tctataagca
7980

aagcctcagc tatgaaaaca acttttaaaa ttgacattca aaagataaga tcaacaacag
8040

gtgccctttg ctttcggcag taggtgtgtt cttgaaaagt tggttgtcaa tatttccttt
8100

tcccattaat ctaatcacat tctaaatata acagggggct ggatgtttga agaaatcctg
8160

tagttgatct cattgcaaag tccttttaga gaggtaaaaa tagtttgttt tattttcaaa
8220

aattaggttt tcttagacta aaacttcttg gaagtggggg catatcatgt tcctgtatat
8280

tgcttacagg taaagaatgt aagaatgtat ggcttagtca cagagagcta gggcttcaga
8340

aatgctatat gtatcaaata ttttatatct atatttactc tgaagatatg acaatagctt
8400

ttaatgtggg tcgagagata tctaactctt cgcttctctt tcaaccttcc aaaacctgc
8460

aatcctcatt ccttcttttc aataacctcc acatgcccc ttcatctgtc ccacccaaag
8520

gaaactcttt taacttacca ttttaattgat caactctgat gacctttccc taagctctct
8580

gtgctaaaat aatctggtac tttgggctga agaagtattc tcctaacgat ggaggtaggt
8640

ggcagtcaaa tgacgcagag attatgatag ggccctgtct aggctgtcac ataaggtaac
8700

attagccaca tgtggctact gtgtacttga aatgtgatga atgccattga gaactgaatt
8760

tttaatttaa tttaaaatg gagtcgcatt tctgttttag gagaacttct aagtatgtct
8820

gaaacaactt gatcatgttt caagtgttaag ttttattaga tctaaatcta aatcaagtat
8880

ttcgaaaaag tattttagca tttgaataga gatgtgctct aagtataaaa cacacagtgg
8940

attttgaaga ttagtaagaa aaaaatgaat gtaaactatt tcattaataa atttttatat
9000

tggttacatg ctgaaatgat aatattttgc atatattggg tgaaattata tatattattt
9060

aaattaactt tccttttttt ttttaaactt tgtgtggcta caaggacact gaaattagaa
9120

acatggctcg aattatgctt gcactgaaca aggctagtct aaggtagtca attgccctgc
9180

ttcttcacta ctttcaagtg taaacaaata tgtcagcatt atccctagaa ggagggttta
9240

ggggttttta gttgaatact cttcagctga agaaaatcca ttaatctcca gttttattct
9300

tgagtttagct ttcaaacta cattgcgaca ataattactt tgggatcttt gaattttctt
9360

tcatgccttt ctgttctatg aagccacat gcgtagttta tgcattttcc aatttaggct
9420

agaagtccaa tgtgctgtcc attgcaccac agagcctggt tgctatgcat tttaagagt
9480

aaaattttca gaaaatactt aaattaaaa gatagatttg gaaagcatat ccaaaatggt
9540

gctttgaact tcttattctg gacaattcct ggactgggga gaaggcatgc ccaagtacac
9600

attgcttgca gtggcacctc caaaccaatg ctggtcggga ctcttctat tacctcgcaa
9660

cagcaatgac tgggccccat aatacagcca cggtcattat aaaagatcat tacaagcat
9720

caaattaact gctattaatg aaagccagaa acacgtcatt caggaggctg aatggctggg
9780

atggaacata cagggttgaa taatggtagg agcctggggt tggaattaga aaactgggac
9840

aggaatccta cttctgctat gaatccattg caagtgactg ggcaagttgt tcctcactcc
9900

ctactgcctt gctcatctat cagctgagaa tgtcaaggca gaaggtctcc agggctcctg
9960

caggctccta caagccagat ctcaatccgt cttatcctac cctcttcctg acctctagct
 10020
 acaagaacag caaggcacag ccccttggtt ttgagaatgc taatgcttgc cagcattctt
 10080
 gtcctgcagg gcatgaagta tagtcctgca caaccagcag gctctgtgac tgcctaacct
 10140
 catccttggt cactgacctc tgtcttagta gcagttcgtg ccatacctacc tagctttgct
 10200
 gggtaagtg cggtcacatc ttaaaaccgt ttgaagcgtt tgtctaatta acatgcagaa
 10260
 agcatatact gatttttatt ttcaccacca tgtatgtgag ggctattttt atggcatgca
 10320
 aagcatctca tttttatttc aattctctgt cttgtaactt taaatgaggc tgaaatttcc
 10380
 ctccatatgt ttattagcat cttctatttt atctcttata cactgactgt ttgcactttt
 10440
 tgccccattt ccctgaatgg tttttttctc aatgatttgt aagcaactct tatatttaaa
 10500
 tatgttaact ctttctgtc acatatatta aaccctttct taaaaacggt gttttccggt
 10560
 ttgtgactta ctttttaatt ttgctcataa attactttta acttacagat agtaatat
 10620
 ttgaaagggt acattcatca gtctttgtat tgttaaaaag agtctttatt gttatacttt
 10680
 taaagtagct tttccactcc tagtcagata ctttcacata taatagttgt ctactttcct
 10740
 cctgaattaa tggggattta ttttggcata tgggtggaag taagaagcta aattcttttt
 10800
 cttttctctg catcatcaat gatctcttac actgttaaata aaactagtct tttcctctg
 10860
 cttccacac ccatttaacc tagccactg ttgacatcgt cagagaaaaac ttttcatcag
 10920
 aaagtctgaa cagcacaatt atactgtatc aatccagctt aaattacaca tgattgttag
 10980
 gatgggtaga aggtatgata aaatggggtg aaagtaaaat gaatgatcat cgatatacat
 11040
 gatgtcaaaa atttgggtat aactgaagaa ttttagtaat ttaaagattt cagaatttaa
 11100
 ataataaat atctgatata ttcaattgga aatttggttg agatcagctg acttgtaaat
 11160

tggacttcac tgataaacat ttgccttaaa atctgtagaa tctaaagcag ggtataccta
11220

gtatgcgagt aagacatcat ttctggaatc aagggggctt ttcagtgttg accagatggt
11280

tatcactgca tgggcaaatg gccaaggag tatggttgtg gcttggctta tttcttctta
11340

acgcagatgc agttgcttcc tttctgggtg gcagtgtgga ggataactgg ggtcgggtgg
11400

gacataaagt ctcacaaggg gaggaagcc cctctgtggg ctctctgggc cagccacctt
11460

ccttacaag tgggatgtca gcttctgtc tgaccagagt caaagtaggc atcagctatt
11520

gccattactg ttgttgata atatggtgaa taaaggaaat gtgttatata tatatttttt
11580

tcttattatt atactttaag ttttagggta catgtgcaca atgtgcaggt tagttacata
11640

tgtatacatg tgccatgctg gtgtgctgca cccactaact ggcatcttag cattaggtat
11700

atctcctaaa gctatccctc cccctctccc cccacccac aacagtcccc agagtgtgat
11760

gttcccttc ctgtgtccat gtgttctcat tgttcaattc ccacctatga gtgagaatat
11820

gtggtgtttg gttttttgtc cttgcgatag tttactgaga atgatgattt ccaatttcat
11880

ccatgtccct acaaaggaca tgaactcatc attttttatg gctgcatagc attccatggt
11940

gtatatgtgc cacattttct taatccagtc tatcattgtg ggacatttgg gttgggtcca
12000

agtctttgct attgtgaata gtgccgcaat aaacatacat gtgcatgtgt ctttatagca
12060

gcatgattta tagtcctttg ggtatatacc cagtaatggg atggctgggt caaatggtat
12120

ttctagtctc agatccctga ggaatggcca cactgacttc cacaatggtt gaactagttt
12180

acagtccac caacagtgtc aaagtgttcc tatttctcca catcctctcc agcacctgtt
12240

gtttcctgac tttttaatga ttgccattct aactgggtgt agatggatc tcattgtggt
12300

tttgatttgc attcctctga tggccagtga tagtgagcat ttttttcatg tgttttttgg
12360

ctgcataaat gtcttctttt gagaagtgtc tgttcatgtc cttcgccac tttttgatgg
 12420
 ggttgtttgt tttttcttg taaatttgtt tgagttcatg gtagattctg gatattagcc
 12480
 ctttgtcaga tgagtaggtt gcgaaaattt tctccattt tgtaggtttc ctgttctc
 12540
 tgatggtagt ttcttttctg gtgcagaagc tctttagtta atgtgttata ttttaagagc
 12600
 tcagttatct gttacctatt tgccagtgc aggaactcat gtcttctgag gcagttggc
 12660
 tcatttatgg tagatataat taattattat ctaattatta ttgcccttga attatttgaa
 12720
 gatcatca ggccttctct taaacacttc cgttcatgtg ctttagcttc tttatatggc
 12780
 cagccactct ctgtggatat gtccagggtt gttgatgtcc ccttgacatc agccaacct
 12840
 aacttaacac tggatgagag ctaagcagtgc ctaaacttat cacttcctc gaatgtgtca
 12900
 ttatactttt attagtgtg cttgggacta gatttccttt ctgacagcca agtcatgtt
 12960
 tgggctcata tgcagcttgc agctaagaaa gccttgagtt gtttctctat aaaccgtctg
 13020
 aaactatttt tagaattcaa gttacaactc tgaagtcctg ggtcttactt aatcaatgat
 13080
 accttctgag ataaatactg aaaatccttc aagtatagaa caaggacatc ctaactctgg
 13140
 gcatctgatg tctgttaaatt gtatcacatt ttatttattt gtatttttta aaggaagtga
 13200
 gaaggagaag tgagatgctg agcatcaagc tgggcagagt tgtcatgttg tgatctgaga
 13260
 ttctgtcctg ttcttcagga ttttaaggagt cgctaaccac tgccaagcat gggctcatcg
 13320
 ggtgcagctg gggaggcaat tacttcaact caggagtttt caaaaacttt ctttatgcct
 13380
 gctcacagga tttcatgggtg ttttcttct ttgttggtac tttgtggaa aaggaaagt
 13440
 accagtgaac aatctctgtg accttctct caatcccaaa aatatgcttg tgataaatca
 13500
 gcagtttatt ttaatgcctc atttgaggc attgcttttc caagaatgaa tctttgcata
 13560

ttatctgaga catcattttg gtgtaaggta gaattgtacc aggtatgtat ctttgtggcc
 13620

aactatactg tcaaaactga gtttcagtca tccaataaaa aggaagctgt ctctgaattc
 13680

agatattcct ccattctaag ggaactgaag tataacataa aataacttaac attctttctg
 13740

tatatgtgaa aaaaccatac ctaaggaatt ttattattca gcccatgttt ctttattgta
 13800

tccagtataa cataactgaa gatgtatata tacacatata tagatatata aagcatctat
 13860

acatctagac atgtatatat acatatatat atacacacac gtatatactg tacttcaatc
 13920

atctatatta aattgtgata atatcacatt accacatggt gatggcatto ttgtaatctg
 13980

cctgaaatcc atgaaaacac atgcagtggc ctggctcatg ttgtctcagg gagtgcattc
 14040

tggtctccag tgatcatcaa ttacttttat aaaagtccac caatcacttg ttggtttccg
 14100

gaaaacaaat ttaaagcaaa cattattaat gtcaaactta cacattttgc tttctggaat
 14160

caactttttt attttaaaaa ttgagaatat aatttctgtc ttttgtttct ttgggcattt
 14220

ttcttattct ccacaattat tcaaagatcc ttagccatag gtctgaaagt gtttgtttca
 14280

gctctcatcc cccttccttg caggtaaggc tggcattcca cagccctgga taactcctct
 14340

gtgctaggtc tatgcctgct cctggggacac tgaacactcc taggtttgca ggggccacgg
 14400

tgtccttcgt gtccctccct gccagcctc agcaggagcc ccctgctcct agaccactcc
 14460

ctactgagt cctgtagatg ttccagcaca cacagtgcct cccttctcaa ggccccagcc
 14520

ccagccctt ctcaggatcat ttgggggctt acatcaactg ctacctgtgt ggagatggag
 14580

gtctgtttgc cagcactgcc ttgcttttcc tcctgcccga ctgcatttaa tcaactgtgat
 14640

aagaggaaac aaagagggaa cgtcattggc ccatttctc tgaatcttga ccacgaaaat
 14700

atcttgaatt gcctccgtgt cccaccccaa catcgctct tgcactgcta cctctgctg
 14760

agtgcacatc agccaagagc caggaaccac tgcccaacca cgacccaggc acagaccctt
14820

atctctagcc tagagaagag gtgtgtttct ttctaatacat ctggtatgta tccaaggaaa
14880

ccataaattg atcattatct ttgtatgtgg aattagaaat attcatatag acatacatct
14940

tttggttttt atttttttat catccctaaa ataaatagca gatgttctca aaatggttta
15000

ttaatttcag tcctaagatt gtataattga ttgaaacat aaatggcaga cagaagaaca
15060

gcaacagtga aaaccaagat taattctaata aaattcatta caatggcagg aatcaagtgt
15120

actttaagct cactgtaaag ttttcagtgt gttgtggatt cagaggcagt aggtgatgac
15180

ttttattata tattcttaag tacctggttt tggagggtga tggcttgttg gagggaatga
15240

tgtctgaagga acataataaa acaatatgat aaatacaggg agagaagaaa gacaaaggaa
15300

aagaggaaac aaatgtcaag aataagatga ataggtaggg gagaagaaaa aataagttgt
15360

cataatttta tattccaagg actacgctag tccttggact acgctggtaa aaagaaatct
15420

aatcatgagg cagagaattt tatttatttc ccatttcctt ttggctcgtg gcttgagttt
15480

gctctggctt tcattgaata catgggcaa ggcaggatga gcttctgcct gaaccgctgg
15540

gatgttatto ctgatctcat ggaaagacat gtcgaggctt ctgaccctgt gcctctgtat
15600

cactgtgttc cctgtatttc tgtgcccttc gtattagttt tcacattgct aatgaagaca
15660

taccagagac tgggtaattt ataaaggaaa gaggtttaat ggactcacag ttccacatgg
15720

ctggtgaggc cttacaatca tggcagaagg cgaaggaaga gcaaaggac ttcttacatg
15780

gtggcaggca agaaagagca tgtacagggg aattcccttt tataaaacca tcagatctca
15840

tgagacttat tcattaccag gagaacagta tgggggaaac cgctcccat gattcagtta
15900

tctccacttg gccctaccct tgacacatgg ggattattac aaaatcaagg tgagatttgg
15960

gtggggacac agtgagacta tatcaccctt gtatccctgt gtctctatat tctgtatcc
16020

ctggctgaat gggtaaagt aatgatcagc ccacagatat gcatttgcct ctgctgagaa
16080

tagaaaggaa tataaatgaa gtctttttct tcatggagct tattaactag atgaagagag
16140

aaagcctgtg ggaaaagaca aatagaacct tcagggttcta tgcagtgagt gccaagtgg
16200

tgattgtgaa aatgaaagtg cagccctgca gatgagcaga agctcatggt gggaccttct
16260

atgcaagcta tgcagaggag agagcatagg atttgggggtt gggacacaca gattcaaact
16320

atggtcctgc cactttcctg ctacatgaac ttgacctgt tattttacct ccatgaagtc
16380

tgagttttca gatgtgtaaa ataataccac taatgcatac ttttttaaaa tttaaattta
16440

aatttttttt tgacactgag tctcactctg tcaccaggct ggagtacagt ggcacgatct
16500

tggtcactg cagcctccac ctctgggtt caagcaattc tctgcctca gcctcctgag
16560

tagctggaac tacagggtgtg taccacctg tccagccaat tttgtattt ttagtagaga
16620

cggggtttca ccatgttggc caggatggc tcgatctctt gacctcgtga tccgcctgac
16680

tcggcctccc aaagtgtgtg gattacagat gtgagccact gcgccagcc taatgtgtac
16740

ttttaaaagc acttttcaaa tacttcatct tgtttaatct tcttggcagt tctgtaaagt
16800

gtgaagcatt gtacaaagt aactctcatt attatattgg aaccagggt tttcagagga
16860

gaatcctcaa ggagataggg ctttcctgcc ttgtgcccat ataaggttgt ggatcccaag
16920

ggacgaaaagc attttcatca gactgggaga ggcaagaaaa ggtcactggc agtatagatt
16980

cttctggagt tcactggcaa cttgtccac ttcagttctc ccttaattct tccagatgtg
17040

cagatgtata gcctgggtgg atgagccaaa atgaattcaa ggctaaacag gtcagctggc
17100

tcagcccatg ggaagcatgc tcatcccaaa gcaaaggctc ctggcagaaa tttctctgca
17160

cttatttacc ctgctgtctt cacattgggg tgcggagtct aattcaaaaa ctccaaggca
17220

tgagtccaag ccaggagct gttagagggt cacatgggag atcttagcaa gagaaaaata
17280

tatgttttaa accctagaaa gaagaaagta aatgcctctc atttacttta aacaaataat
17340

atggccaagg tcacatagca ggaaagtggc aaggggtctg tcctcttttc actgttcaat
17400

catgaccata ttgtcctgga ggcctagggg caggatataa tgaagaaaca tccaattttg
17460

acagagtggg taaatgggat tgtgggactt gtagtgagct gatctgtctt cattcatgta
17520

attatcaata tcatttcaat gaccatgatg ttgatatgga aaatgatgag gatgatagta
17580

cagaaagaac atgaatatta gctgatgtct attgagtgtt tcctgggtgc cagacaccac
17640

cgtaaagtgc tgaacaagga ttatagaatt taattctcta gtgacctttt gagatagggt
17700

atatgattgt cagctttcca caaatgagga agccaatatt cagaaggata cgggtgctgaa
17760

gcttatgcag taagaaatag tagaagcaga attcaaactc aggcaattga acaccaaata
17820

tcaagctctt aattactatt tctagaacct tctgactgac tgattcctcc ctcaccacag
17880

ccaacttttg catccccat ttctggagag catctcccca catggtatat actcaaagg
17940

cccatagtga tcctcttccc aaagagcgtg aatgaggggt ttgaaactgc cagaactctg
18000

gtccagagct tggacttttg atgcgaatga tggagaaacc aagaggttga tttcctcagc
18060

cttcaagaga gctcatcctc agctctctaa agctgaccag aaagaggaga gtgggacct
18120

tgaaaacagc aggagatggg agggcttgaa ttcctaatagc tataaccata ggtcattgtg
18180

ccatggtcct ctgcaccccc aacaccccca aactgatcc tgaaggcttc aacaatcttc
18240

atctgectct ctcctttcac cttctttaca ctctcattga tagccttggg tttttctata
18300

cctccagtat tcccattccc atgatagcca tttaaggaca gggcacttcc tggtcaccat
18360

aaggaatatt gtgggaatgt tctgctcatg gatattgcat cagaataggt ctggaaatga
18420

aacacaccaa caaagtcaac agaggagagt ccaagcaaag cgtctgttta caaagttgtg
18480

gaaggtttag gggaactcat tatgagatag ttgaagcaac ccagagctag gaacagtggc
18540

gctcttacca caccaagggt tgaaggagct cgggtttctg gaaccagac atagcagcag
18600

cagttgtgat agaggagtgc ctggcagaag ccatagcttt cagttgagaa ggtggtcact
18660

gccaaccca ggctcagctg gaaggtaacc agagaataaa ttcccaaac tctttctctt
18720

cctttctgt gctctcttat tgggtgcctcc cattgagaga acccaacaa agccagctgt
18780

aagggaaact aatgaggcat ttccatatag ctcagcctgc tgggaaacat gtactataga
18840

aaagggtgga gagcggatct agagaggcaa aaaaattaaa caacggcacc gatgtttaca
18900

tttctttctg atgtagagac ctcataccac tttctgtac atcttacttt aatgtgtaaa
18960

aaggcctata ggctctttct ctaggttcag taaaagatga actaggaaac atatgccaca
19020

tttgagcttc ctcacagctt ggccctgca aagctgtccc cttctaagaa ctaagtgcc
19080

tgggtattca gtgctgtgtt gtctggtgac ttcgcagccc tgtcaacatt gttttctgt
19140

ttgtttgttt gtttaattat actttaagtt ctagggcaca tgtgcacaac gtgcaggttt
19200

gttacatatg tatacatgtg ccatgttggt gtgcgcacc cgtaactcg tcacttacat
19260

taggtatgtc tcctaattgg atccctcccc catccgcta cccacgaca ggccccagt
19320

tgtgatgttc ccaccctgt gtccaagtgt tctcattgtt caattccac ctatgagtga
19380

gaacatgagg tgtttggtt tctgtcttg cgatagtttg ctcagaatga tggtttccag
19440

cttcatccat gtcactacaa aggacatgaa ctcattcatt ggctcctcct cagaaagggt
19500

gaaaagcaga acctcttaaa ctgctccagg tttctggag agccttgctt taaaaaggc
19560

cctttttttg ggggtggggg gacaccaatg ccgtgactcc ttcaagaatg tgattctagc
 19620

aagagactga gctctattct tttctttcct ccaccctctc ttaaggcaaa ttgagtcaaa
 19680

ccatctcagt ctaccacac catccttctt taattttctgt agttaagagc ataaatccct
 19740

tccctattct gagtgtggtt catcgacctt cagcatcagc atcacctggg agtctgtcag
 19800

aatgcaaaa tcttaggctc cagcccacat ccctgagcca aatctgcctt ttgcaagatc
 19860

cccaggtgat atacatgtac attaatTTTA ggaggattga tgtagcctgt atgccatggt
 19920

aagaggcatg caatgtaatg ttgagtttct ttcttttagaa tttatagtag cagccgggac
 19980

cggtggctca cgctgtaat ccagcattc taggaggctg aggctgggtgc atcacttgag
 20040

gtcaggagtt ggagatcagc ctggcctggc caacatggca aaaccctgcc tttacaaaaa
 20100

aatacaaaaa ttagccaggt gtggtggtat gccccgtag tcccagctac tcaggagggt
 20160

gagatgggag agtagcttga acctgggagg cggagggtgc agtgagccaa gatcacacca
 20220

ctgcactcca gcctgggtga cagagtgaca tagtgagtcc ctgtctaaaa aaaaaaaaaa
 20280

aaaaaaaaaa aaaaaagaat ttatagtagc caaaaagcc tttttgagta attgaagaaa
 20340

ttctgaagtt attgtgtgct ttgaaattta gtatttattg aatagtttat atctttcaaa
 20400

aagcttttat aactcattgg agtctcgtca taactccatg aaatacgtaa ggtagataat
 20460

aggagccaca ttttcacat gaggatctga ggcatgatta agtaaaatca gtcctctgct
 20520

aggtggaatg gttttgagac cagatactag ggctcagaac tcatttttaa gctgttgtgg
 20580

cagctatcaa ttctgaagaa agtcaacatg gcatagcatg cctcttacca tggcatacag
 20640

gctgtatcaa tcctoctaaa actaatgtac atgcacacca cctgagggtc ttgcaaaagg
 20700

cagatttggc tcagggatgt gggctggaac ctaagatttt gcattttctaa tagactccca
 20760

ggtgatgctg gtgctgtagc tccctccgtc ccccgcttt ccccatcag tatttctgaa
 20820
 cacttgatcc acaccaatc tgcctcaaaa tcacttgtgg tggaaaatgc ttgttaaaca
 20880
 ggtagattgt tgtgccccat accaagacat aatggatttc ttatggcttc taaatgtgag
 20940
 aatcactttt ttgtattggc ttgaatgttc tgatttaaga gtatgggatg agggaaatct
 21000
 aggttcattg tgttcacatc agtaggactt aaacctatgg aagactttgc agcagaatag
 21060
 gaagattatt ctcccagcct acccaacagt tgaaaaacga tcagaaggac ttggaagggg
 21120
 gcaatggagg attgatgaca atggcaaaga ctgagcagtc cccaggcgtg acaacaaaaa
 21180
 ggaaagcaat cagctgtgga gagagaaatg gtgttgatga tattatgaaa gtcgggatat
 21240
 aaacaagaat gagaaagggg ggaatgggga aactgaatga gacagaagca cagaaaaaag
 21300
 aagaaggaa aaaagaagga aaaaaggagg aaaagtggga gaagagacag atggagagtg
 21360
 agagaaggga ggcaggtagg taggcagaag aatgaggaag aaagcaaagg agttgagact
 21420
 gggaatggac agaaggaaac ctgtacatgc tgtgatctgg gctttctatg tccctgatc
 21480
 cagggtcctg agatcttggg gtactcgcca ctgcattgta agtttctgtg gatgaagctg
 21540
 gcatataaat acatctgtaa agctgtaact cctgggcatt gtccctgaaa ttttactgat
 21600
 gtctagattc taggtatggg attaaatggg tccacatttt gtgtttttct gagtcaccaa
 21660
 ttacaacaaa acttgctcat atacttatgc ctatcttaaa aaataattta ttttaataaa
 21720
 ataataataa ggcaagaaaa aaagacaccc agatttgaaa aagcaaaaaa tgtctttatt
 21780
 tgcagacaaa atcatgtatg tagaaaatcc taaggaatct acaaattagc tactagaact
 21840
 aataagaaaa tttaatgatg tctcaggata cagggtcaat atttaaaaat tattgtatct
 21900
 caacttataa gtgatgaatg tttggataat ggggtttaaa taatccattt acaacaacat
 21960

ccaaaatata aaataactag caacagactt tataaaaaaca tacaatgaaa attgaagcac
 22020
 gttgctggga aaaagaaaaa agaaaacaaa attcataaaa agacattata ttactgaac
 22080
 agaacactta gtatttttaa gacgtcagtt cttatcaaat ttatgtgtag actcaattca
 22140
 atcccaatca acattccaac atcaacattt ggctagagag caaagtggg gtcagattgt
 22200
 gtagggctgt gttaaagagt tgagcattac atgagataat ggatgtgaag aaagtgtta
 22260
 atagggggtc gtgcagagga agtgcacaag aaatagtatc attgtcatta ttaccattac
 22320
 tgtattgttg ttattattta acaaatgttt tcaagccac tttggcttg gtatgggtg
 22380
 aggaaaacgc taccaggga gaaagataaa acttattctc catgcctctt gagaaaagaa
 22440
 gcagcaatcc ccagatgtag ggaggagggg agccctgctg gcttttagtc agcctcacac
 22500
 aggaaaggag aggtagagat tctgcagcga acaagaatca tgtccgggac ttcacagaca
 22560
 gggacactca aaggttgat gatgtcaat tgcagcatcc acctgcaggg gcctcaccac
 22620
 acttctcact taatgattgt ctctgctgtt agggagccct tgaaagccag gtgtttgtc
 22680
 tccttcccaa agcctcagaa gatagacggg ggccatgagg gaggtggggg cagggagggg
 22740
 ggagtctctg tctgttcttt tctgcccttc cctgtgctct ccacctctat cccctcttc
 22800
 ttcccaaaca ggtggcagga atgttaagac tatgaattcc atcaagagca agcagggttc
 22860
 cttagggcac cagttggtta gtgtaagagg cctcatccgc ttgcaccagc agcagaaaga
 22920
 agctgcagga agagtctctt ttgggcatga gaagctaaag gaggtggtg gttttctggt
 22980
 taactgcaca aggcctgcaa gttctacaga tgcctttga ggaccagagc agcctccttt
 23040
 cctccacac tctccagaca tgggaggcag taggagttca aagcacttct cagaaattga
 23100
 gcctttggtt ctgctgtcag caaaaggctg tgaagctcag tccccgacac tgggtaggag
 23160

aagggtgaat gaagactagg gggctctggg aaggaagtgg ggaacaaac tcttcatttt
23220
taaagagagt gacctggcaa aacacagaga agatactttg ggtttggagg aacgtgcaag
23280
aggtcgagag tcaactgccga ggggttttga tgcggcactg tgctccggca atggaaggg
23340
ctgtggtcag gaaacaaagc agcagggtgtg gcccatgtgg aattgtcagg caggatttac
23400
ctgtggtggc ccgtgccga cagcctggcc caccttctgt ccacctctgc tttccttctt
23460
aggacctcct agaaagaaag ctggaccttc cgaaaagtag cttgggaggc aaataacatt
23520
tatcgaatgc atcaatgtct taggcctcaa gcaagggttt aacatgagaa tctctttgag
23580
ttcccttagc atcctatgag gtagacacta ttaatatccc ccaaggctca gagaagttaa
23640
gtcacttacc aggaacaca actattaaat caggcctgcg tttgatccca gggatgttaa
23700
atttgaggcc tgagctcttt ccatcact gcaagtctc aattgtggca tcaactgcaa
23760
agaaaaacca actgaacgtt gctgtcacat gcagagagag agacaaagaa gagaaagggg
23820
gatgcagcta aatgctgggc tttttgctct tgcaggacca tgccacagag tccaccgtct
23880
ggggagcagg gcctgtgagc tacagattct cgaacaatac aatcatgggt caatgcctgg
23940
gcagagaata ggttctcttt gaagttcatt atctccagaa agcttttcca gcttccagtc
24000
actgcttcac tccagttttc tttcttcttt tttcatcttg gcttatttta gttggcttct
24060
atgcccata ctatcacct caaccaaaga cagagaggct ccaatgcagg aggcagacag
24120
atttgggctt gaaccaatg ccgcttctt gctgtgtgag ctttgtcaag ttattctatt
24180
gcctaatttt tggggattgt attatgttat gttacatgtt atgttttggt acatgatatt
24240
gtgttatatt actatcttct tcatgggtgtt attatattat taagattgat gagataatgt
24300
gtgtcagatt tctagcagg tttttggcat attgcaacca ccaatacact gaacatatat
24360

tgtctatgta actttctcag actcagagat ttttgtctct tttatctccc acaattcttt
 24420
 tcttttcaaa atgtcacttc cagaattttt tcccagaaga atccagtcca ttaaataaga
 24480
 atgttttttc actaaaattt tctttctgca tactcttagg attgtatgat gttcgtctgt
 24540
 ctccccacc tacaacctcc agcagagaga agagggtga gccagtagtg acaaggggtg
 24600
 gactttcaca tctcttttat tccccctgta gctttttgct aagggttaac tgatttcact
 24660
 taaaaaaaa ttttttttta agaggccttc catcctggtc agacttcacc tttttgcacc
 24720
 ttagtattct cggctagaaa atggaaataa taagtcctcc cccactcact ttacagggtc
 24780
 atcataaaaa ctcgataaga tatattgttt aaaatattta aaaaccttgg aatgcttcat
 24840
 gaatataagg cattgttatt caaatacaga gcttcttaag agctttactt tttagctctg
 24900
 tgatcagtta acttttgatg tggagcaggc aaccacgaat tctcagaggc gtgcaacaat
 24960
 aaacatttat ttctcacaca tctgtggagc tgagctgatg taggctgggc ttgcttctcc
 25020
 atctgggggt ggtgtcgttt ggctgaccgt gttgggcctg gctgggttag tatggctcca
 25080
 catgtttctc atcctcctcc caggattagt gggcagccca agcatgtctt tcttcaaat
 25140
 gatgctggat acataagaga gcaagcaaaa acatacatgg cttcttgaag cccaggcttg
 25200
 gaattggtat gccattgctt ttactggcca agcaagtcac atgggtggca tcagggtatg
 25260
 gcttatgcat catctactgt gggaggagtt ttctaagtat gatgagggca cagatacagg
 25320
 gagcggtgac cattacgaac attaatgcca atcctccaca ctgtgtttct tgatctgcca
 25380
 ggacagggtt ctctgaaggc cagtagaggt tcagtgttc caaaaaagga tgctgggctc
 25440
 aggtaggtag acagagttca tcacctctat catgactgat tggctctaag gcccaatgac
 25500
 aaaactcttt acttctgtaa agtggttagac agtgaagaag atgaaaaggc catcatgaaa
 25560

ttcagtctct gcttttagtt tttgtttgaa attcggctctc attggtgtgt ttgttttggt
 25620

ttgttttggt ttcccatatt ttctcccaag gggtattact ttcacaacag tggaaaaata
 25680

ttctttatgt tctctccctt ctcttccca ggattccttt agattcctgt gtgtgcctga
 25740

gggagctaga agagaaatac ataaaacata taactgtggg actcttgtct gagcccatct
 25800

ccattcttaa gtccattca cccccacaaa agctttcctt agagccaatt atgttcctta
 25860

ggctaccata tccccataaa cagacgtgag ccactgtggg ttcctagtaa tgagattttg
 25920

tttggcagtt ttatttgcaa atgactccaa agatctgagt aaaacatctt agagcaagaa
 25980

attctgggtt agaaattttc aaaccagaac acttcttcaa aggaagatga ggttgagaga
 26040

ggggggtggc gtgaggcagg gattgatttt aatgggccag tatccagggc tccagatgtt
 26100

tcttcaaaag gaatctcca aattattgag cccaagttac agaataagcg agaaagaaaa
 26160

tgggtctacag aatctttgat tgggtccgca tgaattcatt aaaagtgggt gtcagggaga
 26220

tgtgaggggt ggggaggctt gcacatatgt atattaacac acacacacac acacacacac
 26280

gcacacgcac gcacacgtac acaaagagct tgggtgtactg gtcttgtcaa atgctttcct
 26340

ccctctgtcc ctccccacga aattaaatcc aaggtctgga gttttgccct ttaggttaaa
 26400

aacaaaaaaa acaacaaaaa aaacaaactg ctgcttttcc aaaaagcatt aaagagagaa
 26460

aaaatataac agcagctact gacttgaatt gaacttaaaa tttccagga agaacttgct
 26520

cagagccgaa ctccagaaact tgttgcaagc tccccagaca tgagaaaata aacttctatt
 26580

aagctggcac tgaagtgcct ggaaggggtgt tcaaccctta gaacaactct tagctttccg
 26640

aggagtttca gggttggaag aacaactttg ccaattctac cctcaacttg gaacaaagca
 26700

ttatacatct ggaggagtct ggtgggttaa tgctttcttt tccatccaag attattcatt
 26760

cattcaataa atatttagtg agaacatact acctactggg cacggtgcta ggtgttgaga
26820

tgtggtgatac aggggtgaaca aaacagacat agtcattgtc ctcttggaat ttgcacttta
26880

ttggaggtag gggagcacia ttaactaatt acgtttactt gagataagt tagttaaaga
26940

aacaaacaag atccagagac agaaagtagc aaggacttac aggatgatca ggggaagtctt
27000

ttccatgaag gtaacattta aactacataa tgctgatatt ttgaatctga acctgaggat
27060

ctcagaatth tccacttctc cctctgaaac agaggagggtt ttggtccaca ggagaggcta
27120

agggttgatt cagtcaaagt ctttgagggt cagattagcc ctgcagtta gctagaagt
27180

gggatggaaa cttaatcac ttgaattcac acacctccgg aagtatttat ggctacttct
27240

ggttctgacg gcaaggtcag gtgaatcaac ttggtcttat gaatatagtt gttttccata
27300

tcccttggtg aaaatagaag cagggttgaa tcaacgtaaa gaggaagttg ttaaaccaca
27360

tgagaaactt ccaaatacagg tgtgtcatgg actctttctc tacctgagat aaagatgagt
27420

ctgaataaaa taactgagaa gtggacttag gtgaaaatga gaagatagta aagatgattt
27480

gtcacaaatt gtaggaaca aaatccaaca ccacactaag aaaataatat actatgacaa
27540

ggtgggacct attacaagaa tgccaatttt gttcagtttt agaaaataca ttagaataat
27600

atacattagt ataactcagg aaaaatattg tgattatctc cagagatggt gaaaaagcct
27660

ttgagcaaac tgaacaccta ctcttaacaa aagcactcaa gaatatataa attaattggct
27720

gtttctttga caagataaga aatatatagt ttggtcttaa caccagtatc tgtgctaata
27780

agaaaatatt aggggcataat ttattaagat cagaaataaa taatatctc tctcccttt
27840

atgtgacatt gtactggaga tattagccaa tgcaattaaa caaggtaaat taaaggata
27900

aaaattagaa acaaataagt taaactgtca gtatttgcag atgatatgat agtatatag
27960

gaaaattcaa tcaattataa aatgaactca aacaataaaa tatttttgcta aatttagaaa
28020

agataagaga atttaaaatt aatactcaga aatcaatagt tcttatatac aaaaccgaac
28080

accagttaga ggatatgata gggaaaacct catttaaaat ccaaagggga gaactttaaa
28140

gcttccttaa tgaaacaaaa gtagacctga acaaacagaa agacaacttt ttgctttttt
28200

tttttttttt gcttaggatg ttttaatatc ataaggttgc cattgttctc ttagctaact
28260

tataaattga ctacaatcct aataatcaca ttaaaatfff ttatagtga attgatactg
28320

aagttatatg gaaaaaataa gaataactag agaaatatta agaaattaaa agagagctac
28380

aatgaaagat taactctgcc taaatttaaa aggcattcta aagtttctgt gattgaacta
28440

gtaagaacct tgtgcgtaag tggaaagacc aactaatgaa ataaaatatg atctgaaaga
28500

gacccatatg catgtgaaaa tttagtatat aataaaaatg caagttctaa tcacttgata
28560

gtcattcctt taaagagata caatgaaatc cataccttaa accatataaa agaataaacc
28620

tcaaatgtat gaaagatagt atatatagaa aataaaacct tataagtcca caattctagt
28680

tagaagaaaa tatcattaat ttctgacttg atctgacctg ggtctaacta taatttgga
28740

tcctgtaata ataatgata atagtaatca aaatgatgat aaagaataaa atttgactag
28800

gtaaagttaa gaaacaacaa atttgaatgt tcaaaaatat tatcagcagt aaagtcaaaa
28860

gccaagtgc gaactggtag aaaatgtttc tatatatatc agagataaaa agttaatgtc
28920

cccaaagtat aaagaatfff tgaaattgag aataagcgga aaaccaaaaa ccaatataaa
28980

tgggcaaaaag atacgttcat ataatttaag tgggccctaa atatataaaa atatcttcaa
29040

ctttcaaata tccttaaaca tataaacata cattcaaata tgtttcttgt aataagagga
29100

atgaaaatta aatataatga gataccattt ttcacctatc agaatggtaa aacaatfff
29160

tttaaactct taacacactg ttgacaaagc tgtgagaaga aatactctca gtgttaggag
 29220
 gaatgaaaaa taatgcaact ctttgagagag gaagttgcaa tctctaacac aaatacatat
 29280
 gtgtttctct cttgaggtgg caatccctat tctagaaatt tactctgaag ctatacctct
 29340
 aacaatatga atatgaaaat acatgtacac aaagtatttc actgcagcaa tgtttataat
 29400
 tgcaaatat tagaaacaac ctaaaatttc catatgtact agactgaatg tgcaatgata
 29460
 catccactca atggagtagt atggagttgt gaaacaaatg aggaagatct ttatgcttga
 29520
 cacagagtga ttttcaggat attttattga gtggaaaaag caaagagcaa aagagtattt
 29580
 ctgcaatgta tactatacta cattttgtgt aaaaaagaat atatacaatg tatgctatgc
 29640
 tacgttttgt agaagaggaa gacagatgaa aatagacata ttacaatat ctgtctaatt
 29700
 ttgctggcag aaacacagga agaataaacc agaaatggat acaatttgtg atctatagag
 29760
 gatgtgggga aatgggatgg aaaggacgag tgggtggcgt aggggacggc acttctctga
 29820
 atataccttt tgtatcattt ttagttttgg aaccatgctg acattttacc taccaagaaa
 29880
 aagcactaac tatatcataa atgaagaaca ttaccacact gaactaggga aaaaatttaa
 29940
 aaactaatac aagtaacatt tgaactcaat tctttgacta tagaccatca gactaaatac
 30000
 aaaaggaaac ataaacatta aaatgtagtt aacagttttt gtctaacaca tagtgtagc
 30060
 aattcagaaa ccactttatg tgaaacgtag aaatgagtga gtgagcaa atgcataactat
 30120
 attaaaggata aaggagagta ggtttctgac tgttgaagac aattaaaaat atgtacagag
 30180
 ggccggggcac agtgactcat gcctgtaatc cctgcacttt gggaggctga ggccgggtga
 30240
 tctcttgagg tcaggagtgc gagaccagcc ttgccacat ggtgaaactc catctctact
 30300
 aacagtataa aaatttagcc ggccgtgtgg cacgtgctg taatcccagc tattcaggag
 30360

