

K I V O N A T

Szacharóz-szintetáz gének manipulálása a szár és a mag minőségének javítására

A találmány tárgya új, izolált szacharóz-szintetáz nukleinsav és az általa kódolt fehérje. A találmány eljárásokat és készítményeket biztosít a szacharóz-szintetáz szint módosítására növényekben és különösen a növények szárában és/vagy magvaiban. Találmányunk tárgyához tartoznak továbbá rekombináns expressziós keretek, gazdasejtek, transzgén növények és antitest-készítmények.

Szacharóz-szintetáz gének manipulálása a szár és a mag minőségének javítására

A találmány tárgya általánosságban a növények molekuláris biológiájával kapcsolatos. Pontosabban, a találmány tárgyat nukleinsavak és azok növényekben való kifejeződését módosító eljárások képezik.

A növényi anyagok kémiai összetételét és mechanikai tulajdonságait döntő mértékben az határozza meg, hogy miként vannak ezek a növényi anyagok felhasználva. A sejtfal anyagtartalma és összetétele a felelős a növényi szövetek mechanikai erősségében megfigyelhető variációk legtöbbjéért, de ugyancsak a sejtfal összetétele fogja meghatározni a szilázs minőségét is. A sejtfalak a növények vegetatív részeinek fő fogyasztói, ami például körülbelül 80 tömeg % a kukoricaszár esetében (1. ábra). A teljes kukoricanövényre – a szemterméssel együtt – vonatkoztatva a sejtfalak a száraztömeg körülbelül 35-40 tömeg %-át képezik.

A cellulózt – a Föld legnagyobb mennyiségben előforduló szerves molekuláját – a plazmamembrán termeli és közvetlenül beépíti a sejtfalba [Ray et al., *Ber. Deutsch. Bot. Ges.* 89, 121-146 (1976)]. Inter- és intramolekuláris hidrogénhid-kötések segítségével a β -1,4-glukán láncok parakristályos mikrofibrillumokba rendeződnek, amelyek végső fokon szalagokat és rostokat képezve a cellulóznak igen nagy szakítószilárdságot biztosítanak (Niklas: „*Plant Biomechanics: An Engineering Approach to Plant Form and Function*” University of Chicago Press, 1992). Parakristályos természete miatt a cellulóz aránytalanul nagyobb mértékben járul hozzá a növényi szövetek szakítószilárdságához,

mintha amorf szerkezete lenne.

A kukoricaszár sejtfala zömében cellulózt és hemicellulózt tartalmaz; lignintartalma kisebb arányú, körülbelül 10 tömeg % (1. ábra). Egy vizsgálatban, amit három, ellentétes tulajdonságú hibridpáron végeztünk, azt találtuk, hogy a cső alatti szárrész egységnyi hosszának cellulóz-koncentrációja korrelál a szál szakítószilárdságával (2. ábra.) A cső lekonyulása nagy problémája a kukoricatermesztésnek, mert jelentős veszteséget okoz. A cellulóz-koncentráció emelése a sejtfalban mechanikailag erősebb szövetet eredményez, ami csökkentheti a cső lekonyulásának problémáját.

Elsősorban a cellulóz szintézisének sebessége szabályozza a sejtfal-tömeg képződését, mivel a cellulóz a domináns alkotórész (1. ábra). Az UDP-glükóz – ami a cellulóz-szintetáz (CesA) szubsztrátja – két úton keletkezik a növényekben: az egyik az UDP-glükóz pirofoszforiláz (UDPáz), a másik a szacharóz-szintetáz (3. ábra). A szacharóz-szintetáz (SuSy) az alábbi, reverzibilis reakciót katalizálja:



Így ez a hasítási reakció biztosítja a prekursor a keményítő és a cellulóz szintéziséhez. A SuSy a szacharóz glükozid-kötésének energiáját használja fel az UDP-glükóz szintéziséhez, és a folyamatban fruktóz szabadul fel, amiből az UGPáz-úton át szintén UDP-glükóz keletkezik (3. ábra). Bár a szacharóz-szintetázról korábban azt tartották, hogy a növényi sejtek citoplazmájában működik, Amor és munkatársai kimutatták, hogy a gyapotban a sejtek teljes SuSy-mennyiségének mintegy a fele szorosan kötődik a sejthártyához és a szacharózból készített szubsztrátot közvetle-

nül adhatja át a CesaA-nak [Proc. Natl. Acad. Sci. USA 92, 9353-9357 (1995)]. Így tehát egy fogyasztó szövetben, mint amilyen a növekvő kukoricaszár, a szacharóz-szintetáz gazdaságosabb útja az UDP-glükóz szacharózból való keletkezésének, mert az UGPáz-út több energiát használ fel nukleotid-trifoszfát alakjában az UDP-glükóz hexózokból való előállításához.

Találmányunk megírásáig csak két szacharóz-szintetáz gén volt ismert a kukoricában, a shrunken-1 (*Sh1*) és a konstitutív szacharóz-szintetáz (*Sus1*); mindkettő a 9. kromoszómán helyezkedik el [Huang et al., *Plant Physiology Rockville* 104, 293-294 (1986); McCarty et al., *Proc. Natl. Acad. Sci. USA* 83, 9099-9103 (1986); Werr et al., *EMBO J.* 4, 1373-1380 (1985)]. Ezek a paralógok rendre az SS1 és SS2 szacharóz-szintetáz izozimeket kódolják.

A membránhoz asszociált SuSy-t megtalálták a sárgarépában és a kukoricában is [Carlson et al., *Mol. Gen. Genet.* 252, 303-310 (1996); Sturm et al., *Plant Mol. Biol.* 39, 349-360 (1999)]. A SuSy mindkét ismert formája a fejlődő endospermiumból készített plazmamembrán-frakcióhoz asszociálva található meg a kukoricában. Érdekes módon az *Sh1*-nek nagyobb szerepet tulajdonítanak a sejtfal képződésében, mint a konstitutív szacharóz-szintetáznak (*Sus1*), amiről azt állítják, hogy inkább a keményítő képződésében vesz részt [Chonrey et al., *Mol. Gen. Genet.* 259, 88-96 (1998)]. A SuSy-ról ismert, hogy reverzibilisen foszforilálódhat egy kitüntetett szerilcsoportján [Huber et al., *Plant Physiology Rockville* 112, 793-802 (1996)]. A nem foszforilálódott forma felületének viszonylag hidrofóbabb volta miatt inkább kötődik a sejthártyához [Winter et al., *FEBS Letters* 420, 151-155 (1997)].

A szacharóz-szintetázról feltételezik, hogy szubsztrátot továbbít a mátrix poliszacharid-szintetázoknak; ennek alapja a Golgi-készülékhez való asszociálódása, valamint egy közlemény a cellulóz szintézisében való részvételéről [Buckeridge et al., *Plant Physiology Rockville* 120, 1105-1116 (1999)]. A szerzők nagyobb mennyiségű cellulóz termelődését figyelték meg *Acetobacter xylinum*-ban, ha a baktériumban a mungóbab szacharóz-szintetázát fejeztették ki. A baktériumból hiányzik a szacharóz-szintetáz, így az UDP-glükóz termelése csak az UGPáz-ágon keresztül folyik (3. ábra). A szacharóz-szintetáz expressziója a várt módon magasabb UDP-glükóz- és alacsonyabb UDP-szintet is eredményezett a baktériumban.

Ha a Susy kifejeződését az antiszensz módszerrel leszabályozzuk sárgarépában, akkor csökken a növekedés sebessége [Tang et al., *Plant Mol. Biol.* 41, 465-479 (1999)]. Az UDP-glükóz és a cellulóz szintje csökken a fogyasztó szövetekben a vad típusú növényhez képest, amiből újra az tételezhető fel, hogy a Susy szubsztrátot termel a cellulóz-szintézis számára. A TUSC (Trait Utility System for Corn; lásd az US 5,962,764 számú szabadalmi iratot, amit referenciaként tartunk számon) susy mutánsal dolgozva azt találtuk, hogy a konstitutív szacharóz-szintetáz „kiütése” a cellulóz-koncentráció csökkenéséhez vezet a sejtfalban, és csökken a sejtfal teljes mennyisége is (8. példa).

Az UDP-glükóz szacharózból történő bioszintézise feleannyi energiát igényel, mint amikor a hexóz-cukrokból keletkezik (3. ábra). Még ha a közvetlen szubsztrát-átadást – amit Amor és munkatársai (*id. közl.*) tételeznek fel – figyelmen kívül hagyjuk

is, a szacharóz-szintetáz részvétele a cellulóz-szintetáz szubsztráttal való ellátásában olyan előnyöket biztosít, mint a fokozott produktivitás, ami különösen hátrányos körülmények között érvényesül, mivel az így megtakarított energia más sejtfolyamatokban hasznosulhat. A szacharóz-szintetáz túlkifejeztetése egy szárspecifikus promoter szabályozó hatása alatt fokozott cellulóz-szintézist fog eredményezni, amitől erősebb lesz a szár. Így tehát a szakterület igényli az olyan szacharóz-szintetázokat, amelyek ilyen körülmények között túlkifejeztethetők.

A szacharóz-foszfát-szintetáz is részt vehet az UDP-glükóz anyagcseréjében, de szerepe inkább az energia szétjuttatása lehet a fogyasztó szövetekben, mint a cukrok felhasználásának gazdaságosabbá tétele (3. ábra). Így tehát – feltételezve, hogy az összes fruktóz-6-foszfát és UDP-glükóz a SuSy-útból származik – legalább egy ATP szükséges a szacharóz szintéziséhez ebből a két szubsztrátból, amit azután újra a SuSy fog elbontani. A másik szélső esetben, azaz amikor az összes UDP-glükóz és fruktóz-6-foszfát hexóz-cukrokból származik, a szacharózt a szacharóz-foszfát-szintetáz állítja elő szacharóz-molekulánként 3 nukleotid-trifoszfát felhasználásával: kettő kell egy UDP-glükóz készítéséhez egy hexóz-cukorból és egy kell a fruktóz foszforilálásához. Más szavakkal, a szacharóz-foszfát-szintetáz részvétele a folyamatban azzal jár, hogy 1-3 többlet nukleotid-trifoszfát fogy el egy, a cellulózba épített szacharóz-molekulára vonatkoztatva, ami azt jelenti, hogy a folyamat nettó 3-5 nukleotid-trifoszfátot fogyaszt.

Négy nukleotid-trifoszfátra lenne szükség szacharóz-moleku-

lánként annak UDP-glükózzá alakításához még akkor is, ha a szacharózt előbb invertázok hasítanak el és csak a hexóz-cukrok lennének jelen. Az invertázok éppenhogy elveszítik a glükozid-kötés energiáját, amit a szacharóz-szintetáz az UDP-glükóz szintézisére használ fel. A szacharóz-foszfát-szintetáznak azonban fontos szerepe lehet a szacharóznak a hexózek feleslegéből való előállításában, amely szacharóz azután a fogyasztókhoz, például a fejlődő csövekbe kerül. Ez akkor lehet fontos, amikor a cellulóz berakódása a szárszövetek sejtfaiba már lelassul.

Minden hexóz-cukor molekulából – a glükolízisen, a citromsav-cikluson és az oxidatív foszforiláción át történő teljes lebomlása során – 36 ATP-egyenérték energia szabadul fel. Amint fent írtuk, minden hexóz UDP-glükózzá történő aktiválásához 1 ATP szükséges, ha a folyamat a SuSy-úton át, és 2 ATP szükséges, ha az UGPáz-úton át megy végbe. Az átalakításhoz felhasznált cukorhányadot – feltéve, hogy az összes többi folyamat állandó – az alábbi egyenlet fejezi ki:

$$\frac{p + 2q}{36}$$

ahol p a SuSy által termelt szubsztrát aránya, q a hexóz-cukrokból termelt szubsztrát aránya, és $p + q = 1$. Ha az összes UDP-glükóz a SuSy-úton át keletkezik, akkor a cukor 2,8 %-a használódik el a folyamat energiaigényének kielégítésére. Ha viszont az összes UDP-glükóz hexózekből keletkezik, akkor a reakciósor energiaigényét a cukor 5,6 %-a fogja fedezni.

Ha a cukrok bármely n hányada a szacharóz-foszfát-szintetáz úton fut keresztül, akkor a hatásfok még tovább csökken, mert az

ebben a reakcióban felhasznált nukleotid-trifoszfát hozzáadódik azokhoz, amelyek az UDP-glükóz szacharózból vagy hexózból való előállításához használnának el. A hatásfok csökkenését az alábbi egyenlettel becsülhetjük meg:

$$\frac{(p + 2q) + n(p + 3q)}{36}$$

Ha a cukor 50 %-a a szacharóz-foszfát-szintetáz úton fut át és ennek az enzimnek a szubsztrátja egyenlő arányban (azaz $p = q = 0,5$) származik a SuSy- és az UGPáz-útból, akkor – figyelmen kívül hagyva a szacharóz-foszfát-szintetáz út működéséhez szükséges energiát – a cellulózzá konvertált cukor 4,2 %-ának megfelelő cukor használdik fel a folyamat energiaigényének fedezésére. Ha azonban figyelembe vesszük a szacharóz-foszfát-szintetáz út által felhasznált energiát is, akkor ez a szám körülbelül 7 %-ra emelkedik, ami 70 %-kal nagyobb energiafogyasztást jelent ahhoz képest, amikor nem halad át cukor ezen a reakcióúton. Szemléletesen szólva ez azt jelenti, hogy közel három vékéval több cukrot kell elégetni minden 100 véka cukor poliszachariddá konvertálásához.

Így tehát a cellulóz termelésének a növények számára leggazdaságosabb módja a szacharóz-szintetáz reakcióút. A szakemberek tudták, hogy a szacharóz-szintetáz részt vesz a növények cellulóztermelésében. Találmányunk azonban arra tanít, hogy ez az enzim a cellulóz-szintézis szubsztráttal való ellátásában játszik fontos szerepet (8. példa).

Mivel a szár összetétele számos, a kukorica nemesítésében fontos minőségi jellemzőt befolyásolhat, a szakterületen igény

van olyan termékekre és eljárásokra, amelyekkel a sejtfal cellulóz-koncentrációja módosítható és ezáltal megváltoztatható a szár minősége, például javítható a szilárdsága. Kívánatos lehet a szacharóz-szintetáz túlkifejeztetése, előnyösen egy szár-preferáló promoter vezérlése alatt, hogy javítsuk a kukoricaszár erősségét.

Egy másik fontos tulajdonság a szem kezelhetősége, például a szemtörés csökkentése a betakarítás, szállítás és tárolás során. A gabonaszemek, például a búza és árpa esetében a mag szilárdságát a maghéj (perikarpium) biztosítja, ami megengedi a lágyabb endospermiumot. Kívánatos lenne ezért a perikarpium cellulóz-tartalmának fokozása a szacharóz-szintetáz túlkifejeztetésével egy perikarpium-specifikus promoter vezérlése alatt.

Találmányunk más előnyök mellett ezeket biztosítja.

Azonosítottuk egy harmadik szacharóz-szintetáz gén, a *Sus3* mindeddig ismeretlen cDNS-ét egy saját genom-adatbázisban (lásd *ZmSus3*; 9. és 10. példa). A *Sus3* az 1. kromoszóma rövid karján helyezkedik el (1,04 bin). A gén kifejezett szekvencia-részeit (EST-k) különböző szövetekben találtuk meg, bár jóval kisebb gyakorisággal, mint a *Sus1* génét, ami arra utal, hogy ez a gén is – a *Sus1*-hez hasonlóan – konstitutívan fejeződik ki.

Általánosságban találmányunk célja a szacharóz-szintetáz 3-mal (*Sus3*) kapcsolatos nukleinsavak és fehérjék nyújtása. Találmányunk további célja a találmány szerinti nukleinsavakat tartalmazó transzgén növények, valamint eljárások nyújtása a találmány szerinti nukleinsavak expressziójának befolyásolására a transzgén növényekben. Pontosabban, találmányunk célja a cellu-

lóz-koncentráció befolyásolása a sejtfalban, hogy ezáltal megváltoztassuk a mag és/vagy a szár minőségét. Találmányunk további célja a szacharóz-szintetáz kifejeződésének megváltoztatása egy növényben, hogy ezáltal javítsuk a szár minőségét és/vagy szilárdságát. Találmányunk még további célja a szacharóz-szintetáz kifejeződésének megváltoztatása egy növényben, hogy ezáltal javítsuk a mag minőségét és/vagy szilárdságát.

Így tehát találmányunk egyik szempontja egy izolált nukleinsavval kapcsolatos, ami egy olyan csoport tagja, amely tartalmaz (a) egy polinukleotidot, amely meghatározott szekvencia-egyezeit mutat egy találmány szerinti polinukleotiddal; (b) egy polinukleotidot, amely komplementer az (a) pont alatti polinukleotiddal; és (c) egy polinukleotidot, amely meghatározott számú, összefüggő nukleotidokat tartalmaz egy (a) vagy (b) pont alatti polinukleotidból. Az izolált nukleinsav DNS lehet.