gctgaggaag gagaatcact tgaacccagg aagcggaggt tacagtgagc tgaggattgt
 30420

gccattgcac ttcagcctgg gcaacagagt gagacttagt ctcaaaaaaa taaataaata
 30480

aaataataac gtacagagga aaatctcgaa tgtaccaata gtaataatct tgggttaaaa
 30540

ttgttggcat cagcaaaaac tcatggtttt attaaatata ttaacattta ataaataata
 30600

tgtattaata tataatttca tatatacgta catctatatg tatcttatat gtattttctt
 30660

tttatatata acatctgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtacaccct
 30720

tcagtcctat gaatagctgt gttcactgaa ggaagcaatg gcacttcaat aaaaacgagc
 30780

acatctagtc ccaagacctt gtttctaagt accattcctc acaaaaaaga atcctcaaca
 30840

aaaagacatt tcctgaagaa atgtctgatg ccataggctg gttggagcaa agaaggccta
 30900

acatgatcgt gaattacctt attgtaccag aaaaaaaaaa aaactatgaa aaaatgatgg
 30960

cgatatatca aaaagactca ggatccagtg tgagcagctc ttattgggca agtcttgga
 31020

atttgaacaa caaaatcaac aacggttagt tacacattaa catgtttgac aaaataagaa
 31080

tccatgagtc cagactgata taaattaatg agtgaataaa tgaatagggtg attgagaaga
 31140

aaggaaagct tctcttcaca atagaatgcc aattaacaaa tacaaaagga acaatggagt
 31200

tagaaaagag ttagaacact accattttgc agcaattaca ataataatta gtttaggtaa
 31260

gaatcatgaa tgggtgatac atttacggag actgtttgat gaggaactgg atattacata
 31320

gcctcacagt gtttctcata ggctatttct aattctagag gaaaaatagt aacttcacag
 31380

taagaagcat ggtggacacc accttaacca aatgattcaa gttatcaata acaacgggac
 31440

aatctgacat gtgtttcca atatacagat gaaaaagagc atagcatcac aattttacaa
 31500

cataaatcta agcaaaagga aagaaaagac aaaccaatt gaaggagagc cacattaagt
 31560

aattggcttg aactcttcaa aatgcgaagg ttatagaagt caaaggaaag actggggttaa
31620

tattcttcta acactgtaag tctcagccta aatccaaagg cctgagaacc aggagtgcag
31680

acatccaaga gcaggagaag agagatggcc tggctcaagc agagagtcaa tttgcccttc
31740

cttcattctt ttgttctatt caggccctgg agggattaaa tgatacccag ctttcaactgg
31800

ggaaggtaat cttcattact cagtctgctg attcaaagt taatctcttg cagaaacacc
31860

ctcacagtca caccagaaa taatgtttta acagctacct ggatcatcct tacctcagtc
31920

aagtcgatac ataattaacc atcatagagc cagaacatgc aataacttg aaataattca
31980

gtcattgtgc aaggagttac ctagactctg ttttaggctg gaaggaggga gtgagatata
32040

aaatgcatac agtgatcagg aggctacgag tactagaaag aactagaaag ctttaactct
32100

ttcatggcaa taataaaatt caagattgtt actaggaata aactgtacca gagttatgtg
32160

agacccaaac acagtcagaa ttagactttg gaggtatttg ttgaattgaa ctgaacccta
32220

tagtcccca ttgttcatag agtgcctaca acctgtggcc actgcactag gcattcagga
32280

ttcagaatgt tgagaggact atgagaagta acgtggctaa gcatggaacc agagagagtt
32340

ggcccaaccc tttcatttac aaatgaggaa actgatgcc acagtgagac tggcctggcc
32400

caacattaca caaccagttt atggcaaagc aaagagaaac acctgtttgg gtagtgaggg
32460

ttgcctttta ggtgccctga cccattttt accaccaagg attgaggagg aattgcagta
32520

caaagagaca agagatagaa aaaatttgag cgttgctgct gctatagacc ctccatccag
32580

caaagccttg agagagatat tctaccaccg ggctgggact gtaaactagt tcaaccattg
32640

tggaagacag tgtggcgatt cctcaaggat ctagaactag aaataccgtt tgaccagcc
32700

atcccattac tgggtatata cccaaggat tataaatcat gctgctataa agacacatgc
32760

acacgtatgt ttattgtggc actattcata atagtagact tggaaccaat ccaaatgtcc
32820

atcagtgata gactggatta agaaaatgtg gcacatatat accatggaat accatgcagc
32880

cataaaaaag gatgagttca tgtcctttgt agggacatgg atgaagctgg aaaccatcat
32940

tctgagcaaa ctattgcaag gacagaaaac caaagaccac atgttcccac tcataggtgg
33000

gaattgaaca ttgagaacac ctggacacag ggtgggaaac atcacacact ggggtctgtc
33060

gtgggggtggg ggaaggagtg agggatagca ttaggagata cacctaattgt aaatgacgag
33120

ttaatgggtg cagcacacca acatggcaca tgtgtacata tgtaacaaac ctgcacatgt
33180

accctagaac ttaaagtata ataataaaaa taatatatat tctaccaccg ggctgctaac
33240

ttggcaaggg ctgaaagaat ccaaccgaga agaaaatagg ccatcattgt gccctctcc
33300

acatcaaaat gcaattactg agcttttctt aagagctcag cactgtgcta gatgttgtcc
33360

atgttgatct tcacaactgc agaggctgaa ttactatatg tttatgtgat aactaaatat
33420

tctattatgc ttgcatttgt ctttggtgaa tattattgaa gtcgtttact gctacggttt
33480

cctctgcagt tttctacta gtccatcaca ctcttctttt tggccctatt tttaccata
33540

ctcagttagt ttcttgtggc aggctacgtg tgtgtgaggt tctgtgtgta tgtgtgtttg
33600

cgcaaacatg gttttgtgtc agtatactc atagatatga aaacagtttc tgttcaaaca
33660

gtaccacac ttagactcca tattaatctt ggctcaattc tgtgaaaact ggccctgaag
33720

tctgtgacct cccttactca cctctttacc tcattctaac caaaagaatc cacttcaaga
33780

acaccttcg cacaagctct tcaaatatct caagttctct tttccagtat ggtgttgcag
33840

ggtgtagtga aaaatgcatg agtttggagt caaacagatt cccgttccat ttctggttgt
33900

atcctatatt accacttgga gttgttgga gtcttaggtg aacacaggag gcaattcttt
33960

ttttacagtg actgacaaaag aggtgagcaa gaaatgtgag tctcttattt attcccagcc
34020

cctctatctt aacttggtga aggtaaagac cagtggatct gaatttagat tatttaccc
34080

cccatgtgtg gggaacacac acacgtgcac acacacacca ggctggaatt tcttcagttt
34140

gttcctccgg gtctactcta tacttcatgg aagctgatct tatgacttac atcaatggc
34200

tcccttgccc tctggcttct attagatttg gcaaattggg agccctggca gatcactgaa
34260

gggaggacag aagttgtagt atttattcct gtgatattct ctccctgggtg gtcaccttgg
34320

gctgtctgcg tcccttgaca gaaggtcacc ctcttgctcag gcatcacttt tctcacagtt
34380

ctctctgtct ctgggttcca acaccactcc ctcttcttga cccttaacgc ctaggggtga
34440

ttatagctcc agctgttatt agcacggagg actgcacgat cccttatgat ttactacac
34500

ccagtaaata gaacctttat taagcttttc ccaaataacc caatttcagt gcaactgtctg
34560

ttttctgctt ggactttcac tcacgtatgt gcatgcacac acacacacac atacacacac
34620

gcatgcacac tcataccac gtactctcta tatacgtgta cattgagtgc tatggaattt
34680

ttagggttct ttgaggacaa cttcctggaa gaagaataat gggagatgtg ctgagaggct
34740

ggtgagcgtt taattgacct ggttgtagc acaatgggaa tgcttttctg aaccttccat
34800

ttgcctccaa tttagaataa ctgcttagtc agcaacatgt ctccaattaa catgacagat
34860

tttttgacc acagatctct ttgggaatct tatgtaagct gtgaaacttc tcccttaaaa
34920

aacctacata ttttaattctg tctgtaattt taggacttca catatcccca aagaccattc
34980

atgatccacc agggttaaag acttgtttta agtaaagaaa atattcacag cctctggagt
35040

agtagacatc tccagagaaa gtctggcatt ctttggtaac tcatggcatt tattacatta
35100

gctttggatg gaacaacccc caacccccaa gtctgagcat gtggtttgat gctgcttct
35160

ttccattttc actctgctct cttggattct atgtagcaaa aaggtagcctt ttctagggtc
35220

acatgtccac ttctcaaact ctgtcttgca ctggggattt gagggtaggac tatcagagca
35280

taatggaaga ggcaggaaag gatagcacta attagagaac ttgataggcc ttttgctact
35340

gataaagatt ctgtcctgta tcatccatga ggtttggatt tgcacttcta aaaaatagct
35400

gtgtgttttt gacaatagcc actccctttt tgctcctttc tgggtctctt cctaaatgac
35460

tttccaggat agtataaaat aagcaatcta ttttagaggaa caaaacaaag aactacacca
35520

ttagcattaa aatatttcta gggtagttta ggtaacatag aactttgtag ctcagcagtc
35580

ttggaatcac tagactctaa gccccatgta gtcagagact gcctgctttt ataattcatc
35640

cccaacagca ggctattgca taacacattg ttaccaccca caccattttg ttgagtaaat
35700

gaatgcatgc atgaacgaat cagaggattc tgttgtctaa ataaactttg gaggtgactt
35760

agtcttctct tccctctgat gcttagatct tctttatctc tcagttcaac atatattgct
35820

caaatgggtc ctaggtgcaa gggatctgga agttattaaa atgaaataga ttttggaacca
35880

tgaccttgac aaattatgat ctactaggag agataatatg gaaatcagta taatactaca
35940

gagacagtaa aaaggaaaaa gaaatgtttg cacttaataa ttattccacc attgttccaa
36000

aagtttcaat tatagggaac taagttccac tactaataaa agtcaatacg cattgtatgc
36060

ttactgcatg ctaaataatat tatgtcatta taatatttat gttcaggtgt atactatgat
36120

tgtatgtatt ttacgacaag gaaattgagg ctcagagagg ttaaagaact tgctcaaggt
36180

gacagagcca gtaggtggta gaacaaagac tcctacattc tgacctagtc catatgtaag
36240

cagatttatt gctagaaata tcttgatggt gcattgaaat ctgcctctcc ttattttcca
36300

tttggaaact attccactct cccgagccag aaaaaccaag cctctcatta cttgaaaaac
36360

caattaatga gttaaattac atcaagttag tttaacagtc agcaattccc tgttgtaatc
36420

atagttaggc tgctaccatt tttgtcccaa tcattggccc aactgaatt ttgtgtctaa
36480

ttcattgctt tctgtctcct taatttcac ccttgccccc agagatacta tgctgagtat
36540

tgatgctgca gcaacatatt aatttttatg ttttctgact catttacccc aaagccatag
36600

tgcatcaaat gcttttgggt ttgtgaccat tttatctttg gtttcagtca ctgtttatct
36660

ggtcgcatct ttgttcattt aacaaatatt tgctatgaag tgcaagtta gtgggtagca
36720

aacactgctc tggcaagggg tatatgaaga aagacatcat tccttcctcc aagactagtc
36780

ttgtttggaa gctggacaaa caaacagtta caatacaatt aaaatagggt taattaagaa
36840

atgtgccaac tgtcaagggt aatcaaaggg gaccctttca gtctgtccca aggatgctag
36900

gtaagtcttc acagagaaac tggattagg gctaagatct agcggaggac taggggcttt
36960

ctagacagac aatggaggca ggacacttct gagagaggaa tgaagccttg aggtgtgaga
37020

cagcatgtgg tagagcatgg agggttcagg cacttgagg agttgggtat ggggagaaat
37080

acagagcacc tgtgagagcc tggggcatgg ctggagaggt aggcaggggc cggcttgcaa
37140

agggctttac ctgtcatgct gaagggctaa ctgctgttct ggagaatttg cagagccact
37200

gaaggttttg aagcatccta atgttaggat tcacttcaag ttttagaaag tttagtcttt
37260

tggtctcaagg gtggggaggg aataaaagaa gcgagagtga aggcattggga ccagttaga
37320

gactagtga gtctgatgaa aaggagatga gagaagccta aagtggacaa gcagctggac
37380

acggcggctc acacctgtaa tcccagcaat ttgggaggac gaggtgggag gatcacctgt
37440

agtcgggagt tcaaggccag cctggccaac atggtgaaac cccgtctcta ctaaaaatac
37500

aaaattagct ggggtgtggtg gcagacacgt gtaatccac ctactcagga ggctgaagcg
37560

ggagaatcgc ttgaacctgg gaggcggagg ttgcagtgag ctgagttggc accactgcac
 37620

cccagcctgg gcaaaaagag tgaaactctg tctcataata aataaataga gtggacaagg
 37680

aaatactgat gggggtgaga aaacaatttg agagacattt gagaagtagt gagataacat
 37740

ttggttactg attgaatatg gagtgttagg ataaccctgg gatcctcatt ttgggcaatc
 37800

acaaagctac tggctacagc aaaaggaggg gaagaaaagt ggtggtggga gggaggcatg
 37860

agaaatttgc tgctggccat gttaagctcc acctcctgtc accaagtggg gatgctcagg
 37920

agatgtttgt gtccacagga ggagagtgat ctgggtggga ggtttacatg tccctccttc
 37980

catctcccgg tccttctttc cctcctccct ctttctttgt ttggtacttc ctttcttctc
 38040

tcttttcttt ttctgccttc ttttgagaat gaaacctaag agtgtttata tgtagaggga
 38100

aaaagtggag ctagaagttt aagacacagg aaaggaggga aaaatgtcct agtgtagagat
 38160

tccaaggagg aagaagggcc aggataaggt ggggagaggg atgcctcttt tgctcattgc
 38220

ctttatctgt ctttgtctgt cctggattga ttgtcctatc cttttcttcc acttcgtgtc
 38280

ttctcccgtg tttctacaca aggcacagtc aaataagggc tggcacgtgc gcatctttgt
 38340

gcacacatgt gtgtgcacaa gcattgaggg atgcacaggt gcacacacct tgtagagcct
 38400

ctagtagctc aagtcacatt ttttcgcatt tccagattgt aagccttact gaggggcaga
 38460

tgggacgttc tcacaggaaa aaggctgacc tcaggagaa agatgtggcc cccacctaca
 38520

aaccaatgat gtctactccc aacaggggac gtctttcccc tgcgtgaagc aacaggacag
 38580

atttggttag gtgtggcatc gatgtcctaa tctcctatga gcacacaatg aatgaagggtg
 38640

gggggctctg ttaactggag ggaatcaggg gctgccctgc aaagcatgtg caaagaaaat
 38700

aatgcaatt ggagtgttgt ctacacatcg tcctcagtgg gcacaaaaga ggcacagtgg
 38760

gatggggctc tgaccagtt tccatttggc cagccgactg gaaacagcta ctactgggc
38820

cactgggttc tttccaaaga acagtgaaga acagtgggtt ggggtaggtg ggggcagggg
38880

ctggtggggc aagaatgtca gaagcaaagc caggcgtggc tggccagta gttccagta
38940

caaagaagcc gctatctgga agagtgggac aaccatcct aaacatggga agaagaaaca
39000

ggaaatttca aaagatacaa ctactaatta aatgaaagca ctcttaacc accccttga
39060

taattagatc atcttctcct agaatgagaa ggaatattag tctgactttg ccctccatat
39120

ggagaaacat ggcaggcagg ctctctgtgc ctggagggtc tggccagact ttgtctggca
39180

tgggtatttt gactccagcc tatttttttag ggatcctcct gcagctggtc ctcaagaggt
39240

ctgaactttg ccgtctcatt tctaaagcct ctttcagcag gtactggggt tcccagggtc
39300

agggaagcct ggtcattgct agaacacatg gcaacatcct tcaggaaact cagctcagac
39360

tgtctggaaa gatgctggaa aggtgaatc catgcaagta ttatctaacc agccccaagt
39420

catgttttca cacttagatt tctacaggat attcataaat tctggaaaca aagatattgc
39480

cactttattt tattttttaca cattgagaag tttccttggc aacataacat gtatttgcct
39540

atgtttccaa ggtttatgga tgaatgtggg tggccaatgg gattcattca ttttgccaag
39600

caaacgatat gtttgaagcc atattaagtg ccaggtagca tagaggatac agatgttcat
39660

aatgtaaact gccctctagc agcttacaga tactacctag gaataatgct tacactgcag
39720

taaacaacat aaggcagaat atatatatat gtctttgtac tcataacaaa atgtaaaagt
39780

caaagtcagg gctcttttaa ctttcttagt agcagaaacc caaattgaac ttaagcagaa
39840

ggggaattat ttatttaaaa cacaaagaaa ttatttattg aaagcaaaca ggaaattgat
39900

tattgcttct tacaactcat gagctaaaag acagagctgg gttcagatac agatatatcc
39960

agagttgtga tgctatcagg acgctgtctg gtaatgaaac caaccacagc acagagcctg
 40020
 gagtggaggg acagaccaac tctcttcaga tgccacatcc tcaccagca gacctctcc
 40080
 tctccagcac ccattcatca aaccttcgag aagacctgg ttgctactgc aaggctttca
 40140
 gcaccaaagt ggaaccaact cctcctcctt ggacatggaa tacttgaact ggccattctg
 40200
 ggacaggggc ctgcccattgt gaatggaggg aggaagaagt gaaatcctta aagggttgct
 40260
 gggtcagcat atcacaatt tcccttgccg tgagccagat ctgatgggat ggtattagtc
 40320
 aataggtaga taataggggg aagagggggg gaagaatgag atggttctac aaaaagaaag
 40380
 gcaggggtgcc agaaggtaga atttgggcaa tagtgagtca cagaattgat tgtgtcttgg
 40440
 agcccatgga agggagctga ggcacctacg gttggagggg gaggtgggga cccttgggca
 40500
 cagctaaaga gtctgctttt gggtaggatg tacagagcaa ttagaattat ttagcagtga
 40560
 cttaaagcca agagaacaaa tctggcttca gagtgttgac atggaagaaa ggaagaggca
 40620
 aatggaagag gctattccaa agggagaagg tctgagcttg gtgggtgcctg ggaaggaagg
 40680
 aactgcagc caagctgagc ctgtgtccag ctgaacacgg acatttactt ctgcttctga
 40740
 ggggaaatat tgtcatgtct tatgggtgta cccagcaaaa gatgtttcca ctaaaagact
 40800
 atagagaatt gcttatttct aaggactgaa ttccgaatgt ttttggtcag gttctctctg
 40860
 aacagagaac cattgcccct gagcagaagc ggtcttcaca caactagcaa ccatgtagtg
 40920
 ttttctcatt caaatggcgt ttggtattca cctgcctgtc ctggaacaaa tgcaaagtca
 40980
 gggttgtttc tagctctccc ccaatattgg aaaaatgctg atgaactgag cctgggtcca
 41040
 ttttacacct gacaaaagga actcagatga gagtgttggg aactgggttt aaggaggagt
 41100
 cagaggtgac tgctgaaatt tctgccagac tacaactaaa ttaaaaaagg aattcggttt
 41160

aagaagccaa tcttcaacta atttgaacat gaggatgctc aacttagaag cctcccatgg
41220

ggccctggag cccagaaggg actgcatact ctttgctcct gcaggtaact cagggctttg
41280

ttctgggtgc tctgacaggc ttctaacctt ctgcaattct gttgccacaa tttctgtggc
41340

accaggagaa aattcatgca tctaggctag ggtggtgctc tcatgtgtct ccaggagagt
41400

gtggcctggg ttagcttaag actaggaatg tgctgacagc tgggaagaaa gtcaaaagaa
41460

ggcatggcat gtggttatgg ggccggaggg accgcttggg gggagagggt gaagcgctga
41520

catgagcttg ctctgcatcg ctggttccca gaggtggcct gataggagag tgccagctgg
41580

tggtatgccg tgggctctgg ggaaggggct agaattgtgt tgcagcccag aggggctggg
41640

ctagttagga gcaaaagaaa acataagctg gggggaaact gcctaataga ccaattagga
41700

gacaggggtg tgccctaagg gatgggtggg agcagggggg gctagaagta gacagagagt
41760

aactctcagc tctgaacagc gtccacata ggaggaagaa cactgatgtg aggccaagt
41820

gcttggaagg gcgttgacac cccattccct acaacaaagg gtttggcaaa ctcatctcat
41880

ttgctacttg tccatatagg aaggatggac accaaaccca aagagacttt tgggatgtgg
41940

ataaaaattg aaaactttgc ctctcattaa tcccttggag aatatagcat ctcttcaaa
42000

ctggactgct ctcttgatc ccttgtcttt acctgaaaat atctttcaag gcctagctca
42060

tgctatcttg gccataggaa ctttccctat cacctccatc aatgctgttt gtccctccac
42120

tagaatcttc tgttttggtg ctctcttac tcagcctcaa gttaaattta tgtgtactgg
42180

ttttaacccc agtaccaggt aaaggttctg aatacacgtg atgaaataag tctctcctat
42240

cccatggcac ctagtatagc atttagttgc tatactaatt attataatta gcatttatta
42300

cgggcttact atgtgcaaag tgccattatt ccatttatta ttcacaaaaa tatttagtct
42360

ctactattct ctaatttatt agataaccta atataaatct tttttgtttt gattgttgct
 42420
 agctgattga aggaaataga attgtggtgt ggtagtaggg gatgttgagt gagtcttcct
 42480
 cattcttaat atttccttcc actaggataa cgcccctggg ctctcttcat caggggagga
 42540
 ggagtgactt catctgatag attcgggaagg gaaaatgtta cccagtcgct tctcagtcct
 42600
 ctttctcttg gagcgagccc atctctggag cacatacact ctcccctgct cctgtaggtc
 42660
 tcgagttctc aggggaaggg tttgccctgt aagagagctt gtccttctcc agtgaggcct
 42720
 gtcactggga agataattct gtacagtgcg gggattcaaa tttagaaagc ccacttgccc
 42780
 cataattaag ccacattcac cagctttaac aatgacagag gtgatattga gttctccagt
 42840
 gatattgagt tctccatctg ccttatgctt ttctgcttat ttccccattg gttcagaatt
 42900
 ctggtgggaa aaggggtgct gtttgtggga tgccttccca cctcatgtg gccctccata
 42960
 agttcttcat ccctcattgc attcaataac ttgggtgtgg tgtgggtgga gtgcatgtgc
 43020
 catatgacct tctcccttga gtctttgaaa cctatttctt actttatgac tttggacttc
 43080
 cttttctttg ggagtataaa agtggatacc accattccac cccttaccg tcatcaggag
 43140
 aagttagttg ggaagcaaag gcggagtgac gtggagggaa gagtctaaaa acagctctac
 43200
 acttagcttt gtccttgcat tgatgacaag cgatgttgga tcctgaaagg tggggatggg
 43260
 aggtaggaag acccaggaag ttgctgtgaa ggctaaatcc taaagacagt cactaaaagt
 43320
 ttccagaagg ggcattctatc actccacagg caagcaagat gagaccatgt tgaaaaggtc
 43380
 ttaaagaagc caggaagag gtgaaacatc tctctgtgtg ctgaggacca taccaggtcc
 43440
 ttcattatca tattctacac attgaaaatg aaggtaagag agcaagcaac ttgccctagg
 43500
 ctaccaggg aatagggagg aaagtctgga ttgaaaccta ctgctgcctg actccgaagc
 43560

cctgttcttt ctctcacct ggaggatgga tgagggtggg atggggcagt taaaagttga
 43620
 ggggtatctc taggagaaag atgagacaca tgaccattgg gtctaaggag ccagagagag
 43680
 gccattgtgt ctgagcctgg acagccaagc ttcctagaga cagaggatgg gatgagatga
 43740
 cttctggctg actctgcctc cttgagctct ttaggatccc ctgcagaatc tcccaaatg
 43800
 atgtcattg gaaacagaga gagcacattc aagtaagtat ttgtgccaca ctatgattgt
 43860
 tccaatgacg tcaagtctgg gatggcaaca gcagcggttc catttctttt cttcaggacc
 43920
 tggattcttg ctccaaggtc ttccacctcc ctctcattt tcttccactc ccactgcagg
 43980
 ctcttctt ggaggagctg gggcaggag taggggtgga ggttggtgtg gcacttgctg
 44040
 acatgcatgc ctgtctctta tgtgttgctg agtcaggaat gggcaggaag gagaggcttg
 44100
 gagacgaagg cagccctggg caatctggtg atgaactga aacgggagag ggtctattgt
 44160
 gcatttgac tgactctgtg caggaagggt gggggcaggc caaggaggac tggcttggt
 44220
 gagacatgag atgtctcagg ggagggttg gctgaggcag tggaagaatg cttgtcggga
 44280
 accggcctgc ctggaattgt aacctccct gctgtctctg caccacactc agagtacagc
 44340
 tgtatgcctc ttagctggtg gattcctttt ctagtgggg acaccgtgga caggacagt
 44400
 ggggaagagc attggactca gagacaactt gacatcaagt cctgcttgga atcctatttg
 44460
 ggtttgtaat tcgttgaagg cagacacaga cataaactaa tccttagttt tcccatccct
 44520
 tcaataagga gtacaacacc tatcacacaa ggctacaaga ataaaagaga gaactgtac
 44580
 aaaggcaggg cctccacttg gcacatctta gatcctagac aaatgctgaa caaggcagca
 44640
 tgggtgtgatg gtgaagggtt acagccagac tgactgactg gtctgtctt cgtttgaagt
 44700
 ttgatattt tgttctcgt gggttttttg catttttact atttatttat ttatttattt
 44760

atttatttat tgagagacag agttttactc ttgtttccca ggctggagtg caatggcgtg
 44820
 atctcggtc actgcaacct ccgcctccca ggttcaagcc attctcctgc ctcagcctcc
 44880
 tgagtagctg ggattacagg catgtgccac cacatccggc taattttgta ttttagtag
 44940
 agatgggggtt tctccatggt ggtcagcctg ttctcaaact cccgacctca gatgatccac
 45000
 ctgcctcagc ctcccaaagt gctgggatta cagggtgtgag ccaccacacc cggccaattt
 45060
 tttttttttt aaaatatggc attaaaatgg tatttatctt gactactgag tttttggta
 45120
 accccttaca ttttgtgcc aagacatctc actcatgtta tcttaggcc agcatggctg
 45180
 gccttgtctc ttaaattcaa atttttaaat ttttcattcc tcaggcttct cacttgtaa
 45240
 atagaaataa taatagcatc tccccacag gcctgtgctg agagtgaat gtgttactac
 45300
 acgtgcctca ctttcatttt caacaggacc tattgccagt aagtgtcgt taatgttagc
 45360
 cattattctt ggcattgctg ttgttattag tgtcagccat cattctctca gccaagagat
 45420
 ccttcagaa gagaagttct ccatggagg ggatgggaga tgagctgaac cctggaaca
 45480
 gaaatggatc agatcaggga gcctcagtgc cctgaagcca gggcaggggt gcatggggct
 45540
 tctgcaaaac ccagggccaa gcatgcaggc aagaggcagc cttgccaagc cctgcagaa
 45600
 aatgcttgca gtgcaactcc tctctggaga cgccacaagt tgccccgaca gcttttcctt
 45660
 ttggggcttg ttaacattt caaacaataa taaaataaat aactgcagtg tttatggaa
 45720
 cctaagacag aaatgaagac gacttagaag aaagttaagt ggaagtggc caggagggg
 45780
 tttagaagat gccccttcc ttgggtgttag cagttttgtg cttcagtttg tcagctcttc
 45840
 ctactgctct ccctagcaa ctctctgcc agtgtgcaca aatgcatgtg tgcacacatg
 45900
 tgcattgtga tgcatgtaca cacacatgac acacacacac tcatccaac gagaggcaca
 45960

catttcagac accatctcac acacacacat ttctctcttt ctggaaatgt gtccaccac
46020
tttctccttt atcttcttca ctgtgttttt ttgttttttt ttttttttct ttttagacaa
46080
agaatctaataaactcattc tggccagatg gtgtacatta ctttaactcc ttcactcctg
46140
tgaaaatttt tggcctcaac cttcactctg aatagaccaa gccaaactcg ttttatgtgg
46200
ataagccatt tccaagtcaa agttccgtgg caggaagtgg gagagtgggc agaggagca
46260
ggtccttaca agcaaagact tgtataatgc tttgcctaga tttgcttacc agtatgttaa
46320
atatatgtgg actattttat tcttcacct tttattttgg tatcttcttt gtctatccag
46380
tctcactttg ccagactcca ttatgcaaga gaggctttgc cagaagcaag gatgctgtag
46440
aggaaagtaa tgaaaatgtg aggatttctc agaaggccac atacaatctc taggtctttc
46500
cagtacacac agcttcaaaa gccagggtgct gtgttcccca agcagggtcag ggcatttcca
46560
ggagccctgg attccagaag ctcaggacac ctcagtggta aagagaccct aatgccggtg
46620
acagcacaga ccctggaaac cctagattgc ctccgtgtgt tcctcttggc ctcagcaggg
46680
ctggtttccc cctcctgtag gagcagagga ggctctctc tcttgagccc agggcctagc
46740
atagtacca taaggaagca tagaagatcc agctgcagat ttggaaaaca gagccctcat
46800
ctctggaaga gccaaaactg atcctcaaca tgtgccatgc acatcagata tcaaacgctt
46860
cattcathtt gccttcgctg ggtaaccagt gaccttgaga agtcacctct cttgagtctc
46920
aattttctca agttgggaac attacatgac aggacaggac tgtgaggatt aaatgtaaac
46980
aacatttttt ttttttctat agaacatact gtggggccag caccatgcaa ggtgctttta
47040
tctacaattt atttctttcc ctcccatgc cttctgcctc tttgaataac tgctgcaatg
47100
gctgtagaga gataaatgca gtagaccaag acttttagtg atttttaaac aatgttctag
47160

ggcaaataca tgggtatggt tacatactcc acaaagtga aaaattatgc tagtcagtgt
 47220

gcaaagtgtg tttttccaga cctgcaaaaa ctgagatcaa cttccatatt tccaacttcc
 47280

atattccctt catcccttaa gctaccaggc agaggatggg agataaaatt tggccataaa
 47340

cttgtcatgg gtccatcttg gtccctaatta gaattatcta cttcctaata cattaactgc
 47400

atggatagat agtctgaggg cctcaagggc aagagatgac ttctaccctt ccacccatcc
 47460

agccctgacc cacccttctc atatacattg gagaaagctc tgctctgtgt aaaacacatc
 47520

aagtgcagga catggttctt ttcttcaact atctacagaa atgaatagcg gccctagaga
 47580

aatgcaaaag atgcctgcaa tccaaggcca aattggagaa catttcaagg agaactgggc
 47640

ctttattgta gagtttgga aatgagtaac atgtagatct gtgaatagat atgtctgagg
 47700

aagcaccagg cagcggaaat agctggttgc aggagcagtt tgcctgccc acctgaatct
 47760

tcaggttcct tcagctaagc cttaggagag gaggctggag agctgctttg ctggatcatc
 47820

tcctttggcc cccctcagac tagaactgca cctcctctct acctagctct gtgctccaca
 47880

ggctgcctct gggctgcact gatgggctcc cttgactgct ggctcctggt cagggttgga
 47940

ggaggaggat gaagtcattg aatagtgtcc ccatgctggg ctgcaggtag agttgctgct
 48000

tctctactga agaccctga cagcaaccct gtcttacagc cagggtggtt gtagttgctc
 48060

cctctccct gtccctttag agtgggtgat gcactatatg ttagatatgg tcagagcagg
 48120

agacatcatt ttcccttaat tctattctgc cttgtaaccc ctttcttaga ctctctgctg
 48180

agtctctgct tatgtgttgg gattgcattg atgcagatgc tttagggcca cgttggtggag
 48240

gatacaagat gatgagacag gaggcagcgt atggtgaatc acttctcaca gctcctaact
 48300

aagttctggg cacatgaatg ggtactcagc acatacctac tgaccaagcc tatggagagc
 48360

tgaaaactca taggctgact ttgaagcgag gcctttgaaa ccatcgctcc aggggaaata
 48420
 atctagggtca aagagacggc tgtggcaagg gtgcgagctg gcaagcatga gggctgacaa
 48480
 agcaggcagc actcctcctg agcagatgac cagccccgca gtgctggatg ggtacgacta
 48540
 ttccaaaccc tgggtgtaatt ctgtcctcta acgtggctgg tttgactcag cactgctggg
 48600
 gagccacaga gcctttcttct ttagcattta tgagctcatt ataagcatcc aagaaaacat
 48660
 caacaaaggg tgctaagtcc ctctagaact gccttttaaa tttttctttt gtaaaaagaa
 48720
 tgatgcttta agtaaaacca caacttacag aattcagaca aaaccaggca gccccaatg
 48780
 ttgttatccc agctcctctt ggaataaatt tgggaatgaa caaaaagtca cagaagagaa
 48840
 aactccagcc cacacccaca ccttttctga catagccgcc tgtgcagggt tttgggacca
 48900
 ggtctgccat ctgtattccc aggctgggct ggagcacctg aggctactat gggtacagaa
 48960
 aggaacctgg catccgaagg tataggctca aatcccagct ctcatgattg ctagctttga
 49020
 ggtttcaggt tcatttctta agtgctctga gctttgggtc cctcctcttc aaaatgggaa
 49080
 tgaggagcat gtctagctcc ctgttttggt gcagtgaata gctacatgaa gaacagctaa
 49140
 accagaacag aaagaaggca aatggaagct acagaaacca aggatttcct tgttgaatcg
 49200
 aatcttcctt caatcttcct tcaccacact agtggatctc cctgtgggag ggatgttgag
 49260
 agtgctccgt gttttttttg tttttttttt tttttttttt gtgatggagt ctcaactctgt
 49320
 caccagact ggaatgcagt ggtgcgatct cggctcactg caacctctgc ctcccagggt
 49380
 caagtgattc tcctgcctca gcttcccaag tagccaggac tacaggcata tgccatcaca
 49440
 ccagctaat ttttgtatct ttagtagaga cgggggtttc ccatgttggc caggatggtc
 49500
 tcgatttctt gaccttgta tctgcgcacc tctgcctccc aaagtgctgg gattacaggt
 49560

gttagccacc acaccagcct tctctgtgtg tttttaatca atagcaatat gtacatatt
 49620
 tagtagggac tataaagaaa ttgtgctaga aacacctggt tattttacag aactttgcca
 49680
 gtgagtacaa aaatatgcgt agcccacaac ctctctcccc tagaaaatac ctctgtggg
 49740
 tagaatgaaa tgggtagggt ttcttaccat tcattagcat ttcaacctgg ctatcttctt
 49800
 cctagaattc cactactccc aattgcccac gtgtttccaa taaaaccaat tcaattgtgt
 49860
 accaacttgg tgacaaaacc tttatgtagt tcttctatca ttttgtgggg tttcaaaga
 49920
 gggttgctct ttgatatggt ttgatgctgt gtccccaccc aaatctcatc tctaattgta
 49980
 atctccatgt gtcaggggag gggcctggta ggaggtgatt gaatcatggg ggggggactg
 50040
 ttcttgtgat agtgagtgag tactcacgag atctggttgt ttgaaagtct gtggcacttc
 50100
 cccctttgct ctgtctccct ctttctctcc tgccaccatg taagacgtgc cttgcttccc
 50160
 cttcaccttc tgccatgatt ataagtttcc tgaggcctcc ccagccatgc agaacagtgt
 50220
 caattaaacc ccttccttat acattacca gtctcagcta gttctttata gcagtgtgaa
 50280
 aacagactaa cacactctta aaagtaaag ttattacca ctttttacag attaagaaac
 50340
 ttaatattac atcaaggaat tagtttatag cactggatga tctggcctca aatctccagt
 50400
 gtgatcagta tcaagctttt ctactgtat caacagagtt ttcctaacaa atcataagac
 50460
 acaagtagtt tctccaactt aatctattat acccataatt cacttgctc cagggaacaa
 50520
 aattagatga aacagactgt agtggtttg tgtaaacag cagctccatt cctttagccc
 50580
 tattctaaaa caatagacca tgtgtgata gctagctaac ctccattcag tgcttacta
 50640
 tgtgtcaggc atttgttcca aacacatctt acatggatta tccatttaa tctttacact
 50700
 aaagtataat actcttatta cccatttta aagaagagaa aactgaggca cagtgagatt
 50760

aaggaattca tccagtttct gctgcaagta gggtagctag cattgtaatc tgggcatttg
50820

ccatgaagct gtgcccttaa ccactccact atgtgctata aggtggagct gacttcatgg
50880

gggctaattc aaccagggga gtgtgctcat attaacagat ctgtgaactc cctggtttgt
50940

ttgtgagaca attatctatt tgcttacctg aatcacagca ataatagtta tacctccac
51000

ttactgcatg cctacttggg agcatgtacc tggacagaag gttttcattt ttatgacaaa
51060

cctgtaaaaa atggcttctt agcctgggtt ttcagatgag gagtttgatt ctccatccaa
51120

agtcatatga taaagaagtg gtgaacagg aattccaacc caagtttggt ttgttttgtt
51180

tctgaaatcc atagactgga gtagaatctc aaattatatg acatttgtct atgcttattt
51240

tttagtttcc aaggatgatg attatttaaa aaccaaacga ttttgaatgt aaaacaagct
51300

taattcaatt tacttataaa caatgagcac tgactcttgt atacatttaa gttatgagac
51360

actccaatat taattcgacc cacacttttc tttctataat ctattgctcc caaatattct
51420

atactatctg catgtctgcc ttagatacat aaggatccag ccctttccca tgtgtctctc
51480

aggatggcac ctttgctctg ggccgatgc catgggggtc tcagtgatgc gcactgaata
51540

atgcaaatat gaggacataa tcacagaagg aagcccttga aattctaaag ataaggccca
51600

cagaccaggg tactttgaat aagtcagtca tccttttcca aggttatact ttttcccct
51660

aaaataataa ctgcagacac agttctcaaa gcaaacaggc ctggtatagt agaggcattt
51720

ttagacaatt tttttctatc tacttggtta gcacatttta aaatactgat tgccaataaa
51780

gggctttcca tagaactcaa aattccaagt ctctgtccca agcttctggg agcattcctc
51840

catttttctg tttttatttt ttgagatgga gcctcgctct tgctcaggct ggactgcagt
51900

ggagcgatct cggctcactg caacctctgc ctcttaggtt caagagattc tcccacctca
51960

gcctcctgag tagctgggat tacaggtgca tgccaccatg ctcggtaat ttttgtattt
52020

ttagtagaga tggggtttca ctacgtaggt caggctggtc tagaactcct gacctcaagt
52080

gatccactca ccttggcctc ccaaagtgtc gggattacag gcgtgagcca ccccggtgg
52140

tgcattcct ctatttagtg ttaataacct aagactcccc agctgaccac caaatgtagg
52200

caggtggagg gactgacctc cagggggcac atgtctcagg ttgacaaaca cttttctct
52260

tgatcctttt tttcagtatg atgaagtatt ctacaagact ctcatgaac cttctgatat
52320

cacattctaa ccatttatat aaaagtcaaa agctccttgt cttcaggatt ccaaagttct
52380

tagacaaaag gtccaggaag tgtctgaatt ctttgggtgc tggttcagta tggatgttcg
52440

ggtccctctt agacttgaga tagaacagga aagagatgaa tagacaatga cagaggtaga
52500

gaaagtacca acccttccca ggctgtcggc ttatttgact gtgggaatca ataaaaagta
52560

tttagtgact aaaatagaat ttgaaccag ttcagatctg taagctaaca ttttttaggc
52620

aacaccatac agcctcttca tcattccctg gcctctttca tcctatccct tttctcccc
52680

cttcctcttt ctctccttta tttctcctta atcatcttgc ttctctttct tatatcttct
52740

cagtaagtaa ataatccaca acctcacagg taaattaaag ttgccattcg cacatctgtt
52800

ggctcccagg tatggttcca ggatactggg aatgttaggt agtgagaaag attcagatga
52860

tattaatttc ctgcaaatgg aatattcatg gcaggcacag tctggccaat tgctcatggc
52920

ctggcgagta cgactagcct cgattgtacc tggctcttgc gaatgtggtg gcatgttcac
52980

tggcatcaca tactccctct catctcctga atggctgacc agtgcttgct tcaagttaga
53040

ttagagtttg agctttgtaa cataagtccc cttctgtctg atcaatgaaa tgagggatag
53100

gtacttgccc gagagttatg ggacctggca tttaacgtga gaacatgcat cttcagaagc
53160

actggtgaga cttcagaatt cctgaaagca ctattcttgc aaggcttgaa tgtaccttct
53220

ccactccacc atgtcctctc ccaaccctac tagcatgata tcaagcttgg gtggagagct
53280

cgcttgcaac tttgtcagat acccatgagc cactctttga atttctgaca atgaacaagt
53340

atgtgggagg ccaaccccat ccaacctgac ccatctgaac ccaatcacat gttttctatt
53400

ttgaaataaa caggcagggt ctcaggaatt ttaagtgagt cctctatata tttctctatc
53460

ttccttttct tccactaatt attacagaat tgatctgccg gtaattacta ctaaattgaa
53520

tttcaggggc agtgctgtat ctggcccaa taccctaagt ccagtaatag cctggaatgt
53580

tgctgagttg ttaccgtcct taaagggtga agccacaatg atactttcta taacttctat
53640

ttggtctttt taagcccagc acttagaagg attcacttgt ggagacctt tcatttcttt
53700

caactctaca gcttagtgcc tgaaaggcac atattgggtct ctgtgaatct ctacgttaag
53760

aatctgattt gcagcactgg aaaccttgct agtcttagct ctgggctgaa tcccatgaag
53820

ccacgccatg cctagggctt tcatgctgta gtccatacct ctgaagaaaa cggagagaaa
53880

catcttctc ctttcttccc cccaagtgtg tgtctgttct cgtcagattc ttggcagggt
53940

taccaactct cttctaagca acttggtatc acaacagaat gattcctttt caatgccaat
54000

ctaaacaaat gccatcaact cttatcccag ggagtatgaa aaacagaaga gtcctttgtg
54060

actctgattg gctttgtgac ttccttgggg ctggtgttgg tctcaggctt gtgtctcgca
54120

gtcatccaaa atgtgggtgg gccatgtccc accttaactg tggttgtgaa cgcagaggat
54180

tcactcaact ttcagcatgc aacaataatc aaatgcagtt ttatggttca cagagggctc
54240

caccttcattg atttcatttg atcctcctaa caacctggg agctcagaga gactaaggaa
54300

cttgtctcat atcacatacc caggagtga ctgcacggca tetgatgcca gcttctgtac
54360

cacgtgtcct cctggtgata ctagagtaaa taaaacataa ctagacaagt aaatattttg
54420

tacaaagtaa tctgctaaaa atgtggatgg catttttaggt tatttggttg tttgggttat
54480

gcctcttctt attttgtaga taactgaaag ctaagagcat tgtcttattt atttatttat
54540

ttatttaggc agagtctcac tctctcacc aggctggagc gcagtgggtgc gatctcagct
54600

cactgcaacc tccacctccc aggttcaagg attctcctgt ctcagcctcc caggtagctg
54660

ggattacagg ccctgccac cacacctggc taatttttct attttttagta gagacaggg
54720

ttcaccatgt tggccaggct ggtctcgaa tcccacctc aggtgatcca cctgcttcgg
54780

cctcccaaag tgctgggatt acagcacttt gcaatcccat gcgtgagcca cgcacccgg
54840

cccattgtct gtcttcataa gcagaaggct gtggcacagg gtaggagaag aacatgaaga
54900

caaagttgag gggagttgct tatctcatgc ttcttgctgc tctgcctgaa gcaaaaatgc
54960

tttatggata cttcacttat gatttaaaat aatagctcac attgtgtact tactgtgtgc
55020

caggaactgt gctaattctt tatattatgc ccataaaacc tgcttctgaa gcattggggg
55080

gtttgcactt gccgttcct ctgtctggaa tgccctcccta tgccctacct catgaacctc
55140

ttccctttac tttcccatc tctgttcagg tgtcacattt ttagttattc tttcctgacc
55200

gctcagtata aagtagctca cacctctacg ttatcacact ccctcttacc tgctctattt
55260

gtctttatag ccctgctcac aattatagtg tacatggatt tcttaatgtg tgctttatct
55320

tccccctccc atcggaatat aaactccttg agagcaagga tgatattttc ccctgctgtt
55380

ttctcagtac caagaagagt gcctggcacg tagtagctgt tcaacaaata tttgctgtgt
55440

aaattatctc attttataga taaggaatct gaggttctaa gggagtaaatt cattgacttc
55500

aggggacaca acaggcaaga ggtggaacac agtcatgagc taaggcagtc ccagatcctg
55560

cctgaccctt cactctcttc ctcatcagtg gaatcgtaag gcacttgggc tgtttgagca
55620

gtgcccttat ccataccctg cccagtcagg gcctcccaga accaggtaaa aggatctggc
55680

ccctctgcc atcctcccac tggcgccctt cccctcccac ccacgcaac ccccatctag
55740

ggagaggaaa ctctttttcc acctaaattc cccctctgtc tagttcatcc aagtcatttc
55800

ccttcccacc aactgttta cacttgtggg tctgtgttga tgcatatcac agccccctta
55860

gtcattgttt agccattata ctttgagggt gtaattttcc atccctctgt ttatatTTTT
55920

gcttctcctt catgtgtcc cctcccact tcttaagggt aaaggaagac atgttcagtc
55980

atacaagcaa taatagagtt ataattaaga tgggaacatt ttgcaaaaag gaggatttag
56040

ttgtgatttg ttgaatgcca tctgagtg atttttaag agcaaattca agctctttca
56100

tgagacttcc tgaacagtgc agtcctctga gcctgctgtc tctatagtct ttgtgtcagc
56160

tttgtggggt caggattgct cgtgaaggat acatgttttag caaggcctgg ttgtgaaaa
56220

agcttaaaaag gaaacttcag ggactatcta ggaagatctt ctgtgatctg ttctctgatt
56280

ttgtgtgtc ttttgtggtt agccattgct cttggtaccg tgtgggactt agcacctgt
56340

gcctcaagggt aagaacacac ttccagccca gcttcagttt aatgttcact atgagagtct
56400

ccagtaacaa cccacactct gcgtgggatt cactgccttt cccatgtcct ggtgttgagc
56460

atacatactg gacagagtta ctgccttttc ccctagtaat taggggaaaa ggaccacatt
56520

ttcccctagt aataggggaa aagggtgtcca gtatcagagc taggtagagc tgaatttggg
56580

tatgaacttg agctgtacc catggtgtct tttcttgag tctagcctc cagcgcttt
56640

agaatttgca ttctgtttgg tgggatgagg acggtggcag tggagacagg gtgagaggag
56700

catggatggg tctaagcct tctgtcttc agcagcatct cgtgctctgc ggctctgctg
56760

tgtgcctcag gggctgcccc ttttcaggtt gcatttatta gcttcctgga tgctctggct
 56820

tccagtaggg tttagcctgt gggaggcact ggatggagat gggagggagg gagagaggac
 56880

aagagaagtc atgttatttc tcctttgcct tcccagcctg ggaggtttct ccagtagcag
 56940

ctgcctctct tcttggaatt cagacccac tagacagggtg tgctgtggct ccagcttctg
 57000

tctgggtgtg ctggcccctg gttcacctga ctccagctcc tcctcctcca tgcagctggg
 57060

tcctgtgtgt gcctccctct gctctgcctt gctgtccctc gtcaggcctc tcagctcctc
 57120

catcccctgg gttaccaact cgctatactg aaatccttct gttttagaaa ctcatatgtg
 57180

attttcattc ttcttggttg gacactgact gatattgagg gatttgtctt agaagtaaag
 57240

ctgtgtgagc attttctctt ggtaagttct cattttccct taatcattag aagacacggc
 57300

acagtaaacg cgagctcatg agcagaaggc tacatagttg tatacatgtg tgtgcttttt
 57360

ttcacaccag agtttgatga agcaaaatac ccaaaggctt tttaaaagta aagaggtttt
 57420

aagagtaaag aaatttaaga aaaaaaacia aaagaagtaa agatgggttag gtctcctgtt
 57480

tgctctcctt ttgcaatcag tctgtctcgc tctgtgagct gactatgctg cagacttttc
 57540

ttttacctcc agcctttcgc ttgtaggggt tgaggctagg ggcggggaga gagagttcag
 57600

cccagcaggc agcaaaccct atgtgcaaac acgatgtgtg aacattgtgt gcttggggga
 57660

actgggatta tttagctgag aagacacgaa atctcctccc cttatttggg ctttgggatc
 57720

tgagttcatt ggaagctgaa atgtaagaga aagaagaaag ggggaaggag tgggacaaag
 57780

gagaggggag tgggagatca gactgggcac attctccctc acagagggaa gcgatttctt
 57840

tgttaggaaa gcagttctgt tttgagaccc ggatgccttg tttatgctca ggtgtgaagc
 57900

tggcttagga attgagttta gcttgctctc gtgcttttat gaatgtgttg gctgctcagc
 57960

aaatttagta cagaaacatt ttatgtaagc agaagagaaa attatatattt aaaagttgtt
 58020

ttaacagaag gtcagagtat gggattgaga attaaaaaaa aaaaaaaagg aaattctgtg
 58080

ttctaggctt taagacagag agtcacacac actaatttag accagctggg gtgatctcac
 58140

agggactcca tgggccagag ctgccgttcc tccatggaga gcacatttcc cctagaaatg
 58200

tgtactcagg tgcagctgaa gggctctgagg tcatggctgc atccctggcc agggctcctcc
 58260

caggtctcag cctctgcttc ctctgttgtc cattgggtcta gagccatcct tcgagaccag
 58320

agccagacca tgagcgagcc aaatgtgcat gaggacagca tgaattatcc ttggagacca
 58380

gggggccctg ttcattccaga cattgtcttt ctttttaaag aagctaaatg cctctttttc
 58440

cctctttctt cttgcatgac ctctgggagc ttgggcagta aatttctctg ggacagggtg
 58500

tgaggaacct cttcaaagaa gcagagatga gacaaactgt ttatcagttc actctctacc
 58560

cctgctgccc tcccaatgct cagggtggcc ctagattcat ccagaagcca ggggtgtgtg
 58620

cttcatgaaa gggccaatac ccctaattgtg ggacctatac aataaaaggc aagggcagaa
 58680

agagagaaga aataaaagga aaaggggtgt cttgagggga gtaacatgata gactcctgtg
 58740

cctggaaggg actaggaaac ttatactgac tggttcactg catttgctca aagagtattc
 58800

atcttcttga gtttgtcagg tgactgtaag agctgggttc aagttctttg gttcactatt
 58860

atatgcctag ctgagcacia tgccctgggac ttagaggga ggtctgggac ttcaccagga
 58920

ttcccaaat agaacaagc atttgttatt tgcaaaaggc tcacaaagcc tgaaaaaata
 58980

tgtgttgtct ctggtggaaa agtaaggga tatagcttaa gtacctacta cgctctgact
 59040

ctttggattt cacctgcagg gcttttagag gtggataatc agggccatct ttttcggggg
 59100

tggggggcga aaacaaccac tgaagaggtg aagaatcttg tagttctaaa aggatcagag
 59160

agaggcttaa aacaactttt gcaaactgaa atatgccagc cagcttaaat tagcaaagct
59220

ggaactgtgg cagaatgaat agcccgtgtc tgagtgagag ggaaggaaag ggtctactgt
59280

gtgataggca atggttccaa agtaggagta ggggcctagt gtggtcatgc cttctgatat
59340

gtcaagaaaa ggcaaaaagt tgatatttat ttgaaatatt ctgattttat aggttggtga
59400

tcaattaaaa aactaacaaa tacagttaac aaacatgttt taatgcttgc agacaggatc
59460

cttcctgccg atgtttggcc tttgtcttta atggaatttg ggctgcccc taacagcccc
59520

ccctaagcaa aacttctttt ctgcctttct tttatattgt tttttatgga atgtagaggg
59580

atttttctgg aatttttgga acggggaaat tagtgacaat ggtggtgctg agattcctgt
59640

ttatttcttg gcttttcccc aattccctga gggaactcat ttcttttgga tcttccataa
59700

catcggtcac atagttgggc acagagaggg cccacacttg gtacatgttg atgaaataaa
59760

tgtgtcttcc tctaagtcta gctgaatgga gaactgaaat attaggggtg gtgaaagtta
59820

cctaaattgt ttgtgccaa attttgctgt gttctttgtc tccagttctg tgaggtgtgt
59880

gttagatacc tattgctgta taacaaatca cccccaaaac atatctgttt aaagtaacaa
59940

acatttatta tagcacacat ctgtaggtca ggaattcaag tgctcttag ctggtactct
60000

catggggctg taaaggaaat atcagttggt gctgcagtct catctgaagg atccactcca
60060

aattcattca tgtggttatt ggcaggattc ggttcctcac cctatatattg actgagggcc
60120

tcagttcttc actgattgtt tgaccggagc tctccttcca ttctttgcca tagggcctct
60180

ccatagggca gccccgaaca tgggatctgg tttccattag agcaaacaag caaaggaaag
60240

cagtgaaga cagaagctgc aatgttttgg taacctaatc tcaggagtga cctcctgtc
60300

actgcttcta tgctgtattc tagtcattaa aggagaccat ctcacactca aggggagaga
60360

ctacacaagc catagtgctg ccatggcaac acaaccatta ggatgatgcc ttgagccttc
60420

cttcttcata tttcttagac ttcctttctg acctagagga gccagaagga cttgggactt
60480

gtcataccgt ccaatcccaa gaaatgacag gaatgaagga atttttaaaa attcaactca
60540

tttcaacaca cacacaaaaa atgagagaag tgagtcaatt tgagaactga aattctggta
60600

gacttttggg gagacactag aagggccatg ttcattgagta ggaaccatag attctgagtc
60660

aattgggcat gaaagttttt ctcttggtgt cccgatcttg gtcgtatagt gaccaagcat
60720

ccaagaaggg gataacacaa tgttcttaca gtgttttctc tgttgattta ataactcagg
60780

catcagccag tggaaatggc aaagctaagt gaaggtttga atgccagaca gattctccat
60840

tggctaaacc caatagacat gaataaggaa taaggccagg aatcagcatc cagatatatt
60900

aatttatgaa tctgaggtta ttccaaaaag caagagaaca ggagcaccat tggctgagac
60960

ggacaccaat accaagctct agagagtaat ttcaggcaga aaggaaacct taattaccag
61020

aggttcaagg gataggcaca cagtaagtcc aaataggcag aaagaaggca tcagggataa
61080

tctggcagag actgcattca aggtcaaggg caggccatcc cagctgctgt tgagactggg
61140

ctccaaggaa gtctccctgg agagactgaa aaacaatcct gggccttaat cactgaggac
61200

tgaacacaa cataggagaa aactagggct taagagaggc ttggtaaga aataggcaag
61260

aacctgatta ttgggaaca agagatccac agtcaggctg gaccagtagc aagacagcaa
61320

gagtagacaa aaagttgacc actgccagaa accaagattg gtctctgagt ctggcatggg
61380

gctgaaacta cttgttgaga tcaggctgtc tcacttgcca ggggtagagt catgctggaa
61440

gtggagccag gacagacatt gaaccctggc taccctgtga tcaattgcac tccataacct
61500

ttgatactc atgggaaaat cagacaagtg ctgttctgaa aggcaattag gtttcatgat
61560

ttcctcccc atccggacat cagcatgaaa ttgggggctt ggcttccaag cccatgttct
61620

tcctttttgt gaagggcagg gtgagacaaa ggacaccaca ggccagagca accctctgct
61680

tcctacacagg ccctagaaga gaagtgagga cgcgggtgac ccacacggcc ttctgtgcct
61740

cttcaacagc tactctccca gcctccatt tctgccaagc cacattctcc acaggactca
61800

aacatttagc tttactttag acttttccag gaatcttccc aaatctcctt tcctcatccc
61860

ccctcacctc ttctcacca cccacagccc cctctccac ccccatgtac aacatccggc
61920

ttttactcat gcttgtttct aattatgcat actaatttat tttctttcct tttttcctga
61980

cctgcctata ttttctact cattctttgg cctccactac cgaaaaatgg aatactctc
62040

ctgtacaggg tcagggtttt tccttcccca gtgcctaaag ctggattcag tccacagtga
62100

cttctctaataacacagagtgt aagggactcc ttggttgaaa agaaggaaaa accccaactt
62160

gcctttcagg gtctcttcaa tgagcttgat ttgtgttaaa tttaggggct caggaatttc
62220

tgccgatttg cagcacttac tactctacc cttcccaaca tttgtgcttt atcgaaacaa
62280

ctaaattagg agcagcttga gggtaggggt cagagcacag acattttctg tattgcctgt
62340

agaaactagc acagtacagg caccctctaa gtgcaaaaga agtccttgct gaccagatg
62400

gaattttcct gaagtgagt ctttgtgtct atcttttagaa ctagactttc ctgataaaat
62460

agtcaggaga tttggctgaa ttttcagta ttcaactggc tatgttggg tcatgtagat
62520

acaaccttta agaacataat gacctttaa atatctattc ctgaagtcag gtttggctcc
62580

tagtctgtgg gcagatgact gttttcatgg tatgctttgt tcgtgtcctt cctgggtggg
62640

ggaggatttg gggggacagt ttgactctag ttatctacaa agttgagtta tgagaccagc
62700

ccttttcaac tttacctctt gcaaattgat cctgtaataa attgacatga cgaaaacacc
62760

aaagcaacca gagaaatctc acaaaggcaa gaacaccatc acacaagtcc tcaaaatctt
62820

tccaagatc cactaggggt taatgccact gatctgatac tggggcacag cctaggaggt
62880

tttgcctgtc ctgaacgcac atgtctccag cttacagctc acaggccagc tttgtttttg
62940

aacagattgt gttgggggaa gggaggcttc tgagacagag ctgagagtcc ccagcaatgc
63000

cacccctggc ttctggccca tctgaatgcc ttactcttcc atgtaagcaa gacatgaggt
63060

cctgctgttt gagttggctt tgccctgctc cgtgatttca gttccttgct tctccagcgt
63120

gtgggctcat ttccgtctc tgagcactat tgtttttctg acttgaacag tcatccagaa
63180

aacctcagta ataaggcttt tgtatcatta tttactttgc taattatata tgtctttccc
63240

atccccagtg cttgaagaca tcaaattgca gttgaagtaa cggcagaggg aaaagcaagg
63300

gcaacaacaa caaaccactt cactcctctg ggtggaaaca tgcaattgta ccaagaggga
63360

taaggaaaaa actgctcatg tttgcagaaa gaagttggaa taagggcgtg aaagcaggcg
63420

ggctacttta gcatctgatg ccagggatat gagagagtag cccagtgtgt tcattcaaaa
63480

agtaaggctc tgaggccagg accaggaaag atcctggagt cccagtcac aatggcatga
63540

agcaggatga tgcagagagg ctaagggggt gattcagtgg gttctttacc tataaaaatg
63600

atggaaacc caacatcagc tgggcaagtg gccctgttca gagctcctgg actcgccatg
63660

actaaaacac tctgttttat ttcgtagttg atatctattg ttcttgccct ccagcaagca
63720

atggaacagc acctttactt tcctttgaga aactattcct tccccattct cagggttgt
63780

gaggatcaag tgaggtcaca taatccagat ctggccaatc agcccatctc atgctactgg
63840

ctacactaat tggtttaggg atagtaaagt actcaaattt gtttactgaa catttagccc
63900

tggaattttg gtgatgggag gtgtagaggg taaactcttg atgatcatcat ttgtgtacct
63960

ggggccagcc atgcctgaac ttttcaatta cttgaccag taaatttctt tattcattta
 64020
 aatgttagag ttgagttttg tctottataa ctgaaagagt gacatgtgtt tatattaagt
 64080
 gcaggctact aatatagtct attggtttcc ccaggcttgt tctccttttc tcaaatagga
 64140
 ctgatcagat agtaagtgtc ccatacctgt aggatacata aaagagcaaa tgaacaagga
 64200
 gcaggagata tgggtggctga cggttggact ataatttttt cttgattttt ctgactttat
 64260
 tttggctgtg atctaggaag tgaaatccag gtccttctaa accccaattt aactgtggct
 64320
 cctgagggag attttcagat ctcaaagat ctgtaagttt aaggaggat gattgggtccc
 64380
 tgtacactct ctgtgagagt tcctttggaa tactggtttc agaatgtaca gccacacagg
 64440
 taactggatt agggacggcc agtggttttt tatgaatgct attgaagaca ctgaatggaa
 64500
 gcaattctca ccagaactct ttttctgtc atctttcagg ggctaagttg acactagaac
 64560
 caatcaatga tccctgaact gggagaaatg ggtggccttg gctgtgattt ttggttgtgc
 64620
 tgcacatttc aagatgacaa tattgtttta aattaacctg ttagagtga cttttcctaa
 64680
 agaactttac cttgatgtct gagtaagggc agtgcaggag gtccccaac ttaggggact
 64740
 gatcatgttc caaaggctcat ttctcagtct ggatgttttc atgagataat gaattccaag
 64800
 atattgtaag gttcaatatt ataaatgggt cttagtttac taagataatg cattgcctct
 64860
 ttgtgaagcc ttctctgggt accctaataa aaagtaactt ttctttcttc ttatatgtca
 64920
 cacagtatta aatgagttgt tttttaaaaa catcttaaaa tcttatactt aagcaattat
 64980
 gtagattatc ttggagtgtt gagcttattg ctctcattag atgttagtaa atggaaactg
 65040
 agtcacaaat ttccacactg ttctagttct ttttcattct gcttcatttt ctgtgtgtc
 65100
 ctgccagggtg tctgtgctaa ctcatcttgc ctgaggaccc ctaccccatg tctcactggg
 65160

ctgcttggt tacttetgaa gccagcacc acctccctc agtggctctt aacaataaac
 65220
 cctgacacag aagttctggc tggccacct attcctttga gtcaggtcac aggtttccct
 65280
 ccagctaaact gatggatgga ctgtccttgg ctcaggggtc cactcttggt caattagata
 65340
 cattttgggg gcagggtcag aattcataaa atgcaggcat ttgaagaaca gggaatcaag
 65400
 gtagagagtt cctagaagac agagcaggga ggtaaaatga aatattctag taaagtatta
 65460
 atatttggtg caggatatgg tcttttttct ttctttcttt tcttttgaga tggagtctta
 65520
 ctctgtcacc caggctggaa tgccctggca taatcttggc tcaactgcagc ctccgcctcc
 65580
 cggtttcaag cgattctcat gcctcagcct ccctagtagc tggattaca gatgtccgc
 65640
 actacacct gctaattttt gtatttttag tagagacggg gtttcacat ataggccagg
 65700
 ctggtctoga actcctgacc tcaagtgatc ccccgcttg gatatggtct taactagaaa
 65760
 tgaagagatc tgagttcaaa ttcagcattt gtcacatgta ccctctctgg tccaaggaa
 65820
 atcaagtaat atctccaatc ttcagatctt tcacatgaag aataaagagt attgatcacc
 65880
 atatctctac acctccttcc agctctaaga cctccaggac tcacctagaa ccaaattcat
 65940
 cccacacag tggcggcagg gtcgtgttgt catgatgatg gcctctttcg gtgaagagag
 66000
 gttgtcagac acctggagca atgccagaga cattatcctt ccctaggaga gctactgata
 66060
 agaaaccact ggcagctcct cattcacttc tccatcattc actcttttct cctttctttt
 66120
 tccttagagg cataaacaag gtgtagatca tcctgttata cataatgaag gaagatttta
 66180
 ggatattcca gcttagatcc aggtgtcccc agcttctata ctactctga agaagtgtac
 66240
 tttatctact ctttttacac catttatccc attttgcca tttctgaaag gaagatccta
 66300
 gaagccaggc ctcttggaag aagttggagc tacagcattg ataggaaaac ctgacaccta
 66360

gaaaggcaag tcagttcttc ccatgcagtg gctttgggga ctgtaattta acaaagctct
66420

tttttacttg tgtaaagtga tctccattcc acttttttct tgcactgatt ctaattaagt
66480

ggaagcaatt gttatcatca tttaggatta tagttaaaat acccctcaaa ttgtttccat
66540

ttggaatggg accttgctcc attctgctct tgctgaagga agtcttcatg gaatcacagg
66600

actcaacaaa cttttctcat taaaaattct ctcttttagg caacagtggc atagaaatcc
66660

ttatgttcta ctgtctgacc cgtatgagaa cagcgcaatt ctcaatttat aaaagaggca
66720

ttttccctcc tcttgggaaa aatattttctc ttggtcactt gagtctagca tctaagtcac
66780

ataatgctat tatagccttg ttttgctttt gatttaggtc atatcttctt tttctgatct
66840

ttcttaaagt gaagggaacca tttgctatgc tcacaaaaat tctcaaggct ggaactacag
66900

catgctctgg tttctaagcc atacacatcc aggtagaagg gtcaatttaa ttgccaatga
66960

tctcagcctt ttctaacact gttccttaag tttctttggc actgcatggc aagcaactga
67020

gtcatgcatg aaaatttctc ttctagacaa cccataccta cagagggagg atttggttta
67080

ctgcagatgc tgagtgtctc tgatggcaca cattccttct tattaatatc cccctggccc
67140

ataatgggtt gtgtggacaa cagctacttc cttcaagggt cctttgcttg gtaaatatat
67200

gcctccagga gctttctggg tagaagcagt cacaaatcac aggggtatct atgaatggca
67260

cggcccatgg aagggtcac actgtcctga tttttctacc ataagagcta aaataaggga
67320

ataggtcttc aggaataag tcagatgggt tgtctgacta agcacagggg cagggattgg
67380

aggtgagatg acagagaaca gtcagcattg tctccgttc cgctccaccc tgcagttctg
67440

agctgaacag ataatttatc tgtggcacia caaagtatat gacttgaag aatgttaaag
67500

atcaagtcca ccttctcctt gtattatatg gctctgagag ccagagaggg gaggtgacct
67560

tcccagaatg tgaatgtggt agagcctagc ttacaattca tgtccaactt aggtgtgaag
 67620
 ttctatgaga ctatgctgcc ctctgcctca caatgctttc tcccatggga gaggaggctc
 67680
 agagggtacaa agcagctgca gaaagactct ctcaagctat tgactcttgg atgacaacaa
 67740
 ttaatgttca acttggcttt accocccaaca tgcactgcac agatgcctgc aggtctacaa
 67800
 tgccgagacc aagtttgggg ggcatttccg cagctctgtc tattgtcctc ctttcctctt
 67860
 ctgtggcat tcttcccaag ggataggcct agtccagggt ggagaaaaac aattcttggg
 67920
 ggtaggggca ttggaattgt atgggcaggg ttctaggcca actttggcct ttcacaattc
 67980
 tgctaagcat tgaacagaat tttcaccta atttcacctt cccttccatc ccttcctaa
 68040
 ttttgctagt gtgtatggag ttgatttgat gagacctgt cttttttttt ttctcccttg
 68100
 taacgtgggc cctatcttca agcctatctt catagccttt cttttccaga ccttgccaaa
 68160
 aagcagtaga agagagggtta caaatataa gaggaaacct taaaacaaac agggtagaac
 68220
 tggagggaat cttaaataat gcctaacca actccatatt ccactgatag gaaactgaga
 68280
 tcttgagaca ccatccaaat tctgatgtaa tcagcagaac ttaatatgcc actttgagca
 68340
 cagaaccaag tacttttagag gtaagagagg ctatttgttt ccagcacagg ctgaaacca
 68400
 catgagaaa ttggacttta agtgaccata gtcacattct gataaattct aaaatgtctc
 68460
 tctcattttc tctgacatca tattgtcatt tttcatacca caaatgtgtt acagtgaaaa
 68520
 ccaggtagat gtcaatagat agtaccaaga aaaagaaaag ttatcattct ttaatgctct
 68580
 ttagtttgaa aactttatctt atttctttat tttgatggag gagtagagtg ggtcaagcta
 68640
 agagttgcct tataaacctg ttgttctcag caaatgagtc agctgagata aaaacattag
 68700
 tttccttggc gtgtcagcat ctctcccaaa ggagggtatt agttcatggg ttggcattgc
 68760

tatcacaaaag aaggctctgg cagcacctga ttggcttggc caagcttttc ccagttggcg
68820

tctgggcatc ttcctagcag tgggtcactc acccagcagt gatttggcct cctgttgaca
68880

taaatatgcc taattcaaag cagaaagtca atttgttatt ttgcctaaag aaatgtaact
68940

accaagcagg ctagattttct gtcttggcat aaccaggaaa tctttaccaa taatttatga
69000

tgtggtcacc cttgttgtag gcattatgga agatggaaag atatataagg aatggcgtct
69060

tgcaaaaaga actggacttt tttttttttt tttttttttg tagcaatgtg gagagtaaaa
69120

gccagggggt caaatggtaa ttatagaagg agattcattt aaagaatgtc ataaggaaat
69180

agaaaataaa ggaaacattg caggaaacca tagggggcta gctgattcat ggaaaatgtg
69240

aggttatggc tgggtctaag aataagtaaa tcattggtga ctatggttgt agaggttagg
69300

acagagaaaag ctggagcaac aagcaatcat ctgccagata agaccccggt ctgggtgaat
69360

gcaaggaggg agcggacaat tttgtcttcc aactctgtgg agctggagcc ttatcttagg
69420

agtcagtggt gcctggaggc ctgatcacca accagtgtaa tgttagaata agacttcct
69480

aagcaaatca atgtgatgat tgatacaaac agccatcact ggggcctcct ctatttccaa
69540

tctcttcctc agtttccag tcttttcaga acctcactgg gaagttcagg gatatatgcc
69600

tgcaggctgg cctggcgag atggatgatg aagggtcctg ccttgagtgt agggtagagt
69660

gattctgagt gctttccaac agactgtaaa ttaaagacac tccaaaaagc ccctgggtgg
69720

gatagatggg gagaaggggt atcactaagt ctcatcttt tcttctaggg ccagggtcag
69780

tggttcaa at aaacatggtg cccacagag atggcctgtg gccaaagtagc tctcagtggtg
69840

ttgactcaaa gtcacctgga caggcctggc ggaaacgcag gatggcccag ccctttccta
69900

atagggtttc cagagttgat ccatcattct gccagaaaa ggggagaata aaaaatagat
69960

cacgagcccc tatgactgtg acttgtgctt ccttggaat gcaaacaggt gaaggagaaa
70020

atgacacagc taggggtgta gggggcactg gaagtcaaca ctttctggg ttcaaatgag
70080

aagggtgagg tgctggcatg gaggaactg gcactggcgg tctccaggca ggactttctg
70140

ggcactgggc aaggccatgg ctttgggagc tgccccaagc tgctggtcgt gccagtgtt
70200

cattcacctt gtgaaatgta ttaattagt agtggtttt gctcagagag agttgtgtg
70260

gcttcagatg gatatagcag aaaaaaggca aatccatcta caggcgctc tgagaaatct
70320

cagctcttgt gagcctgaga agagcttcca tagagagaag gcgagacccc agcaaagttg
70380

gacggagaaa attgagacac ccacaccaat tagtatgagt ctatcaatcc agtgccaccg
70440

aaacacaagt ctaagcctaa cagaatctgt gctgcgaata agtaagattg ttggaacag
70500

aacaaaaaaa aactttgcaa atggtgcttc ttttaagctt tgcagttctc ttctctcact
70560

cccttttcaa aaatggctgt aaaaaaatt tttggtctgg tttagccaga gcgctcaga
70620

gctggtagat gccttcaaaa ctgtgatctc cttttaattg tcagagggta ggctgaggtc
70680

cagggaggta aacagacttt tctaagctc ttagctggc tactggcagg attgagggaa
70740

ggtggaacaa gcctcttcat tccaattca ggtttcttc ttccatccca catctcatgc
70800

aataaacatt tccaaatcta cctgccagc ggggctaaga gcagagtggc ctccagaata
70860

agtggaatag catctcagg actgctggc tgaacacgta ctccgaacat cacagctgct
70920

gccaagcggg cagacgcctt agccccagc cctctccctg ggctttctg ggtcttgtc
70980

acacttctca catcactctg ccttaaattt ttctccccct agaaagagat aacttgcctt
71040

tcattcatgc ctacctaaact cagttcaatt caagtacata aatcctgctg acaaccctgg
71100

cagaagcttt atgagtaggg agtaaagatt gtgagaaagg cagggaatta atttataagg
71160

aaggaacaat gggagaggag agaaacaaag agaaggatgc aaagcaatta atgttaattg
71220

atgcttaata tatgccaggc ccagtcaccc tagttatttc atttaatctt taaaataatc
71280

cttgtgttgt catccatgcg ttttcagttg cagttccatg atttaagaaa agcacacagc
71340

cagaaagtgg cagagaaaaa atcatatcct gagtgcacgc tttaccgggc cagtgccttc
71400

cattctaaca aatagtagaa agagggaag agagaagaat cccttgggtga taggaataag
71460

gtagagagag gctgtgaaag agagaaacag ggctttgggg gatcctgggt tctgatgctg
71520

acctctgtgt catttctgta tccattgtat gaatgtttat ggaaagtcag ctacgaagca
71580

aattgttcta ggcagggggg tttcagcaga gcttgggaga gcagatgggt ggaggagtgc
71640

atccatagga tagcctgggc tgtgggaata agagaacagc aagtcagggt gagtgcctga
71700

aagacaatat aagagagcga caggcctcag agaacctggg aaaaggaatg aatgggggtca
71760

ccttaggaag ttcattagga ggggcctctc tgaagagggt acctttgagt cgagacctga
71820

ggcactggag ttacaggaaa tgtgatgttg cacgtggaga atttcttgga aatggagcag
71880

caatgtagtc tgacctatt gttgggccc ttaacttggg caatgtctca ccaagggtga
71940

ggatgggtgac aggtcctcct gggatatggg ccatggcagg acgaggggac aatgtctgta
72000

aaggatttaa aaagagtcac caattctaac gatgtcaatg ataaaatact tctccatatg
72060

cagggggcca tttccatgac ccagctcata aggagtggag agagtgctat cttcctctag
72120

aggcagctcc aggaagcagg agctggtaac ctgggggctg ctctgaaacc ttcaaaagg
72180

caaatcatgt aggagtctct ggccactttc ttttctgtaa gagaagattc actcaaggtc
72240

aggggaccct aaggaccca gggaatcacc caaaaactcc agccagcccc tagaggaaga
72300

caatccacag agctggcccc atcccttcgg cctgtggcta aggcccatct gcccatggtt
72360

ccttcttgaa gtaattgctg taagaatttg tcattcacia aaggaagtct ctattgtctc
72420

ttttcggatg tcatgtaacc aatgcgtggt aaaggtcagt tgcacagtt ctcttaaatg
72480

gcaattaagc tattgtctga gcaaaccaga ggcacccgga gtgcaagctt cgctcaaacc
72540

tagggacctg ggaaaggctg ggaacagagg agcaagtagg acaccacagg ggagggccgt
72600

ggcttcagga aacgtgggga acaccctgaa ctgtaaaagc agatgcgtca gcagacaaaag
72660

ggggtcattt cactaaagcc tcagtgaagc agagaaattc cagactttgc ccatcaccta
72720

ctggcaggct ctccctccct ctgtccctc cccacctct cctcacctt gtctttctct
72780

gcttcattat ctacgctct ctctttctct cccgcagtt ctctgtctct cagtctcgat
72840

ttctcacatc ctgtttctct ttcttcttta ctctttctcc ctctccttca cctctcctgg
72900

cttcactcca ctccccccac ccactctggt ccctctcaca ttctcatttg gggatggggg
72960

gcggctgaca aaccggattg ggggggtctg acaaactgga tttaaatttc agctccgcag
73020

cgttctcagc tcttgggtctt ggaaatgttc cctgagcttc gggttttctc atctgaaaat
73080

gaaaaaaata atacctagct ccaagtattg agcttggtgt gaggaatcaa acagataacg
73140

tttctaaaat tccaaacaca gtgcctgggc cacggcagat gcttgactgg tgacagcttc
73200

ctttcccagt gagggccagg gaagtaggca cagtgccttg caccctgcaa gcatgcgaca
73260

cagcttcaca aaatacccac agggagccag gagcaccgct tgaattttat ggatgacaaa
73320

attgtgggtc aaggaaggta aaagtaccct ggtaggacac gtaattttta aaaagtctat
73380

agctattgat attactatat attaattccc attctgtccc tccggtttct cccctgaaca
73440

ctaatttcca aatccccaac actctatggt ttaaatgata aacgattcac ccaaatactg
73500

tataattttt ttttctctt tgggtcaggg ccatgctctg tcaccaggc tagagtgcag
73560

tggtacaatc gcagctaact gcagcctcaa cctcccaggc ttaagtgatc ctcccacttc
73620

acagcctccc aagtagttgg gactacaggc atgcacaacc acaactagct aatttttaaa
73680

atttttggca aagacagatc tcactacggt gccctggctg gtctcgaact catggactca
73740

aatgatcctt tctcctcagc ctcccaaagt gcttggatta cagatgtgag ccaccgcaca
73800

tgcccatgta caattcttga agaagctgaa actcatcttt ctttctcggt ataagaaaaa
73860

tcaggatatg tttggacatc cctcccatc tgccctccca gcctcctctt ctccaggctg
73920

aaccagatcc tgctgcatct ttctcaaaga ctctcttaca tctcattat ggtttctgac
73980

ccagtcctcc atgcttctct gaagtttaga gtcaaaatgc ccttttttgc tttaggggtg
74040

agaccttctc tagcaccgtg tgtctacaac acagagaggg caaacagaga tagtgcttgg
74100

tattttatct ttaaaatctt taatcttgat tatattttgc aatgtttata ccattgcaca
74160

acacacggaa gcattagtca caatttatat agctgatttc atggagctga attctatggt
74220

gtttatagcc gttgagtgtc ttctccaga aattggaaag atgccccatt tgctgtaaca
74280

cacacgacct ccacccatt cttttccttt gtaacccaaa ccacaatgtg taattgtgtg
74340

tttattttgt gtttaatgtt tatctcttcg ccaggtctct aaggctctat gagggcaggg
74400

atcacatctg tttagtctac ccagatttca ttatatatcc agcattctac actctgtaac
74460

tgcttagtct caactctgac cacacattaa gtcacaaagg gagttaacag aaaaaaaaaa
74520

aaaacaataa acaagcaaaa tgtttactaa tgttccacaa gctcttttga atcaactggt
74580

ccctgtggca cctggggatc cgtagtcatg aaaatcctgc cagggattct tgatccggga
74640

tctggtgatc ctgggcagga caactcacca ttctgttcgg ttgcccgctt gcctcctccc
74700

gctgggtcca gcttctgtgc ctacaatgaa ggctgtgccc attccctcac gtgccgcgcg
74760

tggcctgaca ccttcctggg cccaccctaa cacattcctc cgttgctttt aacgactgtt
 74820
 tacagctcgc tacctctagg acgtggcttc tgatgaaaac tgtccagacg ctctgcctcc
 74880
 ggtgtaattc agagcccagc tctcctcctt aaaattcttt tacgctattt tccacaagca
 74940
 tgtgcttctg ttggcagtgc cataagtgtg attgatcttt accataagtt ccttgaactt
 75000
 tgctgtgacc tggatgctaa ggtggcggga gctgcctcca tccccacaat ttgactggta
 75060
 caacccgaaa gtaggatgat agtctgttct ggtcaaacad taacgtgggg cagggacaat
 75120
 ggagagaggg tttggtcacg tggcatgctg cagactgccg gaggagctgg aaaggctgag
 75180
 tatgggaaag gggagggatt tcatggcctc ttcgggcttt gagtatttga agaaactgcc
 75240
 tatggaagac cagacctgtg ccctgcctcc caagaaacac acagagctta gacggtggaa
 75300
 gctatctggg gcaaattcag ctccctgtga ggaaggctct ctagtgtgta ggggtgtctt
 75360
 ggtaggtcag ggagggtttt ccagcatcag tcattttcat accaatctct cggtttttgc
 75420
 catattcctg taccaccctg gcttttattt acctaatctt ttttatttac cgaattttta
 75480
 acctaatatt tcctttaaaa caactttaaa ttgacacact ttgtttcatt cctgttttta
 75540
 aaaagtacct atggcaaaaag atatgaatag acacttctca aaagaagaca tgcaagccgc
 75600
 caacaaacat atggaaaaat gttcatcacc actaattatc agaaaaatgc aaaccaaact
 75660
 cacagtgaga tatcatctca caccagttag aatggccatt attaaaaagt caaaaccaac
 75720
 agatgttggg gaggctgcag agaaaagaga acacttatat attcttgaga atttaaatta
 75780
 gtccagccac tgtggaaaga agtttggaga tttctcaaag aacttaaaca gaactaccat
 75840
 ttgaccagc aatcccgtga ctgggtatgt acctaaaaga aaacaaattg ttcaacaaa
 75900
 aagacatatg cactcacatg ttcacatag cactattcac aatagcacgg ttcattggaat
 75960