Találmányunk további szempontjai a következőkkel kapcsolatosak: (1) rekombináns expressziós keretek, amelyek egy találmány szerinti nukleinsavat tartalmaznak egy promoterral működőképes módon összekapcsolva; (2) egy gazdasejt, amelybe a rekombináns expressziós keret be lett juttatva; és (3) egy transzgén növény, amely a rekombináns expressziós keretet tartalmazza. A gazdasejt és a növény adott esetben kukorica, búza, rizs vagy szója.

Az ábrák rövid leírása

1. ábra. A kukoricaszár kémiai összetételét bemutató kördiagram. A cső alatti második és harmadik ízközt 60 nappal a virágzás után begyűjtöttük, megszáritottuk és megőröltük. A szerkezeti szárazanyagot úgy különítettük el, hogy az elaprított anyagot

pufferrel, majd metanol/kloroform eleggyel mostuk. A cellulózt Updegraff módszerével (az őrölt vagy a szerkezeti szárazanyagot ecetsav/salétromsav elegyben forralva), gravimetriásan, a lignint Klason módszerével mértük meg. A hamvasztást 600 °C-on, 4 órán át folytattuk. A fehérjetartalmat körülbelül 3 tömeg %-nak tételeztük fel. Az oldható alkotórészek mennyiségét úgy kaptuk meg, hogy a szerkezeti szárazanyag tömegét kivontuk az összes szárazanyag tömegéből. A hemicellulóz mennyiségét úgy becsültük meg, hogy a cellulóz, a lignin, a fehérje és a hamu tömegét kivontuk a szerkezeti szárazanyag tömegéből. Ismert, hogy az egyszikűek érett sejtfalaiban csak igen kevés pektin van.

2. ábra. Kukoricahibridek cellulóz-tartalmát bemutató oszlopdiagram. A cső alatti második és harmadik ízközt 60 nappal a virágzás után begyűjtöttük, megszáritottuk és megőröltük. A szerkezeti szárazanyagot úgy különítettük el, hogy az elaprított anyagot pufferrel, majd metanol/kloroform eleggyel mostuk. A cellulózt Updegraff módszerével (az őrölt vagy a szerkezeti szárazanyagot ecetsav/salétromsav elegyben forralva), gravimetriásan, a lignint Klason módszerével mértük meg. A hamvasztást 600 °C-on, 4 órán át folytattuk. A fehérjetartalmat körülbelül 3 tömeg %-nak tételeztük fel. Az oldható alkotórészek mennyiségét úgy kaptuk meg, hogy a szerkezeti szárazanyag tömegét kivontuk az összes szárazanyag tömegéből. A hemicellulóz mennyiségét úgy becsültük meg, hogy a cellulóz, a lignin, a fehérje és a hamu tömegét kivontuk a szerkezeti szárazanyag tömegéből. Ismert, hogy az egyszikűek érett sejtfalaiban csak igen kevés pektin van.

3. ábra. Az UDP-glükóz bioszintézisének lehetőségeit bemutató

reakcióvázlat. Rövidítések: ATP = adenzin-trifoszfát, CesA = cellulóz-szintetáz, HK = hexokináz, PPáz = pirofoszfátáz, PPI = pirofoszfát, SPP = szacharóz-foszfát foszfátáz, SPS = szacharóz-foszfát-szintetáz, SuSy = szacharóz-szintetáz, UDP-glükóz = uridin-difoszfát-glükóz, UGPáz = UDP-glükóz pirofoszforiláz, UTP = uridin-trifoszfát.

4. ábra. A *Sus1* EST-k kukoricaszövetekben való megoszlását bemutató kördiagram. A *Sus1* 230 EST-jét találtuk meg a genom adatbázisban, amely megközelítőleg 400.000 EST-t tartalmaz.

5. ábra. A *Sus3* EST-k kukoricaszövetekben való megoszlását bemutató kördiagram. A *Sus3* 26 EST-jét találtuk meg a genom adatbázisban, amely megközelítőleg 400.000 EST-t tartalmaz.

6. ábra. A *ZmSus1* genom-klónjának megjelenítésén a keskeny vonalak intronokat, a széles vonalak exonokat jelentenek. A nyilak két, egymástól független Mu-inszertum allélre mutatnak a 12. exonban (a pontos helyüket a fenti szekvenciavázlatok mutatják meg).

7. ábra. A *Sus1* mutánsok sejtfalának cellulóz-tartalma.

8. ábra. A kukorica szacharóz-szintetázok aminosav-szekvenciáinak illeszkedése.

9. ábra. A kukorica szacharóz-szintetázok polinukleotidjainak illeszkedése.

10. ábra. A 13. szekvenciavázlat szerinti szekvencia: cirok-EST, génbanki azonosítója BF481989. A nyílt leolvasó keret első metioninját kódoló ATG triplet félkövér betűkkel van kiemelve. Az aláhúzott szekvenciát használtuk fel a *Sus3* teljes hosszúságú szekvenciájának levezetéséhez.

11. ábra. A kukorica és a cirok *Sus3* szekvenciáinak kombinációját

használtuk fel a 11. számú szekvenciavázlat elkészítéséhez. A 13. számú szekvenciavázlatból kiemelt cirok-szekvenciát és az 1. számú szekvenciavázlatból kiemelt kukorica-szekvenciát külön mutatjuk be. Mielőtt a két szekvenciát egyesítenénk, az 1. számú szekvenciavázlatból származó szekvenciából el kell távolítani az inverz nyomtatással kiemelt nukleotidokat.

Áttekintés

A. A találmány szerinti nukleinsavak és fehérjék

Az 1. és 11. számú szekvenciavázlatokban megadott polinukleotid-szekvenciák, valamint a 2. és 12. számú szekvenciavázlatokban megadott polipeptid-szekvenciák képviselnek egy találmány szerinti polinukleotidot, illetve polipeptidet. Egy találmány szerinti nukleinsav egy találmány szerinti polinukleotidot tartalmaz. Egy találmány szerinti fehérje egy találmány szerinti polipeptidet tartalmaz.

B. A találmány példaszerű alkalmazási területei

Találmányunk lehetőséget biztosít olyan példaszerű alkalmazásokra, mint a cellulóz-koncentráció befolyásolása sejtfalban és ezáltal a növény szára minőségének megváltoztatása például a jobb állóképesség érdekében. Kíváncos lenne a szacharóz-szintetáz túlkifejeztetése előnyösen egy szár-preferáló promoter vezérlése alatt, hogy ezáltal javítsuk a kukorica szárának szilárdságát.

Egy másik fontos tulajdonság a szem kezelhetősége, például a szemtörés csökkentése a betakarítás, szállítás és tárolás során. A gabonaszemek, például a búza és árpa esetében a mag szilárdságát a maghéj (perikarpium) biztosítja, ami megengedi a lágyabb

endospermiumot. Kívánatos lenne ezért a perikarpium cellulóztartalmának fokozása a szacharóz-szintetáz túlkifejeztetésével egy perikarpium-specifikus promoter vezérlése alatt.

C. **Példaszerűen előnyös megvalósítások**

Bár a különböző előnyös megvalósításokat a leírásban bemutatjuk, a példaszerűen előnyös megvalósítások közé tartoznak a következők.

(i) **A szacharóz-szintetáz gének expressziós mintázata**

A *Sus1* gént körülbelül 230, a *Sus3* gént körülbelül 26 EST képviseli, amelyeket a Pioneer Hi-Breed International, Inc. tulajdonát képező adatbázisban találtunk; az adatbázis magában foglalja a számos, saját tulajdonú nukleinsav-könyvtár adatait, amelyek különböző fejlődési stádiumú növényi szöveteket képviselnek. Ezek a fellelt EST-k egyfajta elektronikus Northern-blotként tekinthetők, és azt bizonyítják, hogy a *Sus3* sokkal alacsonyabb szinten fejeződik ki, mint a *Sus1* (4. és 5. ábrák). Mind a *Sus1*, mind a *Sus3* sokféle szövetben kifejeződik, és emiatt konstitutív szacharóz-szintetázoknak tekinthetők. A *Sus3* azonban mintha inkább a magbelet preferálná, ahol az EST-k 50 %-át találtuk. Ezzel szemben a *Sus1* EST-knek csak körülbelül 10 %-át (ha a kalluszt kizárjuk, akkor 15 %-át) találtuk a magbél-szövetben. A két gén között egy éles különbség van: a *Sus3* – úgy tűnik – egyáltalán nem fejeződik ki a kallusz-szövetben, míg a *Sus1* EST-k körülbelül felét a kallusz-könyvtárakban találtuk meg.

(ii) **Promoterok**

Az előnyös promoterok közé tartoznak – nem kizárólag – a következők: a rizs Actin-1 promotera [McElroy et al., *Plant Cell*

2, 163-171 (1990)], a rizs tungro bacilliform vírus promotera [Yin et al., *Plant J.* 7, 969-980 (1995)], az *Agrobacterium rhizogenes* RolC és a kukorica Sh promotera [Graham et al., *Plant Mol. Biol.* 33, 729-735 (1997)], az US 5,986,174 számú szabadalmi iratban – amit referenciaként tartunk számon – leírt, szövetpreferáló promoter, a kukorica vagy lucerna S2A promotera [Abrahams et al., *Plant Mol. Biol.* 27, 513-528 (1995)], a kukorica Adh2 promoter-elemei [Paul et al., *Plant J.* 5, 523-533 (1994)], A CoYMV promoter [Medberry et al., *Plant Cell* 4, 185-192 (1992)], a bab grpl.8 promotera és annak regulátor-elemei [Keller et al., *Plant Mol. Biol.* 26, 747-756 (1994)], a paradicsom proszisztemin-promotera [Jacinto et al., *Planta* 203, 406-412 (1997)], a kukorica Hrgp génjének promotera [Menossi et al., *Plant Sci.* 125, 189-200 (1997)], a kukorica Sus1 promotera [Huang et al., *Euphytica* 103, 17-21 (1998)], a kukorica P-rr génjének promotera [Sidorenko et al., *Plant Mol. Biol.* 39, 11-19 (1999)] és az US-A 09/666,179 számú szabadalmi iratban leírt mZE40-2 kukorica promoter.

Meghatározások

Találmányunkban az egységeket, prefixumokat és jelöléseket az SI rendszernek megfelelően használjuk. Hacsak mást nem állítunk, a nukleinsavakat balról jobbra, 5' → 3' irányban, a fehérjéket balról jobbra, N-terminális → C-terminális irányban írjuk le. A találmányunkban megadott számszerű értékhatárok magukban foglalják a határértékeket és minden egész számot a határértékek között. Az aminosavakat a közismert, három betűs szimbólumaikkal vagy a IUPAC-IUB Biochemical Nomenclature Commission által ajánlott, egy betűs szimbólumokkal jelöljük. Hasonlóképpen a nukleo-

tidokat is az általánosan elfogadott, egy betűs szimbólumaikkal jelöljük. A szoftverekkel, elektromossággal és elektronikával kapcsolatos kifejezéseket a „*The New IEEE Standard Dictionary of Electrical and Electronic Terms*” (5. kiadás, 1993) értelmező szótár szerint használjuk. Néhány kifejezés értelmét a leírással kapcsolatban az alábbiakban pontosan megadjuk. Az egyes fejezetek címei nem korlátozzák találmányunk különböző céljait és megvalósításait.

A „sokszorozás” kifejezés nagy számú másolat készítését jelenti egy nukleinsav-szekvenciáról, vagy nagy számú, az adott nukleinsav-szekvenciával komplementer másolat készítését jelenti egy nukleinsav-szekvenciáról, ami legalább egy példányban mintául (templát) szolgál. A sokszorozó rendszerek közé tartozik a polimeráz láncreakció (PCR), a ligáz láncreakció (LCR), a nukleinsav-szekvenciára alapozott sokszorozó (NASBA, Cangene, Mississauga, Ontario), a Q-Beta replikáz, a transzkripcióra alapozott sokszorozó (TAS) és a szálhelyettesítéssel sokszorozó (SDA) rendszer. Leírásuk megtalálható például a „*Diagnostic Molecular Microbiology: Principles and Applications*” című kézikönyvben (Persing és munkatársai, szerk., American Society for Microbiology, Washington D.C., 1993). A sokszorozás termékét „amplikon”-nak nevezzük.

Az „antiszensz orientáció” azt jelenti, hogy egy kettős polinukleotid szekvencia – amely működőképes módon van egy promoterhez kapcsolva – olyan irányultságú, hogy az antiszensz szál íródik át. Az antiszensz szál kellőképpen komplementer egy adott endogén transzkripciós termékkel ahhoz, hogy az endogén transz-

kripciós termék transzlációját gyakran gátolni tudja.

A „megváltoztat”, „módosít” vagy „befolyásol” szavak a nukleinsavak vagy fehérjék kifejeződésével kapcsolatban olyan módszerekre utalnak, amelyekkel az expresszió fel- vagy leszabályozható. A felszabályozás például úgy érhető el, hogy fokozzuk a szóban forgó gén transzkripcióját vagy transzlációját olyan módon, hogy a gént működőképes módon összekapcsoljuk egy promoter-szekvenciával, ami fokozza a transzkripciót; egy anyagcsereúthoz szükséges szubsztrátot hozzáadunk vagy túlkifejeztetünk; az antagonista molekulákat blokkoljuk; vagy más, a szakemberek által ismert eljárásokat alkalmazunk. A leszabályozás például antiszensz technológiával [lásd például Sheehy et al., *Proc. Natl. Acad. Sci. USA* 85, 8805-8809 (1998) vagy az US 5,759,829 számú szabadalmi iratot], RNS-interferenciával [Napoli et al., *Plant Cell* 2, 279-289 (1990); US 5,034,323; Sharp, *Genes Devel.* 13, 139-141 (1999); Zamore et al., *Cell* 101, 25-33 (2000); Montgomery et al., *Proc. Natl. Acad. Sci. USA* 95, 15502-15507 (1998)], vírussal indukált génelnyomással [Burton et al., *Plant Cell* 12, 695-705 (2000); Baulcombe, *Curr. Opin. Plant Biol.* 2, 109-113 (1999)], a cél-RNS-re specifikus ribozimokkal [Haseloff et al., *Nature* 334, 585-591 (1998)], hajszál-hurok szupresszióval [Smith et al., *Nature* 407, 319-320 (2000)], vagy más, a szakemberek által ismert módszerekkel valósítható meg. Az említett fel- vagy leszabályozás előnyösen irányítható, azaz úgy, hogy bizonyos szövettípusokban, adott környezeti körülmények között és/vagy a növény fejlődésének bizonyos szakaszaiban nyilvánuljon meg.

A „kódol” kifejezés egy adott nukleinsavval kapcsolatban azt jelenti, hogy az egy meghatározott fehérjébe lefordítható információkat tartalmaz. Egy fehérjét kódoló nukleinsav tartalmazhat nem lefordítható szekvenciákat (például intronokat) is a lefordítható szakaszokon belül, de ezek a közbeékeltek, nem lefordítható szekvenciák teljesen hiányozhatnak is belőlük (például a cDNS-ekből). A fehérjét kódoló információ a kodonokban testesül meg. Az aminosav-szekvenciát a nukleinsavak tipikusan az univerzális genetikai kód formájában tartalmazzák, azonban ennek bizonyos változatai is használhatók, ha a nukleinsav bizonyos növényi, állati vagy gomba-mitokondriumokban, a *Mycoplasma capricolum* baktériumban vagy *Macronucleus* csillós egysejtűekben fejeződik ki.

Ha a nukleinsavat szintetikusan állítjuk elő vagy módosítjuk, előnyös ismerni annak a gazdaszervezetnek a kodon-preferenciáját, amelyben a DNS-t ki akarjuk fejeztetni. Így például, noha a találmány szerinti nukleinsav-szekvenciák mind egyszikű, mind kétszikű növényfajokban kifejeztethetők, a szekvenciákat az egyszikűek és kétszikűek specifikus kodon-preferenciájának és GC-arányának figyelembe vételével módosíthatjuk, mivel ezekről tudjuk, hogy különbözőek [Murray et al., *Nucl. Acids Res.* 17, 477-498 (1989)]. Ennek megfelelően például a kukorica egy adott aminosav számára preferált kodonja a növény ismert génjeinek szekvenciáiból levezethető. A kukorica kodon-preferenciáját 28 gén elemzése alapján Murray és munkatársai adják meg (Murray és munkatársai, *id. közl.*, 4. táblázat).

A „teljes hosszúságú szekvencia” kifejezést egy meghatározott polinukleotidra vagy az általa kódolt fehérjére alkalmaz-

zuk, és azt jelenti, hogy a meghatározott fehérje tartalmazza a természetes (nem szintetikus), endogén, biológiailag (strukturálisan vagy katalitikusan) aktív forma teljes aminosav-szekvenciáját. Az eljárások annak eldöntésére, hogy egy szekvencia teljes hosszúságú-e vagy sem, jól ismertek a szakterületen; ilyenek például a Northern- vagy Western-blot technikák, a primerek meghosszabbítása, az S1- vagy a ribonukleáz-védelem [Clark (ed.): „*Plant Molecular Biology: A Laboratory Manual*”; Springer Verlag, Berlin, 1997]. Az ismert, teljes hosszúságú homológ (ortológ és/vagy paralóg) szekvenciákkal való összehasonlítás szintén használható a találmány szerinti teljes hosszúságú szekvenciák azonosítására. Emellett a konszenzus-szekvenciák – amelyek tipikusan a mRNS molekulák 5' és 3' le nem fordított végén találhatók – is segíthetik annak eldöntését, hogy egy polinukleotid teljes hosszúságú-e. Így például az ANNNNAUGG konszenzus-szekvencia – amelyben az aláhúzott kodon az N-terminális methionint kódolja – segít annak eldöntésében, hogy a polinukleotid 5' vége teljes-e vagy sem. A 3' végen ugyanezt a szerepet a poliadenil-szekvenciák tölthetik be.

A „heterológ” vagy „idegen” kifejezést akkor használjuk egy nukleinsavra, ha egy idegen fajból származik, vagy – ha ugyanabból a fajból származik – emberi beavatkozással jelentősen módosítva lett a természetes formájához képest összetételében és/vagy a genombeli elhelyezkedése tekintetében. Így például egy promoter, ami működőképes módon van egy heterológ struktúrgénhez kapcsolva, származhat más fajból, mint amelyikből a struktúrgén származik, vagy – ha ugyanabból a fajból származnak – egyik vagy

mindkettő jelentősen módosítva lehet az eredeti formájához képest. Egy heterológ fehérje származhat egy idegen fajból, vagy – ha ugyanabból a fajból származik – emberi beavatkozással jelentősen módosítva lehet az eredeti alakjához képest.

A „gazdasejt” kifejezés egy olyan sejtet jelent, amely egy vektort tartalmaz és támogatja az expressziós vektor replikációját és/vagy expresszióját. A gazdasejt lehet egy prokarióta sejt, például az *E. coli*, vagy eukarióta sejt, például élesztő, rovar, kétéltű vagy emlős sejt. A gazdasejtek előnyösen egyszikű vagy kétszikű növényi sejtek. Egy különösen előnyös egyszikű gazdasejt a kukorica gazdasejt.

A „bejuttatás” kifejezés magában foglalja egy nukleinsav beépítését is egy eukarióta vagy prokarióta sejtbe, ahol a nukleinsav beépül a sejt genomjába (azaz kromoszomális, plazmid, plasztid vagy mitokondriális DNS-ébe), egy autonóm replikonná alakul át vagy átmenetileg kifejeződik (mint például egy transzfektált mRNS). A kifejezés magában foglalja a nukleinsav sejtbe juttatásának olyan módjait, mint a „transzfekeció”, „transzformáció” vagy „transzdukció”.

Az „izolált” kifejezés olyan anyagot, így nukleinsavat vagy fehérjét jelöl, amely lényegében vagy teljesen mentes azoktól az alkotórészeketől, amelyek a természetes környezetében hozzá társulnak vagy kölcsönhatnak vele. Az izolált anyag adott esetben olyan anyagot tartalmazhat, amely a természetes környezetében nem fordul elő; vagy ha az anyag a természetes környezetében van, akkor emberi beavatkozással, szintetikus (nem természetes) úton lett megváltoztatva az összetétele és/vagy a sejten belüli

helye (a genom vagy egy sejtszervecske) olyanra, amelyben az anyag nem található meg természetes körülmények között. A szintetikus anyagot eredményező változtatást akár a természetes környezetben, akár onnan eltávolítva elvégezhetjük az anyagon. Így például egy természetes nukleinsav akkor válik izolált nukleinsavvá, ha nem természetes, szintetikus eljárásokkal megváltoztatjuk vagy előállítjuk, vagy egy olyan DNS-ről írjuk át, amely nem természetes, szintetikus eljárásokkal lett megváltoztatva. Az izolált nukleinsav létrehozható az adott gén egy vagy több allélje egy vagy több részének szintetikus átrendezésével („megkeverés”) is. Hasonlóképpen, egy természetben előforduló nukleinsav (például egy promoter) akkor válik izolálttá, ha a genom egy másik helyére építjük be. A fenti meghatározás szerinti „izolált” nukleinsavakat „heterológ” nukleinsavaknak is nevezzük (lásd például az US 5,565,350 és a WO 93/22443 számú szabadalmi iratokat).

A „nukleinsav” egy dezoxiribonukleotid vagy ribonukleotid polimer, vagy a kettő kimérája, ami lehet egy- vagy kettős szálú alakban, és a kifejezés – hacsak nem korlátozzuk – magában foglalja az összes ismert analógot, amely a természetes nukleotidok lényeges tulajdonságaival rendelkezik, azaz hibridizálódni képes egyszálú nukleinsavakkal a természetes nukleotidokhoz hasonló módon (például peptid-nukleinsavak).

A „nukleinsav-könyvtár” kifejezés izolált DNS- vagy RNS-molekulák olyan gyűjteményét jelenti, amely tartalmazza és lényegében képviseli egy meghatározott szervezet, szövet vagy egy sejttípus genomjának teljes átírható részét. A példászerű nukleinsav-könyvtárak, például egy genomiális vagy cDNS-könyvtár el-

készítésének módját ismert kézikönyvekből tanulhatjuk meg (Berger és Kimmel: „*Guide to Molecular Cloning Techniques*”, Methods in Enzymology, 152. kötet, Acad. Press, Inc., San Diego; Sambrook et al., „*Molecular Cloning - A Laboratory Manual*”, 1989; és Ausubel et al., (eds.): „*Current Protocols in Molecular Biology*”, John Wiley & Sons, Inc., 1994).

A „működőképes módon összekapcsolva” kifejezés egy funkcionális kapcsolatot jelent egy promoter és egy második szekvencia között, ahol a promoter-szekvencia indítja meg és működteti a második szekvenciának megfelelő DNS-szekvencia transzkripcióját. A működőképes módon való összekapcsolás általánosságban azt jelenti, hogy a nukleinsav-szekvenciák közötti kapcsolat folytonos, és – ha szükséges – két fehérje-kódoló területet egyesít folytonosan és ugyanabban a leolvasó keretben.

A „növény” szó magában foglalja a teljes növényeket és azok származékait, a növényi sejteket, a növények részeit vagy szerveit, mint amilyenek az embriók, a pollen, a petesejtek, a magvak, virágok, magkezdemények, virágzatok, kukoricacsövek, levelek, hөvelyek, szarak, ágak, gyökerek, gyökércsúcsok, porzók, kukoricahaj, és a többi. A növényi sejtek – korlátozás nélkül – a következőkben található vagy azokból nyerhető sejtek lehetnek: magvak, szuszpenzió-tenyészetek, csírák, osztódószövetek, kaluszszevetek, levelek, gyökerek, hajtások, gametangiumok, sporangiumok, pollenek vagy mikrospórák. A növényi sejtek közé tartoznak a módosított sejtek, így a protoplasztok, amelyek az említett szövetekből nyerhetők. A találmány szerinti eljárásokban használható növények csoportja általában egybeesik a transzfor-

mációs technikákkal megközelíthető magasabbrendű növények csoportjával mind az egy-, mind a kétszikű növények körében. Egy különösen előnyös növény a *Zea mays*.

A „polinukleotid” kifejezés egy dezoxiribopolinukleotidot, ribopolinukleotidot, ezek kiméráit vagy ezek analógjait jelenti, amelyek rendelkeznek a természetes dezoxi- vagy ribonukleotidok lényeges tulajdonságaival abban az értelemben, hogy szigorú körülmények között képesek hibridizálódni egy lényegében azonos, de a természetben előforduló nukleotiddal, és/vagy ugyanazon aminosav(ak)ra fordíthatók le, mint a természetes nukleotid(ok). Egy polinukleotid lehet egy teljes hosszúságú, természetes vagy heterológ struktúr- vagy regulátorgén, de lehet annak egy rész-szekvenciája is. Hacsak másként nem adjuk meg, a kifejezés magában foglalja a meghatározott szekvenciát és annak komplementerét is. Így tehát azok a DNS-ek vagy RNS-ek, amelyek gerincét stabilitási vagy más okokból módosítottuk, a „polinukleotid” kifejezés általunk értelmezett körébe tartoznak. Ide tartoznak továbbá azok a DNS-ek és RNS-ek is, amelyek rendhagyó bázisokat, például inozint, vagy módosított, például tritilezett bázisokat tartalmaznak, hogy csak két példát említsünk. Tudott dolog, hogy DNS-eken és RNS-eken számos módosítást lehet végezni, amelyek mind valamilyen hasznos célt szolgálnak. A „polinukleotid” kifejezés az általunk használt értelemben felöleli a polinukleotidok kémiai, enzimes vagy anyagcsere-úton módosított összes formáját, valamint a DNS és RNS minden, a vírusokra és sejtekre – köztük az egyszerűbb és bonyolultabb sejtekre – jellemző kémiai formáját is.

Az egymással felcserélhetően használható „polipeptid”,

„peptid” és „fehérje” szavak aminosavakból álló polimereket jelentenek. A kifejezések alkalmazhatók az olyan aminosav-polimerekre, amelyekben egy vagy több aminosav mesterséges kémiai analógokkal van helyettesítve, de a természetben előforduló aminosav-polimerekre is. A természetes aminosavak mesterséges analógjainak lényeges tulajdonsága, hogy ha egy fehérjébe beépülnek, akkor a fehérje továbbra is reagál azzal az antitesttel, amely a csak természetes aminosavakat tartalmazó fehérjére specifikus. A „polipeptid”, „peptid” és „fehérje” elnevezések magukban foglalják az olyan módosításokat is, mint például a glikozilezés, lipiddel kapcsolás, szulfatálás, a glutaminsavak γ -észterezése, hidroxilezés vagy ADP-ribozilezés; a felsorolás nem kizáró jellegű. Továbbá, találmányunk magában foglalja a találmány szerinti fehérjének mind a metionint tartalmazó, mind a metionin nélküli N-terminálissal rendelkező változatait.

A „promoter” kifejezéssel a DNS egy szakaszát jelöljük, amely a transzkripció kezdőpontjától „felfelé” helyezkedik el, és részt vesz az RNS-polimeráz és más fehérjék felismerésében és megkötésében, hogy ezzel megindítsa a transzkripciót. Egy „növényi promoter” olyan promoter, amely képes megindítani a transzkripciót növényi sejtekben függetlenül attól, hogy növényi sejtből származik-e vagy sem. A példaszerű növényi promotorok közé tartoznak – nem korlátozva – a növényekből, növényi vírusokból és a növényi sejtekben kifejeződni képes géneket tartalmazó baktériumokból, például *Agrobacterium*- vagy *Rhizobium*-fajokból nyerhető promotorok. A „szövetpreferáló” promotorok a transzkripciót elsősorban bizonyos szövetekben, így levelekben, gyöke-

rekben vagy magvakban indítják meg, míg azokat, amelyek kizárólag vagy csaknem kizárólag egy bizonyos szövetben teszik ezt, „szövetspecifikus”-nak nevezzük. A „sejttípus-specifikus” promóter elsősorban egy vagy több szervben, de csak egy sejttípusban – például a hajtások vagy levelek vaszkuláris sejtjeiben – serkentik az expressziót. Egy promóter lehet tér- vagy időspecifikus, amikor csak a térben vagy időben megfelelő körülmények között képes megindítani a transzkripciót. Az „indukálható” vagy „represszálható” promóterek működését a környezet szabályozza. Az indukálható promóterek által vezérelt transzkripciót kiváltó környezeti hatások közé tartozik például az anaerob körülmények vagy a fény jelenléte. A szövetspecifikus, szövetpreferáló, sejttípus-specifikus és indukálható promóterek képezik a „nem konstitutív” promóterek osztályát. A „konstitutív” promóterek viszont a legtöbb környezeti körülmény között is aktívak.

A „rekombináns” kifejezés egy olyan sejtet vagy vektort jelöl, amely egy idegen nukleinsav bejuttatásával lett módosítva, vagy egy olyan sejtet, amely egy így módosított sejtből származik. Így tehát például a rekombináns sejtek a szándékos emberi beavatkozás következtében olyan géneket fejeznek ki, amelyek nem találhatók meg ugyanilyen formában a sejt természetes (nem rekombináns) alakjában, vagy olyan természetes géneket fejeznek ki, amelyek egyébként csak korlátozott mértékben vagy egyáltalán nem fejeződnek ki a sejtben. A „rekombináns” kifejezés nem foglalja magában a sejt vagy vektor azon elváltozásait, amelyek természetes események következményei (ilyenek a spontán mutációk, a természetes transzformációk, transzdukciók és transzpozí-

ciók), és szándékos emberi beavatkozás nélkül jönnek létre.

A „rekombináns expressziós keret” kifejezés egy rekombinációval vagy szintetikus úton előállított nukleinsav-szerkezet jelent, ami egy sor meghatározott nukleinsav-elemből áll, amelyek lehetővé teszik egy bizonyos nukleinsav transzkripcióját egy gazdaszövetben. A rekombináns expressziós keret egy plazmidba, kromoszómába, mitokondriális DNS-be, plastid-DNS-be, vírusba vagy nukleinsav-fragmentumba lehet beépítve. Egy expressziós vektor „rekombináns expressziós keret” része egyéb szekvenciák mellett tipikusan az átírandó nukleinsavat és a promotert tartalmazza.

Az „aminosav-maradék” vagy „aminosav” kifejezéseket felcserélhető módon használjuk a fehérjébe, polipeptidbe vagy peptidbe (együtt „fehérje”) inkorporált aminosavakra. Az aminosav lehet természetes aminosav és – hacsak nem korlátozzuk – azok bármely ismert analógja, amely a természetes aminosavakkal azonos módon képes funkcionálni.

A „szelektív hibridizálódás” kifejezés egy nukleinsav-szekvencia szigorú feltételek között történő hibridizálódását jelenti egy meghatározott nukleinsav-célszekvenciával kimutathatóan nagyobb (a háttér legalább kétszeresen meghaladó) mértékben, mint egy nem célszekvenciával, egészen a célszekvenciától eltérő nukleinsav lényegében vett kizárásáig. A szelektíven hibridizálódó szekvenciák azonossága tipikusan legalább körülbelül 80, előnyösen 90 és legelőnyösebben 100 %-os, azaz komplementerek egymással.

A „szigorú feltételek” vagy „szigorú hibridizációs feltételek” kifejezések olyan körülményeket jelentenek, amelyek között egy „próba” kimutathatóan nagyobb (a háttér legalább kétszere-

sen meghaladó) mértékben hibridizálódik a célszekvenciájával, mint más szekvenciákkal. A szigorú feltételek szekvenciafüggőek és különböző körülmények között különbözőek lehetnek. A hibridizálási és/vagy a mosási körülmények szigorúságának változtatásával felismerhetők azok a célszekvenciák is, amelyek 100 %-ban komplementerek a próbával (homológia-vizsgálat). Másrészt a szigorúság feltételei úgy is beállíthatók, hogy a kisebb hasonlósági fokú szekvenciák is kimutathatók legyenek (heterológia-vizsgálat). A próba általában rövidebb, mint 1000 nukleotid, adott esetben nem éri el az 500 nukleotid hosszúságot.

A szigorú feltételek tipikusan olyanok, hogy a sókoncentráció kisebb, mint körülbelül 1,5 mól/l nátriumion, tipikusan körülbelül 0,01-1,0 mól/l nátriumion (vagy másféle só) 7,0-8,3 pH között, és a hőmérséklet legalább körülbelül 30 °C a rövid próbák (azaz a 10-50 nukleotid hosszúságúak), és legalább körülbelül 60 °C a hosszabb próbák (azaz az 50 nukleotidnál hosszabbak) számára. A szigorú feltételek létrehozhatók destabilizáló reagensek, például formamid hozzáadásával is. Példaszerűen alacsony szigorúságú feltételeket biztosít egy pufferoldat 30-35 % formamiddal, 1 mól/l nátrium-kloriddal és 1 % SDS-tal (nátrium-dodecilszulfát) 37 °C-on, és a mosás 1X-2X SSC-vel 50-55 °C-on (20X SSC = 3,0 mól/l nátrium-klorid + 0,3 mól/l trinátrium-citrát). Közepes szigorúságú feltételt biztosít a hibridizáció 40-45 % formamid, 1 mól/l nátrium-klorid és 1 % SDS jelenlétében, 37 °C-on, és a mosás 0,5X-1X SSC-vel 55-60 °C-on. Erősen szigorú feltételek az 50 % formamid, 1 mól/l nátrium-klorid és 1 % SDS 37 °C-on, és a mosás 0,1X SSC-vel 60-65 °C-on.