caacgtaggT gccaccaac aatagattgg ataaagaaaa tgtggtacat atataccatg
76020

gaatactatg cagccatgaa aaagaaggaa atcacgtcct gtacagcaac atggatgcag
76080

agggaggcca ttatccaaag caaattaaca caggaacaga aaaccacata ccacatgttc
76140

tcacttataa gtgggaacta aacatcaggT actcgtggac ataaagatgg caacaatagg
76200

ccaggcacgg tgctgatgcc tgtaatcca gcacttgggg aggctgaggc ggggtggatca
76260

tatgaggTca ggagtttgac accagcctag ccaacatggT gaaagcccat ctctactaaa
76320

aatataaaaa ttagctgggt gtggtggcac acacctgaaa tcccagctac tcaggaggct
76380

gaggTgggag aatttcttga acctgggagg cagagattgc agtaagccaa gaaggTcca
76440

ctgcattcca gcctgggcaa cagagtgaga ctctgtctcc aaaaaaaaaa aaaaaaaaaa
76500

aaaaagatag caacaataga cactggtaac ataccaggT aacaacatac ccaatatatg
76560

catgatatac ccacacaaca taccaggta aaaaacctgc atttgtactg cctggTgggt
76620

gtggtttgtg aaccggtatg tgtggtctgt ggtatgtatg gatgtgtgta gtttgtgaac
76680

tggtgtgtgt ggcatatggT gtgtatgtgg gtaactggca taactggggT atggtaggta
76740

tgggtatgtg tggTatgtga actggtaagt gatatgtatg agtgcgatg gtgtgtgttg
76800

cgggtgggtg gtgaactggT gtgtgggtgT tggTatgcat cggTgcatgt gatgtgtgaa
76860

ctggTgtgtg tggTatggta tgcTgggtgT tgtgtagtgt gtgaactgat atgtgtggtt
76920

tgtgaactgg tatgaactgg tgtggTgtgt gatgtgtggT atgtatgggt gtgtatggtt
76980

tgtaaactgg tgcTgtggT gtgaactggT gtgtggTatg catggagatg tgtggTttgt
77040

gaactgggtg gaattgggtg gtggTgtgtg gtatatatgg atgtgtggta tgtgaactg
77100

tatgtgtggT gtggtttgga tgtggTgcat gtggTatgtg cagtatgtgg tgagagtgtg
77160

gcgtgtggtg tatgtggtgc atgtggtgtg tttggtatgt ggtatgcgta tgggtgtgtg
77220

ggtgtatggg gcatatgtgg tttgtgtgtg gtgtggtgtt tgtagagtgt ggtgcgcatg
77280

gtgtgtgtgg tatgtggtgt gtgtgtggca tgtggtgtat ggggtgtatg tggtttgtgt
77340

ggggtgtgtg ggatacgtgg ggtatatgtg gtggggtgtg tgggtgtctgt gatgtgtggt
77400

atgtatgggg gtatgtggtg tgtgaattgg tgtgtgtggt ttgtaaactg gtgcatgtgg
77460

tgtgaattgg tatgtggtgt gcatggaagt gtgtggtttt tgaactgggtg tgaactgctg
77520

tgtggtgtgt ggtatatatg gatgtgtggt atgtgaactt gtatgtgtgg tgtggtttgg
77580

atgtggtata tgtggtgtgt gcagtatatg gtgtgtttgt ggcgtgtggt gtatgtggtg
77640

catgtgtcgg gtatgtagtg tgtgtgtggc gtgtggtgtg tgtggtatag gggatatatg
77700

ggtttgtgtg ggggtgtatgg ggtgcgtgag gtgtgtgtgg tgtgatatgt ggtgtctgtg
77760

gtgtgtggtg tgtatggggg tatgtggtgt gtgaattggt gtgtgcacca tggatgtat
77820

gggtatgtgt ggtctgtgaa ctggtagggg tgctgtggtg gggagtgggtg tgtgtggtgt
77880

gtggcatgta tggatgtgtg tgggtgtgtgt cttgggagtg agggtgctta tcatgagcag
77940

gtgctgcaga agccctcagg gaagatgact tctactgggag ctgagatgac ctcactgggc
78000

gccctccttc agcagatgcg aggacaggcc cacaggccta aagtcaatta gaggcagatt
78060

tccaagcaga atctacttct cctgacctct gattcagtgc cttttcctga caattttato
78120

cctgcgaccc cagccttgag aatgcaaaaa actaccccca acttcagcct gaggtcatga
78180

aaccaaccag gagtgcaggc cgtctagggg aggtgcgagt gccacctac tgggcagcca
78240

ctctgagggc aaagtccacg ggagacacag cttcagccac caccggggct gcgacctgtc
78300

ctcaggatcc tcctcagccc cctccaggag tcctcagctt tctgctttta gcccacaaacc
78360

cagattcact ttcccaggcc attcacgctc tgaggcagcc cagtcgctt tgtccttctt
78420

ccttcccgaaggctctcagc ccttgggcag tatcagaatt tcactttgat ctccttttcc
78480

ctgggagcca gaatagctta gtctcttcaa aaggagagtt agaaagtgtg ccgttcatct
78540

aataaaaaatt aactcctgag agttatccaa tccaataca ggtcatttcg ctggaggcta
78600

ccccgactag gaaacacatt ttaggacatc agcagaggta aagatgagag caatttgctc
78660

aggcttttac ttgggggttaa ggatcactag gccaccagga ccagggatgc cagacaacct
78720

caggttttctt tccaagcttt ctccaaacta ggctcactt ttatcatttt cctaccctg
78780

ggctctatga ttttgtggtg ttgcttttgt gttttcaaaa aactgcatt cattcattag
78840

ctctgcactt tccctcatgc aatgcacaca gaactgtaaa gacaaatgat aaagcccttt
78900

cccgcttga tcttgtgttg ttggccaggt tcattcaaaa ggccattggt tgatggtgat
78960

taagtgtggt tgttttgaat attcacctaa acaaaaatga tcaatttaca caattcttgg
79020

aattcatggg tgtcttgga aaactgaag ccacaaaaaa tgttcaatac ataaatgcca
79080

aggacctcct gctgtcaatt aagacctca tgatagtctt ttgcttattt gctcgtaaat
79140

tcatgtaaaa gtggtttgtg agcaaattaa gcatcagttg gcatgccgga catggacaca
79200

tagatacagt ttgtggtctg gaggggaggt ggaaatgaga gtctcacagg ctgccctcgg
79260

cccagaactt ctgtgtggag gttttgcata aaaatcagaa tttctggatt cttataaaaa
79320

ccctaaagtt ctagcccat tagggccact gttccactca cagccctgag taagcctccc
79380

ccatcagccc aggcattggac ctttgtcct gcctctgttg tcactctccc gactgtctca
79440

gacctaacca gctcctcttc tgtttgtgac ttgcccttat gggagcctga gggctcctac
79500

ctactagatg aagctgatgg ggacattctc ccaaatagaga aagcgatgga aagttgtgac
79560

catcagagga agatcagaga tcaagccttg ttgatgtaaa aggaattaaa gttgagaaac
79620

catttcctgc aaccagaatt gtgaaagcat acaacgtagt accaaaggaa aattaccgaa
79680

gcctccttcc ctgaagaacc ttcaggaagg ggaggccagc cgtctgacgt gatcacagt
79740

atgtcttgcc ctgaaacaag aagcattctt cccaggtcct ttccgccgga ggattcttat
79800

tctgtgatca atgctgagct atgtggatga gtgtttgcgt ggatttcttt taaaactata
79860

ttgctaaata atcttgcccc ctggcccat gcacgttctt tttttttttt ttttttgac
79920

ccagataaga aaaagacttg aatgtatttt tttctataaa atgcagtcaa tattgttctg
79980

tgagggggaa agaagtagtg gaggatgaag atggaacagg ggatgagggtg gtgaccagca
80040

tgccctcatg gatagacact gagcaggag aggcaggat atgcatggag taagccctct
80100

ctctctctag gacaaagcaa attatgtctt catgccagt gactctgcat attcattgaa
80160

aaaaataaga tttaattaat ttagggaagc aatatgggct gaccagacag ccaagtctt
80220

caagatagga cagtaaaaca ttttatgtga gagctggcca atgtctggat ggaagagga
80280

gggccaggga gactggagag atattgagg ggtgaggagg tttgaggaat ccacagtgg
80340

gtgagcttca tgaccagtgg aatgaaataa gggctattct ggatggacga ccaatgtcag
80400

tgcaaggccc tcaatctctt gggataacac tggagttgga atacatatga aaaaaaaaaat
80460

ccacaagtat aaaagaagta agcagaggcc gggcatggtg actcacacct gtaatcccag
80520

cactttggga ggccaaggcg ggcagattgc ctgaggtcag gagttcgtga ccagtctggc
80580

caacatggtg aaaccccgtc tctactaaaa ttagccaggt atggtggcgt gcacctgtaa
80640

tcccagctac tcaggaggct gaagcaggag aattgcttga cccaggaggg tcgaggttgc
80700

agtgagccga gatcgacca ctgcactcta gcctgggcga cagagtgaga ctccatctca
80760

aaaaaagact ggaatagagt acatttgtca taggccaact gcctccatgc tcttgtctca
80820

cagagatggt tataacaagc ttgagtagct acaactggat acttttgtct atatatgctt
80880

tctctgtgat tctttttttc ctctttctta gcttaaaca atgcacacag agggaagtgg
80940

ataaactatt agcagagtaa gcacttgcca agaataccgg gggcaccaaa aacacataaa
81000

aacaggcttg gttgggagcc cgatgtgggc agatcacaag gtcaagagat cgagaccatc
81060

ttggccaaca tgatgaaacc ctgtctctag gaaaaataca aaaattagct gggatatggt
81120

acgcacgcct gtagtcccag ctactcggaa ggttgaggca ggagaatcac tgaaccag
81180

gagctgcagg ttgcagttag ccgagattgt gccactgcac tccagcctgg cgacagagt
81240

agactctgtc tcaaaaaaaaa aaaaaaaaaa aaaaagcgag gggatcttgg ggaacctagg
81300

aagaagttac ttaaatttca gcacatgcta tatagttttc aggattagtc cctctctagg
81360

tattcatctg gggctggcgc ctcttggct ggcaggggct ggtgatcggg actccccagc
81420

gagacagatt gcacccagag cagttaggcc ccgagcgctt cctggagcct ggggaagatg
81480

gcattcattt ttaaagtcca tttattgtgc agttgttgaa agtgggacac agaaatttgg
81540

actgatatgc ctaggccaca tgactactga ttgggccagg gtaaatgcag tctttttgtt
81600

gatggagaag gtcttttctg ggcccagcac aaagcaatgg atcttcagtc ttaagaaatg
81660

cattgaaaac catccttgtg cagtttgggg tccctggtag gccaagagct ggaaaggccc
81720

ttagactctg agtgctgtag ccccccttct aggtgatggt caaaacttga aaacggctct
81780

agactctgga cagaaatgca tgagaattga gtgttcaaag cttggtgctt tgacttcttt
81840

taactttctg tagactgaaa tatgaataag aggaaggcat gaggccctc tcagtttatg
81900

gagccatagc tccaaacctc ccactttcca acccagcaac cctgtgtgtc ctagggaattg
81960

tagcgggtgt cccagaatgc aaaaatcact ctgctagaac tggaaaagga gacgcctgcc
82020

ttttctagca cgatcttttag taaatagatt caggaaatgt gggctcctggg cagagctggg
82080

gccacaagct gagtgccttc aggcagcaga cttagcctct tcaagtccaa tctcatcagt
82140

agtatcaggg ggct
82154

<210> 6
<211> 99791
<212> DNA
<213> Homo sapiens

<400> 6
ttcagtgagt cgacaccgcg ccactgcact ccagcctggg tgacaaagca agactctatc
60

tcaaaaacaa acaaacaaac aaacaaaaaa aaaaacagag agagaattac taatatatttc
120

ttctccaacc ttaatgctca ctccctgagc tatgtggatt atactagtga aagtactgcc
180

ttaaccctact aatgtaaaca gattcctttt aagaagcctc cctatcgaga atgggtctgta
240

agagacaacc atgtcaacaa gtaaagttcc ttctgccttc cctcctgggg tgggctaaga
300

tggacccttc tcaccattcc caattacctg caggaaggca gcaacagaat agagaaggag
360

ggtaacacag tttcctgtac tttaaatacc tccctgcctc cgcaagtaaa acttgagcca
420

tgccaatgag ggaaaaagtg gccgcgattt cccaggtggc ttagtaatac acagttttct
480

gacacactgc aactccattc tttcctaaaa gttccttcca agcctgtagg cactgtaaga
540

caaacagcta agcacttctt gagtgcctac tatttgccag gctgtatgca gcaccaggaa
600

cacaaaaata aaagctatgg ccattcagga acaagatttc aaaatgaaaa ataggaggaa
660

aatctaacat ggatggaaaa aagccacact tttccatgct tcaagtttag tttccatttc
720

taaaatgtct gtcctcatat gttagccaga aagccagttt ccagcaagac aagcagaaat
780

aaatattatg caagtcaaac ctacaagaat gggctctcaa gaaagggaaa taattttcgg
 840
 tcaaaatcta aaaatccata ttcaacaaaa aaagcacaga ttgagaaact aattcaatag
 900
 caaaaaagga aaattaccat tgatccaggt tgccattttt aataaaaagc agtaaggtct
 960
 taaagaaaat attagcattg aagttcaacc aaataaatta ctttgtcatc aaaatgaaga
 1020
 atgccagaaa gttctctgaa atacaatttt taacaaaact gagaaagcag aggaagtatt
 1080
 ttttaatat acaaattttt gacaccaaca taagaattcc atgatatcct aaagagcaat
 1140
 gaaccagggt aatgcagttt ctagaaattt taaaacacta ccaaataata ttgcttcttg
 1200
 cattaagtat agcattgggt tgctagaata atcctaacaa tgacaatgca atgtacttta
 1260
 aattggggtc aaaaaacttt tttcttaaag taccagatag tcaagctttt aagctctatg
 1320
 ggtcatccac agacacactg caactataca aactatatca cccctggtgc aactacacga
 1380
 aaagcaggaa agcaaccaca gaagacagac aatacataaa caaatgagta tgactgtgtt
 1440
 ccaataacat tttattaata aaaacaggca actggcccat gggtggtaat ttgctgaaac
 1500
 ttggttttaa ttattattag tatctacctt tttttaaag atgatgtttt ttagggctct
 1560
 tccatcatga aaattatggt tttatgaaag tgatttgga agtattaaaa tacagaataa
 1620
 agagcaagcc ttgggctatg aagtcaaata ggactgggtt ccaggactgc tgtgtgactt
 1680
 tgggaaactg acctgccatc tctaggtgtt agtttcctca cttgtaaaat taagatgctg
 1740
 acagccctca cctcataggg tcaactactg caaggaccac acaagtgaac atgcatacag
 1800
 cacttagcat agttatggat aaccctgtac ctaccttatt atagcaattg ctatcaactt
 1860
 tcttctagct tatcattaat gatactataa aaacacccag cataacgcca gcataaggat
 1920
 gtgaatgaat gagcaaatga aaggaggag ggcaatggca aaattttacc ttggtttcca
 1980

agctaaagta taacctgagt tttttaatgc tttacaaaaa ggaaaggcaa tcagaaccat
 2040
 aattaatatt taagctgtac tgacagggtt tccagggaaa attctataga acttggtatt
 2100
 ctccaacttc caaaaataag caataaatTT tttttaaaga gctggtattt taagacttta
 2160
 tttgaaccaa agaatgtgtt tatggtaaca acatacatgt tgtaaattaa tcagtttcaa
 2220
 tggtataaaa taatacttcc aagaataaat gtatcatata atctatgtac ctcatgtcga
 2280
 taatttaaatt aaatacatta aataaatata tcattccgaa gtccgcttta tcattggttg
 2340
 taaaatccaa ttgtcatact ctggtatttg ttctccaaaa gaatcatcac tttcatattc
 2400
 actacagcat gagatgagat atttaataaa taataggaca tcataatcca gggattaagt
 2460
 agtttctttt agcaaaattt aatatgaggg tagcatatgc cagccgttcc tagttgagca
 2520
 acaattttctg tgttgggaga ggataacaaa gccaaagtatt ttcccataag aagttcacac
 2580
 tctaaaaggc caggcgcggt ggctcacacc tataatccca gcactttggg tggtgaggc
 2640
 aggcggatca ctaggtcagg agatcgagac catcctggct aagacggtga aacccgtct
 2700
 ctactaaaaa tacaaaaaaa ttagccaggc gtggtggtgg gcacctgtcg ttccagctac
 2760
 tcaggaggct gaggcaggag aatggcgtga acccaggagg cggagcttgc agtgagccga
 2820
 gatcgtgtca ctgcactcca gcctaggtga tagagcgaga ctccgtctca aaaaaaaaaa
 2880
 aaaaaaagaa gtccacactc taaaatgaag agaaatgaga aatgtctatg tgcacacaac
 2940
 ttcccacata cacacattag ttagagacca cagtcatgcc caccctccag ctcacgtgac
 3000
 atttcaaagc agagatgaga cttaaaggat ggaagaaatg gcgtttgttt catggatggg
 3060
 gtaacggagg cttttagggg cttatgggat ggacaagtaa aggatcaatg agacagaaag
 3120
 gcaaggcaag gcaaggcaag ggactcgatc tcaggaagtg gtgagtggag atgaaagaag
 3180

tccaatggaa tggcagtcac tggagatata tttaaagaag agggctcttt tgtttgtttg
 3240
 tttgtttgtt tgtttgagat agagtcttgc tctgttgccc aggctggagt gcagtggcat
 3300
 gatcttggct cactgcaacc tctgcctccc gggttcaagc gattctcctg cctcagccac
 3360
 ccaagtagct gggatttcag gcccgcgcca ccacgcccgg ctaatttttg tatttttagt
 3420
 agagacgggt ttcaccatgt tggccaggct ggtcttgaac tcctgacctc aaatgctctg
 3480
 ctcatcttgg cccacaaaag tgctgggatt acaggcgtga accaccgcac ccggccaaga
 3540
 agagggctctt tttaaagaagc aagataggaa gatccgttaa agagagaact gcataggaag
 3600
 gaaaaggcct ggctaggatt ccctgaacct tctcatttca ctactctctt gttagtcac
 3660
 aatgcagtta agaatcaacg caatcaaadc gaatcacata gtgattaatt tgtgccacca
 3720
 atagctctaa ttttggttaac ttctcattcc aataaaaata atcatctggt taatccattt
 3780
 cttcagcggg attacttggt tgcttgtttc tttagctata ataaactctg aatgccagag
 3840
 tcaccaaaact aagcattaaa caaactgatt gacaattcgg aagcatagga ataattacag
 3900
 atgatgctag gcaatatcac tagattgggt ttcatgaag ctctgactg gattttgcc
 3960
 caaaaatcac taactgcttt ggcttccaga gaatggcctc atggataagc cccaagttcc
 4020
 aaaaactaca aaatgtcaaa atcacattca actaattaat aattttatgt cacaataaga
 4080
 attaggaatc tctagaaggc tgggtacagt agctcatgtc tgtgatccca gcactttggg
 4140
 agatcgaggc aggcagatca cttaagccca gcagttcaag accagcctgg gcaacatagg
 4200
 gagacccgt cctacaaaaa atataaaaat cagctgggca tgggtggcaca tgctgtagt
 4260
 ccagctacc caggaggctg aggtgggagg attgattgag cccaagagat ggcggctgta
 4320
 gtgagccatg atcacaccac tgcactccag cctgggtgac agagcaagac cctgtctcaa
 4380