A specificitás tipikusan a hibridizálódás utáni mosásnál derül ki; a kritikus tényezők a mosóoldat ionerőssége és hőmérséklete. A DNS/DNS hibridek esetében a T_m jól megközelíthető Meinkoth és Wahl egyenletével [*Anal. Biochem.* 138, 267-284 (1984)], ami szerint

$$T_m = 81,5\text{ °C} + 16,6(\log M) + 0,41(\%GC) - 0,61(\% \text{ formamid}) - 500/L$$

ahol M az egyértékű kationok molaritása, $\%GC$ a guanozin és citozin nukleotidok aránya a DNS-ben, $\%$ formamid a formamid aránya a hibridizáló oldatban, és L a hibrid hossza bázispárokból mérve. A T_m az a hőmérséklet (adott ionerősségen és pH-n), amelyen egy komplementer célszekvencia 50 %-a hibridizálódik a tökéletesen hozzáillő próbával. A kevertség minden 1 %-os fokozódásával a T_m körülbelül 1 °C-kal csökken; így tehát a T_m , a hibridizálás és/vagy a mosás körülményei úgy állíthatók be, hogy a kívánt azonosságú szekvenciák hibridizálódjanak. Így például ha ≥ 90 %-os azonosságú szekvenciákat keresünk, akkor a T_m 10 °C-kal csökkenthető. A szigorú feltételeket általában úgy választjuk meg, hogy a hőmérséklet körülbelül 5 °C-kal alacsonyabb legyen az adott szekvencia és komplementere olvadási hőmérsékleténél (T_m) az adott ionerősség és pH mellett. A igen szigorú feltételek esetén a hibridizálást és/vagy a mosást a T_m -nél 1, 2, 3 vagy 4 °C-kal alacsonyabb hőmérsékleten végezhetjük; a kis mértékben szigorú feltételek esetén a hibridizálást és/vagy a mosást a T_m -nél 6, 7, 8, 9 vagy 10 °C-kal alacsonyabb hőmérsékleten végezhetjük; a kis mértékben szigorú feltételek esetén a hibridizálást és/vagy a mosást a T_m -nél 11, 12, 13, 14, 15 vagy 20 °C-kal alacsonyabb hőmérsékleten végezhetjük. A fenti egyenletet hasz-

nálva a szakemberek meghatározhatják a hibridizálás és mosás körülményeit. Ha a keveredés kívánt foka miatt a T_m alacsonyabb lenne, mint 45 °C (vizes oldatban) vagy mint 32 °C (formamidos oldatban) akkor előnyös az SSC koncentrációját úgy megemelni, hogy magasabb hőmérsékletet használhassunk. A hibridizálás és/ vagy a mosás legalább 10, 30, 60, 90, 120 vagy 240 percen át folytatható. A nukleinsavak hibridizálásához kimerítő útmutatást találunk Tijssen: „*Laboratory Techniques in Biochemistry and Molecular Biology - Hybridization with Nucleic Acid Probes*” című művében (Elsevier, New York, 1993), és Ausubel ismert kézikönyvében.

A „transzgén növény” kifejezés egy olyan növényt jelent, amely egy heterológ polinukleotidot tartalmaz a genomjában. A heterológ polinukleotid általában stabilan be van épülve a genomba és így átadódik az egymást követő generációknak. A heterológ polinukleotid önmagában vagy egy rekombináns expressziós keret részeként lehet jelen a genomban. A „transzgén” jelzöt használjuk bármely sejtre, sejtvonalra, kalluszra, szövetre, növényi részre vagy növényre, amelynek genotípusa megváltozott a heterológ nukleinsav jelenléte következtében, ide értve az eredetileg megváltoztatott növényt és keresztezéssel vagy vegetatív úton szaporított utódait. A „transzgén” kifejezés nem foglalja magában a genom kromoszomális vagy extrakromoszomális megváltoztatását a hagyományos nemesítési módszerekkel vagy a természetben előforduló események által, mint amilyen a véletlenszerű keresztbeporzás, a nem rekombináns vírussal vagy baktériummal való fertőződés, a nem rekombináns transzpozíció vagy spontán mutáció.

A „vektor” szó itt egy olyan nukleinsavat jelent, amelyet



arra használunk, hogy egy találmány szerinti polinukleotidot juttassunk vele egy gazdasejtbe. A vektorok gyakran replikonok. Az expressziós vektorok lehetővé teszik a beléjük inszertált nukleinsav transzkripcióját.

Egy találmány szerinti polinukleotid vagy polipeptid szekvenciájának viszonyát egy referencia polinukleotid vagy polipeptid szekvenciájához a következő kifejezésekkel jellemezzük: (a) „viszonyítási szekvencia”; (b) „összehasonlítási ablak”; (c) „szekvencia-azonosság”; és (d) „a szekvencia-azonosság százalékos aránya”.

(a) A „viszonyítási szekvencia” egy meghatározott szekvencia, amit a találmány szerinti polinukleotid vagy polipeptid szekvenciák összehasonlításához alapul használunk. A viszonyítási szekvencia lehet egy adott szekvencia egy része vagy teljes egésze, például egy teljes hosszúságú cDNS vagy egy gén egy szakasza, vagy a teljes cDNS- vagy génszekvencia.

(b) Az „összehasonlítási ablak” egy polinukleotid- vagy polipeptid-szekvencia egy összefüggő és meghatározott szakasza, ahol a polinukleotid- vagy polipeptid-szekvenciát egy viszonyítási szekvenciához hasonlítjuk, és ahol a polinukleotid- vagy polipeptid-szekvencia hozzáadott nukleotidokat és kihagyásokat (hézagokat) tartalmazhat a viszonyítási szekvenciához képest (ami nem tartalmaz ilyeneket) a két szekvencia optimális illeszkedése érdekében. Általában az összehasonlítási ablak legalább 20 összefüggő nukleotid vagy aminosav hosszúságú, és adott esetben 30, 40, 50, 100 vagy több nukleotid vagy aminosav hosszúságú lehet. A szakemberek tudják, hogy a hézagok beillesztése a poli-

nukleotid- vagy polipeptid-szekvenciába megtévesztően magas hasonlóságot eredményezhet a viszonyítási szekvenciával, ezért tipikusan egy hézag-büntetőpontot alkalmaznak, amit le kell vonni az egyezések számából.

A szekvenciák egyeztetése az összehasonlítás céljából jól ismert a szakterületen. Az egyeztetés optimálisan elvégezhető Smith és Waterman lokális homológiát kereső algoritmusával [Adv. Appl. Math. 2, 482 (1981)], Needleman és Wunsch homológia-egyeztető algoritmusával [J. Mol. Biol. 48, 443 (1970)], Pearson és Lipman hasonlóság-kereső módszerével [Proc. Natl. Acad. Sci. USA 85, 2444 (1988)], vagy ezen algoritmusok számítógépes alkalmazásaival, mint amilyen a CLUSTAL a PC/Gene programban (Intelligentics, Mountain View, Ca.), vagy a GAP, BESTFIT, BLAST, FASTA és TFASTA programok a Wisconsin Genetics Software Package-ban (GCG, Madison, Wi.). A CLUSTAL programot többen is jól leírták [Higgins és Sharp, Gene 73, 237-244 (1988); CABIOS 5, 151-153 (1989); Corpet et al., Nucl. Acids Res. 16, 10881-90 (1988); Huang et al., Computer Appl. Biosci. 8, 155-165 (1992) és Pearson et al., Methods in Mol. Biol. 24, 307-331 (1994)].

A BLAST programcsalád, amely adatbázisokban végzett hasonlóság-keresésre használható, az alábbi programokat tartalmazza: a BLASTN nukleotid-szekvenciát hasonlít össze az adatbázis nukleotid-szekvenciáival; a BLASTX nukleotid-szekvenciákat az adatbázis fehérje-szekvenciáival; a BLASTP fehérje-szekvenciákat az adatbázis fehérje-szekvenciáival; a TBLASTN fehérje-szekvenciákat az adatbázis nukleotid-szekvenciáival; és a TBLASTX nukleotid-szekvenciákat az adatbázis nukleotid-szekvenciáival [lásd

Ausubel, *id. mű*; Altschul *et al.*, *J. Mol. Biol.* 215, 403-410 (1990); Altschul *et al.*, *Nucl. Acids Res.* 25, 3389-3402 (1997)].

A BLAST szoftver nyilvánosan hozzáférhető például a National Center for Biotechnology Information honlapján át. Ez az algoritmus először megkeresi a magas értékszámú szekvenciapárokat (HSP-k) úgy, hogy meghatároz egy rövid, W hosszúságú szót a kereső szekvenciában, ami megfelel egy hasonló hosszúságú szónak az adatbázis szekvenciájában vagy egy bizonyos pozitív T küszöbértéket meghaladóan illeszkedik ahhoz. T neve „környező szavak találati küszöbértéke”. Ezzel az első szóval indul a keresés az olyan hosszabb HSP-k megtalálására, amelyek azt tartalmazzák. A megtalált szavakat ezután mindkét irányban kiterjeszti a szekvencia mentén mindaddig, amíg a kumulált találati érték növekszik. A kumulatív találati értéket a nukleinsav-szekvenciák esetében úgy számítja ki, hogy összesíti az M (értékpont az egyező párokért, mindig nagyobb, mint zéró) és az N (büntetőpont a nem egyező párokért, mindig kisebb, mint zéró) paramétereket. Az aminosav-szekvenciák esetében egy értékpont-mátrixot használ a kumulatív találati érték kiszámításához. A szavak mindkét irányban való kiterjesztése akkor áll le, ha a kumulatív találati érték egy adott X értékkel esik az addig elért legnagyobb értéke alá; ha zéró vagy annál kisebb értéket vesz fel; vagy ha a program eléri bármely szekvencia végét. A BLAST algoritmus W , T és X paraméterei szabják meg az illesztés érzékenységet és sebességét. A BLASTN program (a nukleinsav-szekvenciákhoz) beállított paraméterekként az alábbiakat használja: a szó hosszúsága (W) 11, az elvárás (E) 10, a leállítás 100-nál, $M=5$, $N=4$, és az össze-

hasonlítást mindkét szálon végzi. Az aminosav-szekvenciákra a BLASTP program a $W=3$ és $E=10$ értékeket, valamint a BLOSUM62 értékelő mátrixot használja [Henikoff és Henikoff, *Proc. Natl. Acad. Sci. USA* 89, 10915 (1989)].

Amellett, hogy kiszámítja a szekvencia-egyezés százalékos értékét, a BLAST algoritmus statisztikai elemzést is készít két szekvencia hasonlóságáról [Karlin és Altschul, *Proc. Natl. Acad. Sci. USA* 90, 5873-5877 (1993)]. A hasonlóság egyik mérőszáma – amit a BLAST megad – a „valószínűség legkisebb összege” [$P(N)$], ami annak valószínűségét adja meg, hogy az egyezés két nukleotid- vagy aminosav-szekvencia között véletlenül áll fenn.

A BLAST kereső algoritmus azzal a feltevéssel indul, hogy a fehérjék véletlenszerű szekvenciáknak tekinthetők. A legtöbb valódi fehérje azonban nem-random szekvenciájú területeket is tartalmaz, amelyek homopolimer szakaszok, rövid ismétlődések vagy egy vagy több aminosavban dúsabb területek lehetnek. Az ilyen, kis bonyolultsági fokú területek még akkor is összehasonlíthatók, ha a fehérjék más területei egyébként teljesen különböznek egymástól. A probléma megoldására kifejlesztett szűrő programok kiküszöbölik a kevésbé összetett területeket; ilyenek a SEG [Wooten és Federhen, *Comput. Chem.* 17, 149-163 (1993)] és az XNU [Claverie és States, *Comput. Chem.* 17, 191-201 (1993)], amelyek önállóan vagy együtt is alkalmazhatók.

Hacsak mást nem állítunk, a találmányunkban megadott, nukleotidokkal és fehérjékkel kapcsolatos azonossági és/vagy hasonlósági értékeket a GAP programmal (GCG Version 10), alapbeállítási paraméterek használatával számítottuk ki.

A GAP (Global Alignment Program) arra is használható, hogy egy találmány szerinti polinukleotidot vagy polipeptidet egy referencia-szekvenciával összehasonlítsen. A GAP Needleman és Wunsch (id. közl.) algoritmusát használja az egyezések megtalálására két teljes szekvencia között úgy, hogy maximálja az egyezések és minimalizálja a hézagok számát. A GAP figyelembe vesz minden lehetséges egyezést és hézagot, majd olyan illesztést készít, ahol legnagyobb az egyező bázisok és legkisebb a hézagok száma. Lehetővé teszi a hézagkészítés és hézag-kiterjesztés büntetőpontozását az egyező bázisok egységében mérve. A GAP-nek „profitot” kell elérnie minden egyes bevezetett hézag hézagkészítési büntetőpontjaiból. Ha a hézag-kiterjesztés büntetőpontját zérónál nagyobbra választjuk meg, akkor a GAP-nek úgy kell profitot elérnie minden bevezetett hézag után, hogy annak büntetőpontjait megszorozza a hézag-kiterjesztés büntetőpontjával. A hézagkészítés és hézag-kiterjesztés büntetőpontjainak alapértékei fehérje-szekvenciákhoz rendre 8 és 2 a Wisconsin Genetics Software Package 10. verziójában. Ugyanezek az értékek nukleinsav-szekvenciákhoz 50 és 3. A hézagkészítés és hézag-kiterjesztés büntetőpontjainak értékeit a 0 és 100 közötti egész számok közül választhatjuk ki; így például ezek egymástól függetlenül 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60 vagy nagyobb számok lehetnek.

A GAP egyik tagja az illesztőprogramok családjának, de az összes többinél jobb minőségű. A GAP négy számmal mutatja be az illesztés jóságát: ezek a „Minőség”, az „Arány”, az „Azonosság” és a „Hasonlóság”. A Minőség egy mérőszám a szekvencia egyezésé-

nek mértékére. Az Arány egyenlő a Minőség osztva a rövidebb szegmentum bázisainak számával. Az Azonosság a ténylegesen meg-
 egyező szimbólumok százalékos aránya, a Hasonlóság pedig a ha-
 sonló szimbólumoké. A hézagokkal szemben álló szimbólumokat fi-
 gyelmen kívül hagyja. A hasonlóság akkor kap értéket, ha az ér-
 tékadó mátrix értéke egy szimbólum-párra nagyobb vagy egyenlő
 0,50-dal, a hasonlósági küszöbvel. A Wisconsin Genetics Software
 Package 10. verziója a BLOSUM62 értékadó mátrixot használja (He-
 nikoff és Henikoff, *id. közl.*).

Szekvenciák többszörös összehasonlítása a CLUSTAL illesztő-
 programmal (Higgins és Sharp, *id. közl.*), alapbeállítási paramé-
 terekkel (hézag-büntetőpont = 10, hézaghossz-büntetőpont = 10)
 is elvégezhető. Páros összehasonlítás esetében a CLUSTAL alapbe-
 állítási paraméterei a következők: KTUPLE = 1, hézag-büntetőpont
 = 3, ablak = 5 és megőrzött átlók = 5.

(c) A „szekvencia-azonosság” vagy „azonosság” két nuklein-
 sav- vagy polipeptid-szekvencia vonatkozásában azt jelenti, hogy
 a két szekvenciát alkotó csoportok azonosak egy meghatározott
 összehasonlítási ablakon belül. Ha a százalékban kifejezett
 szekvencia-azonosságot fehérjékre értelmezzük, akkor fel kell
 ismerni, hogy a nem egyező aminosavak különbözőségét gyakran a
 konzervatív cserék okozzák, amikor egy aminosavat egy hasonló
 fizikai tulajdonságokkal (például töltéssel vagy hidrofobicitás-
 sal) rendelkező aminosav helyettesít, ami nem változtatja meg a
 molekula funkcionális tulajdonságait. Ha a szekvenciák konzerva-
 tív cserék miatt különböznek egymástól, akkor a szekvencia-
 azonosság százalékos értékét felfelé kell korrigálni. Az egymás-

tól a konzervatív cserek miatt különböző szekvenciák esetében „szekvencia-hasonlóság”-ról vagy „hasonlóság”-ról beszélünk. Az eszközöket, amivel ezt a korrekciót elvégezhetjük, jól ismerik a szakemberek. Tipikusan azt csináljuk, hogy a konzervatív csere egy részértéket kap a teljes különbözőséggel szemben, és ezáltal növekszik a százalékos szekvencia-azonosság értéke. Így tehát ha például egy aminosav-egyezés 1 értékpontot kap és egy nem konzervatív csere miatti különbözőség értékpontja 0, akkor egy konzervatív csere értékpontja 0 és 1 között lesz. A konzervatív szubsztitúciók értékpontjai például Meyers és Miller algoritmusával [*Computer Applic. Biol. Sci.* 4, 11-17 (1988)] számíthatók ki, ami a PC/GENE program része.

(d) A „szekvencia-azonosság százalékos arányát”-t úgy határozzuk meg, hogy összehasonlítunk két, optimálisan egyeztetett szekvenciát egy összehasonlítási ablakban, ahol a polinukleotid-szekvencia hozzáadásokat és kihagyásokat (azaz hézagokat) tartalmazhat a viszonyítási szekvenciához képest (amely nem tartalmaz hézagokat) a két szekvencia optimális illeszkedése érdekében. A százalékos arányt úgy kapjuk meg, hogy meghatározzuk az egyező nukleinsav-bázisok vagy aminosavak számát mindkét szekvenciában, ezt osztjuk az ablakban található összes összehasonlítható elem számával és az eredményt 100-zal megszorozzuk.

Alkalmazások

Találmányunk – egyéb dolgok között – készítményeket és eljárásokat biztosít a találmány szerinti polinukleotidok és polipeptidek mennyiségének módosítására (növelésére vagy csökkentésére) növényekben. Pontosabban, a találmány szerinti polinukleotidok

és polipeptidek olyan időben vagy helyen, például fejlődési szakaszokban, szövetekben, és/vagy mennyiségben fejeződhetnek ki, ami nem jellemző a rekombináns módosításoktól mentes növényekre.