aaaaagaaaa gaatttctcg gaaagtctta atttttatct aaacctatct gaaaattcaa
4440

tgatgaattt tatgaactta gccctgaact gacattcctg gtgaaaactg caatcattgt
4500

aggacaagag ccgtcgtggt aaatgctaag gctattttat acacatccaa tattttggga
4560

gcactgcaat taaaagtctt ctacactaa aaccaatgca caaagcacct ctgaggttta
4620

agcctcacta ggctgatgcc tgaaacacag caaatgctc atactcttta tacacatgct
4680

caaaggctta gctatcaaaa gacagataaa aaggagcatt ttcccccaac tgatggaagc
4740

aagtccaagg tgaagaaaga ggatagaacg agaaggtaaa ccatctctca aaggtagaaa
4800

atcaacacag agtcaaccaa tgtcttctatc tcaacccaac ccaccaagca atgtatcaga
4860

ttggattgga gacagattgg agactctatg aaagaactgg catgcacaac ccaagaacct
4920

cctggattct tggcccttct tttaaagaag gggggtttct agacaggaat agaaccaca
4980

gagtcattga agccagcccc ctaccataaa ttgttcataa ccacaagctt attggtcagg
5040

tttataaagt gattcagact ctttagaggc tacaggaagg gtgcagataa gaggacactt
5100

atgctggggg ggggaattat tggagaggcc tccaacaatg agatatcacc tataattaca
5160

ttaaagattt aaaaacatga ttgaatcaca tgtaacaat ctctaatta gagctctgac
5220

tctccttaga atgtccaaca tccactgcat ctaagaagca ctagtattta ttatgtagta
5280

ttgtttgtgt caaattccca aaaaagggtta tcttgaagggt taatggattt aagtagctgg
5340

aaacctgat tagagttgcc catagaaatc tgaattctgc tcaaacagtt caaagctgac
5400

aacaaaaacag tgatggactc tgaaagtga gagaccata gagttcacat gtaatgttaa
5460

aatgttcggc acttcacttc tcaactgggg tacatgtaca cctaagacct acatcttaga
5520

agtttccaga gttacagaat gaactggatc gtgtctcaga ttaacaattg attcaaagaa
5580

tggggataga aaggagacca tgacaaaata aaggaggggg ctgttttgta tcacagcaaa
 5640

caaggagatc ctgtgtaaaa gaaattatca aaattgttaa tgccagaggg tagtggtca
 5700

tgcttgaat cccaccattt tggtaggcca aggcagaatt gcttgaggtc agaagtccaa
 5760

gacaggactg agaaacacag tgagagtga accccatctc cacacacaca aaaacatctt
 5820

tgcattaaaa aatgaaacta gaggttttaa agaactgctg caggctacgt gacttaaaga
 5880

atataagttc tctgccttct ttgtcatcag gagtgcctct gtgattacgt ataggaaggc
 5940

aaaccttaaa ggaggtaact gaaatcttta cttcaacata ctcaattgta tgcacacagc
 6000

ttggtgaagca aaataagtca caaataagca atattgcttc aaaaacctt tttaaaaaa
 6060

gatacataat caattatctt catattaata gagcaaagga ataccatgaa taatccaaga
 6120

taccaggtaa tccagaagtc agatactgag aagtgaata catcagataa tttttacagt
 6180

gtgtctatga tcttaattac attttactaa gcctaatact agtttttctt tcttttttag
 6240

aattttctaga ttacattatt tgaaagtaaa taatggcttc aaatcatgtt aggcaacaga
 6300

attcccagga aggagtgaag gtaaatcaga tgactgaact atattacaat ttaagctgtg
 6360

atgtttgcac aactataaca atattataaa tgagcttgag gtttcttact gaagatgttt
 6420

aatatttgat ccaaaatata gagcagtttt gactaagcac tgtaagcaaa gaaatacgct
 6480

atcatagcag gacacctgta agtttccatt gtataagaaa atgtttcctt tgtttaattg
 6540

gagagtcaat taaacataca aatatctgaa gacattatta tgaaggctga tttggaggaa
 6600

atgtataatg cagatacatt ctgggtgtaa gtattaaatt catacaacca ttctggaaaa
 6660

taattcaacg ttgccaagtc aagagcattc octaaaacc atgagttcct ctccaagta
 6720

aagacatgag aacaattctt gcacatgtgc aacagcgtgc atacacaagg gtgtttgcag
 6780

gaccgcgcca tgatagcaaa gcactggtaa aaaatcaa atgtccagaggg agtataactca
6840

gaccatggaa tactgtacag tggcaaacca aaagaaagac acctgcatcc acttagaaga
6900

atattacaaa actaatattg aagcggggaa aagcaagtca gaaaataaaa tatataacat
6960

cattccactt ataaagaaac caggaatatg tacaactaag caatatatta ttaggggata
7020

aagagatgta tggtaaaatt ttatataaag aaaaaacaag ggaataataa atgtaaaatt
7080

caggaaagtg actatctctg ggggcaggaa gggagggagt tgcaactgag aagcgcacag
7140

agagaacttc aaaattattg ataactttca acaaagttgg atgttagaca aataggtgat
7200

cattttatta ctattattta aacatgctca taatttaa attcacatgtat agctatatat
7260

tatgtattat agatataaaa caattcagag gtataaaatg caataaaaaa gtgtacaagc
7320

ccaaatctct agcatactgc caatcaaaag atgattgtta ttcataaagt atctcaaaaa
7380

ggaacattag actattttcca gaaatcctgg ttgctttatt attcatttat tattttattac
7440

cccaggaaga taaatttttag gaagtttcta ctttgagtgt ctttcttcta agttgcaaat
7500

acaaataaat gcttattttg aattataact atactggaag ctcatgtttc aaataaagct
7560

atztatatat agcttcatca ttgtttgaat gcttatcatt gcagcctcat tacaagaac
7620

atctaaaaga tttttttcca acctaaagaa aaacagtgct ataataaggt atcttacatt
7680

catgggtaaa aatactcctt ttagagaaca taatataatt agtttctatt ccaagtctgt
7740

cagtgtgcac tgtaacctc tgacaattat gcatgaacct tttaattagc atgttatctt
7800

tgggtgcaatt aataacttgt acttcatgtt ataggcaaat taaagttttc ccactaaaac
7860

caaattcttc agcatagttc taaaagtgtt ttgccattta aatgaaattt taacattcag
7920

acctcttgct cccaaaacac ataagccatt ttttccaaag ctagacagct tttctctatc
7980

cacgttcttt caaatatttt acctaagaca gagaatctca tattaaattc aatttcacac
8040

ttgccaatca cacacgcca ctcacaattc ctccagatga tgcaatgtgg cctaattgtg
8100

atgcttatgt atgtacatgt gcacacacaa acaggcataa tgtttaaata catataaaca
8160

cattcattgc taggtattgg ttaaaacaaa caggcttaca gcataccaac ttattaaaca
8220

taatgattga tattagtagt ttcattaatg taattaacct ttagatttga attatcaagt
8280

tcaaagtcta agaccgagag cagtgggtct caggccggta caataatggc aagtcattgg
8340

attcaagaga tgagtctaaa aggccctca gtcctgcccc taccagaaac cagagatatt
8400

tgctgctggg aataatgagg aggggggttc taagtacaag atcagctcta ttctttcatt
8460

ccctggaaat ctgccctcgc ctctgatacc tcaaaaaata tctgttgcac ttccaaggga
8520

aagaacaatg atataagcct tgctgttaaa aaaaaaaaaa aaaatgccac tacttacaag
8580

ataaaaggca aggtgttttg tttttttccc ctttctctta gactaggcat gaataatctc
8640

aaaatatttg tttgtcaata aagtgattaa aataaagtga tagacaagac ataatgatta
8700

gagaaatgaa tcaaagtaat atttgggacc agggctcttac aaatctagag taagtatgtt
8760

gagcagaaat agatactaaa aatttacatt agaacaaaat aaaaatggat ttagtaatgc
8820

tggattttaa ttggaaatag ctgttttaac tcatgacctt aaagaaaatt tgtcttggcc
8880

aactgcggtg gctcacgcct gtaatcccag cactttggga ggccaaggca ggcagatcac
8940

ttgaggtcag aagtttgaga ccagcctggc caacatgggtg aaatcccatc tctactaaaa
9000

atacaaaaat tagctgggca tgggtgtaca tgctgcagt ccagctgct tggggcgctg
9060

aggcaggaga accgcttgaa cctgggaggt ggagggtgca gtgagccgag atgggaccac
9120

tgactccag cctgggaaac agacagagaa ggaccctgtc tcaaaaaaaaaa aaaaaaaaaa
9180

aaagaagaag aagaagaaga aaaagaaaat ttgtcttctg gctctctccc taaaatgaac
9240

taactgggga atcatttcta ccttcagaa gcatttgaat ggaacagctt gatatcatgc
9300

agagcaccta aattttggat gccactatgc tctatcaaag gaatcagagc tcctcagaga
9360

atagctgact ccatgtattg ggtcatgaga aaatcaagat ccaactggaac ttgtgggtcc
9420

cagaaaacaa cagtgccttc aagaatgcct gggggtaatg ctaaggggac caaagaggac
9480

atttaaaggg actcccgag gtcatgctgt gaccatctga gcatcgacat gacaatatga
9540

attttaatta attagaacaa gtgcatccca taaaagctga gagtccagga tactcaaaga
9600

gaaacagtaa ggctggctgg gcacggtgac tcatgcctgt aatcccagca ctttgggagg
9660

ccgagaaggg tggatcacga ggtcaggagt tagagaccag cctggccaac atggtgaaac
9720

cctgtctcta ctaaaaatac aaaaattagc tggctgtggg ggcaagcgcc tgtaatccca
9780

gctactcggg aggctgaggc aggaaaattg cttaaacctg ggaggcggag gttgcagtgg
9840

gccaagattg tgccattgca ctccagctct ggacaacaga gcaagactcc atctcggggc
9900

aggagaagaa aaggtaaggc ctactgaaaa cacatttgta tcataatacc aaagaagagt
9960

gactataaga tgtcctaaaa atagatgggtg aagtatttaa agattaggat actctctttc
10020

tcctgtgaag tatgtgcttc tttatgagga aagaatcagc atttttcttt ttgtgtgttt
10080

gtaggacaga atgaatgcag agaattctaa cccaaatagt tccagaaaag taggaatata
10140

ccatgaaaga tggcagcaat agtttgaaga gtgagtcaga aaacaagtag ttgagaacc
10200

gttagcctat attcctaatga ctgaaagt tagtgcatgg agaaaaaata tttaatgtaa
10260

agtcatgttt gcctttataa tttcaaagaa tacttaaagc agtatgggaa atgcctaaat
10320

ggacacattc agaactcctg agggaacgag cagaacacag agggcaagag aacttggaat
10380

cagaaattcc acgttcaagc accacgttgc ctattcctca actttaagaa tgatcaatcc
10440

tgccatatct gcctcccagg gtaagtctga ggaactctcc ctgtggttta tgtttgttca
10500

actgtatttg caggteccatc ttttaagatag tttgcccatg aagatagcaa ggtacttttt
10560

atgaccctat ctcaaaatca ggacatattc tggtaaagac acataaatat tggctcaaat
10620

ggattttggt ttacctgaat caaagggatt ttgttttacc tgaatcaaag ggatgaaaat
10680

ttgccccaga aatagttgac caagtaccag gaggtttcaa aggaaaaact aggaagctga
10740

cagattgaat gttcaaaatg attccaaccc aaaatgctct attaccaga acttgactcc
10800

agagataatc caaccaacat gaacctctgt tttctattta ctgagagcat agttgctgct
10860

caatacatcc attcatgcaa caagtgccat attcatgac ccattgccagg cacgttctat
10920

gtgcttaaga tgcagttagt gacaaaagag gccaaagtcc atgcccttgg tgactgggga
10980

caggcaaagg tggatggcaa taaataagaa tattagatga aggaagtaaa ataagactat
11040

gtcacttaaa atgacttggg gggtagcata cttttgaaag gcttctcaga agaggtgaca
11100

tttttgagct gagacacaaa tgacaaggag ctggcattac aaagagctgg ggaaagaatt
11160

ccttgacaga gaaagaatca atggaaaagc tctgagttag gaatgtggct ggcatgttat
11220

taaaaaaaaa gaagaagaag aaaagaaggc ttaaacaaag tgaatgccaa gaagtgttaa
11280

gatgtgagct tgcaggggaa ggaaagagct agagaacacg gtgtgagtaa cgtaagcgaa
11340

aacaataaat aatacctttc ttcacatcca aaagctgac agatttcttt ttccaatgcc
11400

aaaatacatg taagtacata atgtcacggt gcataccact gcggtggaac acaaggaaaag
11460

aagagctata atgcaaacia acagcaactg actctttggg gagcttaaga aaaaagaagt
11520

ccaggcacca acacaaaagg ctgaataaat cacaacctta ctcaaagaac aatattaagg
11580

gcataggctg tagaaaaggc aactggttca atgcttctca tgccctctga gagctatcac
11640

aattaatggc atcatcttcc aaacaacagg taatacaatg ttacttctgt gttatagaaa
11700

catttgggtt ataactggag atgtatggta tattcatcac aaagatgata atcatttcct
11760

ctttgtgtat tcaatggaca gaaatctcta agactaggcc tctgcttccc aaatattcat
11820

gcaactggaa ctgttacaag ggtaaaggca tgaattaaaa ttatagaaat ctgatgaga
11880

tggaattga aaatacacta aaacctaaag tctcctatag ttaacatggg ataatatatg
11940

catagttcct aacagcgtac ctacaatcct ataagggtg agtaaagtgt agtttgcttt
12000

ccaaagtaaa agtgacctga ttttctgtg ttgctcccc ttgagatcta aagctaaacc
12060

tctactgaat aaaagcagag tgacttaca ttgagatagg ccaatctaaa actgaaaagt
12120

ctaaaatgtc aaacttttaa aataaggga tgttctactg ttatttgact cattaataat
12180

aaaaatacct agtgattgtg aatttaaaac tgtagactta tgatttgtgc attttatata
12240

tgttattctt caattagaag ttcaagaaat aataatggga ggctgtatt aaaccattta
12300

tataaaagtt tctcaacat aatggaaaga tcaagatgtc tagcataatt tgaaagtatg
12360

tttctatttg acagaaacct tggcttgga aagttaaact ctaaataatt gtacagaaag
12420

aagtattctg ggtttcattc attaaaagtc attgaccact gcagaaaagg gagagcaaga
12480

gatttacata tatttttctt gacgttttat atgcattagg cctggcaatg aacttgaggt
12540

aggtattact atctccttat ttctaaggga atcagaatac tgtaaacaaa gacatggagc
12600

taaataatgg agtcaaatcc tggcttacta ggctccacat ccatgttgct ccacttatca
12660

gacggccacc caaacatata agtttaacaa attccaaatt attggggcct aaatagacac
12720

tttttggtt caatattgcc cttcttcctt tgtattctga atatcatgtc aacctgaat
12780

cataagccaa tgcttacaaa tttaacttca cattatgagt ttgagccca actgatgcct
 12840

aaccactaaa ccattcatat aacacactga aaagaagcca tttaagaaat gagaaaaata
 12900

ggtattgttt tctctgtatg gctgtaaaag ctgagataag aacatttttg caagttcaat
 12960

agagccgaaa ggcttgtgaa ggctatggga ggatgagtat attggaagaa gatcagtagg
 13020

tcaagaatct aagaagagct atttatgtgg gagacaggaa ttaccttgaa ggaagagcag
 13080

agccccagaa gaagaaaaac agaataatac agctaactat aaaaggaccc agggccaggt
 13140

gcggtggctt aggcctgtaa tcccagcact ttgggaggcc aaggcgggtg gatcacgaga
 13200

tcaggagatc aagaccatcc tggttaacat ggtgaaacct cgtctctacc accaatataa
 13260

aaaatcagcc gggcatgggt gcaggcacct gtagtcccag ctacttgga ggttgaggca
 13320

ggagaatggc atgaaccgg gaggtggagg ttgcagtga ccaagatgac agccactgca
 13380

ctccatcctg gaccacagag cgagactcca tctcaaaaaa aaaaaaaaaa aaaaaaaaaa
 13440

aggacccaag agaataataag tcccatgaaa acagagctcg agtctgtctt gttttgtcac
 13500

attacctatt tgatactcaa tatgtaattt aatttctgag tgaataaatt aagcaaatga
 13560

ataacttaaa agtatagcta tgaatgttaa aaaaaaaaaa agaaaaagaa aaagaaagaa
 13620

aaaagaaaaa attaatggg gagaggcatg cctatggcct tgtccttggc cctgaataac
 13680

tgtttggcca gatagtatgg tttcctaacc ctagctcatg gtgttcaaat tctactagt
 13740

tttctttgtt tgaactttcc atctccatga aactaaaaag tataattttt aagctgcaga
 13800

acaaatcatt ccattgtcta tttctgagg ttacatagct ggctgcatgt actttcaaca
 13860

attctatcat gctttgacac ttagaaatat tttccctacc agctttggga agctgggagg
 13920

tttgtgttta ccatactctc ctgatgttac gctttagttt aaaaataaca gtcattatca
 13980

ctcaattatt cagtgatcaa gtgtaacacc aggtgcaaac aaaagtctcg agaacactgg
14040

aacatgctct tgttactgca aacgatttaa gagctttgaa tttatgaagc cagtgatcat
14100

gttactctaa attagaagca ggtaaaacag aagaatggca taagctaaat tccttctttt
14160

cgtttcttta cttccttttt ccttttaagg atatctatag caaactatat tagctttcta
14220

ttttctccac tatttgact agcattaatc aacatccaaa aagtgtacaa attattttat
14280

ctttataatc aatcttctga aatgctaact tgaaaatata aacagcacia ttactttttt
14340

ctcttttctt tcccatttga aatgtgtgct tgtgtgtgtg tgtgtgtgtg tatctgaatc
14400

tgtacatata aatcaaatac acacacatat acttgtttta aaaaaatcct caaagaatca
14460

atcacatatt acagaactga aggaatattg gagttcatct ggtccaaagt gttggtattg
14520

taaaatcact ttcaaatttc acagttatcc aaatggtaat gaatattaat acctaagaaa
14580

atcaatcaaa gattaacatt aagacattag attttaaaaa tcattgcatt taatattttt
14640

tatccatttt caaatttcag aaactaggtc tatattgtcc tatagatcta acaagcacta
14700

taaaagaaaa attacaaaga tttccaata tacagataga cacataaatg aaaactttta
14760

attgcaagag aaagggaact ggattcatca gttacatgca tttagaaaaa tctgtctcat
14820

tcatcagcag caatgaagga gaaaagttgg caacagagcc aagtatactt actttctaag
14880

tagccaataa ttggaggcat ttaaaaatct ttttactaa aatatacatt ttttgctcaa
14940

actccagatt tctgacattt ccttattcta ctgataatga cttcatgtga gcaaaatgga
15000

atgaaagatt gaaaaaaatc agatatatgc taccataaa ggtttatcaa tatgatatat
15060

ctaaacattt atgtacttat tggtttaact tagttgatat aaatgaatat gcatatcata
15120

cacatcatgt gaacagaaga gcagatcagt ggctgggggt aggggattaa ctgcaaaggg
15180

gtacaaggaa acttgtggtg atggaaatgt tctatattta ttatgtggtg gttgttcaca
 15240
 actgtacatg tttgtcaaaa cttaccaata gcacatttaa aattggcaga ttttattgta
 15300
 tgtaacatat ctgactcccc gacaaaaaaa accaaaactc tgttatacca tgacacctaa
 15360
 aaataattcc tacagagact tacctagaac atcatggaca ctcaaattt aaccatgacc
 15420
 caagaaaaca ctataaagta tgcataaaca gtgattctcc tataaaatct acacacatac
 15480
 acaatatgtt aagaaatgag tacaaaaatt aggggagcta ccataatata tggcagaaac
 15540
 agatttcctg cccctgccca taactatcat ttaccaacat tcaatcgaat ttaagaattg
 15600
 ggggaatttc agagagtatc ctaattcact tcctaattca ctctctctag ataccaccac
 15660
 tcaaagattc ccgtggttcc cagagccttc acttgacatt catacaactt tgaccaggaa
 15720
 agaagggtgc gccacctcca cctgtgcctc tgcctctata ctagaaagga ttaggcagag
 15780
 aaagagattc aaatactgga aggaaaagaa ggtattggga agccaggatg gtttgttggg
 15840
 gcaaagtatt ctcttgtcat acagatggct ctcttcttgg tggtgcccca gccaccctc
 15900
 aagggtgtct ataagtagga gaagagaaaa gcactaatga tgattataac ataattaaat
 15960
 aatttatgaa ggataaaata tcagagtcac gagagctctg aaagaatagg tagactacat
 16020
 gcatcttacc tacacattga gtattaaaaa tcctagctgc cctctgcagt gatgccccac
 16080
 ccagaagacc aatccattgg aactattgca atctgatcag attaaacaca ccaaggtcac
 16140
 tctgaaagta tggaggaaaa caccaaatta agaacaaaat aatactttta aaatataaac
 16200
 cagtttaaaa gtatcaaaaa cgtttctatt tagtccatct ctgatctctc tcccaaacct
 16260
 ctgaccgaaa gagacacaga tttttgcttc taaaaacaaa gtatatttta ccaacaggat
 16320
 aatataagtg ctttcaattt tatctaataa aaagtagggc aaagtgaata acaggtgcct
 16380

ctacagatgg agtatactaa atacatgtaa cacttctgat aatgttgaac caccacctcc
 16440
 ccttttggaa tttgaaacat cccgcataag atctgcattt ttcaaatatt gctaaatatt
 16500
 tagagtctaa aaagatgaga ttcattatit tgcacagtaa cacatctcct aataatattt
 16560
 caggccagtt ttgaagaact gaaattgata tcttttccgt atatacttaa cctcctacag
 16620
 atttatttga ttctcaataa ttactattct ctggtaaaag gcatatgtta ttacagtaca
 16680
 ctgcactccc acaagctgcc cattcagtggt attttaataa atctctgccg gagctgtata
 16740
 aatttgatcc tgtttaaatt taataaggat cagaatatat caaagcacia ataccatgg
 16800
 gagaataga tattttgcat agttattgaa tggcctattc agggatacaa tatgtaagac
 16860
 ccagaccatc aaccttgaat taaaagaga gattaatgcg ttcaaagccg gctacttggt
 16920
 aaataacata tgaaattagt atttaataaa acgtcttctc tatattgttg ctaatggat
 16980
 ccactctagt acagctttta tatcaatatt ctcatacaga attctaatat ttacaagaat
 17040
 tgtcagcaat gtatttctta aatatctaca ttatcttgat aaaatgcaac actgtatgga
 17100
 acagagatga agaatacctc atagtaattt ggaattgttg agtttaatgt tcatttgagt
 17160
 aatataaata aaaagaacia ccactttgct ggatatagtt ttcacagatt acaagcctga
 17220
 tgatcaatga actgctttcc cattctcttc ttagcttgga tatgtttctg atttgaatcg
 17280
 tatttcaata atcattaata tcctgtgggt gagcaaatca tttcaatttt ttccaaccaa
 17340
 aatctacact tgagtgtaaa caccaatggc aatgattaat gagcctggat ttagaggctg
 17400
 ttagaagata caggtagcaa aatgtcaggt atcaggaaat ccttcctcat agtaaggacg
 17460
 actcacaaga caactattgc caggaagatc aaagggaaat ggtatcatgc gaggcattga
 17520
 gaaccaacct gttctagcgc taaatgcacc agacagatgg atgggctaatt ccgtcttcag
 17580

ttatctacaa atactgcggg ggccacaaat tgtttaaagc cactgaaact gacattcact
 17640
 catgcatgtt atatttgaat ctacacctca ttattcccaa ccactttcat cttattctta
 17700
 aactagaaat gtaagggtta gtgcattcat cagacaaact gataaaagac actatctggt
 17760
 tagcctgtcc agatccatta gaaacctcac cagaagatta ggtcttactt tcagataggt
 17820
 agttatcagg taatgaagtc tctcctcaat gcgcctggct cccgagcgag gttggcactt
 17880
 catcctccac tacatcttct atgagtatct aaatcctctc taccctcttc tgactcttct
 17940
 gtcctctgct ggatgggggt cagtgagaga aatgcatcct agctgagcag acgaaaaaaaa
 18000
 aaacaacatg gttgggcaat cagctatttg tcacactgag aagacatcac ttcttaccaa
 18060
 gcaagaagta gtagaatggg ctgggcgcag tcgctcacac ctgtaatccc agcactttgg
 18120
 gaggccaaac cagcaggagc tcttgggcct gggagttcaa gaccaacttg ggcaatgtat
 18180
 tgagagctta tctctacaaa aaaaaaatta aaaaattaga tgagcagggg ggcacgcac
 18240
 cacagtccca gctactcaag aggttgaggt gggaggattg cttgagcctg ggaggtcaag
 18300
 gctgcggtga gcacagtga cactggtcac accactgcac tccagcctgc gtgacagagc
 18360
 aagactctgt ctcaaaaaca aaaaaaaaaa aaaaagaaga agaagaagga gaagaaatag
 18420
 tagtagatga gctacaaata caatgcacca atggctgaag gaaactgaag tatggatgca
 18480
 gatatgacat gtcccatata tctaatttta tgaaatatgg gtatataaaa tccaaacata
 18540
 aaaatcagat ccataagact aatgagtgtg gttaaaattt cctgaaatag tatttctaaa
 18600
 tatatttaaa aactagcatt ggttgaacac cttcagtttt ccaataaaat tgtgtaattt
 18660
 attatgctat aatttttctt taaaactttt attttaaaag cccaggcttg gcagcctggc
 18720
 caagactagc accttatctt acacaaaata cagtgcctct cctcttaaca aatatagatg
 18780

taaccactgc cagctgttta caacaagcaa aagctgggaa gccctattag gtgttatgtt
18840

tgttctgctc gtgttaatag actcatgttc aatgtcagct cttcttccta gatcagtgtc
18900

tgccaaatth cagttatthg tatatcactt tcatgcattt tgccatttaa agcaggtcca
18960

cactcttcct tatccaaaac cttagggcca gatgtattht acagttgaaa gthttcaaga
19020

ttaggaaagt aatatggtgc atattccctt attacataat acctctagtg gtctgaggca
19080

gcaccccata gtcaaacatg tggacattht thtagcaaaa tgtctgaata ttcacataaa
19140

gataaataca ggaaaaaaa gthtcatgtg aattcaggtc aagthttgat gcaaatgtgt
19200

aacctaaac atthththth atthgtthth cagagcatth agaathtgga attatggcta
19260

aggtatagta ttgtgagcct ttatththct ttaattcatg catttgcata aaaattaata
19320

tgctththtg ttaacaaaag gacacttht ttcatacatg gaaaactggt atthctthct
19380

ataaatagca cataactatc aaaataaata caataaacac aaaataatct atthaattct
19440

agcaaaccaa agctacagaa ggggtgtaac tgggtcctgt atgatctcct ccgtctctcc
19500

ctctctctct cttccctat aatgtgtctc aatctctctg ctaaaaaaga gattagcaag
19560

tattaggtgt taaagacata ttcgcaccaa actgagatth tctcactgat gtaataagaa
19620

agcctgaaag gaaaatggaa atgctthcta tgtaattcaa tatggthtaa aaacatatcc
19680

ttgtagctaa aaatcaagct acagcctacc aatggtagaa gccagtagc acctgcctag
19740

cttccatcta tcttatgcca accagcacac catgtthgt catgtccctg gccagthcta
19800

ccttctctgc cagatathg gaaacataca aggggtgtag tathcagctg gaaaaaaa
19860

aagthcagtg gtataagacc acaggggtcc aggaagtgt tcatthtaca taaagccat
19920

thtacttht gccctthaa tatgtthtg cthtgtgcat ggccatctgt gccgatgaca
19980

ctggttaccac tagagtacct aaatatgagt gcagtgatca actgtgcgta cacagcttca
20040

gtggatgcaa gtatgatctc tccgatgac ctgggtccct tcagaggcct ttggatgggt
20100

ctgagcagta gaactcaaac tgtgatccca ggaaggcagc atcagcatca cctagcaact
20160

tgttagaaat gcaaattctt gggccctacc ccagacctac tgaatcagaa actctgggtg
20220

tggagcggca acatggattc taacaaggcg ttcaggtgac tctcatgctc agtgaagttt
20280

gaaaagcctg tgggtaacga gcacagtctc tggaacctgc ctggtgattt gaaggaatcc
20340

cagatcaatc acctcctcac ctgagacttt gaacaagtta ttttaatttca gccatgcttc
20400

catttccggt tctgttaact gggacattaa tagtacttac cttatagagt cgtggaggat
20460

taaaatttatt aatacagata gagtacttag aatagttcct gaaaatactg cacaaatatc
20520

aatagtattc catgtaagct acataaaacc tttacttaca gctaaataat agtattttct
20580

ccaaataata aagacaatta ataactacaa gcttcatttg atgctcagaa tgaggttcat
20640

aggctttgtc tcctaataaa gtggaagtgt ggaacttcag cagagacagc caaatgtcag
20700

gccacaggcc agatgcagac tttagacagt ttccaaatag tgctaacatt taaaaatgtt
20760

aaaatatcac agaaatcaaa atttctgggt ccccttgaaa aatcaggacc tcagacaaca
20820

ctactctagg attccctaga gctactggca ggcagagtgg aggagcaacc acctcctttg
20880

aagaagtctc gtctcttgca gtctcacaca aatgggcagt tcagttgcag tcatgtacca
20940

tcttctcac acctatgaca caggggtttg tgagccctgt ccttctgcaa actctaaggc
21000

ctctaagtgt gtctctgtg tgcaggacag ggtgacagct gtccactgg gaaagcaagg
21060

gagttcccaa gacactggtc tttcatgcaa caacaggag agtcccagga aacctgggaa
21120

aaaatgggtca ccctatctgt gcagatgggt cccaagtcta caatccaagt gtctctccta
21180

tctgggccct gtttccaagt accatgagaa gcattttttt gtttgtttgt tttttgtttt
 21240

ttgcgaatcc cacttttgcc tcaaaacaca ctttccttca cgtttatcat gtgctattca
 21300

agcacttcat ggatgaaggc ctctctata gcaggcacag tcctggacac tggagatgaa
 21360

gcaacgaaca aaagagacag aaatctctgc cttcgccagg ttatgttcta gcatggagga
 21420

gagagacaat gaataaataa ggcagaggga atgttgaatg gtagtaagtg cctggagaaa
 21480

acggaagcaa ggcaggagga cggagagtat cagggaaagc aatgaggtgc ccatcaaagt
 21540

cctccctgag aaaagggaca tctgaggaaa gacttgaaga gaatgaacaa ttgagccatg
 21600

caaatatctg atgcaacggt atttggggca gaggaaaaca ggacctgggc cttgaggtgg
 21660

gcgtgtgcc a ggggtttcca agtgatcaag aagagcgggg tatgtggtgg gagggaagtc
 21720

agacaggcct tgatgaggtc aaggaaataa tgtccaccac tgagttaaga ccaggctttc
 21780

ctctgaatga atcaggagc cattggagag ctccaagcag aggagtcaca gggacctagc
 21840

ttacttattt ttctcaaata gccatcgaca tctagttggt ttgacttggt attcaaggag
 21900

gtcagtccaa ggaaactaga ttccccacct tctataatct cagcatcccc tctcattgaa
 21960

gacaggaacc acgtttaaca tcaatgagtc ctgtgggaat aaagactgac agggaaccaa
 22020

acagcaaaag ggaagactta cctcagttta tttatgggga tttatatact gcaaactaat
 22080

actctatttc actggaatat ttgtgtttat gaagacatcc tgatgaacga gacttttttt
 22140

ttcagttcat gaaattttca ccataaaaga accctgaatt cctattttga aatccattgt
 22200

gccattttta ttctgcctgt atttgaatat agcatcaata tttaactact cattaaatga
 22260

attgtgactt ttatagtga aatgtcagtc tttgcagttc attagattca tttaaagcat
 22320

agctgcttga tatttttttt ccctggaaga ataaaagaac tgtgaattga agacataaac
 22380

tttaaattta agcatgtgtt ttcttcaaaa ttaacattca ggatctgtaa cattattcat
 22440
 gagattgagg ggctgtttgc atctttctgc atttacttag tgactgtcgt cttttatgga
 22500
 gtatagagct gtttatcaca gttggatatg accagtggaa tgtcactgtg gacattaaga
 22560
 tgccataagc tattatgaat ctctacaatt ctgcatgtag tattaaaaac attatcttga
 22620
 actatggcat cagaaaaaat aaacctgga agatgtcaaa ttaataact cactgaatcc
 22680
 aatagggcct aaaagtggc tatttgagtt ttaaaaaata attctgataa gactttagac
 22740
 tgccatgaaag aaaatgatat attaaaaatg gaatttgaga aatgaattca aatgaatttt
 22800
 ttagatgttt gcaattcttt ctctctaaaa ttactatgat ttacatgtat ggagcagtaa
 22860
 atgcaaatga tttccaatca ccacgggtac gtgtactcaa ttacaaagtc tcttaatgag
 22920
 cattgaaaag atttcagcat tcaactgaaca agacataaag caaatgaact cttcagtgcc
 22980
 ctcccttccc cacagagggc tgaaggggac tcactcttgt tacaagaaca cgccactccc
 23040
 tccctctctt cctcttccca gtccatccct caccocagtt tcagccacca tccaaaaaga
 23100
 aatcgacga tgcccttga agcacattga gcttaaaagt gcctagaaga taaagcctaa
 23160
 acttcagcac gacaggcaag acctttcaaa atctggcctc gaccatctc tctgggtctc
 23220
 cttctgaga acccctatct tctgcctttt cacaactcca cacttatgca tgatgctgtt
 23280
 ttccctccct ggctcaccca cccagtggt cttcttctgt gatggtaacc tcccagaaa
 23340
 gtctcttgag ctagagtgat cattattaaa ttaaaaaaaaa aaaaatccta ctttgttcta
 23400
 ttctgatttc aaaaacctga aaggcaagat atttgaggga aaaaagagaa tattgttttg
 23460
 ttgcttttta aaaaataaat ttgcaggctc tgacaatccc gtgctttgtg tgtgctgac
 23520
 acactgcaat ctaatcaggt acagtaaata aaatgacttg agaacccttc tgacgaatgc
 23580

aaaggctgat gatattttgt gggaggagg ctattcatgt atcttctttg agtatactaa
 23640

agatccattt aattaagttt ggggagggg gttctctatc ttttctctg tccttttttc
 23700

ctgaagctct cctactggaa agggaatgga agaaaacètt atgtggagtc cggagtgçct
 23760

acctaaatta ctaagacaat aaaggacata caaaagaaga taatcaaagt ttactttggg
 23820

tacttgaaca cttgctaaga gcatgcatcc tgcagtcagt aacattacca tctatactca
 23880

gagggcaaac gctaatttca aatccagagc aatgtcaagg atttatcact gcaacccaaa
 23940

gtatctttgc tatcaaagac agtgggggca tgaactacaa aggcaataac taaaaagagg
 24000

ggctcatctc acacctctac tcctctccaa tattgggaag tccatttgag ttctagagag
 24060

tgcacatgcc aaaatgcaag tgcgtgcatg cctgcgtaca cacacacaca cgcacaccat
 24120

tccaacattt tcccatattt tgtccactcc ctctcttctt tgttatgggt agcagcagtg
 24180

ataatgcact ttaatggaaa cctgcctaata gaacccaaaa ggaaaacgat taggaaaaaa
 24240

atatgcaata catcgcatg aaagggtgcc ataattaaat tacttggttac cgtttttaaa
 24300

tagtgcaaat attagtatgc ccacaaagaa ctaaaggttt gtttgtttta atctagatct
 24360

tccagaattg gctgttcagg aactgcccat gttgcccåaa gtagctactt tttataagaa
 24420

tttccactg cctttgaaat attctctgac attcccattg gcttaagtat tctcatactt
 24480

tctgcctaa atttgtagtt tccaagttt aattacccca gaacaatctg gtcatcaata
 24540

accaaataaa agaggatggg atttttttcc ctcataaaac tattcatcaa gattcgacca
 24600

agcctgcctg gaatgtccag cgaaagcttc aataatcagg gagaaatgag ttgcagaaac
 24660

tcaaagagaa cagaatagcc aaggcaggag attacgaagc caggagatga gggcacactg
 24720

aattaaacat tcttaagcag acttacatac ataatgaatt ggtcaattct gtccaatgtt
 24780

tctgaaagag attacccttc caacaagcca aatgagtgtt caaggatggt agctcggatg
24840

gattcctgaa caaaataagt gttaatcact ttatctcccc caatcttttt tgtgaagaag
24900

agcaactttt catttccata aagccttgaa atcgttggtc ggcaggggat gggggtggtg
24960

ggggagggct ggtagagga tacagacaat gaagtccttg aaaaggggga tttcctgaca
25020

ataaaccctt gatcagagcc agagagggga cattgttgtc ttcagcactc ccctctcctt
25080

tgggaattcc aggcgtagta gagagaagag ggtcacagtg acaactgctt gggaaggctc
25140

taaccaggta gtctttgcga ttttctgttt gggaaaagct ggcctacaac aatacatttt
25200

cttggaatgt gttcttttaa acaaattgct acagactttc ttaagacatc ctgtcaattg
25260

ttccaaaaat gtgtaatcag cttatgaaat ctttaagcaa agaacattaa atgttagcat
25320

aaaatctcat aactagtatt gtattgagat tctcctacac accaggcaca gtgctgggca
25380

ctaaacacat agcagcgaac aaacgaaaca cagcggtcac tactgtcttg aaatcacagt
25440

taggaaagac aaaactgctt cttaaatctc tgtgcactcc atgccagaca cagtagccat
25500

tccataaata tttgaggaat gaatgaataa aaagacttca gagcttgagg ttcatgtgac
25560

tcaagattta aaaataaaaa aataaaaaag taaaagaaaa gaaaaagcta gcaggtatac
25620

tagaaccaaa cacacaaaat aatgagccag agtgtgctcc cacaatatta ggagaaatgg
25680

aacaagaaat caaaaccctc gacatattgc tcctatttat tattttaata tcatcgttgc
25740

acagaaccag tcattggtgg tgctttaaca attcctccta aatagctttg tgacttagtc
25800

caagcaccca gagacagatg agccccaatt caggcaagtc ccctgttgaa ggacaactat
25860

tttctcaaat gaacaaagat cttctcttct caatccgcc gatcacacag ctttataagt
25920

ttgaaaagtc tttttgctaa cctagaggta ttctctaaca gaaagaggga gatggaaaaa
25980

gaaaacatct gcaagggtgct gaaacacgcg tagattatctt tacctgatcc aggagcctcc
 26040

ttccctgggtg actacctact cctcagaatg agtgctggcc ttcttagcta ttctggggaa
 26100

cacaatgatg gaggtcctt ttagctccat caggccctag agaaataact gccctatatt
 26160

tagatagatg ccctttgtaa caaaggaata attcatccca gttagcactt tctaaatgac
 26220

agaaccagaa agttttcgaa aatattcttc tagctcagtg gctctgaacc aggggacgtt
 26280

cggcaatatac taaagacata ttttggtgtc acaactgggg gtgctgctgg catctagtgg
 26340

gtagaggcca gggatgctgc taaacctcct atgatacaca ggaaagtcc ccaaaacaaa
 26400

gaactattta gcccaaaaca tcaacagtgc tgagatgaag aaatcctgct taagttatac
 26460

attaaataaa gaacagcagg gcattctactg gtgacccag atagttctgt taaaatattc
 26520

tgaagtagga ttcttctccc agagtctatg ttcagcaact taaccagatg ttaaaaactg
 26580

accgatctt aggatgttaa caatatttta aaataatact gatggtaatg gtaacattaa
 26640

tgaaatgctt ggtatacaga gagttatagt gcataattac aatgacccta tgttgcaaat
 26700

gttattagtg tcctctagaa ggtaactgta ttaatcacat acaataactc tatggagtag
 26760

gtgggtttat aaccccttag gaagtgtgtg ctgtgggcaa gacttgatct gttttttttt
 26820

tcaccaccat ctgtccagca ccagcagag tgaaatctcc atttcagaaa tgatgatgct
 26880

gagaggctca ggaacatgcc agagatcaca gggccaataa atggtggaac ccagttcaa
 26940

accgggcaga ctgcctggtt ccctagcctc tctctctcgc accgtgccag ccggccaact
 27000

gcttcatatg tggagtcatt gaaacctatt agacctcttg agtctaacct tagataagca
 27060

gaaccaacct aagggggaag gaattaggct gctttccaaa gttacaaaaa gtttcccaat
 27120

attcaagaga ctgattaaac aaaccagtga cccaagaca atggtgaagt acatttgtat
 27180

tgccatgaga cacttgaaag caatctaaga ggggtggctgg ttttataaga ggttaactgc
 27240

ttctcagttt gagcaagact catgtggcac caggcactag gcaccgggca ggcttcctga
 27300

agctaaaatg gccactttct caaaggacaa atgccctgca tgagtgcaga aactctttgc
 27360

aaccactagg gagctaactc agggatcatg ggatttacat gtggaccaga caacaacttg
 27420

gccaccagtc caatggatcc ccagtgccaa gttaacctta gcacagtctt cgggtaaaat
 27480

ctagatcaat gctgattatg tcaggaggtc aactttcagg aattaatact ttttgaatga
 27540

taagagagtg caaacattaa tttaaagtct cagagagtat actaattttc ctgttttcat
 27600

catttcagta gattctataa ttctaggggg aaatatccca taagaaataa attctagcaa
 27660

gtactgtaaa taaagaattg tatccacctt tataaaaaga ctaaaagaag aagagataag
 27720

agcttgtttt ttagttttac aatattgaca tagtaggctt tatttttttt taaatgtgga
 27780

aatttccctt ttagaacgat aagaatacta gcctggacac tatacaaaca tagtcattgt
 27840

acttataagt gctttaagta tttattcata gcctgccttt ctctacaaag gaagtggagt
 27900

gacttttaaa agcatatgtg attttttaaa ggtaagtaac aggtgaatca aagcaaggaa
 27960

aaagtaaaat tgagaaaata ataaagccta gatggccagc acacaggaat gcatatccta
 28020

ctgccctaga cagctactaa atggggtttcc tgtttgactc caagcttctt ggccatcaat
 28080

gcaaagaaag aacaagaatg gtaagattaa tgcccagtaa gatacagcga agtggatttt
 28140

ttctcagttct tcttgatata gtaacctaa aagattctct gtaaatcatc atttaaaaag
 28200

acaccatgag gtgtaatgag ctatacattg taaatagaac tttttaaaaa aaatgatggt
 28260

caggaacta ttactcataa tcctctcaat acaaagtga gccttaatca gtgtgcaatt
 28320

caataaagca gagagggcct gtgaaggcaa ctgtgaagtg gggtgccaca taatatgcca
 28380

tcgtgagcat tcagctctct cacaagtctg gcagaaatct cttatcaaat taggttatct
 28440
 ggatgaacgg aaggaagcaa ggcattctct tcatgccatt ctccttagat atattatttc
 28500
 ttccaacaga ttttattaac acaaactggg cagcagggac ttcaagaagg acaaaaatcc
 28560
 tgtgtcaaaa ccattacata tgttggcaac atttatcatc aggaaaagcg tcccaggaag
 28620
 ttgaattcta atgtatgtgt ggtgttaciaa tggaaatctt agtgtgcaac tggaatcagg
 28680
 ccatgagttt tgcctgtcat caaatgtatc tgtaactta ggggtgaact gccccacc
 28740
 ccactttttt ttttttgag acagtgtctc actctcttgc ccaggctaaa gtgcagtgcc
 28800
 atgatctcaa ccactgcaa cctccgctc ccaggttcaa gcaattctcc ggctcaatc
 28860
 tccgagtag ctgggattac aggtgcctgc caccacgcc ggctactttt tatatttgta
 28920
 gtagagacgg ggtttcacca tgttggtcat ggttggccag gatggtctgg aactcctgac
 28980
 ctcaggtgat atgccctgcc tcggcctccc aaagtgttgg gaatgtaggc gtgagccacc
 29040
 atgcctggcc cctctttttt tttttcttt tttttctaac tacacccttc tcttttctc
 29100
 ctctacaact gaatggtctt tctccctgat ataagaatct tctagatcac tccatccat
 29160
 tgcattcttc ttggtctctt gttatcactc ttcttactca tcttactagc ttttaacttg
 29220
 actacgtttt cctttacttg gtttctgcct tggacttcca caagcttcac cctaataatc
 29280
 atgcagcctt gaaaaacatg gaggaaggaa aatgctgtca cctgccatct tgtacctgtt
 29340
 taaagctcca ctgttttcag tgtttaaact atgaagaaag gctccgatga ttttaagctgt
 29400
 atagtcaaaa ctacagtgtc ttctaggaat ctaggattga gacaagatgt aaaaagagga
 29460
 acatgtaata tattatttta cccactggc aattagaaat cgttagccaa agccaggatt
 29520
 ggggaggata ttcgaaaagt tattaatgca gacagagcaa ggcttttatg ccaccaagaa
 29580

agttatttat tataaacaat gggttttgag tcacatggac aatagctatt ttaaaaaaca
 29640
 cgaatatatc caggaaaggt gcaaaaatat aaaggcaggc aatcctagtt tgtggtgctg
 29700
 atctattacc actgcaatga ctgtcaatag tgaccctgag cagaactcaa acagttaaga
 29760
 aaactcacca cctgttttcc aaggaatfff atttttgctt aagggtgtaa gtttgatact
 29820
 aaatacggag ggtgttcaat gtcaagtga ggacatagca atttaccctg caggccagtg
 29880
 ttccatagtg ctactattca tattagatta acttaagtgg tgtataaaaa acttttttaa
 29940
 agtttaatat ctatgcattt attgtatggg taatctttac acattgcaag caatacaggg
 30000
 tcgccatatg caatgggatg acaaattaaa aagtgggtca atgtaaaaa ttataaataa
 30060
 tagcagtgat aggtagatac agcagaaatg aatgtggcaa atctgcgaac acattgcaag
 30120
 aactggaag tggtcagaag gaaaatggta gctgtctaag ctatccttta gaagattaaa
 30180
 ccagaagtaa agcatagatt ccattgaagg tgagaaaaga tgacaagcag ggagagaggc
 30240
 ctgggctagg atactgatgg ctgaaagaaa agaaaaagat tatcagaagc ataataaaga
 30300
 agacatgaag actccctgga ctaagggaat gaaagaagcc aagtaactta aattaatttt
 30360
 gaagttttgc acctggaaat acggcactga aaaaaacaga aattcaaagg ggagtttggg
 30420
 gggcagtggg gatacaggtt gccctctggc ttatgctgga ttgaaactga cggcagacag
 30480
 ggaagcggtc atagcctgga gagagccact atgcaggaga tgagaaatgt agatcacagt
 30540
 agagatgagg atttgagacc ctggggtaaa ttgcaaaaat ggccacaatt ctccacacct
 30600
 ctgtctatcc atgcccttct gtaatgtaat tttgtggctc ctcccatcaa gaaatggatt
 30660
 ctgcgtcttc tcccctagaa cctgggctgg ccagtgact tgcttttagcc agtagaatgt
 30720
 agtagaagga acagtgtgcc atttctgggc ctggactcaa gggttcttgt gtactttcct
 30780

gctgtccgaa atactgccat catgaacaag cccaggctgg cgtgctggag gtgaaagata
30840

catgccagag aactgagggg ccccggtgga cagccagcca acctcctgaa gcaaaccctc
30900

cctgccaatc tgcagttgac tgcaaacaca taagagaccc agaataggat aagcactgcc
30960

cttctgagcc caacctaaat tgctgacatg cagaatcaga aacaaaatat gtgtttttga
31020

agtcacaaaag tttagaggcg gtttatcact cgcgatagat aactggtaca gacaacacgg
31080

gtctgcctag aaagtgaaaa acatacgaag gcaaaacagt aagagggtcaa gctttaggag
31140

gatattccca attcctagga agaaggggga aagctgtaga caagacaaca gggcaatggc
31200

aatcagaaac attcaagtat ataatatcag ggaaccagga acatgggggc aggggacaca
31260

gaggataaaa gtgggggtag gagaagtcaa ggactcatca gggaagtata gcaactgggct
31320

tagcagagaa tctgcctaga gaaggcactg gtatttttgc atgaatgatt gaagtttgat
31380

ctgtcaaggc agagctcgct ggagaagagt gagaagacag tgtcagtaac gtaacaagga
31440

atagaagtgg tctatggagt gttaaggagg gcctgggtgg gaaagatatg gaaccctagg
31500

taaagattac ttgtttgaag aatcattcag agagatgttg aagacgaggt caacagagtt
31560

ggacaaaagt gtaaatgtga atggtccttt ctctaggacc ctacaatcac aagtgtggct
31620

gggaccatcc atgccagcat cacttgggaa tttgttagca atgcccactc tcaggcccac
31680

cccagacca ctacctgtct tttatccaca tccccatgtg ccattcaagt ttgagaagct
31740

ctgctctagg aagtacaaag acaaggaact gaagagcatt taatgacatc ttgagaaagt
31800

cattctgttt tgaattccct ttcaatcccc cagataaaat aacacaaaagt cttgccttgg
31860

tactaagcca tcttcgacac ctccctaaca cttgcaaadc tcccttttca acagagttca
31920

acagagagca gagccttagg caagcaaggt ctccagttag atttgaaaaa cactgtttac
31980

acccccttttg tgtttatttt taggattttc ttttatgtgt ggaaagcaat cagggttttcc
 32040
 atttgagta gtcttaattt cacattaatt tattaagag aaaaaagag gtaataaaaa
 32100
 atatgtagca tataataatg tggatgcaaa tgagaagaca gcaaaagttg ttacaagtga
 32160
 ggtggcactc aaagcgaaga ctggaaaatc tatagtggaa gtcagaagac atctgggcac
 32220
 atttgaaggt agaaaggcaa gagacaggaa agatgggagc tgtttgaga acaagaata
 32280
 ttcaacaggc aggaaggact gttggaagga tttatcccag agtagaagct gccctctctg
 32340
 ggcctggaag aaagaaagat ctggccctg aatgggtgag gatgacaaat cagagatcta
 32400
 aggagataa tatcaaattc ttcataagta aaagattgga ggtccctctt aaaatgaggg
 32460
 taggttggtt tgactgggac tgcaaaagag tatgaaagac ttggaatatc catttgggac
 32520
 aatgaccaag agaacaaca gatgaaaagg attccgcagc aaaggcccag ggacagtcag
 32580
 gcagcattaa tttttagtaa ggaaaatgga cacagtttat cccaaagctg agaaatgaaa
 32640
 ctctaaggct gacacaaaaa atacaacagc aaagaaaggg ggagaggag attaggtagt
 32700
 agtggcaagc atagctgaca agtcctcata gtcaaggta catacaaagt caaagggttg
 32760
 ctaggcgagc tggctcatgt ctgtaattcc aactctttgc agggccatgg cgagaggatc
 32820
 acttgaggcc aggagttaa gaccagtctg gtcaacatag tgagaccca tctctacaaa
 32880
 aatattaaaa attagctggg tatggggtgt gcacttgagc tcctagctac tcaggaagct
 32940
 aaatcaggag gtttgcttga acccaggagt ttgaggctgc agtggctatg atcacaccac
 33000
 tgcaactccag ccggggcaac acaacaagac cctgactctt aaaaaaaca aaaaagagtt
 33060
 aatggaaaag gccagtagac tggactctgg gaacaagatt ctcaaggat atagggatat
 33120
 aagaagtgag aaataaacta ggtttgatca aaaaggaaaa acatcatgtt tttaaaaaac
 33180

tggtttttaa aatgtcaact aaatattaac tagtgttcaa aattttgtta ccaagagttg
 33240

ggacttgata aatggcacta caaaaactca ggaaaaattt tgaaaagaag agaaatttgg
 33300

gcctaaaatg caggtaatca attaaaatta ggacataatg atgtgtgggt cttgcaaaga
 33360

acataaatca ctgggtattt taagtgtgcc caaaataaaa tcccaaaact gaagaaaggc
 33420

ctgaaaagat ataaatctta ttacagtgtg tctctctcac caaaagagga tcatccaaga
 33480

agcacataaa aactaaaact ttgaagttaa ggaaaaacaa gaacactaca gagtatacat
 33540

aaaatttcaa acttgagat acccaagaag ggtttctcag acctcaccca tcgctcactc
 33600

acacacagca cacttgctga actacctggg gtcccttctt cttccactcac tccttccaca
 33660

gtaattgaga gtttctactg gcacatggcc agccacaata taaagtacat ttcctgagta
 33720

cccaagtgtg gccatgtgat gaaggctagc cagttagata taagcaaagt ggcatgtggc
 33780

agcttctgga aacctcgtt tgcccttgat ccccttcaat ccccttcgtc cttcttctgt
 33840

gtcaaacat ggaggttgac tagaggtcaa ccttgagcca taaggataag ggtcacagga
 33900

tagctagagg cagcttggt cttcagtgtt taagaagcct ccatactagc cttggactac
 33960

ttccctctag atttggttag ggaaacaaga aatgagtatc tacttcattt aagacgttgc
 34020

tgttcaagat tttttctatt acatgtaagt tgaaccctaa tgtaccctct ttgctattat
 34080

gggggcctcc tattccttga ccatttcttc aatggcctga atagttggat taaaagcagt
 34140

cacttcaata ccacatttcc ctcaactcct gttctctcat ccctgcctag gcaaaatcca
 34200

ccaccaagtc agcttcataa tcattgttct cttcctgcca ccaacacatc aaatactgat
 34260

aaaaaaaatt aaaaataaaa tgtaaaacta aacacttcta atagggatca tcacaactgt
 34320

agtccaactc cttgaaaatg gacaagtgc ttcactgctg cattcctatt gcctagaagt
 34380

gtctgaccca tactatgtca acacatattc actcaaatac taaactacaa ccagatcttc
 34440
 tgctcaacag tatttttatt catctctagt tcacttttta ccacatttta tgtgccaggc
 34500
 acttctctaa atgtgagatg taccatgtgc atcaaagtaa atacaccccc taccctcatg
 34560
 aagtttacat ttcagtggaa agagatcaac aagagcagat gaataaatga ataatacaaga
 34620
 aattcaaata gggtagatgc tatgaacaaa attatgaagg tagagagata ttaattagag
 34680
 agatggggaa aaggatattt tagaaattga gattttttta atggtcagga cataactgaa
 34740
 ggtgatatct gggctcagac atagggtgctg agaagaggcc agccaagtga agacatttgg
 34800
 ggagagggct ctagacagag agaatatcaa gttcaaaaag atgggatgaa ggctgggtacc
 34860
 gttaaaacac ataaagcgag ggtagagggg aggagtggta agaaaatgag gtaggaaaag
 34920
 tacacagga tagatctcag gcttgtcaat caattcccat gccatacat attgttctct
 34980
 ctctctctcc tctcttcccc aacatctttt ctactgaaat agatatacac tatatatata
 35040
 tatatatata tatatatata tatatatata tatatatata catgtgtata tatagaacaa
 35100
 gatatatattc tcactctata tccatacagc cactgagaat cttattccca tagtccaato
 35160
 cactggccct tttccttaac tcttgacatt tttaagagtc aaggctcttg tgtgtttccc
 35220
 aggctggaag tgcagtggtg tgtgtgtctg tgtgtgtgtg tgtgtacata tatatgtata
 35280
 tatacacatg tgtgcacata ttacatatt ttatatatat atgaatcttc agccatcttc
 35340
 cttttctttc ctattatcac aaagaaatag ttaccaaggc ttcatttaag ggctaactct
 35400
 cttttatcta cagtgttcgt tccaacagca actgaccacc caatccatca ttcaatagct
 35460
 ctctcactct ctctctaato tcttcttcaa acatgttcag acattgtctt gaaaatatct
 35520
 tctcattagc ccacttcaa aatgtcacc tctacctata cacacatata tgtatgtaaa
 35580