Találmányunk izolált nukleinsavakat is biztosít, amelyek megfelelő hosszúságú és egy találmány szerinti polinukleotiddal komplementer polinukleotidokat tartalmaznak, amik próbaként vagy sokszorozási primerként használhatók gén-transzkriptumok kimutatására, mérésére vagy izolálására. Így például a találmány szerinti, izolált nukleinsavak próbaként használhatók a mRNS-szint csökkent voltának kimutatására a kívánt transzgen növényekre irányuló szűrővizsgálatban, génmutációk (azaz cserék, kihagyások vagy hozzáadások) kimutatására, enzimek expressziója fokozódásának vagy aktivitása megváltozásának kimutatására vegyületekre irányuló szűrővizsgálatokban, gének bármely számú allélváltozatainak (polimorfizmus), ortológjainak vagy paralógjainak kimutatására, molekuláris markerként növénynemesítési programokban, vagy eukarióta sejtek helyre irányított mutagenézisére (lásd például az US 5,565,350 számú szabadalmi iratot). A találmány szerinti, izolált nukleinsavak felhasználhatók még az általuk kódolt polipeptidek rekombináns expressziójához vagy antigénként antitestek előállításához és/vagy szűrővizsgálatához. A találmány szerinti, izolált nukleinsavak felhasználhatók még egy vagy több találmány szerinti gén szensz vagy antiszensz szuppressziójára egy gazdasejtben, szövetben vagy növényben. A transzkripció vagy transzláció befolyásolására olyan kémiai reagensek is alkalmazhatók, amelyek a találmány szerinti, izolált nukleinsavakhoz kötődve, azokkal komplexet képezve, azokat elhasítva vagy

keresztkötéseket képezve bennük a transzkripció vagy transzláció módosítására használhatók.

Találmányunk biztosít még izolált fehérjéket is, amelyek egy találmány szerinti polipeptidet (például preproenzimet, proenzimet vagy enzimet) tartalmazznak. Találmányunk olyan fehérjéket is biztosít, amelyek legalább egy epitópot tartalmazznak egy találmány szerinti polipeptidből. A találmány szerinti fehérjék az enzimfunkció agonistáinak és antagonistáinak meghatározására, illetve immunogénként vagy antigénként használhatók olyan antitestek előállításához, amelyek specifikus immunreakciót adnak egy találmány szerinti fehérjével. Az ilyen antitestek az expresszió szintjének mérésére, a találmány szerinti nukleinsavak expressziós könyvtárakban való azonosítására és/vagy izolálására, homológ fehérjék más fajokban való felismerésére vagy a találmány szerinti polipeptidek tisztítására használhatók.

A találmány szerinti, izolált nukleinsavak és polipeptidek a növények széles körében, különösen egyszikűekben, így a *Gramineae* család fajaiban, köztük a *Hordeum*, *Secale*, *Oryza*, *Triticum*, *Sorghum* (például *S. bicolor*) és *Zea* (például *Z. mays*) fajokban, valamint kétszikűekben, például a *Glycine*-ben alkalmazhatók.

A találmány szerinti, izolált nukleinsavak és fehérjék használhatók még az alábbi nemzetségek fajaiban is: *Cucurbita*, *Rosa*, *Vitis*, *Juglans*, *Fragaria*, *Lotus*, *Medicago*, *Onobrychis*, *Trifolium*, *Trigonella*, *Vigna*, *Citrus*, *Linum*, *Geranium*, *Manihot*, *Daucus*, *Arabidopsis*, *Brassica*, *Raphanus*, *Sinapis*, *Atropa*, *Capsicum*, *Datura*, *Hyoscyamus*, *Lycopersicon*, *Nicotiana*, *Solanum*, *Petunia*, *Digitalis*, *Majorana*, *Cichorium*, *Helianthus*, *Lactuca*, *Bromus*,

Asparagus, Antirrhinum, Heterocallis, Nemesis, Pelargonium, Pannicum, Pennisetum, Ranunculus, Senecio, Salpiglossis, Cucumis, Browallia, Glycine, Pisum, Phaseolus, Lolium és Avena.

Nukleinsavak

Találmányunk – egyéb dolgok között – izolált RNS- vagy DNS-nukleinsavakat, valamint azok analógjait és/vagy kiméráit biztosítja, amelyek egy találmány szerinti polinukleotidot tartalmaznak.

A találmány szerinti polinukleotidok közé tartoznak az alábbiak:

(a) egy izolált polinukleotid, ami a 2. vagy 12. számú szekvenciavázlat szerinti polipeptidet kódolja, ide értve a találmány szerinti, példaértékű polinukleotidokat;

(b) egy izolált polinukleotid, ami egy növényi nukleinsav-könyvtárban végzett sokszorozás terméke, amihez olyan primer párokat használunk, amelyek szigorú feltételek között szelektíven hibridizálódnak egy találmány szerinti polinukleotidban lévő lokuszokkal;

(c) egy izolált polinukleotid, ami szelektíven hibridizálódik egy (a) vagy (b) pont szerinti polinukleotiddal;

(d) egy izolált polinukleotid, ami meghatározott szekvencia-azonosságot mutat az (a), (b) vagy (c) pont szerinti polinukleotidokkal;

(e) egy izolált polinukleotid, ami egy olyan fehérjét kódol, amely meghatározott számú, összefüggő aminosavat tartalmaz egy prototípus polipeptidből, és amely fehérje specifikusan felismerhető egy olyan antiszérummal, aminek termelődése a fehérje adagolásával lett kiváltva, és amely fehérje nem ad megfigyelhető immunreakciót egy olyan antiszérummal, amely teljesen ki lett

merítve a fehérjével;

(f) az (a), (b), (c), (d) vagy (e) pont szerinti polinukleotidokkal komplementer szekvenciák;

(g) egy izolált polinukleotid, amely legalább meghatározott számú, összefüggő nukleotidot tartalmaz egy (a), (b), (c), (d), (e) vagy (f) pont szerinti polinukleotidból;

(h) egy izolált polinukleotid egy teljes hosszúságú, dúsított cDNS-könyvtárból, amelynek olyanok a fizikai-kémiai tulajdonságai, hogy szelektíven hibridizálódik egy (a), (b), (c), (d), (e), (f) vagy (g) pont szerinti polinukleotiddal; és

(i) egy izolált polinukleotid a következő eljárással előállítva: (1) egy teljes hosszúságú, dúsított nukleinsav-könyvtár elkészítése; és (2) a polinukleotid szelektív hibridizálása egy (a), (b), (c), (d), (e), (f), (g) vagy (h) pont szerinti polinukleotiddal, és ezáltal a polinukleotid izolálása a nukleinsav-könyvtárból.

A. Egy találmány szerinti polipeptidet kódoló polinukleotidok

Amint az (a) pontban írtuk, találmányunk olyan izolált nukleinsavakat biztosít, amelyek egy találmány szerinti polinukleotidot tartalmaznak, amely polinukleotid egy találmány szerinti polipeptidet kódol. Minden találmány szerinti nukleinsav-szekvencia, amely egy polipeptidet kódol, egyúttal leírja az adott nukleinsav minden lehetséges néma változatát is. A szakemberek tudják, hogy egy nukleinsav minden egyes kodonja (kivéve az AUG kodont, amely rendszeren csak a metionint, és az UGG kodont, amely rendszeren csak a triptofánt kódolja) megváltoztatható úgy, hogy egy funkcionálisan azonos molekulát kapjunk. Ennek megfelelően

egy nukleinsav minden néma változata, amely egy találmány szerinti polipeptidet kódol, eleve benne foglaltatik a leírt polipeptid-szekvenciában és ezáltal találmányunk oltalmi körébe tartozik. Ennek megfelelően találmányunk magában foglalja a találmány szerinti polinukleotidokat, valamint azokat is, amelyek egy találmány szerinti polipeptidet kódolnak.

B. Növényi nukleinsav-könyvtárból sokszorozott polinukleotidok

Amint a (b) pontban írtuk, találmányunk biztosít egy izolált nukleinsavat, amely egy találmány szerinti polinukleotidot tartalmaz, amely polinukleotidot egy növényi nukleinsav-könyvtárból sokszoroztunk. A nukleinsavak sokszorozásának módszereit jól ismerik a szakemberek. A növényi nukleinsav-könyvtár egy egyszikűből, például egy kultúrnövényből lehet elkészítve. A példászerű kultúrnövények közé tartozik a kukorica, cirok, lucerna, olajrepce, búza vagy rizs. A növényi nukleinsav-könyvtár elkészíthető egy kétszikűből, például a szőjából is. A *Zea mays* B73, A632, BMS, W23 és Mo17 vonalai ismertek és elérhetőek. Más ismert és hozzáférhető kukorica-vonalak a Maize Genetics Cooperation-tól (Urbana, Il.) vásárolhatók. A búzavonalak a Wheat Genetics Resource Center-től (Manhattan, Ks) szerezhetőek meg.

A nukleinsav-könyvtár lehet egy cDNS-könyvtár, egy genom-könyvtár, vagy egy olyan könyvtár, amit az intron-feldolgozás bármely szakaszában keletkező nukleáris transzkriptumokból általában állítanak össze. A cDNS-könyvtárakat normalizálni lehet, hogy növekedjen bennük a viszonylag ritka cDNS-ek aránya. Adott esetben a cDNS-könyvtár egy teljes hosszúságú cDNS-t szintetizáló módszerrel készíthető el. Az ilyen módszerek közé tartozik az

„Oligo-Capping” [Maruyama és Sugano, *Gene* 138, 171-174 (1994)], a „Biotinylated CAP Trapper” [Carninci et al., *Genomics* 37, 327-336 (1996)], és a „CAP Retention” eljárás [Edery et al., *Molec. Cell. Biol.* 15, 3363-3371 (1995)]. A cDNS-könyvtár elkészítéséhez mRNS-forrásként gyorsan növekvő szöveteket vagy gyorsan osztódó sejteket előnyös használni. A kukorica fejlődési állapotai a „How a Corn Plant Develops” című segédletben vannak leírva (Special Report No. 48, Iowa State University of Science and Technology Cooperative Extension Service, Ames, Iowa, 1993.; <http://www.ag.iastate.edu/departments/agronomy/corntitle.html>).

Egy találmány szerinti polinukleotid (vagy annak alszekvenciái) például elkészíthetők olyan sokszorozó primerek használatával, amelyek szigorú feltételek között szelektíven hibridizálódnak legalább két helyen egy találmány szerinti polinukleotidon belül, vagy két helyen egy nukleinsavon belül, amely egy találmány szerinti polinukleotidot tartalmaz és határol, vagy egy helyen egy találmány szerinti polinukleotidon belül és egy helyen az azt tartalmazó nukleinsavon belül. Egy vektor inszertum 5' és/vagy 3' végeit a szakemberek által jól ismert módszerekkel nyerjük [lásd például a RACE (Rapid Amplification of Complementary Ends) módszer leírását Innis és munkatársai „*PCR Protocols: A Guide to Methods and Applications*” című kézikönyvében (Academic Press, Inc., San Diego, 1990), az US 5,470,722 számú szabadalmi iratot, Ausubel és munkatársai ismert kézikönyvét, valamint Frohman és Martin közleményét: *Techniques* 1, 165 (1989)].

Adott esetben a primerek komplementerek a megcélzott nukleinsav egy alszekvenciájával, de a szekvencia-azonosságuk körül-

belül 85-99 %-os lehet azzal a polinukleotid-szekvenciával, amihez kapcsolódni tervezve voltak. A szakemberek megértik, hogy azok a helyek, amelyekkel a primer párok szelektíven hibridizálódnak, úgy vannak kiválasztva, hogy egyetlen, folytonos nukleinsav keletkezzen a kívánt sokszorozási körülmények között. A primer hossza nukleotidokban mérve legalább 15-től 50-ig terjed, így a primerek legalább 15, 18, 20, 25, 30, 40 vagy 50 nukleotid hosszúak lehetnek. A szakemberek tudják, hogy a célszekvenciához való kötődés specificitása fokozható a primer szekvencia hosszának növelésével. A primer 5' végéhez egy nem kötődő szekvenciát („farkat”) adva például egy klónozási helyet vihetünk be az amp-likon végéhez.

A sokszorozás termékei expressziós rendszerekkel lefordíthatók a szakemberek által jól ismert módon. A kapott transzlációs termékekről úgy bizonyíthatjuk be, hogy a találmány szerinti polipeptidek, hogy például megvizsgáljuk a katalitikus aktivitásukat (specifikus aktivitásukat és/vagy szubsztrátspecificitásukat), vagy igazoljuk bennük egy vagy több epitóp jelenlétét, amelyek jellemzőek egy találmány szerinti polipeptidre. A PCR-val kapott templatokon történő fehérjeszintézis módszerei jól ismertek, eszközei a kereskedelemben kaphatók (lásd például az Amersham Life Sciences, Inc. 1997. évi katalógusának 354. oldalát).

C. Egy (A) vagy (B) pont szerinti polinukleotiddal szelektíven hibridizálódó polinukleotidok

Amint a (c) pontban írtuk, találmányunk izolált nukleinsavakat biztosít, amelyek találmány szerinti polinukleotidokat tartalmaznak, amik – szelektív hibridizációs körülmények között –

szelektíven hibridizálódnak egy (A) vagy (B) pont alatti polinukleotiddal. Így tehát az e megvalósítás szerinti polinukleotidok felhasználhatók az (A) vagy (B) pont szerinti polinukleotidokat tartalmazó nukleinsavak izolálására, kimutatására és/vagy kvantitatív mérésére. Így például a találmány szerinti polinukleotidok részleges vagy teljes hosszúságú klónok azonosítására, izolálására vagy sokszorozására használhatók egy letéti könyvtárban. Néhány megvalósításban a polinukleotidok genomiális vagy cDNS szekvenciák, amelyek izolálva lettek vagy egyébként komplementerek egy olyan cDNS-sel, amely egy kétszikű vagy egyszikű növényi cDNS-könyvtárából származik. Az ilyen növényekre nem korlátozó példaként említjük a következőket: kukorica, olajrepcse, szója, gyapot, búza, cirok, napraforgó, lucerna, zab, cukornád, köles, árpa és rizs. A cDNS-könyvtár legalább 50-95 % teljes hosszúságú szekvenciát (például legalább 50, 60, 70, 80, 90 vagy 95 % teljes hosszúságú szekvenciát) tartalmaz. A cDNS-könyvtár normalizálva lehet, hogy növekedjen benne a ritka szekvenciák aránya (lásd például az US 5,482,845 számú szabadalmi iratot). A komplementer szekvenciákkal kevésbé megegyező szekvenciák esetében tipikusan, de nem kizárólag, a kevésbé szigorú hibridizációs körülményeket, míg a nagyobb azonossági fokú szekvenciákhoz a közepesen vagy erősen szigorú körülményeket alkalmazzuk. A kevésbé szigorú körülmények megengedik a körülbelül 70-80 %-os azonossági fokú szekvenciák szelektív hibridizálódását, és ezzel az ortológ vagy paralóg szekvenciák azonosítását.

D. Az (A), (B) vagy (C) pont szerinti polinukleotidokkal meghatározott szekvencia-azonosságot mutató polinukleotidok

Amint a (d) pontban írtuk, találmányunk olyan izolált nukleinsavakat biztosít, amelyek találmány szerinti polinukleotidokat tartalmaznak, amik nukleotid-szinten meghatározott azonosságot mutatnak egy (A), (B) vagy (C) pont alatti polinukleotiddal. Az azonosság fokát például a BLAST, CLUSTALW vagy GAP algoritmusokkal, beállított paraméterek használatával számíthatjuk ki. Az azonosság foka egy referencia-szekvenciához képest legalább 50 %, és bármely egész számú értéket felvehet 50 és 99 között, a határértékeket is ide értve. Így például az azonosság foka a referencia-szekvenciához képest legalább 60, 70, 75, 80, 85, 90 vagy 95 % lehet.

Adott esetben az e megvalósítás szerinti polinukleotidok egy olyan polipeptidet kódolnak, amelynek egy közös epitópja van az (A), (B) vagy (C) pont szerinti polinukleotidok által kódolt polipeptidekkel. Így tehát ezek a polinukleotidok egy első polipeptidet kódolnak, amely egy olyan antiszérum termelődését váltja ki, ami az (A), (B) vagy (C) polinukleotidok egyike által kódolt második polipeptiddel specifikusan reagáló antitesteket tartalmaz. Azonban az első polipeptid nem fog reagálni az önmaga ellen termelődött antiszérummal, ha teljesen ki lett merítve az első polipeptiddel. Így tehát az e megvalósítás szerinti polinukleotidok felhasználhatók olyan antitestek előállítására, amelyek például az (A), (B) vagy (C) pont szerinti polinukleotidokat tartalmazó, expressziós nukleinsav-könyvtárak szűrővizsgálatára vagy tisztítására, vagy az (A), (B) vagy (C) pont szerinti polinukleotidok által kódolt polipeptidek immunológiai meghatározására alkalmazhatók. E megvalósítás polinukleotidjai maguk-

ban foglalják azokat a nukleinsav-szekvenciákat is, amelyek egy találmány szerinti polipeptidet kódoló polinukleotiddal való szelektív hibridizálódásra használhatók.

A polipeptidek szűrővizsgálata egy antitesthez való specifikus kötődés által kényelmesen megvalósítható a peptid-könyvtárakban. Ez a módszer lehetővé teszi, hogy nagy peptid-gyűjteményekben is megtaláljuk a kívánt funkciójú vagy szerkezetű egyedi peptideket. A peptid-könyvtárak antitestekkel végzett szűrővizsgálata jól ismert a szakterületen. A vizsgált peptid-szekvenciák 3-5000 vagy több aminosav hosszúságúak, gyakran 5-100, és gyakran körülbelül 8-15 aminosav hosszúságúak lehetnek. A peptid-könyvtárak közvetlen, kémiai szintézissel való előállításának módszerei mellett számos, rekombináns DNS-t alkalmazó módszer ismeretes. Ezek egyik típusa egy bakteriofág vagy egy sejt felületén jeleníti meg a peptid-szekvenciát; az adott peptidet kódoló nukleotid-szekvenciát a fágok vagy sejtek tartalmazzák. Ilyen módszerek vannak leírva a PCT 91/17271, 91/18980, 91/19818 és 93/08278 számú szabadalmi iratokban. A peptid-könyvtárak előállításának más rendszerei magukban foglalják mind az *in vitro* kémiai, mind a rekombináns módszereket (PCT 92/05258, 92/14843, 97/20078; US 5,658,758, 5,643,768). Peptid-könyvtárak, vektorok és szűrővizsgálat-készletek a kereskedelemben kaphatók olyan szállítóktól, mint például az Invitrogen (Carlsbad, Ca.).

E. Egy prototípus polipeptid egy alszekvenciáját tartalmazó és a prototípus polipeptiddel keresztreakciót adó fehérjét kódoló polinukleotidok

Amint az (e) pontban írtuk, találmányunk izolált nukleinsa-

vakat biztosít, amelyek a találmány szerinti polinukleotidokat tartalmazzák, amik egy olyan fehérjét kódolnak, amelynek egy összefüggő aminosavakból álló alszekvenciája van egy találmány szerinti, például az (a) pont szerinti prototípus polipeptidből. A prototípus polipeptidből származó, összefüggő aminosavak hossza legalább 10-től a prototípus szekvenciában lévő aminosavak számáig terjedhet. Így tehát például a polinukleotid egy olyan polipeptidet kódolhat, amelynek egy alszekvenciája legalább 10, 15, 20, 25, 30, 35, 40, 45 vagy 50 összefüggő aminosavat tartalmaz a prototípus polipeptidből. Továbbá, az egy polinukleotid által kódolt ilyen alszekvenciák száma 1-20, így 2, 3, 4 vagy 5 lehet. Az alszekvenciákat akárhány, így 1-től a szekvenciában lévő nukleotidok számáig terjedő számú, például legalább 5, 10, 15, 25, 50, 100 vagy 200 nukleotid választhatja el egymástól.

A jelen megvalósítás polinukleotidjai által kódolt fehérjék – ha antigénként alkalmazzák azokat – poliklonális antitestek keletkezését váltják ki, amelyek specifikusan kötődnek egy prototípus polipeptidhez, például egy olyanhoz (de nem kizárólag), amelyet egy (a) vagy egy (b) pont szerinti polinukleotid kódol. Általában azonban egy e megvalósítás szerinti polinukleotid által kódolt fehérje nem fog reagálni a prototípus polipeptid ellen készített antiszérummal, ha az antiszérum teljesen ki lett merítve a prototípus polipeptiddel. Az antitestek előállítása és vizsgálata jól ismert a szakterületen. Példaszerű immunvizsgálati módszerek az ELISA, a kizorításos (kompetíciós) vizsgálat, a radioimmun-vizsgálat, a Western-blot technika, a közvetett immunfluoreszcenciás mérés és a hasonlóak.

Egy előnyös vizsgálati eljárásban teljesen kimerített anti-szérumot – amelyet a prototípus polipeptid ellen készítettünk – használunk egy kizsorításos kötődési mérésben a fehérje vizsgálatára. A mérésben a polipeptid azon koncentrációját határozzuk meg, amely 50 %-ban gátolja a prototípus polipeptid kötődését az antiszérumhoz. Ha a kötődés gátlásához szükséges fehérje mennyisége kevesebb, mint kétszerese a prototípus fehérje mennyiségének, akkor a megállapítjuk, hogy a fehérje specifikusan kötődik az antigénnel szemben termelt antiszérumhoz. Ennek megfelelően a találmány szerinti fehérjék felölelik a prototípus polipeptid allelikus változatait, konzervatíván módosított változatait és minor rekombináns módosulatait is.

Egy találmány szerinti polinukleotid adott esetben egy olyan fehérjét kódol, amelynek molekulatömege – mint glikozilezetlen fehérjéé – 20 %-on belül azonos a találmány szerinti, teljes hosszúságú, glikozilezetlen polipeptidével. A molekulatömeg könnyen meghatározható SDS-PAGE módszerrel, redukáló körülmények között. Előnyösen a molekulatömeg 15 %-on belül azonos egy teljes hosszúságú találmány szerinti polipeptidével; előnyösebben az azonosság 10 vagy 5 %-on belül és legelőnyösebben 3, 2 vagy 1 %-on belül van egy találmány szerinti, teljes hosszúságú polipeptidhez képest.

Adott esetben az e megvalósítás szerinti polinukleotidok egy olyan fehérjét kódolnak, amelynek specifikus enzimaktivitása legalább 50, 60, 80 vagy 90 %-a egy sejtkivonaténak, ami a találmány szerinti, natív, endogén, teljes hosszúságú polipeptidet tartalmazza. Továbbá, az e megvalósítás szerinti polinukleotidok

által kódolt fehérjék adott esetben lényegében hasonló affinitási konstanssal (K_m) és/vagy katalitikus aktivitással (azaz mikroszkópos sebességi állandóval, k_{cat}) rendelkeznek, mint a természetes, endogén, teljes hosszúságú fehérje. A szakemberek fel fogják ismerni, hogy a k_{cat}/K_m hányados határozza meg az enzim specificitását a vetélkedő szubsztrátokkal szemben, és ezért gyakran specificitási konstansként is említik. Az e megvalósítás fehérjéinek k_{cat}/K_m értéke legalább 10 %-a egy találmány szerinti, teljes hosszúságú polipeptidének az adott polipeptid endogén szubsztrátjával mérve. Adott esetben a k_{cat}/K_m értéke legalább 20, 30, 40 vagy 50 %-a, és legelőnyösebben legalább 60, 70, 80, 90 vagy 95 %-a a találmány szerinti, teljes hosszúságú polipeptidének. A k_{cat} , K_m és k_{cat}/K_m értékeket a szakemberek által ismert számos módszer bármelyikével megmérhetjük. Így például a kezdeti sebesség (azaz a reakció első 5 %-a vagy kisebb hányada) gyors összekeveréssel és mintavétellel (például folytonos áramlású, szaggatott áramlású vagy gyors leállító technikákkal), villanófényes fotolízissel vagy relaxációs módszerekkel (például hőmérséklet-ugrással) határozható meg olyan mérési eljárásokkal összekapcsolva, mint a spektrofotometria, spektrofluorimetria, magmágneses rezonancia-mérés vagy radioizotópos eljárások. A kinetikai paramétereket szokásosan a Lineweaver-Burk vagy az Eadie-Hofstee függvényel számítjuk ki.

F. Az (A)–(E) pont szerinti polinukleotidokkal komplementer polinukleotidok

Amint az (f) pontban írtuk, találmányunk olyan izolált nukleinsavakat biztosít, amelyek az (A)–(E) pontok szerinti po-

linukleotidokkal komplementer polinukleotidokat tartalmaznak. A szakemberek tudják, hogy a komplementer szekvenciák bázispárjai az (A)-(E) pontok szerinti polinukleotidok teljes hosszában megfelelnek egymásnak (azaz szekvencia-azonosságuk 100 %-os a teljes hosszuk mentén). A komplementer bázisokat hidrogénhidak kötik egymáshoz a kettős szálú nukleinsavakban. Az alábbi bázispárok komplementerek: guanin és citozin, adenin és timin, adenin és uracil.

G. Polinukleotidok, amelyek az (A)-(F) pont szerinti polinukleotidok alszekvenciái

Amint a (g) pontban írtuk, találmányunk izolált nukleinsavakat biztosít, amelyek olyan polinukleotidokat tartalmaznak, amik legalább 15 összefüggő bázist tartalmaznak az (A)-(F) pontok szerinti polinukleotidokból. A polinukleotid hossza bármely egész számmal megadható a legalább 15-től egészen annak a polinukleotidnak a teljes hosszáig, amelynek az említett polinukleotid egy alszekvenciáját képezi. Így például a találmány szerinti polinukleotidok közé tartoznak azok a polinukleotidok, amelyek legalább 15, 20, 25, 30, 40, 50, 60, 75 vagy 100 összefüggő nukleotidot tartalmaznak az (A)-(F) pont szerinti polinukleotidokból. Adott esetben egy jelen megvalósítás szerinti polinukleotid által kódolt alszekvenciák száma bármely egész szám lehet 1 és 20 között, például 2, 3, 4 vagy 5. Az alszekvenciák 1-től a szekvenciában lévő nukleotidok számának megfelelő számú, legalább 5, 10, 15, 25, 50, 100 vagy 200 nukleotiddal lehetnek elválasztva egymástól.

Az alszekvenciák elkészíthetők *in vitro* szintézissel, *in*

vitro bioszintézissel vagy *in vivo* rekombináns módszerekkel. Egyes megvalósításokban az alszekvenciák nukleinsav-sokszorozással készíthetők el. Így például olyan nukleinsav primerek készíthetők, amelyek szelektíven hibridizálódnak egy szekvenciával (vagy annak komplementerével) a kódoló régió belül, vagy azzal azonos kiterjedésűek.

A találmány szerinti alszekvenciák annak a szekvenciának a szerkezeti jellemzőivel bírhatnak, amelyből származnak. Másrészt az alszekvenciákból hiányozhatnak azon nagyobb szekvenciák bizonyos szerkezeti jellemzői, amelyekből származnak, például a poli(A) farok. Adott esetben egy olyan polinukleotid alszekvenciája, amely egy olyan polipeptidet kódol, aminek egy (a) pont szerinti prototípus szekvenciával legalább egy közös epitópja van, kódolhat egy olyan epitópot, amely közös a prototípus szekvenciával. Másrészt az alszekvenciának nem kell egy, a prototípus szekvenciával közös epitópot kódolnia, de használható a nagyobb szekvencia izolálására például nukleinsav-hibridizáció útján azzal a szekvenciával, amiből származik. Az alszekvenciák a génki-fejeződés módosítására vagy kimutatására is használhatók azért, hogy az alszekvenciákba olyan vegyületeket építsünk be, amelyek kapcsolódnak a nukleinsavakhoz, komplexet képeznek azokkal, elhasítják azokat vagy keresztkötéseket létesítenek bennük. Az ilyen vegyületekre példák az akridin, pszoralén, fenantrolin, naftokinon, daunomicin vagy a klór-etil-amino-aril konjugátumok.

H. Polinukleotidok egy teljes hosszúságú, dúsított cDNS-könyvtárból, amelyek fizikai-kémiai tulajdonságaik alapján szelektíven hibridizálódnak egy (A) – (G) pont szerinti polinukleotiddal

Amint a (h) pontban írtuk, találmányunk biztosít egy izolált polinukleotidot egy teljes hosszúságú, dúsított cDNS- könyvtárból, amely fizikai-kémiai tulajdonságai alapján szelektíven hibridizálódik egy (A), (B), (C), (D), (E), (F) vagy (G) pont szerinti polinukleotiddal. A teljes hosszúságú, dúsított cDNS-könyvtárak megszerkesztésének módját az alábbiakban röviden ismertetni fogjuk. A cDNS-könyvtár legalább 50-95 %, például 50, 60, 70, 80, 90 vagy 95 % teljes hosszúságú szekvenciákat tartalmaz. A cDNS-könyvtár egy- vagy kétszikű növények különböző fejlődési stádiumú szöveteiből készíthető el. Példaként említjük az olyan növényfajokat, mint a kukorica, búza, rizs, olajrepce, szója, gyapot, cirok, napraforgó, lucerna, zab, cukornád, köles, árpa vagy rizs. Egy találmány szerinti polinukleotid és egy teljes hosszúságú, dúsított könyvtárból származó polinukleotid szelektív körülmények közötti hibridizálásának módszereit jól ismerik a szakemberek. A szelektív hibridizálódás eléréséhez bármilyen szigorúsági fokú körülményeket alkalmazhatunk. Egyes megvalósításokban a szigorúság alkalmazott foka a hibridizálódó területen belül legalább 70, 75, 80, 85, 90, 95 vagy 98 %-os szekvenciaazonosságot mutató szekvenciák szelektív hibridizálódását teszi lehetővé. A teljes hosszúságú, dúsított cDNS-könyvtárak normalizálhatók, hogy növekedjen bennük a ritka szekvenciák aránya.

I. cDNS-izoláló eljárással készített polinukleotid termékek

Amint az (i) pontban írtuk, találmányunk biztosít egy izolált polinukleotidot, amit a következő módon állítunk elő: (1) elkészítünk egy teljes hosszúságú, dúsított nukleinsav-könyvtárat, majd (2) szelektíven hibridizáljuk a polinukleotidot egy

(A), (B), (C), (D), (E), (F), (G) vagy (H) pont szerinti polinukleotiddal, és ezzel izoláljuk a polinukleotidot a nukleinsav-könyvtárból. A teljes hosszúságú nukleinsav-könyvtárat a (G) pontban és az alább leírtak szerint készítjük el. A szelektív hibridizálás körülményeit a (G) pontban ismertettük. A nukleinsavak tisztításának módszerei jól ismertek a szakterületen. A tisztítás kényelmesen megoldható szilárd fázisú módszerekkel; az ilyen módszerek jól ismertek és a hozzájuk való készletek kereskedelmi szállítóktól, például az Advanced Biotechnologies-től (Surrey, UK) beszerezhetők. Így például az (A) – (H) pont szerinti polinukleotidok egy szilárd hordozón, például membránon, gyöngyökön vagy szemcséken immobilizálhatók (US 5,667,976 számú szabadalmi irat). A találmány szerinti eljárással készített polinukleotid termék szelektíven hibridizálódik egy immobilizált polinukleotiddal, majd a szilárd hordozót izoláljuk a nem hibridizálódó polinukleotidok közül olyan eljárásokkal, mint a centrifugálás, mágneses elválasztás, szűrés, elektroforézis és hasonló.

Nukleinsavak szerkesztése

A találmány szerinti, izolált nukleinsavak (a) standard rekombinációs módszerekkel, (b) szintetikus úton, vagy a kettő kombinálásával állíthatók elő. Néhány megvalósításban a találmány szerinti polinukleotidok egy egyszikű növényből, például kukoricából, rizsből vagy búzából, vagy egy kétszikű növényből, például szójaból klónozhatók, sokszorozhatók vagy más módon szerkeszthetők meg.

A nukleinsavak alkalmas módon más szekvenciákat is tartalmazhatnak a találmány szerinti polinukleotid mellett. Így példá-

ul egy multiklónozási hely – amely egy vagy több endonukleáz restriktációs helyét tartalmazza – lehet beillesztve a nukleinsavba, hogy segítse a polinukleotid izolálását. Ugyanígy transzlátálható szekvenciák is lehetnek beillesztve, hogy segítsék a lefordított, találmány szerinti polinukleotid izolálását. Így például egy hexahisztidin jelző szekvencia kényelmes eszközt biztosít a találmány szerinti fehérjék tisztításához. Egy találmány szerinti polinukleotid egy vektorhoz, adapterhez vagy linkerhez kapcsolható a klónozása és/vagy kifejeztetése céljából. Az ilyen klónozó és/vagy expressziós szekvenciákhoz további szekvenciák is adhatók, hogy optimalizáljuk a működésüket a klónozásban és/vagy expresszióban, hogy elősegítsék a polinukleotid izolálását, vagy fokozzák a polinukleotid bejutását egy sejtbe. Tipikusan egy találmány szerinti nukleinsav hossza kisebb, mint a találmány szerinti polinukleotid hossza, ami kisebb, mint 20 kbp, gyakran kisebb, mint 15 kbp, és gyakrabban kisebb, mint 10 kbp. A klónozó vektorok, expressziós vektorok, adapterek és linkerek használata jól ismert és részletesen le van írva a szakterületen. A különböző nukleinsavak leírását lásd a Stratagene Cloning Systems 1999. évi katalógusában (La Jolla, Ca.), vagy az Amersham Life Sciences, Inc. (Arlington Heights, Il.) 1999. évi katalógusában.

A. Rekombináns eljárások nukleinsavak szerkesztésére

A találmány szerinti, izolált nukleinsav-készítmények – például RNS, cDNS, genom-DNS vagy ezek hibridjei – növényekből, mint biológiai forrásokból nyerhetők a szakemberek által ismert nagyszámú klónozási eljárás bármelyikével. Néhány megvalósításban oligonukleotid próbákat – amelyek szigorú feltételek mellett

szelektíven hibridizálódnak a találmány szerinti polinukleotidokkal – használunk a kívánt szekvencia azonosítására egy cDNS vagy genom-DNS könyvtárban. Az RNS izolálása, valamint a cDNS és genom-DNS könyvtárak létrehozása jól ismert a szakterületen (lásd például Clark, illetve Ausubel már idézett kézikönyveit).