cacatacata tgcaaatgta ccgagacaca cacatacaca gtttactagt acctactcaa
 35640
 tacaacctag gatctttcac cagacattct agtgtctcca cagtttggtc tcaaccaacc
 35700
 acatctccca ccatggaaag agaaaatctc atatccatag aatatccagg tgacctatag
 35760
 gtttcttgaa ggcagaccac atccaatctc agcctcaaca agaagggcgt tcctaataata
 35820
 aacaatatat atgatattta cattacataa gcacaggctc attttggttt ccacaaagaa
 35880
 taataataaa catagacaat ggaaatctc agaaatcttc aaaatcatca agcttctctc
 35940
 cagttagccg tttcaatgga ctttgctttc ctatcctaca cccacattct ctcatgtgga
 36000
 catcagcact ttgtttttcc tccaaagaac cactgctcca tactcttagc tcctgtggct
 36060
 tccaaagaaga tgacagtgtc cagctcccc actggtatgc ctacgtggct acgagcctgg
 36120
 gcatgggacc caagtctatc caatgagact tgatccagaa aagctgctga aacatgcaag
 36180
 aaagaatctc tccctcttga agaggctgag ctgtgagaat gtaacttgga gctgctgctg
 36240
 gacaacttcc cactacttga gtacagcctg gctgatatga taaaaaaaa aaaaaaaaa
 36300
 aaagaaagaa aaagaaaaaa acagatgaaa gcagagctaa tataagaaag tcttgatagc
 36360
 atcccttgag actctagatc cagctatgcc tgaaacaaaa cttacaatcc cttttctatt
 36420
 taactcagtt tgactgaagt ctttttctca cttgcaacca aagcattttt actaaaacaa
 36480
 tcattttaac cacttttcca ttaaacattt ttattagaat ttttaataa gtcaaataat
 36540
 aagggttgaa tctctcatta aagataagtt ggcacaaagg agcagcacga tggatctacc
 36600
 gagtatattt tattaactgt ttgatatttc tagatttaac tcatcccctt ttgaagaaga
 36660
 tcaacttcca accaatcata catatttaat aaaagttttc tagtcaggtc agcattttga
 36720
 acttctactt cctacaccta ggagattttt taaagaaact gtagaatttt tattattcta
 36780

cagtaaattcc cttgcttatt acaaattagc tgtagaaatt gctggaaaaa atctcacaat
36840

ttggtacagt tatactacat ccatttaaaa taaattgaat tttaaaactc actcttatca
36900

ccatctttga cataattctt ctgcataaac aatttagtca atgtagataa ggctacatta
36960

aatttgcaac atttggaata gaaggacca gcttaaaacc tcactcccc aaaagaatga
37020

gtatatatca ttcaaatacg aatttagagt aaaaactcaa atgtgtcttc atattagccc
37080

atgtgtctat cttcaaattc ttcttttaca aacaaagcaa aattgaatta agattaaaga
37140

tagaatgtta atcatacaag aagggccatt ttccagggcc tggcatgttt tacacaagca
37200

tgaaacatat ttaagagaac atggcctccc tgaggaagag attaaggata caaacctagg
37260

aaacaaagaa aataaaacga tctacctcaa agggataaga ttaacctgtg agaagctcat
37320

gcggacctta aataaaagcc tctactcctt atatttttaa acaggatttt tctttttttt
37380

ttagagcact aattccacct tcaaaatata tttgagtttg taagtatact caacaaaaat
37440

cctgtccttt cttattaacc atactcaaaa cagattctct aaaatagtta tgacaccatt
37500

aacatttcat gaaatccaaa acaaattgtg cttccaaaaa aataaatgtt ttcattcata
37560

aatggctaatt atgggtttta attatgttat cacaatttt gcattcattcc cccttaataa
37620

ccatttaaga atgtttcctt ccaaatagct attaagagag aaaggatgtt tcatcatgca
37680

tacacacaca taaacacaca ctcttggtt atgtcattgt ttttaatggg gcagtaggat
37740

ttccaaattc ttcaagtttc acacattcaa gaatgtttcc ttccaaataa ctattaagag
37800

agaaaggatg tttcatcatg caaacacaca cataagcaca cattccttgt ttatgtcatt
37860

gtttttaaca gtgcagtagg atttccaaat tctctgttcc acatgttact tectaaatta
37920

ctgcctcact gactgggaag ttgaattggc aacaccagcc aatcctcagg tccaaggatg
37980

ttgtccttgc atttcctgtc tccttcctct tgaaggactc aaattcattt cattcatctg
 38040
 ctttactaca aacagaaaga ctccatcttt ggttggtcag ccccaaactc tctcttgagt
 38100
 tccagtgcac gactctgtga gaatatacca ctgaaatata caaaatgagc ttgatctggt
 38160
 tatctttcct caaaaccagc ttctcctctc tccgctctgc ttccactcaa ggtcttacca
 38220
 ttggtccagg ccccaaattc aaacaccctc aatcggttct gcttcctctt tctgccttcc
 38280
 agtcgggagg ttctgtttag tgctttcttt tcttttgcac ccaaattcac atctccttga
 38340
 ttcaaactct tcccatacct tacctcaact gggatgatgc atcctactcc cggctctctc
 38400
 tactcttgcc ggggcctaca ccattctact caaaaggcca gatcaatctt tacaaaagta
 38460
 aaactctaac caatccaccg gtccacttaa aatttttcag gctgggagca gtgggtcaca
 38520
 cctgtaactc cagcactttg ggaggtcaaa ggcaggagga tccactggcg ccaagagttc
 38580
 aagaccaacc ctggtgacac agcaagacc tatctctgca aaaaataaaa tgaaaaagtt
 38640
 agccaagcat agcgaacaca cacctgtggt cccagctact caggaggcct gagtagctgg
 38700
 gaccacaggt atgtgtgcat ctcttgagcc tgggaaattg agactacaat aagccatggt
 38760
 cactgcactc aagcctgggt gacacagcaa gactccatct caaaaaatat ttatgtgtat
 38820
 atatatatc ctggaatgca ctgagttgca gacaccgtg actttccaac agccttcttc
 38880
 cttgctaaca gagcataaca ctaattaatt atgctcacgt agcaatgccc agcctcaagg
 38940
 gatattggat tgggtctaaac tgatcaaagt aatcttcctc ttgctagtg agtagcctag
 39000
 ggttggtacg caaccattt ctgaccaata agataaaaga gacggtctag tggaaaagct
 39060
 tttccttctt gccaaagat gtacaggaag agaaagtact catttgtag cactcccacc
 39120
 accatcacct tgcttcttat ttgtgtgtt ttcattgtgac gatgtgacac ccagagctac
 39180

agcagccaac ttgtagccat gggaagactg ataaaaactcc aaggatagca gaataaaaaag
 39240

aaaatcaaag aagggtcttt caagacatga ttgaaccatc caaaacaacc tggcactgcc
 39300

aacccaaaga cttctttgaa ggtaaacaag tgtcttttat ggtttatgcc accatttggt
 39360

gagttttctg ttctctgcag cccaaagcac tctaactgat aaacttagaa agagcaaact
 39420

cttgacctta gaaggccctc tgcttgatgt gctactcaaa gtgtagtcta cagaccaca
 39480

accttagcat tgctgggag ccagccctg ggcccatgaa tcagaatctc caagggtaaa
 39540

gtccaaaagc ctgtgttata acaagctccc aggtggttct tgtgcatggt aaagtttgag
 39600

aaacttgcc ctacaaaatg atcacaaatt attttttacc atctcagcta ctaccgtaca
 39660

tagcatgaga tgcacttgat ctgcaatttg tctggtggca gtaacctttc cgggtttgag
 39720

catcatttgt tgctccttta gtttcttctc ctcttaatca aaatcttcaa acttgaagcc
 39780

ttcaggtcct gagaggatta caaaggtctg aatatgagac aggctgtgaa acaggccgtg
 39840

atgaactatg gggtagcaggc acatctgaag gacaccgagc agatgttgcc atacaaggat
 39900

attggttcag tgttgccaga tgtctaaaga aaagctggaa tccagattta tatgcaaagt
 39960

tttttagata tagacaacca agcctttttt aatactctgc agattaaaca aattttcaag
 40020

tggattcaca tccagctgag tttgtgacat gaagtagatt ccttgaaata tcgacacata
 40080

ctttgaaaac tgtgaccacc tacatgcatt gacttcattt gaaatgacta aaattaattt
 40140

tctttggcta actatatgct agtctctact gagcacaggg tcttcacag tagcccagca
 40200

tctggagtgg tcaaatagtt tgtggacctc tagttaggag cagcaggact gttttatcca
 40260

ttagtcacta tacaccacac acttagaact gcaagcttct catggggcca caacaatgtt
 40320

taaaatgtga gaaaaaata aactggccac acatcaccag aaaactacaa aattacaact
 40380

aatagatgtg aagctaaata ttacaaaaat gtaaaactaag tcaacggatc catatcaact
40440
cagatatggt atcatatcac attagattaa atgcaggatg tggatgcatt ttaatgtttg
40500
atggggcatg gaaggagct cagatcaata gtgcttagag ctgccaaggt tctaaaaagg
40560
ccctgggtac caacctgaag ccaaccattc tgtggttaagt ctggcttcca caggtgtcac
40620
aaatttcttt acgcagctct actccattct attgaactgt ctacatagtt ggacatggga
40680
gcagaactgt ccacatagtt gcatgtaaac cataaatgac ttattaaaaa gttatgtttg
40740
gatcctcatc tacatatgaa ttttaagaca gaagtacaag aatcagcaca caagcaagtt
40800
ttgagggtcc cacataatga ttttaattat tttttgttaa tttattccaa catttaaaat
40860
tcatgataat tcacataaaa acccagatta tcatcctctt gaaaaaagaa ttgtaaagta
40920
tagagatatt ggacccaaat gttggcatgg caacactcta ctaatgctga cgaaggctat
40980
ccctttaaat ggggtatgtt ctctacttgg ccgcaaacc tgtctccatg taggctaact
41040
tgctgtccca agagacaatt gagtttatgc cccctagttt aaggctctta ggaaaacctt
41100
atttcttttt cacaatgtgg cttcaatatc ttacgtgaag acctgtcgtg ataggatata
41160
ggaatctctc aatctaagaa gagtttcaat caaattcccc ccaatatgtt ctcaatttga
41220
accatgatcc ccaagaagat tgacaaatga caaattaccc tgcagtttct tttgcccttc
41280
aagctagcca aggcagaagc ccaatcaggt ggagtcaaag cactaatata agaattgtctt
41340
aattaatcaa agtgatcaat aaataatata cataataatt agaaagcaaa gtctgggact
41400
gatagaataa aaacaacata gcactaatct aagtaagggt gagagggtcaa tagccataat
41460
tcaaaaggaa aagccctaaa aatatgcaga aatagctagt aactggagtt tagtccaagt
41520
ttgttaacaa cccactgtat gatttcacct cacaccaca ttaaacacag tatcttattt
41580

tgggatggca gaattctgac aactacaggt tgagtatccc ttatctgaaa tgcttagatc
 41640

tagaagtttc agatttaggg gttctttgga ttttggaata tttgcattat acttaccagt
 41700

tcagcatccc taatccaaaa atctgaaatc cgaaatgttc caatgagcat ttcctttgac
 41760

tgtcacatca gtgctctaaa aagttttgga ttttgagta tgttgattt cagattagt
 41820

ctaaattata tttgctaatt gatattctgt gtattgtttt ctactggtg aaaagacaaa
 41880

ttgaactcag gaaagaaaaa ttacaaatat atagtcaaag agtcaccaga tcatttgagc
 41940

aacataatga gtttttattt cttaatatat agggtttaat ttgttaacca caccagaaaa
 42000

agaaatgttc aaacagtcga ctttctactt gaaattaaaa tggaaattag cccttaggaa
 42060

acagtgatga cttaagttt atacatcagt ggctgggtca tgaattatta actgatttta
 42120

atgatcctgg ctatatttta tttgaagtta ttctcctact tgcaacgact tttgaacca
 42180

caaaactctt tcatccattc tccacagaca cagtacactt tcatgagcac tgacatttga
 42240

aaaggatcca tatactcaca gaatgtaaag ataatacaca tgcatataat tgatacaaa
 42300

gactctggtc atttgtaaa gagagtaatg aatgaaacct tattattcac ggtttgtaaa
 42360

taaatttaaa gataagcgct aactttcact tgttaagcct agagccaaga gttcgggtgt
 42420

tcttctaaag tttccagat ggtcattaaa acaagagagg gaagaaacat actttcttg
 42480

aaatccatca tttcccatcc agaaagataa tacatgtcat caaaagttaa gtcttttagag
 42540

attaatat aatgtgagta ttttcaaata aaataaatgc atgtctttat atgctagttt
 42600

ttacattttc agtataaacc tatattgaac agcaccactt cagacacaca aattagctgt
 42660

gaaacaaaat gctataaacc tttaattagc acttgctgct ttatacccat aaagttaata
 42720

ctctctcttc tactatatcc tctgtaaagg aaaattatgg gaaataaaat tgtgtttttc
 42780

acttccgaaa ataaaagtag caaaatatag tattgtggct cagtttaaag atagttatat
 42840
 ttttacttgc aaggaagtea atacaaagat cccttacaat aagtataatt tgattcctac
 42900
 caatattaag cgctatgtct aactttcata acttgggtac acagagcact gatgacgatc
 42960
 agattgctaa ataagtaaga aagcaagtta tttcaggcaa atagctgcat taagaaaata
 43020
 cgacagtaag ccaacttgta tccacactcc tgaagctgaa caagcatctc aataaataat
 43080
 gtctgttttc aagctttata taccctccct gaaataaatg ttctgagctt aaatagactc
 43140
 aaaatagcct ttctcagaaa ataaaaaaga gacacacaca aaacaccccc aatttcctct
 43200
 ggtcccttga agtacccttg ggatcccagt gaacagagat ttgccacgca ttgcagtgc
 43260
 tgcatttgta caatagggtt ctattttggt gtcattcgag tgttgagat cttaaatcag
 43320
 agaaagaagg agttacatta ggcaaggtaa tggcttccaa aaataagagc atgccagtt
 43380
 tatacaaggc ttgcatccac atgctaagtt gtaaggagat aaataatacc ccactggaac
 43440
 aattaggtag attctcagaa aactacata cttaattagg ctgcctaagt cccttactga
 43500
 gcatacatgg gggacttttt aaagcattta ataacctgag taaatccaaa atcagaggac
 43560
 acaaactaat ccttaactct ggctccagcg atcggccttt tcttgctct ctataaatgc
 43620
 ttagcgtagc ccttagaatt cggagagcag gaacgctcca accaccaaag gggttaaagca
 43680
 caggaataa agtcccctcg atctatat ttggggaggc ccctctgtgc cagacgctga
 43740
 gctgggcgat gatgacacgc gagggcgagc aagagagacc ctaatggtct ttacaaaact
 43800
 tcagcactag agttgggaac gcagtataaa caaatagaaa ggaacacgag cctccatcag
 43860
 ccacggggca cagcaaaactg ttcctcactc tagatgcaat tcctcctctt cctgttattt
 43920
 attcattagt ttgttttagc acctttcttt cagcaggctc cttatttcct cctgttatcc
 43980

aagcaatcac aaacaaacta cgatttgaaa gaaaaagaga gggaggaggg aggagcagga
44040

ggaaggggaa agaaaaccac attcaaaagg cacaggccat tcagagcccg gctgtttctc
44100

tgaatctcgc cctcagggtg atgttttcga caaattcggc gctgggggct tcagggagga
44160

cagaggacac cggatgggtg gggggaattc cgcgcggccc gggtagcagg gctgcgactc
44220

tcgcccgcct cgctgctccg agcgacgccc gctgcccttc tgagggtccac ctcgccacct
44280

ggaacgctcg gtccctgaag ttggcgaggg gggagggtga gaagggtggag tgaccccagg
44340

gtgtgggtg gaggtctcca atcccgatct ccccgaaatg caaaacttag ccccccaaa
44400

aaaatgagag gaaagaagcc acttactcta agcggggagg ggcgggcacc ccggttccga
44460

ggtccaaagt gccctaacga aagggcagca aaggccacct gcgggggtccc accacttccc
44520

ccactcccc tgtccaggca tactcaaaag tttgttcctt tccgtccgta agcggccccg
44580

aggccggctc tggagtaggt gggggaggta agaaggactg ggggtggggg agggcaagat
44640

catcggaaca ggagatcagg gaccctgcg tagccccac tcagtccagg ggggctgcga
44700

tcgtgtttac taataacagc aacaaacaat cacacgccag cccttctttc agtcctcggg
44760

ctctcttcgc tttttggggg cagcaaagga gagtctggcg gggagggtggg agggagcggg
44820

ggggctgagg agagcgccgg ctgggggaca aaaagcggcg tgggatgggg cagaaagaga
44880

agcggggaga gctggacaga gacgaaggct caaaggtaga agagacatgg gaggaagag
44940

gccaaagaaa agtgcagcag gcggtgagcc cgagcgcggg gcgcgcccag gatggggagg
45000

tgccccccgc cccccggccc accggcccag cccggtcacc ttgctgtagc cgtagtaccc
45060

caggcactgc gcgaagtcca ggctggcagg gtccccggcc gccgcggggt agaacctcac
45120

atccatgccg aagctgggccc cggggccggg ggccgggact ggggttcgcc ggggccggga
45180

cccgcctcct cgccgccgct agatccaccg tcgagggcgc ccgggggtgg cgcgtgggac
45240

tcgcgggccg agggggcccg ggacccagag cccgaggagc tcgggagccg cggccgccgc
45300

acacaaaggc gcggccacgc gagccgcggg agagcgggag gcggccgggg ggacgcgcc
45360

cgccggggca ccgaggcagc gctgcgcgcg ggccggggcgc cgggggcgcg gggcgcgcg
45420

ctggggcccg ggtcggcgag gcgagttcag gtgcgctggg cgaggttggg acggcgcgcg
45480

cggcgggcgc tggccccgct cctcctcctc ctccccgggc ggactgagga gacgagccgc
45540

ggagacaagg ggccccgcc ctccccctct tctccccctc ctgcctcagc cgccggtccc
45600

ctccccgcgc cgccgccgct ccgcccctcc caccgcgggc agctggcgcg ccgccccccc
45660

cgccggtgcg ctctcggcc cggaaccgct ccgggagcgc ccaccccgcc ccgccggccc
45720

catcaccccc gccgcggcgc cccccctctc ggggccccctc ttctccctac ctgcccccc
45780

ttcccgctcc ggtcctcttt gtccccactg cctcggagcg ctctctctc ccctctttc
45840

ctcgcccgaa gcggacgcca gccctgcgtc ccgtgtgtcc caccctctc ctgcgcgacc
45900

acagcttccc cgggccccag catccccaag ccgtctgctc ccttctgtc gctttctcc
45960

ccatgacccc ctctcttctt catgggtgcc tgtccccggg cccctctctc cgaccccgcc
46020

ttgcctttcg aatgtccct gcgtggacac acctaacca gacgccaacc tgcacgagtg
46080

tctcccggtg gccgggagcc gggaccccaa gttcacctgt aggatttccc agaaggggtg
46140

ttggggcctg acctttcaca ggtaggcaga cttgaagctt ccaggtgaa gcctcgtgtc
46200

ggttctcttc gtattctctt tgctcttgtt tgtattcgaa aggcatttga gtaactggat
46260

ctaggcagat cattcactgc tctgggagag aaaatgtccc tattttccta aaaggctcta
46320

gctgattaac ttccccactc tacactacc caccctctt tattttctta tctttcttaa
46380

gctccatctt tttatttaaat tttttttctc atcggttta ttgctttgag ttaatctgag
46440

agtttgccct tcagccttg cgaagggaag gggctttata ctaaagga acgttgagg
46500

gaggccgtga gatctgcgc cccccagtt tagttatctg ccagctactt taaccagagt
46560

ggcgacattc tgaaatattt cctatgacac tcgctaagag tttcatcccc tggtcacacc
46620

ttgctggccc agggcctgag ctttttcaga caagcatgac atgggttttc ttttcatttc
46680

ttctgttctc cagagtagat cgggcttggt acacctgtgt gtgtaccttc actcaatatt
46740

atctttctgg aacggagcag aaaaaataa acttcaatta ggtgaaagt ttgttatttg
46800

ctcttctctt agcaaaaact tgggtcagga gactactgca tgagagagca gtgtgtgttg
46860

agaaaagaaa caactgccc cgttgtccct ggggtgactc tgatccccac cctccgaaa
46920

aagctttggt tccccctccc tagtctccat ttaacgttga caaccagcg gtgtcccta
46980

atggcccagc aggttgacac aattgacagt gagctgagat attatcagga caacaccca
47040

ttgagcaagt tcttgaaacg cagaagtga ggcctttct ctgggccact gggcgctgt
47100

ggcctttcca ccaatcagag ctggagggtg cagtggccac ctgttgtctc gcaagagaac
47160

atcttttaag tttcgctatt agtatctaac aatggcctac agagcaggat ctggccacag
47220

gagaggagca gtgtgatgat ctggggtacc aaaaaagaat tttttaaaaa aaaaatttta
47280

cagatagctc ttacaaggag ccaaaactct ggaaagctaa gtttttcttt gcacactcac
47340

aatcctcacc tttgtctcat tctgttttgc atctgaagca atctgagctt gattaaagga
47400

tgaagaaaag tataatcaaa gttacagaga ctaaacctga acttgtaatc agacttcgag
47460

gatcaggaat ataaatacgc ttcgcagtgg aaaccgcta gctgcaagaa tgggaattgc
47520

aagtgggctg atgtactgca ttcagctttg ggcggtagag cctagaaaaat gaagcgagaa
47580

ctttaccct aaggctctca ggagtttaga atgctcgggtg acaatcccta actcctggcc
 47640
 attgttttgg aaaagtcttt ctcttaaatt tcccttcctt gttaggatgc aaataccttt
 47700
 gtgaaggga caatactcat ctgtctgcat tttattattt gatttaggat ataaggtaag
 47760
 caaaagatct gagtttagat tttgatttag agggagggga agaatccggg tgatgagctg
 47820
 aagatgtgta tgttgaggca tttagaatga gaaaatcact ggaacagaaa aagggcattg
 47880
 gtgaaggctg ctgacttgac gatttagccc aatctccacg catttgccag ctatctttat
 47940
 ttattgtttt tttttttttt ttttttttga gtacctacca tgtgccagac atgattttag
 48000
 gcacttgatg atgagaaaga ggattcaatt cccgacatga cacatataga agtcctcaaa
 48060
 gaggactatg attttggctc ttggaacaca aaatacacac gaactcacat gcacacgtgt
 48120
 atacatgccc tgaaccatcc cagtcacctg gagatactgt catggcacag ccactagctg
 48180
 ctgttgagag aagcgacatg aaaaacaaac tcacatacag tgtgaaggat gccagcttct
 48240
 acacccccat ggtatttgta taagaaggct caagggatga ggcagtgcac cccttgctac
 48300
 atcattgctg gatcccaaga acctcagacc atgcttgata catagtttgt ttgcaggaaa
 48360
 taaatattca ttaaagaaag aaagttaga aagaacatgg gaaggaagaa ggaaaaaac
 48420
 acagaatggt ttaatagtgt aatttatttg cagggaacct taaggcagga tttgagaaga
 48480
 gtatatccag gacctctgat gtgaattaga acaatctact ttgggtgggg tgggggcctt
 48540
 ttgcaccctg ctgttgtaga cccttcgcag acataaaaac taagatccta tgtggtttcc
 48600
 tttagaactg atctgattac actctccatt gctttcgcac tatcatctaa ttttaagagt
 48660
 tccagtgggt ttgaagacta tccattaaaa actttagaat aatgctttaa aaaattaaat
 48720
 gcaattgtaa ctgtcatggt atttgttatt ttaagggtgt cactattttc agtctgtcag
 48780

cagtcttgggt tctgaaagcc atagacaaag gtttccttta tactggatta gccttgaatt
48840

tgaagaacct tctccaagaa catatcttaa atatgctttg aatttatgct gaacaatgct
48900

aacagttatt gatttagaag ggcagcagaa gaatcaattc tgcagccaaa atacaacttc
48960

tgggcataag aaaattcctt tagggttttg agatgattcg gatcaaaact gtccatctta
49020

ttataatgct tgatgtttct aattaaatag ttgatgtcat ccaagaatcc cttaaaaatg
49080

gggagaaaa acacaaactt tacattgtgg gaaaactacc ttacatcttg acggaacctg
49140

aaaactttct ctactttatc tggtttcaaa ttatgcaaga aagtcccat gaggacttca
49200

ttaaacctgc cagagtttta taacaaaaaa cctcactgga accccactt ttacgaatct
49260

atgaaagtaa ttttaaccct ctgtgttctg cttgcccaag atcagatgca aatactatta
49320

aacagtatat acagtttaat tctcttgagt tttggtgatg aaaaatccta tccaaatatt
49380

atgatccttg caatacgtaa tgtaaccaa ttaacgatag ctccaggtgg aaattaacaa
49440

ggccagggt actcacgttt cctctcatct attoccatag taacagctaa tgtattatat
49500

tgctaaaata tcttatattc agttttatgt ctgaacttgg tgttgtgtca taaagcgtag
49560

cattcctatc cagactggat tcttaatatg tgtatcctca tgaaatttag gagagtttcg
49620

tgtctgcaaa tcaacacttc atgtgcatag tctaaaccac atgcagataa aagcaagtca
49680

gaacctacca agactctcac catagtgaac ctacacatgg tatagaggaa gaattcttac
49740

ctgggagacc tgatgacaca ttaccagtgc aggacctaaa gcgaatcacc atagagagat
49800

ccacacagga attatagtgt gaaaaaaaaat cccaatcat agacattggt agcttcagtt
49860

ttcagtgaat ctaaaactct ctagtatgaa aatgtgagct attttgatat ttttactttg
49920

gtcaaaaaaa cacagaaatc caaaaaatac agaaatctgg gaaatgggat aaaaaaagc
49980

aaacttttcc ttcattgataa taaatatata ttcagttgtg tacatgaggt tgtttcccat
 50040
 actactacaa ttttagtcct attggtctct gtattaccag gcacaagaag ggctctagaa
 50100
 aactaatcat tttgtcagtt gactattact tttacaaaga atagataaca cccataggtt
 50160
 ctgacagata catagtaatt atggaacagt cacctccaac accctggtag caaaaaacta
 50220
 ttaagccaag aattcattac cacaaatatg gaatgcagca agtatttgct tgtgacctaa
 50280
 gtagctcttc aataaagata ttttagaaac caaaacactc cacacaatca ggaggtgtgc
 50340
 aattttcact gtattttacac tctgccttaa aaaaaaattt gtatgttgct ctgcctgttt
 50400
 tctgaaatat attctatttc ttttaaagtg aaagaatgac ctcttttgca ggaacgattt
 50460
 tcagagacag cgtgtttatc aaatgggttc tctgctttca taaaacaat ttattttcac
 50520
 tgaactcttc acgataagca tcaaagccag gaagaatgta caatgaagtt tacaatttca
 50580
 ctccagaaac aaatactggg tagttattat tttgcttaag tgaaaaacat ttttttccc
 50640
 caaggagaca aaggtagtaa tgggggaggg gggtttataca ggagtgaaaa tcaggaagtc
 50700
 cagacagtct tcagcaggag gggccctggg tgttgccac aaagaccacc ggctgaacaa
 50760
 tgaaagctga agaccagta cttctctcgc aaatgggtgg gggtcagtcc acagttttat
 50820
 tcttcgctaa gagacagcta tagaggcatt aaggagagaa aatcattagg cagatataat
 50880
 taaacccttc tatgacagat ctatatacca ttgcacatag atatctctca tagttgggtgc
 50940
 tttggatgac aatcctttct ttgataaaca tccatgtttc taggcttaat ttatatgaca
 51000
 tgttattttt taataaccat agaacaataa aattgtaaaa ataaccaaga aaagaactga
 51060
 gtctaataaa gattctgttt cttagaagaat gggaaacaat tgcatttgat cacagaacag
 51120
 ccaaggacta gcagcaaaat aatctctcct aatttacata gatctgtgcg tgccaaacag
 51180

atgagtgttt gcttttcgag taatagtga ggtgcgtatc ctgcagtga taaatctcta
51240

actcatcaaa ctgtttcgcc agaattggag atgccctctt tttctggaac agtttcaagt
51300

tttctcccca gcattccagc cacctgttta tggagttcat cctcgggaga agatgcatgc
51360

aattagaaac aagaaaatga aaataaatga ccttgtttgt tatttataac aaagaaactc
51420

tgcaaaatta atgctgtgaa gcttgacctc ctgacccagg aaaaagagcc acctcatggc
51480

tttctggcca atggacaata gtctggcagc tcccattgtc cagccaccct gacttttggg
51540

tggtcattaa agccacatat ccaaaatatt aaatgtttta caagttagat tagaggattt
51600

gcacttttta aaatgcgtgt tttaaaaact gaggttagtt agcacaagag attttctttt
51660

taaggacagc aggtaaaaat catttagctt catctactcc atctatcctg tgtcaacaac
51720

atgttgaagg aaaggaaaag atttaaaaat tctgaactag tgctttggga attgaagaat
51780

ggttaaatat ataatagaaa tcaaatagaa aggtgggtctt tgaaagtggc tcagggaaaa
51840

cattgtgttt cctccctgct gaggcctcag ggaggaaaagg gggaattaca atgttctggt
51900

acataaatgc agtgtggatg gtgcccccta gagttgtgag atgaaatagc attgctgttt
51960

ttccttcac ctcttatccc aaccacaca tctagaagcc cagggcagga aggaaatgaa
52020

ggtggtaatg gagttttgaa gactcttcag aactgggtct tgggtggaac atcttgaggt
52080

aggatggcga ctgaaaatga cactctctgg tgtggctgat tctgaagctg gcctctccta
52140

ctgcatcttg cctctttccc catttggctc cctccttcat tcaccacaa caaattccat
52200

gagtgtcttc tattagtcta aaccttagag taaaatccgt gtttgatttt ttccatcca
52260

agttgcttat ttattggcca atctgccag tgctactaca tctttaattg ttttattttt
52320

cccaggagg aagtaactgt ttattcaaga aggttctcta ctgttcctca aaatgtcttc
52380

ttactgtga aataaagtca gtttcataag gcatgcagat tgtcatgcaa acaccaccct
 52440
 gagcattcta gatcatacct cctcaccttc ctttttcttt ctgatccctt gagggtagct
 52500
 cattatatgc cctccctac ccttcttgg gaatgctgca gccataaatc tggagaggtc
 52560
 attagaatta cgaactctgg ttggtgtctc acatacctac aacagcaaca acaacaacac
 52620
 aattatagcc actgtagagc caatgtcctg gttaatgttt gtgtaactgg ccaattctca
 52680
 ggcccagaga aaggagttaa agaaagtttt tgcattctac cagtctaaaa cttcttctgc
 52740
 agtccaaaac cacatttcaa ctattgttta actgaaagac tgagccctct taatctggaa
 52800
 ccaccaacta aattccccct gtgtgcttaa acaccaagtt agagagctgt gttcagtgc
 52860
 agtgagaaaa agtagagtgt gattacccca tttaccacac agctgaaacc tgtgtcattt
 52920
 atttttcaac aaatgctgac tgtgccttcc taactatatg ctctacgtt tccagaaatg
 52980
 ggaaaaagca tgatccctgt ttatactaaa catgacaaga catgaggaaa gtagtataat
 53040
 aatgaaatat tgtaatgagt gccataggaa aagcattctt aggctagtgt agtcaacaaa
 53100
 catttattta cagaaattgt tacaggtaat atgttagtgt gctgcttaga aaaggcggg
 53160
 atcatttatg atcaggataa tcagaggaaa cttcatggag aaggcatttg aaagtagata
 53220
 gcattttggt gaggcaaaag ggatgaaaaa atacactagt ggaggagat tttgaacaca
 53280
 gttattagag agaaaatatt ccatgtgtgg tgggaatttt gggttgcatc agccatttgt
 53340
 tgggtagatt aagcttcagt aacaaataac ctttaatatc ttagtgggtt accagcacag
 53400
 aggactatct ctcagccccg ttccagtcca ccacccatca gctgctgttc atttgacat
 53460
 cttcatctca ggaccaaagc tgaaaaagaa gcttctatct ggggcatgtg ggtctttag
 53520
 cagaggaaaa agagagagtg gaaaccgct acagtgtgca aagctgctca cctcattccc
 53580