A1. Teljes hosszúságú, dúsított cDNS-könyvtárak

Számos cDNS-t szintetizáló eljárást írtak le, amelyekkel dúsított, teljes hosszúságú cDNS-könyvtárak hozhatók létre. A dúsított, teljes hosszúságú cDNS-könyvtárak legalább 60 %, előnyösebben legalább 70, 80, 90 vagy 95 % teljes hosszúságú inszertumot tartalmaznak az inszertumokat tartalmazó klónok között. Az inszertumok hossza az ilyen könyvtárakban legalább 2, 3, 4, 5, 6, 7, 8, 9, 10 vagy több kbp lehet. Az ilyen méretű inszertumokat befogadni képes vektorok jól ismertek a szakterületen, és a kereskedelembe kaphatók; ilyen például a Stratagene lambdaZAP Express vektora, amely egy 0-12 kbp kapacitású cDNS-klónozó vektor. A 95 %-nál jobb tisztaságú, teljes hosszúságú cDNS-könyvtárak létrehozásának példaszerű eljárása van írva Carninci és munkatársai közleményében [*Genomics* 37, 327-336 (1996)]. A teljes hosszúságú könyvtárak létrehozásának más módszerei is ismertek a szakterületen [Edery et al., *Mol. Cell Biol.* 15, 3363-3371 (1995); WO 96/34981].

A2. Normalizált vagy szubsztrahált cDNS-könyvtárak

Egy nem normalizált cDNS-könyvtár annak a szövetnek a mRNS populációját reprezentálja, amelyikből készült. Mivel a ritkább klónokat számban felülmúlják az erősen kifejeződő gének klónjai, ezért izolálásuk nehézkes. Egy cDNS-könyvtár normalizálása az az

eljárás, amellyel az egyes klónok reprezentációs aránya egyöntetűbbé tehető. A normalizált könyvtárak készítésének módja megtalálható a szakirodalomban [Ko, *Nucl. Acids Res.* 18, 5705-5711 (1990), Patanjali et al., *Proc. Natl. Acad. Sci. USA* 88, 1943-1947 (1991), US 5,482,685, 5,482,845 és 5,637,685 számú szabadalmi iratok]. Példaszerű eljárást írnak le Soares és munkatársai, amelynek eredményeként a klónok gyakorisága közötti különbség négy nagyságrendről egy nagyságrenden belültre szűkül le [*Proc. Natl. Acad. Sci. USA* 91, 9228-9232 (1994)].

A „kivonatolt” (szubsztrahált) cDNS-könyvtár készítése egy másik módja a ritkább cDNS fajták feldúsításának. Ebben az eljárásban a mRNS egyik adagján elkészített cDNS-t arra használjuk, hogy egy másik adagban – hibridizálás útján – lecsökkentsük vele a gyakoribb mRNS-szekvenciákat. A cDNS:mRNS hibrideket eltávolítva a visszamaradó, nem hibridizálódott cDNS feldúsul a ritka szekvenciákban [Clark, id. mű; Kho és Zarbl, *Technique* 3, 58-63 (1991); Sive és St. John, *Nucl. Acids Res.* 16, 10937 (1988); Ausubel, id. mű; Swaroop et al., *Nucl. Acids Res.* 19, 1954 (1991)]. A cDNS szubsztrakciós készletek a kereskedelembe kaphatók (például a PCR-Select a Clontech-től).

A genom-könyvtárak létrehozásához nagy szegmentumokat készítünk genom-DNS-ből például restrikciós endonukleázok felhasználásával, és vektor-DNS-sel kapcsoljuk össze azokat, majd az így kapott konkatamereket a megfelelő vektorba csomagoljuk. Az ehhez szükséges végek kialakításának és a nukleinsav szekvenciájának bizonyításához szükséges szekvenálásnak a módszerei az ismert kézikönyvekben megtalálhatók. A genom-könyvtárak létrehozásához

használható készletek a kereskedelemben kaphatók.

A cDNS- vagy genom-könyvtárak szűrővizsgálata olyan próbákkal (szondákkal) végezhető el, amelyek egy találmány szerinti polinukleotid szekvenciáján alapulnak. A próbák arra használhatók, hogy genom-DNS vagy cDNS szekvenciákkal hibridizálódva a homológ gének izolálhatók legyenek ugyanazon vagy különböző növényfajokból. A szakemberek tudják, hogy a szűrővizsgálatban a hibridizálás szigorúságának különböző fokozatai alkalmazhatók, illetve mind a hibridizálás, mind a mosás körülményei lehetnek szigorúak.

Az érdeklődésre számot tartó nukleinsavak nukleinsav-mintákból is sokszorozhatók különböző sokszorozó technikák alkalmazásával. Így például a polimeráz láncreakció (PCR) technika használható a találmány szerinti polinukleotid-szekvenciák és hasonló gének közvetlenül a genom- vagy cDNS-könyvtárakból történő sokszorozására. A PCR és más *in vitro* sokszorozó technikák arra is alkalmasak például, hogy kifejezendő fehérjét kódoló nukleinsav-szekvenciákat klónozzunk; hogy különböző mintákban a kívánt mRNS-ek jelenlétét kimutató próbáknak alkalmas nukleinsavakat készítsünk; hogy szekvenálásra vagy más célokra alkalmas nukleinsavakat állítsunk elő. A T4Gene 32 Protein (Boehringer, Mannheim) a hosszú PCR-termékek hozamának javítására alkalmas.

A PCR-alapú szűrővizsgálatok is megtalálhatók a szakirodalomban. Wilfinger és munkatársai leírtak egy eljárást, amellyel a leghosszabb cDNS egy lépésben azonosítható úgy, hogy az inkomplett klónokat kizárják a vizsgálatból [BioTechniques 22, 481-486 (1997)]. Az ilyen eljárások különösen hatásosak a fent leírt, a teljes hosszúságú cDNS-ek szerkesztésére szolgáló mód-

szerekkel kombinálva.

B. Szintetikus eljárások nukleinsavak szerkesztésére

A találmány szerinti izolált nukleinsavak közvetlen kémiai szintézissel is előállíthatók például a foszfo-triészter módszerrel [Narang et al., *Meth. Enzymol.* 68, 90-99 (1979)], a foszfo-diészter módszerrel [Brown et al., *Meth. Enzymol.* 68, 109-151 (1979)], a dietil-foszforamidit módszerrel vagy a szilárd fázisú foszforamidit-triészter módszerrel [Beaucage és Caruthers, *Tetra. Letts.* 22, 1859-1862 (1981)], például automatizált szintetizáló berendezések használatával [Needham-VanDevanter et al., *Nucl. Acids Res.* 12, 6159-6168 (1984)], vagy a szilárd hordozó módszerrel (US 4,458,066). A kémiai szintézissel általában egyszálú oligonukleotidok állíthatók elő. Ezek kettős szálú DNS-sé alakíthatók át vagy egy komplementer szekvenciával való hibridizálással, vagy egy DNS-polimerázzal végzett polimerizálással, amihez az egyes szálakat használjuk templátként. A szakemberek tudják, hogy a DNS kémiai szintézise legjobban a körülbelül 100 bázis hosszúságú szekvenciák előállítására alkalmas, és a hosszabb szekvenciák a rövid szekvenciák összekapcsolásával kaphatók meg.

Rekombináns expressziós keretek

Találmányunk rekombináns expressziós kereteket is biztosít, amelyek egy találmány szerinti nukleinsavat tartalmaznak. Egy nukleinsav-szekvencia, amely a kívánt, találmány szerinti polipeptidet kódolja – például egy cDNS, vagy egy genomiális szekvencia, amely egy találmány szerinti, teljes hosszúságú polipeptidet kódol – használható fel egy rekombináns expressziós keret megszerkesztéséhez, ami a kívánt gazdasejtbe juttatható. Egy re-

kombináns expressziós keret tipikusan tartalmaz egy találmány szerinti polinukleotidot működőképes módon összekapcsolva a transzkripció megindulását szabályozó szekvenciákkal, amelyek a polinukleotid transzkripcióját fogják vezérelni a megcélzott gazdasejtben, például egy transzformált növény szöveteiben.

Így például a növényi expressziós vektorok magukban foglalhatnak (1) egy klónozott növényi gént az 5' és 3' regulátor-szekvenciák transzkripciót vezérlő hatása alatt, és (2) egy domináns, szelektálható jelző-szekvenciát. Az ilyen növényi expressziós vektorok kívánt esetben tartalmazhatnak még egy promoter regulátor-régiót (ami lehetővé teszi az indukálható vagy konstitutív, környezet- vagy fejlődés-függő, sejt- vagy szövetspecifikus, szelektív expressziót), egy transzkripciós kezdőhelyet, egy riboszóma-kötőhelyet, egy RNS-feldolgozási jelzést, egy transzkripciós befejező helyet, és/vagy egy poliadenilációs jelzést.

Alkalmazhatunk egy olyan növényi promoter-fragmentumot, ami egy regenerált növény minden szövetében vezérelni fogja egy találmány szerinti polinukleotid expresszióját. Az ilyen promotorokat „konstitutív” promotoroknak nevezzük, mivel a legtöbb környezeti körülmény, fejlődési stádium vagy sejt differenciáltsági állapot mellett aktívak. A konstitutív promotorokra példaként említjük a karfiolmozaik vírus (CaMV) 35S transzkripciós indító területét, az *Agrobacterium tumefaciens* T-DNS-éből származó 1'-vagy 2'-promotert, az ubikvitin-1 promotert, a Smas promotert, a fahéj-alkohol dehidrogenáz promotert (US 5,683,439), a Nos, a pEmu, a rubisco és a GRP1-8 promotert.

Másrészt a növényi promoter vezérelheti egy találmány szerinti polinukleotid kifejeződését egy bizonyos szövetben vagy más, pontosabban meghatározott környezeti vagy fejlődési feltételek esetén. Az ilyen promoterokat „indukálható” promoteroknak nevezzük. A környezeti feltételek, amelyek az indukálható promoterokon keresztül befolyásolhatják a transzkripciót, például a következők lehetnek: egy kórokozó támadása, anaerob körülmények vagy fény jelenléte. Az indukálható promoterokra példák az Adh1 promoter, amely hipoxiával vagy hidegstresszel, a Hsp70 promoter, amely hőstresszel, és a PPKK promoter, amely fénnel indukálható.

A fejlődési feltételekre érzékeny promoterok kizárólag vagy elsősorban bizonyos szövetekben, például levelekben, gyökerekben, gyümölcsökben, magvakban vagy virágokban indítják meg a transzkripciót. A példaszerű promoterok közé tartozik az 5126-portok-specifikus promoter (US 5,689,049 és 5,689,051), a glob-1 és a γ -zein promoter. A promoterok működése a genomban való elhelyezkedésüktől függően is változhat; így egy indukálható promoter bizonyos helyeken teljesen vagy részben konstitutív lehet.

A találmány szerinti nukleinsavak közvetlen kifejeztetésére mind heterológ, mind nem heterológ (azaz endogén) promoterok alkalmazhatók. Ezek a promoterok arra is használhatók, hogy például rekombináns expressziós keretekben antiszensz nukleinsavak expresszióját működtessék annak érdekében, hogy növekedjen, csökkenjen vagy megváltozzon a fehérjék koncentrációja és/vagy összetétele a kívánt szövetben. Így tehát néhány megvalósításban a nukleinsav-szerkezet egy növényi sejtekben működni képes promotert tartalmaz, amely működőképes módon van egy találmány szerinti

polinukleotidhoz kapcsolva. Az ilyen megvalósításokban használható promoterok közé tartoznak azok az endogén promoterok, amelyek egy találmány szerinti polipeptid termelődését működtetik.

Néhány megvalósításban promoterként vagy serkentő elemekként szolgáló, izolált nukleinsavak építhetők be egy találmány szerinti polinukleotid nem heterológ formájának megfelelő helyére (általában a szekvenciától felfelé), amivel a polinukleotid expressziója serkenthető vagy gátolható. Így például endogén promoterok megváltoztathatók *in vivo* mutációval, delécióval és/vagy szubsztitúcióval (US 5,565,350 és WO 93/22443 számú szabadalmi iratok), vagy izolált promoterok építhetők be egy növényi sejtbe, egy azonos eredetű génhez képest megfelelő orientációval és távolságra, hogy a gén kifejeződését vezérelni tudja. A gén expressziója a növény fejlődéséhez megfelelő körülmények között módosítható, hogy megváltoztassuk a találmány szerinti polipeptid mennyiségét és/vagy összetételét a növény sejtjeiben. Így tehát találmányunk készítményeket és eljárásokat biztosít idegen promoterok és/vagy serkentő elemek készítésére, amelyek működőképes módon vannak egy találmány szerinti polinukleotid natív, endogén (azaz nem heterológ) formájához kapcsolva.

Ha polipeptid expresszióját kívánjuk, akkor általában tanácsos egy poliadenilezett szakaszt csatolni egy polinukleotid kódoló szakaszának 3' végéhez. A poliadenilezett szakasz származhat a természetes génből, számos más növényi génből vagy T-DNS-ből. A hozzáadandó 3' végszekvencia származhat például a nopalin- vagy az oktopin-szintetáz génekből, vagy más növényi génekből, vagy – kevésbé előnyösen – bármely más eukarióta génből.

A kódoló szekvencia 5' át nem íródó szakaszához vagy kódoló szakaszához egy intron-szekvenciát adhatunk, hogy fokozzuk az érett mRNS mennyiségét, ami a citoszolban halmozódik fel. Egy kivágható intron beillesztése a transzkripció egységbe mind a növények, mind az állatok esetében több mint ezerszeresére fokozza a gén expresszióját mind a mRNS, mind a fehérje szintjén [Buchman és Berg, *Mol. Cell Biol.* 8, 3495-4405 (1988); Callis et al., *Genes Dev.* 1, 1183-1200 (1987)]. Az intron génkifejeződést fokozó hatása tipikusan nagyobb, ha a transzkripció egység 5' végéhez közelebb helyezkedik el. Az Adh1-S 1, 2 és 6, és Bronze-1 kukorica-intronok használata jól ismert a szakterületen („*The Maize Handbook*”, Springer, New York, 1994). A vektor, ami egy találmány szerinti polinukleotidból származó szekvenciát tartalmaz, tipikusan tartalmaz egy markergént is, amely egy szelektálható fenotípusos bélyeget kölcsönöz a növényi sejteknek. A tipikus vektorok, amelyek alkalmasak gének kifejeztetésére magasabbrendű növényekben, jól ismertek a szakterületen; ide tartoznak az *Agrobacterium tumefaciens* tumorindukáló (Ti) plazmidjából származó vektorok is [Rogers et al., *Meth. Enzymol.* 153, 253-277 (1987)].

Egy találmány szerinti polinukleotid kívánság szerint szensz vagy antiszensz irányban fejeztethető ki. Belátható, hogy a gén expressziójának irányultsága közvetlen hatással van a növény megfigyelhető tulajdonságaira. Az antiszensz-technikát kényelmesen használhatjuk gének kifejeztetésére növényekben. Ennek megvalósításához a kívánt génből származó nukleinsav-szegmentumot működőképes módon egy promoterhez kapcsoljuk úgy, hogy az RNS antiszensz szála íródjon át. A szerkezettel ezután egy növényt

transzformálunk, ahol az RNS antiszensz szála termelődik. Kimutatták, hogy a növényi sejtekben az antiszensz RNS gátolja a génkifejeződést azáltal, hogy megakadályozza a szóban forgó enzimet kódoló mRNS felhalmozódását [Sheehy et al., *Proc. Natl. Acad. Sci. USA* 85, 8805-8809 (1998); US 5,759,829].

A génelnyomás egy másik módja a szensz-szupresszió (más néven koszipresszió). Kimutatták, hogy szensz orientációjú nukleinsavak bejuttatása hatásos eszköz lehet a célba vett génnek transzkripciójának gátlására. A módszer alkalmazását endogén génnek expressziójának módosítására lásd Napoli és munkatársai közleményében [*The Plant Cell* 2, 279-289 (1990)] és az US 5,034,323 számú szabadalmi iratban.

A növényi génnek expressziójának gátlására katalitikus RNS-molekulák (ribozimek) is használhatók. Lehetséges olyan ribozimeket tervezni, amelyek specifikusan párt képeznek látszólag bármely megcélzott RNS-sel és elvágják a foszfodiészter gerincet egy meghatározott helyen, miáltal funkcionálisan inaktiválják a megcélzott RNS-t. Miközben a hasítást elvégzi, maga a ribozim nem változik meg és újabb molekulákat tud elhasítani; úgy működik, mint egy valódi enzim. Ribozim-szekvenciák beépítése antiszensz RNS-ekbe RNS-bontó aktivitást kölcsönöz azoknak, és ezáltal fokozza a szerkezet aktivitást. Megcélzott RNS-ekre specifikus ribozimek tervezését és alkalmazását Haseloff és munkatársai írták le [*Nature* 334, 585-591 (1988)].