gctcactttt cactgggcaa agcaaatacag atggccaagc cccatgtcaa tggggcagaa
 53640

agtatgatcc tcccggaggg actggtctgg caaagagtaa ggaggagtcc cctattttgc
 53700

aataatgtag tcggtgccat ggaaagagtg ttttagagac ctacccaaaag tttcaggcag
 53760

caagcaaaat tgtttagtaa aaggagttag tgagtgttga ggggaagagt ttggtgagat
 53820

aagatgccgg gaagtgatag ggggaaggag acaggctggg aagtgaacaa ttataaagta
 53880

ttgaaggctg ggcagctgag ttagattcga tgcaaaagga tgtatgatgg catgaaaaaa
 53940

cgcagctgtg agttctgac taaagtttgt tttgacccat atagagcaga ctcatTTaac
 54000

atttaattag ttatatactgt gtgtcagttg tgggtggcagg tagacatttg catagatcta
 54060

tatttttaaat ttgatccctc atgacacctc tctgggttag ataagtatta ttctctttgt
 54120

taagctgaag gaggtccacc aagggtgaat gacttgctta aggacacaaa agcagcacac
 54180

agaactgtga tgcaaaagta ggtctttcta ctgcatcact gcagggcaaa gaaattaaaa
 54240

tgtcaccttt ctttagtatt gacttggtat tctttatctt ctctttaaT acacagtctg
 54300

aaagcaaggt taatatgatg ctttttctca tattaattaa aatgatgttt atataactaa
 54360

tccttgcata gtacaacttc ataagtgaat aaaaaagaga cagatcttta actgaggtat
 54420

ttattataaa atttcgtggc agagacagtt acttctaacc aaatacagaa tgctctccac
 54480

tatgtatttc ccagtttttc ctgctggttag gtgaggagca gctgaggcaa taaagaattg
 54540

gtgtgcttcc tccatatctt tcttgacctg ctgtggtgtc cttggaagac acatgcacca
 54600

gatggaagaa gacaattcat gcctcagaat tcacatgagc gaaaccatgt cttgatggta
 54660

ttgcactaca agatatccag tgtctatttg ttacagcagc atagcccaac ctatcctgat
 54720

taatacagat ccttctgact ttaattgata caaggaaagt gaccatatct gtcttgacaa
 54780

tcctcttata cctactgctg aatacaggac ctccacata atatatgctc aataaagttt
 54840

gattagtgga aaatgaatga atctcaaatt gagtctggcc tctgtctact atttggttaa
 54900

accacttatt tactgctttt agaaaccagt ggttgatcaa tctttttgat atatcatttc
 54960

atcagatcac tcctcagatt aagactttat tatatattta atcaatttta aactccttcc
 55020

tttgacattc aagggtccca taatctcacc tagattccta caactttacc aatttattgc
 55080

tctaatacaa tctaacctat cactactggc ctgtgacccc tggtcactat attcactggt
 55140

ctagggacaa acatttagtg ttcttaagca aatcatttca agtctctgag cctcagtttc
 55200

tcattctgta aagttgttat ttgtagcaaa agacataatg cataaaaaag tttaagttat
 55260

gcagcattat gaaaatccaa atttattttg tacttggtgt tggatatatat taaataaatt
 55320

gatagatcag aggagtgaca ttaaaaaatg tgttgccatt tcagtagcac atttgcata
 55380

atctcccatg acaatcttgc atggaattgc tgtctagggg gaaatggtaa aggaaaccag
 55440

ggagaccctg ttggttatct actccaaatt cattccctct tttttttctt aactctacct
 55500

ctattttatt taagttctgc cttctctctt gggaccatat ttatctaggg aggatccttt
 55560

acaatggccc tattttctact gccaatgatt actgtagtca tgaactctga tgcgtttctg
 55620

gttgataagt tgtggaaaag agtgtaagag ggcttctaag aagggtttct gtgcttttaa
 55680

aaagaaacac tcagaaaata tataataatt tcttcttctg gacagtgtca tgtctcgaaa
 55740

tgttgagccc agctcatcag catgaaagga tctctagccc gaggtcagag ctgacacgct
 55800

gagatgagaa tggaggggca gaaacacgga aagaacatag atccttgatg atgtcaggag
 55860

tcactgatta actaaccctg gagctccttc tgtcctgatt tctagttatg tgaaataact
 55920

cataatctta tcttttgatc taattgagtc ctagttttct gttacgtgca atggaaagca
 55980

gtgaaaccct cttagactga gtccattagg gtacccgtaa caaaatacta taggctgggt
56040

ggctttcata acagaaattc atattctcat gggtatgaag gctggaagtc ccagatcaag
56100

gtgcagcagg tttagattct cttgaagcct ctctccttgg cttgcggatg gccaccttct
56160

tgctgtgtcc tcacatggcc ctccctctgt gagtgcacat ctctggtgtc tcttcctctc
56220

tcataaagtc acaaggcaga tttgattagg gccacccta agggcctcat tttaacctaa
56280

ccacctcttt aaaggctctg tctccaaata cagtccacatt ctgaggtact gggggttaga
56340

gctttgaaat atgaattttt tttttctttt attttttttt tgagatggag tctcactctg
56400

tcgcctaggc tggagtcag cagtgtgac ttggctcact gcaacctccg cctcctgtgt
56460

ttaagtgatt ctctgcctc agcctccga gtagctggga ttacaggcac atgccaccac
56520

acccggctaa tttttgtatt tttagtagag atgggtttca ccatgttggt caggctggtc
56580

tcgaactctt gacctcaagt gatccgcca cctcggcctc ccaaagtgt gggattacag
56640

gcgtgagcca cgcacctgg cccaaatatg aattttgtgg gggacacagt tcagcctata
56700

acatttgctc aagccaaaaa aaggagtta aaatttcaga aaaaagcctc acagactcca
56760

agtgcacttt atcaaagtca gacatccaga gaactgggaa gccaataggc agctaattctc
56820

ttcatgtttt tccctcttca tggttcaat tttacatttc tctcagcatg catgctttgg
56880

ttttttcttt ctctcacc cttctgctat acaaagctca tcatgagtgc ccgatgatgg
56940

gcttcagccc ctaagtcaca tgacccttat gtttggctac ccttagctag caaactcagg
57000

ctccatgcct ctttgtgaaa ttcttgagat agagaatctg attggcccag tttgtatcaa
57060

gtgtttatcc ttgatccaat aggctgtggc caggggcatg ggtcaggcgg tacacaaggg
57120

ctgcctcttg tgatctgtgg ctaagttagc tcaccagag gacttgaatg ggaataaatt
57180

ccttagcagc tctagtggaa gagggtaatg tggataaaat tatagtgtag tcaagtggat
57240

ttgaagcaga tgaaacagtc atacacaaag gtgttgattc attactgaca ggctgtaaga
57300

agtccttagt aaagtccac tttatcttgc cctcggttct gatttttttt ttaatgttcc
57360

caatatttta ataaacagat taggacatag aaaatatacc tacgaatttt tagttgacaa
57420

aatgatggaa acatttggtta agatagttag aaatgagatt cagactcaaa atcatcttaa
57480

agggctagaa atctgtgtag aatcaataag ataaaaatta gctcacataa atattaactt
57540

ccaaatttgg atttcagtgc tcacttgcaa agtgatggaa tggagttttc aaaatgaact
57600

cttaaataaa ttatagcact ctgtggatat gtgtctctct ttccaagaag gaagtccttt
57660

ttaattagtc gcatgctaaa gttagtgttg ctaagtaaat gttttcctag ggttgggaga
57720

tttctataat ttcagtttcc cttttgtttg cagatcaagt agttttaatg acccagtaaa
57780

taatttggag tcaccacagt taactcacca accttgaaaa tccctgaacc aagaggcccc
57840

ttgatactgt aaggaggggc taccataatt aatagcataa gtctaaccag ggtaccatac
57900

ctttccctat attgctggat attctgtaag tttggctttc ccaaattagg acacagagat
57960

aaggataaac tcactttcca gctatttctg tatgcctttt ttatttcttt caacttttgg
58020

aaatgacctt atgaatttcc tcttattatg aaatatatct ttattgaaaa ctggagtcca
58080

aacattacaa aaaatacata gcataggaag tgaaatgaaa tactgcataa ccaatgtaaa
58140

tgattaggtg actcctcttt cagattacat tcctgagcat tttctaacaa attatcaata
58200

tttttatttt aatgaatata gaatcacacc ctacttattc ttttgtaatt tgcttttccc
58260

atgtaataat atggcctgca ggctgtaatg atagacttgt tttttgcaga acaatccatt
58320

agccaagggc aactagaaaa gccagacaaa atattcgatt gatttggttt gaaggcgtga
58380

gagagctacc atgtctgcaa ggacttgagg ggctaagatt tttgagaaaa gggaaagtaca
58440

aggagtgtgc tcgacattca gtctgctttc ttctcaaggt ctttagctga ttcacaagta
58500

gtggctgaga gcagagaagc tgagctaagc ttctggtagt cttatcatat tgaagggaaa
58560

aattagaatt cagagcccta aaaggaaaag agtccgtgat aagcatccca aggtttttgt
58620

tgacgtccac caagggctat accctagaac ataggagtga accagaaata gaccaaccct
58680

cacaaagaac aaagtctagc tttgattaaa cattatccgt gattggatta ggtgatctgc
58740

ctctatgcta acagtttact aaaaataaaa atgaatcttc cttgaaagaa aataaaatta
58800

ccaagaggct caaattatct taatcttata tacacatctg gcatctaatt tcaaactacc
58860

agtcatttaa ggagacaaga ccacatgact gaagaaatag atgagagggg cggaccctaa
58920

gagtgtcggg atattggagt tgctagacat tgactttaa ataaactgtga ttacatggtc
58980

attctccact cacattactc aagagatact tagtacatgg acttagcaca tgagctgtgt
59040

atgtaatagt cctcaacaa atatttaatg atatgaactg ataatccaaa ataattggtac
59100

taatataatt gcaattttaa octaagtaat atatgggcaa ttggggtttc tgaacaccct
59160

attaagtctt actgtctgag ggctccatta aatcacctca ggttccttga gatgatttaa
59220

tgagagccac agacactaaa ttacatggag aaagtcactc ttttctattt tttttaagag
59280

aaggttctca ctctgttgtc caggctgtag tgcagtagtg caatcatagc tcaactgcagc
59340

ctcaaaactcc tgggctcaag tgatcctcct gcctcagcct cttgagtagc tgggactaca
59400

ggcatgcac atcatatcta actaagtttc taacaaaaat tttgtaaaga cagggtctca
59460

ctatgttgcc caagctagtc ttaaaactcct gccctcaagt gatcctctta cctcccaaag
59520

tgctagaatt atagggtgtga gatacctggc cctttttctg tttcctttta gtcttctga
59580

ttatgccata gagggagtct cagcttggtg ctgacgtgtc ttttatacca tcctaattgt
59640

atttgatttt tttttttttt aacagagaaa gatctcaagc ttctcagtca atgtcttggt
59700

atctctactc tgggtgttctt aaggaaaatg ctatgggctc ctttggacat ctggtaaagt
59760

ctatgaaccc tatatcaaaa taatgcttgt aaattcttac aataaaatac ctagtattac
59820

aaagtaaaga atatgttgaa atacagttcc atccaccatc cttttgccac tacagtttgt
59880

gattgttgta tttattttca tggttacttt ctacttatgg aaagtgcaat tggtttctgt
59940

ttatagtcac gacatacaat ttgcatttta aatacatttg tttaaattaa agaggggaaga
60000

tgaattaagg aaataagtaa aagtaggaaa tatgacaaag ctaaaagtgg tgtgctgtga
60060

ttgagtgacg aatgcagact taacccaaga gatgacttta ttaatgggaa tttcaggcac
60120

cctttcacat tggggactgt actaaaccag gcttgacaca agcctcactt gtccatcagt
60180

caatgtccat tgattgagca gcccaccaa gtaccagaca ctgttccagg tgtgagggca
60240

gggcattgag caaggcaatt tattctcaag gtgtttacat tgtagggaaa ggatatagac
60300

aagaggcagc ccaaaacatt caattaaaag gtaacttcag atactgattt ttgtcctgaa
60360

gggcaaaagg gtgactagca gggtagtag agagggagtg acttggggta ggtgtaggac
60420

agtagcagtg gggtttgtgg cttggcccag ggctgcatta gaaaaagacc aggctccac
60480

tgcagcagtg cccggaaaaa tgcagagacg agccaggatt tctggatcat taccatgagg
60540

aagggaataa atgtgtgttc tctttcagtt gtcctctttg acacctgggt aaatggctgt
60600

ttacacaggc agatcaccca gccgctccct caaaggtaaa cagctgcttt ctgcctcccc
60660

tgctagcaca gattggggcc ctgaaacagt gtgtaccag gcaaaagagg aggtattctc
60720

catttaggtt tccattcatt cattcaggca agatttcttg agtgcctcct aaaatgcagg
60780

caccaaggta ggtactggga acacagtaag gaacaataaa gaggtccctc ctccccacc
 60840
 tctctctcct ctccacttag aatccaagct agagggtgtgc ctcagagtct caggccactc
 60900
 ctctcatgat acagatggaa atatcagaat cagaagaggc aaagtgactt gccaagacc
 60960
 cggagctagc ccacttcaag caagtcaggt ctcaaaacgg tgccagctgc ttctattgcc
 61020
 tactacctga cttgctcaat cttggtcagt atttggttg gcacttttta gtgtattctc
 61080
 cattactaga ctattatggc ttagggaagg agaaaccaga ggtccatcct acaagcacac
 61140
 tgaaatgtaa tccaattcta agtccccgta ttacaacatg gactattgtg actgaccaca
 61200
 tttgccttat atttaccttc ttccgcaggg atatgcccga tttcagtttc ccatgttgaa
 61260
 atagcctaca gttgagacaa acatacatta catttaagca taccttccat tttttataag
 61320
 gtactatatt tttaattggt tcaatcactg tgaaattaac aaatgttcat gagcaccac
 61380
 tgtgggagct cacagtgtgg gagttccgcc ctgggtggagg tactgggatt taatgataaa
 61440
 ccaaagacaa atatatgctc aatctggtct ctgccctcat atctagtctg ggaggtgaagc
 61500
 atttgtcaaa taatcgacaa tatatgattg tgtatgtatg tataaaacta aacgctgtca
 61560
 gttgtttaat ttaaaaaggc caaaaagaag tttgctttga agaattgagat caaggaataa
 61620
 aaattagtcg gttaaaaaaa attacatata taaacataaa tttaccacca taatagatac
 61680
 ttgttttagt gcaggtcctc cgtgatgcac actaagacag gaataaacag gtaagggatt
 61740
 tattggggaa aatgcctgtg agtgaaaatt ggaggtcaag cacgatggct catgcatata
 61800
 atcccagcac tttgggaggc cgaggtaggc aggctgcttg agcccaggag ttcaaaatga
 61860
 gcctgggaaa cacagcaaaa ccccatctct acaaaaaaac atacaaaaat tagccagcta
 61920
 tgggtgccga tgctgtcat cccacctact gggaggctga ggtgggagaa tcacttgagc
 61980

cagggagttc caggctacag tgagtcttga tcacgccact gtgctccagc ctgggcaaca
 62040
 gaccgagacc gtgtctcaaa aaaataaata aataaaataa aataaatgaa aagaaaaagt
 62100
 aaaagaaaat ggtgggggag gctggagaag gccaggaaag ccatcagact gggaaagtagt
 62160
 tctgtgaaga ggagagagga ggatggatgg tcttagattc agtgtcatcc taaggaaagt
 62220
 ctggcaaggc cacatttgcc catcagaaga gtcccccttt ctcaaggatg gaactggctt
 62280
 agttttcctg ctgggcttag ctacaagctg ggaacagcca ttgggaagtg tggctctctgc
 62340
 agacatgtgg tgatggattt ggagtacagc agctggggct gtccatcagt aagggatcta
 62400
 agaattacat gctcatggcc actgcaatgc taaaatagag caaaacacaa taaaaacaag
 62460
 ataaattaga gggacatgat ctggtctggg acaaggaaga ttgtttcaga aaatagtctc
 62520
 tgtcttacat tcacaaggga aaagagcaat gttccagaaa gagtggcgag tgcaaaggcc
 62580
 ctgtggcaga aggcagcatg ggatgttcaa gaaactgaaa gaaggccaga gcagccta
 62640
 ggaaataata taggagagag tggaatggga ggctgagagg tgagcattgc tggggacccc
 62700
 tacaaggcct ttaaggagag gacattgttt taaacacaca tctaaaactc actacttaca
 62760
 tgacagacat agggcagagg gtggagtact aggcgtgtgc ctgggtttcc aggtggctga
 62820
 actgtagcat gactggcggg actgaaggag tggagaggct aaggctctga tgtcacctac
 62880
 ttttgttagg attagaacag gggtcacaaa ctgttggacc ccatggcaaa tgtaattttc
 62940
 tttttcaatt ttttattttt tgagacagtc tcactctgtc acacaggcta gagtgcagtg
 63000
 gcatgatctc agctcgcggc agcctcaaac tccttggtc aagccatgct cccccacctc
 63060
 agcctcctga gtagctggga ctacaggcac atgccactcc acccggttaa tttttttct
 63120
 tttttttttt tgtagagatg gggctctacc ttgttgttca ggttgatctc gaattcttgg
 63180

gctcaagcag tcctcccatc taggcttccc aaaatgctgg gattacagac atgagccact
 63240

gcaccagcc acaaagataa cctaaagatg tgtttacttt gaccaggca gtagtttaaa
 63300

aaagttttaa tttgttggtc acatttaaaa actggacaat ttctacataa aaatctgaat
 63360

tactcatgtc tcttaaaaaa ataacatcta gcaatggtag gccacattc cttcctgaaa
 63420

ataattagct gggaaagagt agggactgac ccttttagac acggtataaa tagcatggga
 63480

gttgatcagt aaatatttgc tgaatgaaag aatacatgaa tgaaaagtca gagccctata
 63540

ggtcagcatg gacggcggta aaggaacctg gctgagcctg aaagagaatg tgatctaaga
 63600

ttaaattccag gatatgctgg taaatgttta acagccaact ctttggggag gaaaaaagtc
 63660

ccaatttgta gtgtttgccg attattgtga tgtaaatact cccatcatga ccaatttcaa
 63720

gctaccaaca tgctgacact gaacttgag ttggaaggag atgaacaggc ataatcaggt
 63780

ctctgagat ggcccaagcc ggccccagca ctccactggt atatatgagg ctagaattac
 63840

tacataactg gaatagcaac tttctggacc atatgcctgg aacacagcag gtgctgaata
 63900

aatgtttgtt gatccaggaa ctgactgtgt tgaagccac agatgggaaa tcagtagaag
 63960

gcaggtaaga gtaaaaagaa gggcagagaa ttgggggtac agaccctga accataagtc
 64020

agaggaatgt tgtacatggt ttcagatccc tccactggtca aatgaaggca aagggttaga
 64080

tctctccaaa tcttttagagg gacatgatgt aactccatta agtaactcag tgattttcaa
 64140

cattaaaaag tgtaattatc ttttcaaact aaatattaca aataaattct gataaaagga
 64200

aatgtacgtc tgttgagacc cctgatacta taactgggtt agtattccct gaagccactt
 64260

gtagtagcct attcacatgg ggactcccca gtgaacattt cctcatatcc acactcttgt
 64320

gaaggtcctt caccttgaat ctggactggg cctatggcct gatttgatgc atagaatgca
 64380

cagatgagaa taggtgggtt tctggcataa gctcaaagaa ggcacagtgg cttccagagc
 64440
 cctgagctac tagtaagaaa gtccagctac cctgctggga aggccatgga gagaggccat
 64500
 ggggagctca tagagaaaga gagggaccct gagactacat ggagctagga agaacaacac
 64560
 ccagccatcc cagtgtccca gctgagcccc actccaaacc catctgcca ctgagcacat
 64620
 ctacaagagt aaccactgac aggccagcac aagcccagcc aagtccagtt caaattgcag
 64680
 attgtaagca aataaagcag ttgatgtttt aagggtttta gatgatctta gagttagat
 64740
 tatgttaagt atgtgagatc cctaaatagc catcaataaa aaaattgatc agtacaatgg
 64800
 agaactgcac aactattaaa agaagaaggc aaatccatgt ttaccgacat gaaagaaaat
 64860
 caccaagatt taccctaaca tgacaaagca agattcagaa gaatctgtag ggtatagtac
 64920
 gggatttgcc attgcatgaa aaagggaagg taagtatttg tgcattgctg catgtgcatt
 64980
 tgtgtgtgtg tgtgtgtgtg tgtgactata aatttgtaaa tgcatagaag atttctggaa
 65040
 agatacataa gacattagaa cagactggaa gacttgaaga agaggggaga ttgtgtatcc
 65100
 ttttgtacat attttcattg tgtatccttt tgtacatatt tttaccttat ttgcaggat
 65160
 cacttattca aatcatagaa ataacatttt aacaaaaaag tgaggcctgc ttatataatg
 65220
 ttgaagacaa ctagctacca cactatggga acacactatt atattaaact tgatgattcg
 65280
 agttttgcct aaattttact cagaatcgaa tacaagaata gaggaagaa tttaaaaccc
 65340
 atctggctaa tgcataccct tgccaactat tctattagaa aggaagatgg cagtggggag
 65400
 ggggttctta caatggaaaa ggtattggct tccctagtgc ctgaggtcgg gttctccaag
 65460
 aagcaaattc agagtgaagg attttaatga gagtaatcta tgtgggaagt ggtcccaaga
 65520
 agcaccaaaa cagaagtggg gaaaggacac agagaaggaa aggaagtcac tgaagggact
 65580

gtgagttacc tttgtgggca actggggttt cgtccactg gggacctctg ggagacactg
 65640

taggacatgt ttcaggggtgt cccatcctaa gtgggaggaa tctggagaat ttatactctg
 65700

ctcccagggg catcaattct ccagtgttct cagtcagctc tgcttgctgg ctgagatacc
 65760

agagcaagtg ctggtctggg gaagttgtca gcgtgtagag gaggtgagta ctgagcagat
 65820

acaggcaggg cactgtcagt gcatttgctt acctagacag caaagccaga gacaaaagct
 65880

cgtgtgctta tgtggcagga gagagagagc gaacctatgca gggaaagtga ttactttatg
 65940

attttttttt tttttgagat ggagttttgc tcttggtgac caggctggag tgcagtggg
 66000

ctatatcagt tcaactgcaac gtctccctcc cagtttcaag caattctcct gccctcagcc
 66060

tcccaggtag ctgggattac aggcgctcat caccaagcct ggctattttt tttttttttg
 66120

tatttttttag tggagatgga gttttgccat gttggccagg ctggtctcaa actcctgacc
 66180

tcaggtgatc cacctgcctc agcctcccaa agtgctggga ttacaggcgt gagccaccgt
 66240

gccagccaa caagttgatt ttttagtttt atcagatgtt ttctaggaga ccataggaag
 66300

ctactttgtc tacacattac gcctctgtaa aagaggaagg aagaaatttc caggggactt
 66360

cggctcctct gcacttctgc attgtccatg taaaggtgct tattgagtat ggattgtggt
 66420

gtcttgacc ttggtggcag caggaaagtc ccggggcagg aagtgagtgg tacataatgc
 66480

aggagagggg agaggaggtg cctgataacc ctgagagtct gggagctgct gcccagtgga
 66540

gctcctccat ctgtaacact ggcggaaact gtagtaagca ggctttacaa cagtgggaaa
 66600

catgcatgcc gatgtcaggg cctcccatcc aaaagcaagg ccaggatgga tccaggggtg
 66660

tgaataaatg aaacctaact cagaacctaa caaaggagac tgcgccaagg ctaagaggag
 66720

tgccctgaca gttcacagca tcttagagca cattgcacca cattatgcta tttaccaag
 66780

gagaaaattc ataaacattt ttaaaagtag attctcaaga ctagacatta ttttaattatt
 66840
 gcttattgat gattataagg taaagaatgt tggtagagcag gtagaaaaga aaggtaatta
 66900
 ctaggacctt attctgggtc tccagaaata gatcatagtg taccaacaaa tttccatttt
 66960
 tgatagaatt actttgctga tatacattga tctaagcatt gcttctggcc atgtcactga
 67020
 taatattttt ctggacaaca tgccaaaaca cagggtggaat gactgtacag tcaaaaatat
 67080
 tttttgacta tacagcagat ttcaatcaac tataactcagc tctctagaga tatgccacag
 67140
 gtgcttgttt ttactcttgt gtttattgta ttcgatattc tttctcaat gatttggag
 67200
 aagatgtagc aggcatactc accacattta tggaaggaat cattggcatg tttgatggca
 67260
 gaatcaggat tcaaattgta ctgacaggtc aggatgttag gtcaacacca accatatgca
 67320
 atttaagagt aatagaatct gtatttaaca ttaaaaagtc agttgtgttg gcacaggatt
 67380
 ggggagttgg attgacagca attgagttgg aaaagacctg tgggttttaa ttaacccaaa
 67440
 gttcaataag atccaacaat gaaacagctt attttttaaa gcaatttttag gctgcagtaa
 67500
 taaaagcctt attctcagct gaagagcagt gatgatccta atagattctg tgaggctaca
 67560
 tctggagtat tgtgtactca gttctgggtg aaagagttag caattacata caggagggat
 67620
 ttcaaagtca tgtcacacaa ggactgggta tcagcaccgg ggataaatgc gaacaaaatg
 67680
 ataatatattt acaggaataa tagtgacctc catctatttg atgagtattc aggtgtaaaa
 67740
 gggagtaaac ctattctttt tgctctatgc tttagtcagc tatggtcaca ttaataccat
 67800
 gttggcttgt aataacatct atttatctct tatgcattca cagataagct ggagtttgtc
 67860
 tgatctaata taggcttggc tgggcttggc tccaggccat gattgtggtt caggtttgct
 67920
 gcatgtgttt ctttttcttt gggaatcagt aggtgccctt gggaatatatt ttcttacagt
 67980

gatagaagag gcacacgagg gtgcatgcat taaaagcctc tgctgatatc atgtctgcc
68040

atatccattg cccaagttaa gccatcccaa atttaagggg cttcagccac tgtgaggcca
68100

tgggaaggga actgcatgta ttctattat ggagtagtgg aacaaataat ttaatttatt
68160

ccactccaga ctagcaggtc aaaatgaggg tcagtggctc gaagctatag gaagatagat
68220

tttggttaata aaaaactgga tgatcgttg aactgtccaa agatggatgc aaacaactga
68280

agtctttgtc ccagaaggga ttagacagag ttcaatgggt catttcccaa gagcatccca
68340

gatggaattc cgggtgcatt tggaaagta gagactgagc cagttagtat ctttcctact
68400

ccaacatcct caaattccat agtcctgttc tatccatac tcttcctctt gactctgaca
68460

tcctgaggct tttactctct gtgccattcc ttggggaacc taattatata tggcttcttt
68520

gtgtctctaa gagattccat gagtatttct tgtcttcta attgctttgg aagcatttca
68580

gtgtgcaggc tgagtcttaa gtttctctta aattgcctat agctattagc tcagtgtcag
68640

gcaccaagcc cttagcatgg cattaaatca tcaccaaca ccagtgagc cctcaaacca
68700

tctttgcagc ctctctccta cttttcccat cttattatcc agccgaattg agatacctgt
68760

ccttcccgat acacagcccc tgcttttcaa gtctttgctc acttatagac cattatcctc
68820

ccttgtctag accccatcca tccactcag tgcaaacc aa gtgctacctt ttccaagatg
68880

ttttcactga tgctccaacc aaaaagaagt ttctctacct cttctgggtg gtcttaaac
68940

ttctcatggt attcgagatg tttgtggtat agaaaagtga ttctccactc taactatcca
69000

tgatgacttg agaatactta aatcatacac gccagggtcc catcccaaga tagtctcata
69060

ttattgatct ggaaacagaa actccaagta atagcagtgt cagcaataaa gcaggaaactc
69120

acatagcaaa gtgtatggaa ataagttatt ccagggttgg tatggcagtt taatgatgcc
69180

accaagaacc taggctcttt ctctctgac tgccattccc aatatacag caatgtcttt
 69240
 cctcatgggt gcaaacacgc tgctatggct ccaagcatca catcctcaca ccagcatacc
 69300
 aggaaggaac aaagcaagca taggaaaagt gttcttctca catatctctc ttgctcgaga
 69360
 aagagaacca agaagtttcc tgcaaggaac aaagcaagca taggaaaagt gttcttctca
 69420
 catatctctc ttgctcgaga aagagaacca ggaagtttcc tgcagacatt tcctcacacc
 69480
 tcattggctg gaactgggtc acatgcacat ccttgaacta tcattggcaa agggaaatgg
 69540
 gtcataaaaa ttgcttaagg ccaagcagga gccattgttg gggttaaact gtcctgaaca
 69600
 aaatctaggc tctgaatgat aggcaagaaa gaagagggga tggccacaac ttctttaca
 69660
 cgggtgaataa caagagcagc ttaagtaagt agatgccctt atggggcagt aatcttgctg
 69720
 tgggcagaaa catgtccatg cacacagaag ctttaggaat cttcagagag gttccatagt
 69780
 cacttgtgaa ccaactgaagc ctctttctcc aagccattgt cttgaatgct gtttttaact
 69840
 gattacaatg catgcagctc ccgtgcccc cttatttttag cttttgagga tgttcatttc
 69900
 cccggaatct ccaaatataa atcaggtttt cacagaaaac cacagagccc tcacagggtg
 69960
 ggaaatcatc taaagatatg ctaatgattt gggtttctag aaaaatgtat atctatgaag
 70020
 gaggcagga ttaagcacag ctagtatgta tgatttttcc ttttcctttc atcacctgta
 70080
 ataaggcagg atggagactg aatgtatcag acaaacacagg gatcaaatac tagtttcaac
 70140
 atttgctgcc tgtttgacct tgggcaaatt ctttaacctc tctgagcctt ggtgccttgt
 70200
 tctctcatg cataaaatga accctcctgc agggatgata tgtagaccaa cactgataat
 70260
 aatgacagta ataccagtta ccctgtcatg ggtgtccatt ccatgacagg cctgactcta
 70320
 ggcattttat acacattctc ttcttttttt ttttttgag acatggctct gctttgttgc
 70380

ttaggctgga gtggagtggg gcaattgcag ctactgtag cctcaaactc ctgggctcaa
70440

gtgatcctac catctcagcc tcctgtctga atagctggga ctacaaattt cagctaattt
70500

tttttttcat tttttttag aggctgagggt tttgctatat tgcccaggct ggttctcaaa
70560

cttctagcct caagtgatcc tcctgctttg gcctcccaaa ttacaggcat gagccactgc
70620

acctggccac gtcacgtca caaaaaatgc aatgagttag gggcttgtgt tgctcccaac
70680

gtttagagta ttgaaggga cagcccaaag tcacaaagct aatgggcaag ggttgaatcc
70740

atatgtctga ctccaaagcc agttctgcag catattaatt gcatatttct tggtttcttt
70800

ttttttctc caattatttg gtgctatgcc aaaatcatca gaattgttcc tgtaggcatt
70860

ctttttctta attacatatt gcattgttct ctttctcact ttcccttggt tgatccaagg
70920

taaattttta gtctttttaa aaatgctcat acattataaa tccttgctaa tatacagttc
70980

tttcagactg tcctgccttg cttgaggaag cagatgtggg tcaggagaa agaaggaccg
71040

atgaactcaa atagcctttt cacttttatt ctctcttttc acttcaaaat gctgtgtcat
71100

cagtgaatcc cctcctcata atctcttctt tcctggagaa taccagccta gtttcacttc
71160

ccttgtaaaa tatccacaga tctttgtgcc ccaatcactt tggtgtatt gcactgggcc
71220

ttgctaattg cacaatatcc ttttcataat aaggatatcc acatgacaca gtagactttt
71280

gtacctggcc agaatcacta cattcatcgc agataccatg tgctgtctat tggataccct
71340

cattttctct tttcaggaag ggctgggtgc agctttctc ctgcctctta tgaaccagga
71400

aatttttctc cctgtatctt ggggctgctg ttttccagcc tctgccaagt ttccattttt
71460

gtgggacatc agttgaccgg cataccgta ccctggacgg ggacttgggt taactgaggg
71520

ttggcattct ccccaaagcc tgttctgcat atatttttga aaacctattg gttttctgcc
71580

taccttcttt cctccatgag tagacttcaa tgaacacatc actctctttc ttcggttaatt
 71640
 acctggtctc acctgctcat cataaaagag ctattcttat tgtagggct gacagattca
 71700
 ctttgcagggt gatttaatat gcatctgtag tgtccacaga acatattcta gaagttaaaa
 71760
 gtacataata taatgaaaac cctgatgcaa cttttgcttg aaatatgagt atagcaaaga
 71820
 catttgagag aagagcctga aaatagggtta atagtagttt aattgtatct aataagagct
 71880
 accattgatt aagcctataa taatactagt aactccccta aatatgctct acacattact
 71940
 tcatttaatac tacataataa cactcagagg aatgtattac catcccagat ttacagagta
 72000
 aatgtattac catcccagat tcacagagaa aattaaggct aagaatgtta ggtaattggc
 72060
 cgagtagcct gcagctagta agtatcagag aacggatttg aaccaggcc ttggctgggt
 72120
 tccaagtcca caagcttagc tgctgctacc cagcttcaact gatttttgtt ccgtgcactg
 72180
 cctctaggat gtttccgtat ttgccttctg gttaatatct actccctcct ccctagcaca
 72240
 ccacattaag ggtgtggcat gggcttttca cttcctgaat gtttgttact cagtttcaac
 72300
 cagctttact ctctttcaaa gtgacataaa agtgacagcc ctttaagctc acagaaaaac
 72360
 aggagctgat tccagtacgc gattgactac catgcaagat tttgtctttc cttttatatt
 72420
 gttgaagatt ccttgatggc aacatggccc acacattatc taacctccag acaatcccag
 72480
 aggatctgca ggaaaccacc aagcacccca gaggggtcac tggcaacaaa tggattaagg
 72540
 aaatggcctg atttatggga ttattggggg tcgggggtgga gaattataga caattttgtt
 72600
 tctactgcag aatcagcatg atagagtga tatggacttt ggagtcaggc ccaggcctgt
 72660
 aggatgacct gggaagcagt ctaggcagct ttggcccact tcagcttgta agatagagat
 72720
 aatacatcat cctagaaata tcgcatggat tatcagattt tgtatataaa gtccccagca
 72780

gacacctggc acacagtaag tgctcccaa accaaagcaa cagttcttag aagcttagct
72840

tggtctcaggc tgaagtctca aatgatctag gggtcactct ccagccattt ctaagccatt
72900

gagtgaact ggacatgtta gttctgggtg tgacttcagt atcctaagtt gggaaataga
72960

actcatcatt ccactctcc ttttaagtat ttgataaga ataaccgaga agtccatgta
73020

gcattttgaa aatataaaat actattagat tggttttgtt ttgttttcat aggaaagaga
73080

aaggaagaag gaagggagag gagaaaaaga gaaaagaatt caggactgat attttcaa
73140

ttccaaagt agcaagattg gtgtttggta ctctttgatc tctccttttc ttaaaaaaa
73200

gggggatgag tgtgttagag tctacccca tctagaatta gagagagagg gaaacagaga
73260

ggtaaaaata tccttctat tatccctcaa attttgccac agagaaaaca actcactctg
73320

cttttgatc aatgacctta ctgcctcagt gtctcatct gtcgaaatgt gtttaggaat
73380

agttgccctg cacacctgtc caaggtgtta gaagagccag atagataaga caatggggaa
73440

acttattttg aaaatcaagt agtattaaaa tcaaggtgga gagaaggaac accaggccaa
73500

gagtgaacaa atctgcattt gaacctgggc ttggatgatc agcagctttg tgactttggg
73560

caagtcatct cacctctcca cactgcaggt tcctcatctt gcaatgtgga aagcagaacc
73620

tctccgggtt ggaggatcct gtgggatccc acgtgagtca gcacctgaa attgcacata
73680

cagggcacat catttctacc atcacctaag gtcaatccca gggtattgct ggccttttgg
73740

tttgctatgt ttccacctt tctcaciaat taaaaatgca gacatgtaa ttctctatct
73800

atgtgcactt caggaaactt caagatccct ttagggttat atctttcctt agcctcttgc
73860

aaacctgaaa gcaaggtgac tgtaggacaa actgtgtggg cactggatgg ggaatgcatg
73920

cctctggtag aaggggaagc cagaattcag agccaaagaa caggcaccat gcaggactat
73980

gagtccagga tgtccaaatc tgactttctaa aggaatgccca gtgatactga ttgttttttaa
 74040
 tataaaatct accattttaca aaatacaggc tcaagaacaa ttgtaaacac tgcacagggc
 74100
 aagttttatt gttttttctta aaagagctga aaaccataat caagtgcato tgccttttgt
 74160
 tgtagattgg ggcttgcaat ctaggagta gagaaaggct gcctgactgg ttagcaattg
 74220
 aaacacagag ctggggccat ctagtccgaa tgcagtgttc atcgttcttt gctaaacaca
 74280
 aggtgaccgt ggttggttg gtttagcaagg tcacctagct tcctaggcag accttattcc
 74340
 accacaatgg gcttcctttg ctagaaagga catggaaaaa ggagtgacaa aagaggagaa
 74400
 agagctagtg aaggaaagag aggaagagag aggaaggtcc ccagttgtta ttacctaggg
 74460
 agtgcccagc atgcctttca gaggaacaca ggtgcctacc cctttctcct gcctgctaca
 74520
 gccaccacc accaataaca actgtgcacc ctactgttg ccccatgca tggagcagag
 74580
 atagagcaca gatagggaga aaagaatttt caagggtgag ttggaatagt gacttctgca
 74640
 agcctggtag gtttgtctct tttccggtat ccttgccact gcaccagggtg aaggggaagc
 74700
 agactgtaaa atactttgaa tattatcaag gttatttggg gaaaaaaaaac agcaacgacc
 74760
 tcacctgatt gtgccagggt tgtgatagga atctgggggg aagttgtcct caatgacctc
 74820
 catttcttga atgaccctg gccacctgc tacctctgag atgagtcagg ggacacaact
 74880
 ttcttgattt ccatcatcat ttgcctgcta ctgctgcaga gactcaagtg gaaaatgtca
 74940
 cttcttgaag tcatggcttt ggtgtctggt gacaggaggt agatagagtt tggacagagc
 75000
 cacttgtag ctgcgtggt taatctctct gaacttcagt gtctctagct ctaaaatggg
 75060
 ggtagcagcc tcaatgagca aattaaagca aaataacaaa tgtaaagcac ttatcccaag
 75120
 acaaggcaca tgagaagtat tcaacaaagg acagctctag ttatgggtggt gatggtctct
 75180

aaagcttcca gactcagatt cctaaccaga cagccatgag acatgggcc cacagtggca
75240

aaagggtagg ctccaggctg cctaagttta aatatactga ctctaccact tcagcaaggt
75300

actaaactct gtgcttcttg cctcatctgt aagatggaga tacctcattg ggctgttgga
75360

aagattgaca ggggtcacc ttataatgtg ctccagcaca tgcttgacc aggatggatg
75420

ctgcttgggt gatgttatca tccagggtgg cactgtggaa cacagactga tacttaagtt
75480

cagcacttgc ctcatggagc ttacagtcca gatagaatag gtgcacatga gatttttctc
75540

agcaaagttg tagtcagtgc tgcccggtgc acccatctgg tctcagcatt gctgctcatg
75600

tgcactgtcg tgacccatct ttctgtgtg gaccttaaga aaccttgcc taaggatatg
75660

tttggggaaa ccaaattatg acagatataa actgaagtgc caagacgaat ctttcaaagt
75720

gctgtgatgt ccagatacgt ggcaagtgc aggaatttgt gtttgggcc actgggagag
75780

aggtcctttt ccccggttc ccgcagtga aagctgctgt atgtctgata cctgtgctcc
75840

aatctaccc tcgagttctt atttggcaac aggcttgctt agcaatgtcc ctggggctcag
75900

ggagccacgg gtgggctgat gcaagaaact gcagtttgt ctctaggctg cgacagcatc
75960

acagctgtga gcttctcaat ttaaaactcaa catctgggca caagaggaag caccctccat
76020

ccaggcaagt tgaagtttct tttcctgtat ctttaacaga gggttaaaatg taaacacaaa
76080

acaaagcaaa acccaactta actctctggg tcctgtttct gttaaaaaca ttttttaaaa
76140

aataataata attagaacaa tagaccagc tgagagaagt ctgtctcata ataacaagag
76200

agactaaaac acgctctccc acaaacagat aaggaaatag tgagctaaaa agtaaaactt
76260

aagagtgeta cttctcaact cagtaattga ttcataata agctctagcc caattttcta
76320

tgaanaataa caacatagag aaaagaaaag ccactcagat gcagctatgt caccgaagtg
76380

ttcactggag aaattttaata tgtgtcagag aaaaatctgc cttctccaaa acattgcccc
 76440
 ctccaaaaaa aatgcttctc gttgtgggat ttttctctcc aactgacag agagtttgaa
 76500
 agtaaagcag ccgactagag tgtgggtgcc cacgccttag aacaaagaga ttaaagaata
 76560
 ggcttttctt caaataaatt cgattttttt gcaacaaatt cagggcctgg acagtttaat
 76620
 gaagaataaa aaatgctcag acccctctct ttcagctttt ccaacaaaac acagtcattt
 76680
 ccaaccactc acaggtgttt ttcttcccaa aacaaacaca gaaacaaata ttacagagt
 76740
 tatgaaaatt caatagaagt cacaggggca caaatttcct cacctgcac caagatctct
 76800
 ggcttgggat tatcaaacag aacgtcacac ctactttcca ttacgggaca ggaaatcttt
 76860
 gagccaaaga gaatcagtct gcagaatgtg ttcgtaagga taaaggcctc atcttttcaa
 76920
 gaaagtgttt gcagctccaa ggaagaatgg actttggaca aaggaaacac cagtgacttt
 76980
 ttctctatct gccatcact gcagaaatgc ccttgaagag ctcatgactg ccttgccaat
 77040
 taatccctat gttgaccaac aacacaaaga gtggtgctga ctaccacact gaccaccagg
 77100
 aataatttcc tacttaggtg tttgcgggaa aggcgtgtgt ttgcttttct caaatgtccc
 77160
 ggatgtgaac aggtctcaaa attaggaaat gcagctggga atggctgcta tcaacaaccc
 77220
 ttatttcac ttttttcttt cttcaataga gctagagaag caggcaagaa aagaaaataa
 77280
 tgcaattatg aaaggaggga ggaagacag agggaaggca gcggggcaga agggaagcgg
 77340
 aaaaaagatg ggagagaaag gaagaaaggg agggcagaat ggggacaaaa gatggcaggg
 77400
 agggagcgag ggaggaagga gcagaaagga gaggaagaa gaaaggaagg aaggagagag
 77460
 gcagagagag agacaaaaag aaagaatgag agagaaacct gctgttgtca gaacctgact
 77520
 ggtttacgct gaggaaccag atagcctctc agtgaaaagc tgagattaag ctcacacat
 77580

aatttcatt atgttaatgt tgctttaaaa tatattgcac agagtgtgc aacaatcaaa
77640

ttatcaatta taatgtctac tcagagtaga aaagtgtgc aaacagaggg ctgattactg
77700

cttcagtaca tccaatctgc atactcattg ttcagtacag ggaaaacaaa ttatcttgtg
77760

ttggccccag ttcaaattta attaaatgag tgcaaagcaa ttataataaa gtacaatagt
77820

gtagtaatga aaggaaaata actcattaac ctattaatct ttatttatgt tcctctcaaa
77880

taactatgta atttctttat caccgcatcc cataagaaaa agtagggtct cgctgctcat
77940

ggcctggttt aacaggttct atttggttaa gatgcctttg tgggcaccat ttatactcac
78000

tttgcttagc ttgggggcca cgcaccccc agctgtctgt catgctccct gccacacgt
78060

atgcgtacct gatctagatt tagcgctcct ctcccacttt tccctgacac ttgtcagtta
78120

cagaaggtaa atctctctgg aagtgagaag tgcttaccac agcatatggg aaagaggcat
78180

cttcacgcat ctggtatgta aggcaatgca aaaaagggga gaggacagaa tcacttctgg
78240

tttagagaga gccagtgct tgacagctga agagattcca ctaaaagcta acactccact
78300

gtgctttcca gttaataaag ctttttacat gtatgatttc actgaggcct cctgataaca
78360

ctgtgagggg gcctgatgta cagatgagga aacagaatga tggagtagaa aagtatggtg
78420

tccaagcctt gctactcatg agtagcagat ccaggattct aatttaggtc tacagccct
78480

aaatcttcat ttcattcccc tagaaccttg ctactcaaag tgtgggccag ggaccaggag
78540

cattggcgcc acctggaagc tcattggaat cttgggcctc accctagacc tgcagaataa
78600

gactccacat ttataagat cctctagtga ttctcacaca ctgaagtta agaagcactg
78660

gaactattgg atagtacact gccttttgag agaatggtga ctcaggcaaa tctcccagca
78720

actacttggg taagcatttc cttgcaattt gcaacacctt gagatgacct tgagatgatc
78780

ttgggtgatc aaggttgatt ctggcttcca ttttcattgt atctgaagac agggctggaa
78840

gtagaggaca tagtaaaagc agaaatgatt tctctttctg gatccacagc agacccgaaa
78900

atcttccctt ttttcttcca aagagaattt gtgaaataga gagaaaggag aggggagaaa
78960

ctgatgaatt aaggaaggaa acagacactc gttaagtctt taccctgttc ttggagggtt
79020

tccatgtatt ccttcattta atcttcacaa caagcacctt ttaagataa ggtgactgac
79080

gacaggttta agtgatatgc ccaaggtcac ccacactgct gggtaacaga aatggatttt
79140

ggcttggtt caaagtttac actttcactg tactctattt tctcttaaaa aaaaaaagg
79200

aaaagggaag gaaatcaaga tgaaataaaa ggggtggtgtg gtagtcatta gtgctgttca
79260

ccaagaatct ccagttttta gtctctgagg tatgtgaagg attatacttc ctgccctgct
79320

tagaagtagg tgggagccat gtaaataagg ttggccaatg aattggcagc agcggtttca
79380

tgtttcattt tcagatcaaa tatttaattg cttgtggaca ccttacagag ctatctctct
79440

ttctctgcc atgaaaattg ccaatattca agatgggtgtg tgcatacataa acctgtatcc
79500

atagtgagtg aatatgtctc ccagagaccc ccaccaacct atgatagaca cataatgaga
79560

gtgataaaaag caataaacct ttgtgttatt ttgagccct gaagtgtgat ggttgcttgc
79620

tactgtagca tctctgagcc catccccact cttccagatg ctaggaagac gtgtgaaggg
79680

aaacagttaa catttctaca ctgcaacttt ttcacttgag actcacagca ttcctatgag
79740

aaaggtaagg aatggctatt atttattatg tatttataaa ttactattta cagataattt
79800

aacactgaca tagcagttta agtgtcctgc ccaagatcac aaagtttata agtgggaagg
79860

ctaggttttg agcacaaggc ttgtaatgac tgatctttct tcaatgcaga acatttcctc
79920

caaaaaagga ctggagcacc attacctgcc cttgggtata ctcagattgg tataatctagc
79980

tttgtactca cactcctctt cccccaccac tcaggaggaa tccaagttat tgggcaaaaa
 80040
 tcaggccaag gaaaaacaag tacatgcagt tgccagcaga tgacgacatt gggctcttgc
 80100
 tagagaaacc accagagaaa gttcacagtg ataaccctca atgaagtcaa ctgtaaggag
 80160
 cttcattgtc cagtttgagt gttttttgtt ttttgtttga gaaagggctt tactgcatcg
 80220
 cccaggtcgg agtgcagtgg cacaatcact gctcattgca gcctcaacat cctgggctca
 80280
 agcgatcctc ctacctcagc ctcccaagta gttagaacca cagggtgcgtg ctaccccacc
 80340
 cagcaaattt aatttttttt tttttttttt tagagacagg gtctccctat gttgctcaat
 80400
 ctctgtctcta actcctggga tcaagcgatc ttcctacgtt ggctcccaa agtgcgtggg
 80460
 ttataggtgt gagccaccat gcctggcctg tgttcacatt tcatgtgatt gtttttgaa
 80520
 gatagaaagg caaaagttag tagtagcagg attgacctt gtaatatgat cttaaaaata
 80580
 ttgtctccta atttgcatat ggaagaatct caactaaata acctcctttt aaccagttca
 80640
 ctaaacaaac caaagttctc cctttttcct ataaaacatg cttctaattg aggttagagg
 80700
 ctcttggcaa gtaggctact gtccagcatg tccaaggctc tgctctccca tctttgagtg
 80760
 ctatgtttac taagagtcta ttaggttcat tcccaggctt gtttatgctc actgtatttc
 80820
 ttattgtcat tatgtacctt aatgtaatgt gagtcagaaa aaaacatatt tctgtgaaga
 80880
 ctggatagaa cgctttggaa agatttgata aaagcaagtt actaaaaaga gtgttggcag
 80940
 acctgtaaca aaaaattgac aaaagttatg gtctatgtgg aacctgccct cagaatgttc
 81000
 ttttagtttc ccaaggcttc acattctgtg ttaaagaaac aaactgaaat tcatggacaa
 81060
 tgcattgcga gtgtgattta tacaagaaag gcaatctaga cgtcttagca atggaatgta
 81120
 accatgtttt ttaaaagcct aagtcataata tcaaagatt gtcaaaacga taacccaaaa
 81180

caatgtatag gttttgtgtt aaaataaaat gaaatattta aagtatggat cgtttttaat
 81240
 catttctgaa ttcaagccct acctatttgt cccaattgct tctgatataa ggactttcac
 81300
 ataattactt ttccctgttg ccattttccc atttcccctt ccacagacat actgtgctac
 81360
 cttaaagtgc gaggtctttc acatatcttt ttttatccaa ttacctcatt caacaatatt
 81420
 tatttttaaaa aattttttta agagatggga tctgcctatg ttgccaggt tggctctaaa
 81480
 ctctggggt caagcagtc tccaccttg gcctcccaa gtgctgggt acaggcatga
 81540
 gccaacacat ctggccctct acaaatattt cctggacaat tactatgtgc caagcactgg
 81600
 cctgggtcct gggcatagac caacgaactg tagagccaac agtagagaaa acagacttta
 81660
 aataatgaga gaatcatcat ttaataataa ctgtggtgtg cgttatggag aaattcagag
 81720
 ttcactctgt ttttatataa ctcagacctg ggatggattt cctacttctg aatggtaaag
 81780
 taagtgtctg agtagatgag ccaccagcct tcttccttgt attctatttt taatgcagtt
 81840
 tctttgggcc aaatgatcac cgtcacctt accagctcca gactgtgatt ccatagtggc
 81900
 tggattgcct tggagcttag agcaggttat ttcagtttac tccatggata ttctctccac
 81960
 tcttctccac ttcatgatgg ttaatatga gtgtcaactt gattggattg aaggatgcaa
 82020
 attattgttc ctgggagtgt ctatgaagggt gttgccaaag gagattaaca tttgagtcag
 82080
 tggactgggg aaggcagact caccctcaat ctgggtgggc acaatttaat cagctgccag
 82140
 cgtaaccaca ataaaagcag gcaaagaacg tgaaaagact agactcgttt agtcttctgg
 82200
 cctacatctt tctcctgtgc tggatgcttc ctgccctcaa acatcagact ccaagtctt
 82260
 caattctggc ttccctgatc ctcagcttgc agacggccta ttgtgggacc tcatcctgtg
 82320
 accgtgagtc aacgctcggt aataaactcc cttttatata tacatctatc ctattaattc
 82380

catccctcta gagaacccta agatatctct attctgttcc ccaagaggtg gacctttatg
82440

gacccaatca actgcctccc tcaccctgtg cctccaggtg gcttggaaag caggaggcga
82500

tgaaggctgt gggttttgtt ctctatctat ctatctgcag aattgtcaag tgttggcagc
82560

ctctctcttg gcaaagccac cacccttgtc aagcagcact ctctttatag ctactctcta
82620

ggtcagggat tgtcaaaca ggcccagact aaggggtggc aagcaaagt ccttggaagc
82680

caaggttacc ctcgaggtta tgcaaagca gggcctgaga gtccattctt tcttaaattt
82740

tgcacctgg acacctcact tgccctaccc ttgtcccggc cttgatgaaa acttcctttg
82800

taaaggacta gagaataaat attttaggct ttgtaggaca cacagtttct gatgcaagta
82860

ttcaattctg cttttgaaat gtgagggcag tcatagacaa tatgtaaaca aatgggcaca
82920

gctgtgtttc aataaaactt tattttcaga aacaggtaca gggccaggtt tgatccatgg
82980

gccatggttt gccaacacct gctctaagt atacagttcc ccatggttgc ttactctggg
83040

acagcacggt cccttggtgt ttcttttaat cctgctcaca tctttgcaaa aagccccttt
83100

actaagctcc cctcgaatta tccaaattga gtgtgccatc tgtttcttgc taggacccta
83160

attgatgcac ttggctttgc atttctccaa cacaggcttt aattgcaaag taacaacaac
83220

aaaaaaaaa agaaaatgga gactacaacc aattactctg tagcaggga gtggtggtgt
83280

ttacatgtat tcatccattc ccaaataatga tagaatattt cacaaagttg cagtacactc
83340

ttaaagtga tgggagctga aataaataag cttttttttt tcccctgggg cctgcactct
83400

caaccctgct gccctacttg tcctcttcta atggaatgga ggtctattct aattaaaca
83460

gaaaacacat cttttcaca aggagaaacg gaggacattg tactctgttc ctcaaagaaa
83520

gcatgcttta ataagaatcc ttttataatc cctgctttta tttgtagctt cccaaagctt
83580

ttgctggtaa taaaaggatt gttaaagttga caaatataag cagggacaaa gaagtttggg
83640

gagataatgc tctagtgaat gggctagtat ttgaaagttt attgatcctg caacacctag
83700

aagatgatca cgaagagcca atgctgccat caatcagggt gatatttcta ttacaaagta
83760

aaagcccttc caggcagggc tatccatttc tcatatcatg taagaattat gctttgtgct
83820

gagaaaaata attcaggcca aaatgcagggt cttggcatct gaagatctgg attcaaactt
83880

ggcttcatag aacttctaac ctacttcttt atcatatgga agcagcaaat agagctgtaa
83940

caaatgcaat gtgttctgtt gtggggttgc agacatctgc tgtcattcca ccactatgac
84000

tatcacaagt agcatccatt actcctggaa aaagccaccc attttcaactg aggttccctc
84060

actctcacac ttaaagcctc acactttcca cttctggtgg agtccatct ctgtggcaag
84120

ggtggaacat gtgaccaagc ctagaccaat taatatactg tataccttgg gcagagagat
84180

ttgtttaatt ataggcaaat gatccagtta cagccaatga gactttctag gaaagtgatt
84240

agacaggctg tcctttctca ctaggcttga atgtgagggc tggaactgct gcagccattg
84300

tggatatatg gggagcctga ggctgaagtc aacagagtag aaggatgaca ttgttcaagc
84360

cccacataga gccatgccta aagctggtat tggttgggag aaaatataga caagttttca
84420

caatttcctt tatctaaacc ctttgtaatg tgactttgca gattctccca ttgagaacat
84480

taagagatgg aaactacttc ctactccttg aatatggact gctctttag ctgctttga
84540

cccatgcatt ccagtgaag tgatggtgtg cacttttgct gtcagaatcc tgcaccatgc
84600

agaaaagcct aggtgaacct gctggaggat gagagacaat gtggatgaga gccaaaggcat
84660

cccagccaac agccagaagt agtgttgctt agctgaccca cagctaactg cagatcaata
84720

agggagccca gttcagacca gaagaaacat ccaactgagc ccagcttaaa ttgctgactt
84780

gcagattttc tgcgaaataa acagttgttt taagctactg tatgttgagg tgatttatta
84840

ttcatcaata actgatacac atcccatccc tgaatttttc atttatatga accgataatt
84900

ttatttttga tgtagtcagt taggaatggg ttttctgtga cttgaaataa aacaattctt
84960

aacacattaa agatgtgtct ttatttcccc cactagatag catcattagg gattctgtct
85020

tatccttata caaagcactt gccctatctg gatccaaagc tgagtgactg atcaataact
85080

acttagtgat tgactatgtc atattgacag attcacaaat cacattttga agagcaaaga
85140

atatatttgc tcttgagtct ccaatgtcat ctgtagggca attgtacttt atttggett
85200

gagcctcaac agaaatttgg aagatacaag gtgcttctcc aaccatcaca atagctgaaa
85260

acaaatacat ctgagacatt gtatctaagg aacccaactc tgccatctcc ctcaaggga
85320

cccagaaaaa tctgggcttt aatcctaccc tccttgctat aaaaacctga aatgatgtca
85380

tctttttctt tctaaaaataa aataaaggca tgtagtgtct ctaggcaaga ggtctagatg
85440

tatctgcctt tcctgtcttc caggacacca tccacttatt gctacaactt gctgcattat
85500

tcctgcaatt tattttcttc tttttgcaca aaaaaaaaag gaagaaaaaa ctaaagcagg
85560

gattccatag aattcatcct ctactttagt aaaaaggtag aatccacctg acctattctg
85620

tatgtcagag tggaatccag cagccatttc totccaatta ttgaaaaatt gtgtgtgtgt
85680

gtgtgtgcgt gtgtgcgtgt gtgtgtgtgt gtagagagag agagaaaaaa aaggtagata
85740

tacaccagca tttttattct taaataagag gataatcagg cttactccca cccaagtga
85800

tttacacttt aaattcttaa atggctggtt tttttttccc tcagatctgc agacgaagg
85860

ccaaaggaag aagattttgt tgaagactga aaagagaaat atacttggtg gacagagata
85920

actcctttca gcacagaggt caggtaaggg aatcacagac tgacatttac ccttctctct
85980

gaactctccc aggcctcaac agctcctaatt ctttcctttc ttcattctcc ctccctctca
86040

ttcacatctc tttccccaag aagactcaag aaaaagtcca agataagaaa ttccatttat
86100

aagaaaagga aggcacaggt taatttggtt tatctctccc aacctcaaca atggatgtga
86160

gttctgacaa ttgcttggtg ccagaaccaa tcaaactttc ctccctcatt agatgacaag
86220

gctgcagcat gaagaagatg ccatattgtg gaattgggtg tcaaggagaa gacattcgaa
86280

ggtgcatgaa cagagtgaag gtggtgcaag caagtgcaga ggaggtgggg agactggagt
86340

gctgggaaaa gggctcagcag aagctagcaa acatgaaaga cacaacacaa aacaagcagt
86400

tgtgcaaatt aggatgataa ggggggagtg gtcaaagggtg tgtttctcat ggctcaacca
86460

gtgcaaaactc agtcttatga ccctacatga ctgactgcac ctggcctctg catcccttca
86520

gcctcttccc caaaaggact cctccttatt gccagccaca caagccttcc ttcatttctc
86580

ctaacaggcc agacttggtc ctggctcccg gtcttggtcat tcaactgtttt ttctgccata
86640

gctggctcct tcccaaacat ggaatctcag cccaaacaat atctcctctg agacaccagc
86700

cttgatcacc tttatacaat acagatctgc tctgcccatt cactttctat ctcatgacct
86760

tacaatgtgc cccttcgtga tagctcacac cagctaaaac catcttggtt gtggattgct
86820

cacttgcccc ctgcctgtct ctagctctag gaggtcggct ccctgtcact gtctgggtcta
86880

ctgtgtcgta ccacccatgc ccagcaccta gccttgacaa ttataggact taatgcaaac
86940

ttcttgactg aaagaaaaat gaagaaactc agccttgga gagaaaacaa ccattaaatg
87000

attatttaat aaaatagagt attaatctat tttatattaa ttattaattt agttcataaa
87060

atattaattg tctttacaac gaatgtgtcc ttagtcattg gtgtttgtgc tttgtgaatc
87120

ttcacatccc atggggccag taatgggggt gggggtgtag tgtgggaagt ggggcaggga
87180

accaatctgg cctgcagggg ttacctgcc catgggtgcc tgcaggcttc atgtggtttc
87240

agtggcatgg caccctgatg ggaagagtgg ctcagaattg gtggagtatg gtgataataa
87300

aaagcaaaac aacttttagt tgattattgt tttttaattc tataaccgact atggcccctt
87360

cactacctgc accctgtgag gagagttttc ccacatctgt cccactccca ccctcacact
87420

ccactggttt tttccccctt cttgagtga cacacgcact ggaagagcag gtgtccttga
87480

agggaggacc agctccatga caggacacat gtcctcccaa acacaaacat acgcaaacac
87540

agaacaaatt agaatctatc ctggctctctg agaacacttg ctctcaggcc actaaagcac
87600

cttgataata aattgtctctg tgcctgctac ctctgtctcc acctgctcct ggcagaaatc
87660

tcttcctca ttccagccac tctctggagg gcaggcaggg ttctgtttat ccatgtccac
87720

aggccatgcc aaccaatggc tcatggcttc cccacttcca ggaaagccac atggcccagg
87780

atacaataga gtagggaatg agtaggagga gaactggtct ctttatccca gcaatatgga
87840

ccttgtgatg gaaagagacc taaagacagt cacatttact tctcatctct acccggggcc
87900

tttcacacct aacagagtat ggcaattggt acatggatag cactgagctg agtgacccat
87960

cattacttcc tgccatcgga cacttatgtt ccatcccagt agggttccct agtcacaatg
88020

aagatccagt ctctatgcta cctccatctg atttcttctt tgatctctac tcctttttatc
88080

ccactaacia attttgtatt attacttttc atttattcaa aagtatacat ttgtttgtgc
88140

acaaaaaaat acaatttctt tttttgtatg gttcagattt aaatcccaca tacatgatat
88200

atgcaagcaa gttgctttgc tggggagatt tgtgaaattt cctcacactc atgttgtaag
88260

ctactttgtg gtcctcatct cttttcacat ggacagcact cttacagaaa ttaagaactt
88320

tatttttttg tagaaaacca gatgtaagcc cattgaaagg agaaaccttg gctatctttg
88380

tcattataat aaatctagta tgaaattggt ggtagaaaa tattaaatga aaggaagaca
88440

gagaaggagg gaagaggaag gtaagaaggg aggagggagc tagagaaaag ggggcgggag
88500

ggagaaaaga aattaagttg tcctgaaact caactgctag gacagtact gtgagtaaag
88560

aggaaataac agtttataga aaaatatgtt tcttattttg aaaaaaaaaa gtttggaata
88620

aaacaaaaaa tacaaagacc cctatTTTTT agaattatga ggagttggag gagtcaacta
88680

aatgtttgtt aaacatcact acttctacct ataaagggga aaatacttat aaagggaaaa
88740

ttttaactag agtgatattt gggacttaaa atgtctggta ccaattttat ccattttctca
88800

tcccagagca ggaagtttgc attgtttatg actgaatatt caagtatctt tccccacta
88860

ggatttactg taggttcgtg gtagtgacag gggatttaaa tgaaaatttt taggtctcgg
88920

tcagtaaaac gacctgtgtc agcctagaga aagaggaaga gaggagacag ggttgaccca
88980

gcgctggacg agtctgaaga tgtgcatggg gagaaaataa ttgagctcat gtgacccaaa
89040

gtcaagtgtg ctctgggttt aggtctcagc tgcagtattt atcggactgt gaggatcatt
89100

tcagctccct gagcctcagt ttactactca tctacagaat ggaattaata acattaaatg
89160

agatgatgca tcactagaca tttagctcag atcatggcat gtattacata tgccacaact
89220

gttagctagt gttatgataa tctttgttat tattatggtc aagatgccta ctgcaagact
89280

aaatatatat caaatgtttt ctgttgagca tacaacttga gcataatttc caacttacct
89340

tacagttcca tgtaccacg tggtcggatt ctgaacaatg accacttcca gaccggactt
89400

acaaaatctc tctccacact ctctcatttc cctcatcttc cgagaagatc caatggagga
89460

cttcaaagtc ctagaagatg gcaaagcgac aagacagaaa gaatgtcatc catgaatgac
89520

cacgtggagc agagctccca ctccctgcc tgctgtccca cattgcacaa gaatgtgaca
89580

tagaaacaca atttcatttc tcttatgtca cggaaatttg aggactcttc cttatagttg
 89640

ctagtattaa ttaccctagc taatgccaaag gaagaccaaa aagtctttcc aatgctttgt
 89700

aaggactaac cagaagacac acataaacac agctttgctc aaattcattt ttgtcacaaa
 89760

tcactttctc agggcctcag aatgtgaatt ccagctgctc tgtcaaccaa tgtctagaaa
 89820

catttttgct gttgagtgc aagtcttcat tcatgagggt gctctccaaa cgcattcatt
 89880

ccattggtgtt catttcttgg tatagaagca tatgaggaac agcacactcc ctggatctat
 89940

atactacatt acctgtatta ataagaatgc taggctgtac ccattcatta tctaccagcc
 90000

cgccagacaa aatgattcac cctagactta caggttgaag tcaattgagt atcctttgaa
 90060

ctggagggag aggcgcttcc ataaagtggg gagagcttta tggcagaaaag ttctattaat
 90120

gacccactt cagtgcctat caggagcagg ccatataatc ccttcccgga attacatgct
 90180

ggtcgggtta gataaataac ccagatctct gtcatttata atccaaaacg ttaacaatta
 90240

atttccaaa tcaatcctga agggggctag aagcctaaat cgctagggtcc aggcagacct
 90300

cactccatga atttctcatg agtcaaaatt taagcctcga aacttgtaac taactgttgt
 90360

caagacacag ctgacttagg caatcaagac aacagtattc agcagtaacc tgggtgaaac
 90420

atactctaag tgtgaattaa taagtcggct gacaccaggc ccagatggt tcccactctg
 90480

gcaatatgcc tgaagtgatg ataagcagat ttacatactt aactaatatg tttatttaaa
 90540

agaaatttcc aaatccaaaa taacatccag atttcttatt tggctgcaga tcaaaagaca
 90600

tcctatcaat aaagacagaa tccaggataa accttttcaa catctgctta aaaccaaaaa
 90660

acagatttga gagttttggg aaaacttaat tgcagacaaa tttcactgcc tcttacataa
 90720

atatatatat tcatgtccct ggggtgggtc tgacctcatc aattttcgtg ttcacaatac
 90780

acaacattgt tcttttacca gcaccaacaa ttctgccttt atggagaatt cccatcccaa
90840

ggttttcttt cagactgaag gataaacaga ctttctctct ggccaggaaa agtcgaacaa
90900

aagtcagggg catgtgctaa cacttggaat gttgtagaaa agaaaaatcg acccctctac
90960

cagtaagatc tttgtttttc tataaatatt agaaagcatt tatttggctc ccatctccga
91020

taaatctgga aggagaatag aaaagtagag atgctccttg agtattacat tttttttttt
91080

tcttttcggg ggtgttgatt cccaaggga gagcttagta actggatcaa agtcttaaca
91140

gatctgtgtg cttgtgaagg agatgataat aggaggggtg ttggccatag tagcacttat
91200

cccttttagtg tgaattagtt cactttcttc aggtcttagc tgtctgaaat tggaccatga
91260

tccttgggca ctgagacatt tttctcccat ctggccagtg ctgcaattga gctaaggatt
91320

tggttcagag actcttgttg cctggggatc tgtttaatgt caagaccgat attagctggt
91380

tgaccctggg gtcgggtgac ttagagagtc agaagtcac catgtttgct gctgctccac
91440

cctttgaact tagtcatgat aaactcctgg gtccaatcct gttacagcca aaatagcaag
91500

aaaagagtgg tagaagcaat cagaaaaatg catagtaagg caccaatgag ataccatctt
91560

cctcttctca attggcatat actttaaaat agcgggtgttc agtcttttgg cttccctggg
91620

ccacactgga agaagaattg tcttgggcca cacataaaat aactaacac caatgataac
91680

tgatgggcaa aaaaaaaaaa aaagaatcac aaaaaatttc tcataatggt ttaagaaagt
91740

ttacaaaattt gtgttgggcc gcattcaaag ccatcctggg ccacatgcag gacaagcttg
91800

ctttaaaata toccatttga agacagttca gatctcccca gttatagatg atttatcaaa
91860

taactatgta tgtatgaaat taatatatca aaagccttaa aacatttcca acctttgacc
91920

aaaaaattgt tcttctagga tgtatcttaa gacataatta tgatttgtgc acaaatttag
91980

ctatgaaaat gtagttttac ttatgatatg acaaaaatat cacaaatgac ctaggtgtca
92040

aacaatagag accagttaat taaatgatgt tccacaactt atataattat taaaaacatt
92100

taggatattt acatttaata aaacccatag tgagggcatt tagattgttt ctaatttgga
92160

cttttatgaa taaagctact atgaaattat tgtatagttt ttctttttat aaatagatgc
92220

tttcatctgt tttgggtaac cctccatgag taaaatagct ggattatatg ggaaataaat
92280

gtttaacttt ataaagaact accaaacttg tttccaaagt gattatatca ttttacaaca
92340

ctgcaacaat gtatgaaaat tctagttgct ccacctcctt ctcaatactt ggtttcccaa
92400

tcattttaat tttagccatt ctggcagggtg tgtagtata tgtcattatc gttttaattt
92460

acattcccct aataattaag aatgtgagca tcttttcattg tgcttattct ccacttgggt
92520

attttatttt gtgagttgac tattcaaac ttcttttcat tttttattag gtcatttgtc
92580

ttattgagtt gtaggagttc ttcatatatt ctggatataa gccctttgtc agatgttttg
92640

caaatattat ctcccaacct gtggttttca ttttcatttt gtttatagtt tctttctttc
92700

tttctttctt tctttctttc tttctttctt tctttctttc tttctttctc tttctttctt
92760

ttctttcttc ctttatttct tttcttttc tttttttttt ttttggtgag acagtctcac
92820

tccatcacc agactgcagt gcagtggcac catctcacct cactgtaacc tccacctgct
92880

gggttcaagt gattctcatg tagcagccc gattacaggc acttgccacc acactcagct
92940

aatttttgta tttttagtag agacaggggt ttgccatgct ggccagcctg gtctctaact
93000

cttgacctca agagatctgc ccacctggc ctctcaaagt gctgggatta caggcgtgag
93060

ctaccgtgac tggcttcac atagtatctt tctaagtaca gaaattttga acttaagttg
93120

cttttgagaa aaagaaatgt tgtggtgctt agtgaaaaga ttcaagagca cctgcctttc
93180

aaccctgct cagcctctgt ttttttactt ggctgccaac tctcttgctg tgtggacttt
93240

tcctcatcaa tccaaccatc tggcgcttac tggtggaagc tgtgagtggg taaaggccaa
93300

tgagggtgtca gggacaaca gtagaaacag aagccaaacc taacctgcc tggagcttcc
93360

catgtgctct agaaagggaa gtgccttgct tgctgaggca ccttctagta ctttccagct
93420

ttaaacgaag catggatctt ttccttcaca attaacaaaa gcccttgctc aggtatgttt
93480

ggcagtgct gtcatgcttc tttgtcaca gcaaaacaag aaactaggga tccttcaaaa
93540

ttctattaga aaatagacct ttcaggaaat gttgcagaat tggctttctg ggatatggga
93600

ccaaatattc cttcattccc tgagaattga cttctttctg ctgtatgctc caaatgaact
93660

ccatatatac atacctcatc atgcctctaa atttttaggg cacacatttc tttacatgct
93720

tacctcctgc aattacactg gaagcccctt gaaagcatga ccatgttttg tacatctttg
93780

tatgtatagt gcctgtctta gtccatttct attgctgtaa gaaaatacct taaatctggt
93840

aaattataaa tcatagagat ttatttctca caatcctgac gactgggaag tccaacatca
93900

gggcaccaga agatttgatg tctgcagagg gctgtctctg ctttcaagat ggtgccttct
93960

tgggtgatcc tcacatagca aaatgacaga gagtaaaaag gactaaacag acttctttaa
94020

gatgttttac aagggcacta attccattca taagggcagg gcctcataac ctaatcacct
94080

cctaattggc ccacctctta ataccattgc attggtgatt aagttttaca aattttagaa
94140

agacacaaac attcaaacca tagcagtgcc taccacagaa ctgatacaga atgagaactc
94200

aatggatact acttgaatgg ctgaatgaat gaagtgaatt aattgttcgt atccccaaat
94260

ctgagagata tctctctggt taataatcta ttttgttaaa actttgctct ggagttggac
94320

agacctggga tcacatctca gctccagttc ttgctagctg tgtaactgag acaaagtctc
94380

cagcttcctt tgagtcccat ggcttattaa aggggaacaa tggttataat tataataact
94440

acctcatgag tgtgtatgga ctcaatgagt taacacaaag tgtgtaagcc agtgccctggt
94500

tcttagtaag tgctctgtaa atgttagctt tggttattta catgctgtgt tttatgtgac
94560

ccatttctta tgaaataatg tttgctacaa gtagaagatg atgggtgatgg ccaaagtgc
94620

attatatatt ttagtggtca taacaaccag ttattattat ccttttttca gttgtaattg
94680

aagtgcacatg cgtgaggtca caaagccagt gataaagctg ggattcagac tcaggctgac
94740

catactttta tctctaaaca tagtcactct gtccttcac acaggaaagc attctctgat
94800

ttggaggaga gaattcctgt agccaaggat ggtggcataa gccatagaca tggggctgtc
94860

aaattgctta gcctagaaga acaagagtca gccaaacttg gggactccct gttctatccc
94920

cctccctctt atggcaatag gattaagttc tgtgtattag tccattttca tgaggctgat
94980

aaagacatac caaagactgg gtaatttact taaaaaaaa aaaaaagagg tttaatggac
95040

tcagagttcc acatggctgg ggaggcctca caatcatggc agaagacaaa agtcacatct
95100

tacatggtgg cagacaagag agaatgagag ccaagtgaag gggggatctc tctttataaa
95160

accatcagat ctggtgagac ttactcacta ccattgagaac agtatgagga aactgtcccc
95220

catgattcat ttctctccca ccaggctcct ccacacacac ttaggaatta tgggatctac
95280

aattcaagat gagatttggg tggggacaca gccaaatcat atcaagtata ttgcagaaaa
95340

catgaataac agtggctaaa acaagatata attttttttc tcttacataa gaaaagagca
95400

gcctaaagct tathtagcct aaagctggta tgggtgggtc cagtcacag gaccagggtt
95460

ctactttttt gtccttgaa aataagcatc caggagaaat gctgttcttc ctcacggggc
95520

ctgagattct cctctctac caagtgtcac ctttctgctg gtctgcagaa actcctctgt
95580

ctcttcaggc agcagcacca gatacaccaa gcagaccaa acaacatcac aaaggctctg
 95640
 aaaattaaac tgtctttgta accacagtcc acgaaactat gccaaagacct ccgtgctaaa
 95700
 cctaaacagg ctgactgcct gctaaaataa aagattaaaa caggatccag agtctcctaa
 95760
 tataatagcc aaaatggtca gtatacagta aaaagccatg aactgagaa ccaagcaaaa
 95820
 cacaagtga ctaaaaaaga ccatcaactg acagccatgc tcagatgaat cagatgctgg
 95880
 aattgcctaa caaggatttt aaaggagaca gaaatgcttc aacaatcaat tacacattct
 95940
 ctgaaacaa ataaaaaaaa tataaaatct gggcaaaaa tagttttttt atttttttta
 96000
 aaaaaggaaa ccacagaact ctaaaatatg cagaaactct aaaataaaac agaaatagaa
 96060
 agggcatgtc tcttagccag atcagccatt caaagaagtg cagaagaact attcaatgat
 96120
 tttcacctaa tttccttggc tactttccag ctgcaaggaa tgctgggaaa tgtagtcttt
 96180
 aagcaggcac accaatgata gactggattt ttgtttacca ggtaaaagga aagaatatgt
 96240
 atcgggttac taagctacc aggtctgcc ctgacacaaa gggctcatc attgggaaca
 96300
 aaatcaaggc tctgggagtt tacattggaa tatgatctta ctgtggataa gaccaaagaa
 96360
 atcctgtgca cacggatgag gattataaat cctggttcag gctggcatga tccacagttg
 96420
 ccacgattcc tgtatgcctg cagaaagaca acttaatgtg gcctaattgt actaacatct
 96480
 cgacaggctc accttttgag ggaaggccat ttaattacct gattaatttg ctcatcttaa
 96540
 ttaacttggt cctattgttg ataacctgt attgattgct ctgctggtag taattagaaa
 96600
 actggtaaag aagtaaggag agaagcccaa agaggggggc agcacagcac acagattgtc
 96660
 ttcctgtctc ctaagaatct actgtgtagg ctgtgagctt agcatatttg cttctctttt
 96720
 taggtaaata tggtatagac acaaatgatg ttctctgttg attatttata accactccac
 96780

gttccttctt agagtatctg cccttttcat tttttcaaag actcagctct catgagcccc
 96840
 acatgacctg ctaccccat caaccacagc tgattggata agcacctgac ctagactgtg
 96900
 ccagtcacat ttgattgttt cagatttctg aatcagagca tgaagaaaca gtaccagtta
 96960
 actttggatg ctgaaactat aggacgatgc agacttggaa ggtggacaag gtagctatct
 97020
 cgtgatagcc accatggacc ctgtattccc accaacaatc agagagcact tgtcagttaa
 97080
 gttcaccaga ctccaatgg cattatagtt ctcccaccaa acttcttgcc ccagggtgta
 97140
 tggactatct ctgattatct gatagaagt tccctgtaca ttagctgact caagttcatt
 97200
 gccatttttt gcaacccaaa gagtcccttc tataacaaag cactctcatc attgacttct
 97260
 ctctatctta gttttaagt tctctggtga aaatgcaaat ttaaaggacc tattaanaac
 97320
 aacagagtgg gagttgctat caaataaagt ataaactgct aaactgctaa ggctctatc
 97380
 aagaccaatg ctgtgctaac tctttccaac ccagtaaaac gaggccttat tgtccaccac
 97440
 taacatgtca ccacctcccc tattgacttc agtcccttgg gacccctaa ctctcatacc
 97500
 ccacccagct tccgagaact ctattgcctc ccagaagtaa tcatagtcaa cacttagtgg
 97560
 tagttatttt ggcttgaggc catggccttg caaagctggt attgaagtga tgctctcata
 97620
 tctaatttga ctgaattgag tgttccagta agactggtga tgttttctt gcttcagggt
 97680
 ccctcctaag gagcaaatgc tgccttactt tcacagaagg ctgcaaatgt agacttgtaa
 97740
 agaaaatctt tacaaacact tcttggtgtc ttagagatca cccaacacaa ttacacatgg
 97800
 ggatagtctg ggaagtacta aaataaacct gatttatcct tcacatatgc ctttcatttg
 97860
 ctattgtggc ttcaaagctc aaaaaaata gagagatgtt tctctaaact ttactatctt
 97920
 attaaactgag aatgtaacgg aagtgaaca gccaaacaaa attaaaagca gcttactcag
 97980

ggatacagag tcatgactat gtggagcccc atatctttac ttctttgtct aaagtccctca
 98040
 taagcactgt ctgagtttgc atgagtccta caaactagaa caactcccaa cttcattgag
 98100
 aagaatgagg acatctttgc aaagaaggca gaggaagcc acatgctact tgtaaacatt
 98160
 tgatggctac ccagtgccaa gggttttttt ttttttttta acaatcctga cagcagaaat
 98220
 ctggaagctt gatcctagaa gactgaagag cccccaacat gaccctgaac tctcacataa
 98280
 gagaggaact cttagttcgt ttgccattaa gagaaaacca cagactaggt catttataat
 98340
 gaacagaaat ttatttttct acagttctaa atgctgggac atccaagatc atggggaccg
 98400
 catccgacaa gagtgttcta gctgcatcat tccatggtgg aatggccaga gagggtgagg
 98460
 gagagaagca aaagggggcc aaactcatcc ctctgtaagg aactccctcc tgagataaca
 98520
 gcattaatca ttcatcaggg cagagccctc acggcctaata caccttttaa aggtcctatc
 98580
 tcttaatact gttataatta aacttcaacc tacatttggg agggaacaaa cattcaagcc
 98640
 acagtagaaa ccagcttctt tttctgttgt taaagcaaca ctttttttct tctctgttac
 98700
 ttgcagccta accagatcct aataaataca caagcttcta tgcgacaaa tatgtctaca
 98760
 atgctgttta aatggcagca tgggtattcta ggatagggaa gggccaccat ttactcaacc
 98820
 agcccattct tggaggatgt ttaagctatt tccaatattt tgctatgttc aacaaagtat
 98880
 taaggatctc cctagtagct aagcaccat gtaaatacct agatatctta ggataaacct
 98940
 gtaaactagt ttatgtgtcc tagtttagct gggcccagct ggaccatcac acatatttaa
 99000
 atttttttga aacttattgc caacatgcat ctccagagag gtagtacacc aaattatgct
 99060
 cctacaaaag ttgcagaaac ttgcctgtct atccatactg tcaccacaa tgggtgattat
 99120
 tactttttat ctgtgacaat ttgatttatg ctgggatttt gcagccagca ccaaggtagc
 99180

atgttgcttc tcagtctaata aataatttcc ctctccctca tcatcatcac cagcactacc
99240

aatttctctg cactatgggt actaacatca aaaatgttcc catcacattt acttcatttt
99300

gaggaaaaga acagttacta ccattcatte attcactcat cacaccaata ttactgaatc
99360

ctaccctgga ccagcattgt tatggtagat cttgataata tagtgacaga caataaccatc
99420

cttcccacca ttacatagat agactgcctg gggctcatag tctactcaca tctcagcaca
99480

tgagatagtc ccattatcac attctcttct atctctctca caggcactca ctcaccagcc
99540

atggagaact taatgggcat taggttctta atcctgctgg gtataattgg agataaccac
99600

cctggctgac tgagcactat accaggctac tcatggatga tttgggtatt agacgtaatc
99660

agagtggccc ttgtatgcta aaaaggagac tcagtctcca tcccatggaa gattacattt
99720

ttctgactgc catccctcag atctagatgg tattgtctcc ctgacatcca aatatcttca
99780

atcatataac c
99791