Számos különböző térhálósító, alkilező és gyökképző reagens használható függelék-csoportként a találmány szerinti polinukleotidokon a nukleinsav összekapcsolása, megjelölése, kimutatása

és/vagy elvágása céljából. Így például Vlassov és munkatársai [Nucl. Acids Res. 14, 4065-4076 (1986)] írták le egy egyszálú DNS-fragmentum kovalens kötésekkel való kapcsolását a célszekvenciához nukleotidok alkilező származékainak segítségével. Hasonló munkáról számoltak be Knorre és munkatársai [Biochimie 67, 785-789 (1985)]. Iverson és Dervan egyszálú DNS szekvencia-specifikus elvágását oldották meg egy módosított nukleotid beépítésével, ami aktiválja az elvágást [J. Am. Chem. Soc. 109, 1241-1243 (1987)]. Meyer és munkatársai kovalens keresztkötéseket hoztak létre egy megcélzott nukleotiddal, amelyhez olyan alkilezőszert használtak, ami komplementer az egyszálú célnukleotid-szekvenciával [J. Am. Chem. Soc. 111, 8517-8519 (1989)]. Lee és munkatársai egyszálú oligonukleotidok fotoaktivált térhálósítását valósították meg pszoralén segítségével [Biochemistry 27, 3197-3203 (1988)]. Keresztkötéseket alkalmaztak hármasspirált képező próbák esetében Home és munkatársai [J. Am. Chem. Soc. 112, 2435-2437 (1990)]. Az N₄,N₄-etano-citozin keresztkötéseket létesít egyszálú oligonukleotidok között [Webb és Matteucci, J. Am. Chem. Soc. 108, 2764-2765 (1986); Nucl. Acids Res. 14, 7661-7674 (1986); Feteritz et al., J. Am. Chem. Soc. 113, 4000 (1991)]. A fentieken kívül még számos olyan vegyület ismert a szakterületen, amely nukleinsavak összekapcsolására, kimutatására, megjelölésére és/vagy elhasítására alkalmas (lásd például az US 5,543,507, 5,672,593, 5,484,908, 5,256,648 és 5,681,941 számú szabadalmi iratokat).

Fehérjék

A találmány szerinti, izolált fehérjék egy legalább 10 ami-

nosavból álló polipeptidet tartalmaznak egy találmány szerinti polipeptidből (vagy annak egy konzervatív változatából), amelyet a találmány szerinti, fent leírt polinukleotidok bármelyike kódol. A találmány szerinti fehérjék vagy változataik akárhány összefüggő aminosavat tartalmazhatnak egy találmány szerinti polipeptidből, ahol az aminosavak száma 10-től egy teljes hosszúságú, találmány szerinti polipeptid aminosavainak számáig terjedhet. Adott esetben az összefüggő aminosavak ezen alszekvenciája legalább 15, 20, 25, 30, 35 vagy 40, gyakran legalább 50, 60, 70, 80 vagy 90 aminosav hosszúságú. Továbbá, az ilyen alszekvenciák száma 1-20, például 2, 3, 4 vagy 5 lehet.

Találmányunk biztosít továbbá egy olyan fehérjét, amely egy olyan polipeptidet tartalmaz, ami meghatározott szekvencia-azonosságot vagy hasonlóságot mutat egy találmány szerinti polipeptiddel. A szekvencia-azonosság vagy hasonlóság foka 50 és 99 % között lehet. Példaszerű szekvencia-azonosság vagy hasonlóság az 55, 60, 65, 70, 75, 80, 85, 90 vagy 95 %. A szekvencia-azonosság foka például a GAP, CLUSTALW vagy BLAST algoritmusokkal határozható meg.

A szakemberek észreveszik, hogy találmányunk katalitikusan aktív polipeptideket (enzimeket) foglal magában, bár nem korlátozódik azokra. A katalitikusan aktív polipeptidek specifikus aktivitása legalább 20, 30 vagy 40 %-a, előnyösen legalább 50, 60 vagy 70 %-a, és legelőnyösebben legalább 80, 90 vagy 95 %-a a natív (nem szintetikus), endogén polipeptid aktivitásának. Továbbá, szubsztrátspecificitásuk (k_{cat}/K_m) adott esetben lényegében azonos a natív (nem szintetikus), endogén polipeptidével. Tipikusan a K_m -érték legalább 30, 40 vagy 50 %-a, előnyösebben

legalább 60, 70, 80 vagy 90 %-a a natív, endogén polipeptid K_m -értékének. Az enzimaktivitás és szubsztrátspecifitás (k_{cat}/K_m) mérésére szolgáló módszerek jól ismertek a szakterületen. A találmány szerinti fehérjék általában – ha antigénként alkalmazzák azokat – egy olyan antitest termelődését váltják ki, amely specifikusan reagál egy találmány szerinti polipeptiddel. Továbbá, a találmány szerinti fehérjék nem adnak reakciót egy találmány szerinti polipeptiddel szemben termelt antiszérummal, ha az teljesen ki van merítve ugyanazzal a polipeptiddel. A kötődést ki mutató immunvizsgálatok jól ismertek a szakterületen. Egy előnyös vizsgáló módszer a kompetitív immunkötődés vizsgálata. Így tehát a találmány szerinti fehérjék antigénként alkalmazhatók olyan antitestek előállításához, amelyek reakcióba lépnek egy találmány szerinti fehérjével, és ezért példaszerűen felhasználhatók immunvizsgálatokban vagy fehérjetisztító eljárásokban.

Fehérjék kifejeződése gazdasejtekben

A találmány szerinti nukleinsavakat használva kifejeztethető egy találmány szerinti fehérje egy rekombinációval módosított sejtben, például baktérium, élesztő, rovar, emlős vagy előnyösen növényi sejtekben. A sejtek a fehérjét a természetestől eltérő módon (például eltérő mennyiségben, összetételben, helyen és/vagy időben) termelik, mert emberi beavatkozással, genetikailag úgy lettek módosítva, hogy ezt tegyék.

Elvárható, hogy a szakemberek ismerjék a számos expressziós rendszert, amelyekkel egy találmány szerinti fehérjét kódoló nukleinsav kifejeztethető, ezért meg sem kíséreljük az olyan, ismert módszerek részletes leírását, amelyek fehérjék kifejezte-

tésére szolgálnak prokarióta és eukarióta gazdaszervezetekben.

Röviden összefoglalva, a találmány szerinti fehérjét kódoló, izolált nukleinsavak expressziója tipikusan úgy érhető el, hogy például a DNS-t vagy cDNS-t működőképes módon egy promoterhez kapcsoljuk (ami vagy konstitutív, vagy indukálható), majd egy expressziós vektorba beépítjük. A vektorok akár prokariótákban akár eukariótákban zajló replikációra és integrálódásra lehetnek alkalmasak. A tipikus expressziós vektorok transzkripciós és transzlációs terminátorokat, kezdő szekvenciákat és olyan promotorokat tartalmaznak, amelyek alkalmasak a találmány szerinti fehérjét kódoló DNS expressziójának vezérlésére. Ahhoz, hogy egy klónozott gén magas szintű expresszióját elérjük, kívánatos olyan expressziós vektorokat szerkeszteni, amelyek a transzkripció vezérléséhez legalább egy erős promotert, a transzláció megindulásához egy riboszóma-kötőhelyet, és egy transzkripció/transzláció terminátort tartalmaznak. A szakemberek felismerik, hogy egy találmány szerinti fehérje módosítható anélkül, hogy biológiai aktivitása csökkenne. Bizonyos módosítások végezhetők rajta, hogy elősegítsük a klónozást, expressziót vagy a megcélzott molekula beépülését egy fúziós fehérjébe. Az ilyen módosításokat jól ismerik a szakemberek; közéjük tartozik például egy metionin hozzáadása az N-terminálishoz, ami egy kezdő helyet biztosít, vagy járulékos aminosavak (például polihisztidin) elhelyezése bármelyik terminálison, ami egy kényelmesen kezelhető tisztító szekvenciát biztosít. Beépíthetők továbbá restriktációs helyek és terminátor-kodonok is.

Fehérjék szintézise

A találmány szerinti fehérjék sejtmentes, szintetikus eljárásokkal is megszerkeszthetők. A körülbelül 50 aminosavnál rövidebb fehérjék szilárd fázisú szintézise úgy valósítható meg, hogy a szekvencia C-terminális aminosavát egy oldhatatlan hordozóhoz kapcsoljuk, majd sorban hozzákapcsoljuk a szekvencia további aminosavait. A szilárd fázisú szintézis technikáit Barany és Merrifield (*„The Peptides: Analysis, Synthesis, Biology”*, 2. kötet, pp. 3-284), Merrifield és munkatársai [*J. Am. Chem. Soc.* **85**, 2149-2156 (1963)] és Stewart és munkatársai (*„Solid Phase Peptide Synthesis”*, 2. kiadás, Pierce Chem. Co., Rockford, Ill., 1984) írták le. A hosszabb fehérjék úgy szintetizálhatók, hogy rövidebb fragmentumokat kapcsolunk össze az N- és C-terminálisok kondenzálásával. A szakemberek jól ismerik a peptidkötés létrehozását a C-terminális aktiválásával (például az N,N'-diciklohexil-karbodiimid kapcsoló reagens használatával).

Fehérjék tisztítása

A találmány szerinti fehérjék a jól ismert, standard eljárásokkal tisztíthatók. A találmány szerinti, rekombináns fehérjék vagy közvetlenül, vagy fuzionált fehérjeként jelenhetnek meg. A rekombináns fehérje a sejt szétoldásának (például ultrahanggal vagy French-préssel) és affinitás-kromatográfiának a kombinálásával tisztítható. A fuzionált termékek esetében egy megfelelő fehérjebontó enzimmel végzett emésztés szabadítja fel a kívánt rekombináns fehérjét.

A találmány szerinti fehérjék – akár rekombinánsak, akár szintetikusak – a lényegében tiszta állapotig tisztíthatók standard, jól ismert eljárásokkal, köztük detergenses feloldással, szelektív

kicsapással (például ammónium-szulfát használatával), oszlopkromatográfiával, immunpurifikációs vagy más módszerekkel (lásd például Scopes: *„Protein Purification: Principles and Practice”*, Springer Verlag, New York, 1982; Deutscher: *„Guide to Protein Purification”*, Academic Press, 1990). Így például antitestek állíthatók elő az itt leírt fehérjék ellen. Az *E. coli*-ból történő tisztítás módja az US 4,511,503 számú szabadalmi iratban van leírva. A fehérje a termelő sejtekből kinyerhető, majd tovább tisztítható standard fehérjekémi-ai technikákkal. A termelt fehérje kimutatása szintén ismert módszerekkel, például radioimmun mérésekkel, Western-blot technikával vagy immunprecipitációval oldható meg.

Nukleinsavak bejuttatása gazdasejtekbe

Egy találmány szerinti nukleinsav egy gazdasejtbe juttatásának módszere találmányunk szempontjából nem kritikus: transzformációs és transzfekciós módszerek egyaránt kényelmesen alkalmazhatók. Ennek megfelelően az eljárások széles választékát fejlesztették ki egy DNS-szekvenciának egy gazdasejt genomjába inszertálására, hogy megvalósuljon a szekvencia transzkripciója és /vagy transzlációja, ami a szervezet fenotípusos megváltozását okozza. Így tehát bármilyen módszert alkalmazhatunk, amely biztosítja egy nukleinsav hatásos bejuttatását.

A. Növények transzformálása

Egy találmány szerinti polinukleotidot tartalmazó nukleinsavat adott esetben egy növénybe juttatunk. Általában a polinukleotidot először egy rekombináns expressziós keretbe vagy vektorba építjük be. A találmány szerinti, izolált nukleinsavak a szakterületen ismert technikákkal juttathatók be a növényekbe. A maga-

sabbrendű növényfajok transzformálására szolgáló technikák széles köre van leírva a technikai és tudományos szakirodalomban, valamint a szabadalmi iratokban [lásd például Weising et al., *Annu. Rev. Genet.* 22, 421-477 (1988)]. Így például a DNS-szerkezetet közvetlenül a növényi sejtek genomjába juttathatjuk olyan eljárásokkal, mint az elektroporáció, a polietilénglikolos (PEG) precipitáció, a részecskebombázás, a szilikonszálas szállítás vagy a mikroinjektálás; ezeket az eljárásokat a növényi sejtek protoplasztjain vagy embriogén kalluszon végezzük [Tomes et al., in „*Plant Cell, Tissue and Organ Culture, Fundamental Methods*”, pp. 197-213; Gamborg és Phillips (eds.), Springer Verlag, Berlin-Heidelberg-New York (1995); US 5,990,387]. A DNS-szerkezetek polietilénglikolos precipitációval történő bejuttatását Paszkowski és munkatársai írták le [*EMBO J.* 3, 2717-2722 (1984)]. Az elektroporációs technikát Fromm és munkatársai írták le [*Proc. Natl. Acad. Sci. USA* 82, 5824 (1985)]. A ballisztikus transzformációs technikákat Klein és munkatársai írták le [*Nature* 327, 70-73 (1987)].

Az *A. tumefaciens* közvetítésű transzformációs technikáknak nagy szakirodalmuk van [Horsch et al., *Science* 233, 496-498 (1984); Fraley et al., *Proc. Natl. Acad. Sci. USA* 80, 4803 (1983)]. A DNS-szerkezet egy alkalmas T-DNS határoló területtel kombinálható és bejuttatható egy szokványos *Agrobacterium tumefaciens* gazdavektorba. Az *Agrobacterium tumefaciens* virulenciafunkciói fogják irányítani a szerkezet és a vele szomszédos marker beépülését a növényi sejt DNS-ébe, amikor a sejt fertőződik a baktériummal (US 5,591,616). Noha az *Agrobacterium* elsősorban kétszikűeken alkalmazható, egyes egyszikűek is transzformálhatók

vele; így például a kukorica *Agrobacterium*-os transzformálását írták le az US 5,550,318 számú szabadalmi iratban.

A transzfektálás vagy transzformálás egyéb módszerei közé tartozik (1) az *Agrobacterium rhizogenes* közvetítésű transzformáció [lásd például Rigby (ed.): „*Genetic Engineering*”, 6. kötet, Acad. Press, London, 1987; Glover (ed.): „*DNA Cloning*”, 2. kötet, IRI Press, London, 1985]; a WO 88/02405 számú szabadalmi iratban az *A. rhizogenes* A4 törzsének és Ri plazmidjának használata van leírva a pARC8 vagy pARC16 *A. tumefaciens* vektorokkal; (2) a liposzóma-közvetítésű DNS-felvétel [Freeman et al., *Plant Cell Physiol.* 25, 1353 (1984)]; vagy (3) az örvénykeltéses eljárás [Kindle, *Proc. Natl. Acad. Sci. USA* 87, 1228 (1990)].

A DNS közvetlenül a pollenbe történő transzferrel is bejuttatható a növényekbe [Zhou et al., *Meth. Enzymol.* 101, 433 (1983); Hess, *Intern. Rev. Cytol.* 107, 367 (1987); Luo et al., *Plant Mol. Biol. Reporter* 6, 165 (1988)]. Polipeptidet kódoló gének expressziója elérhető, ha a DNS-t egy növény reproduktív szerveibe injektáljuk [Pena et al., *Nature* 325, 274 (1987)]. A DNS közvetlenül injektálható éretlen embriók sejtjeibe és bejuttatható a kiszáradt csírakezdemények rehidráálásával is [Neuhaus et al., *Theor. Appl. Genet.* 75, 30 (1987); Benbrook et al., *Proc. Bio Expo* 1986, pp. 27-54; Butterworth, Stoneham, Ma. (1986)]. Különböző növényi vírusok is használhatók vektorként, mint például a karfiolmozaik (CaMV), a gemini, a rozsnokmozaik vagy a dohánymozaik vírus.

B. Prokarióta, alacsonyrendű eukarióta és állati sejtek transzfekeciója

Az állati és az alacsonyrendű eukarióta (például élesztő) gazdasejtek transzfektálhatók vagy különböző módokon transzfektálhatóvá tehetők. Számos eljárást ismerünk DNS állati sejtekbe juttatására: ezek közé tartozik a kalcium-foszfátos precipitálás, a befogadó sejt összeolvasztása a DNS-t tartalmazó baktérium-protoplasztokkal, a befogadó sejt kezelése a DNS-t tartalmazó liposzómákkal, illetve a DNS közvetlen bejuttatása a sejtekbe DEAE-dextránnal, elektroporációval, biolitikumokkal vagy mikroinjekcióval. A transzfektált sejtek jól ismert módszerekkel tenyésztethetők tovább (Kuchler: *„Biochemical Methods in Cell Culture and Virology”*, Dowden, Hutchinson and Ross, Inc., 1977).

A transzgén növény regenerálása

A fent leírt transzformációs technikák bármelyikével létrehozott növényi sejtek úgy tenyésztethetők, hogy belőlük a bejuttatott genotípussal rendelkező teljes növények regenerálódjanak. Az ilyen regeneráló módszerek gyakran alkalmazzák a szövettenyészet növényi hormonokkal való befolyásolását. A transzformált sejtekből felépülő növények például egyes sejtekből, kalluszsövetből vagy levélkorongokból regenerálhatók standard növényi szövettenyésztési módszerekkel. A szakterületen jól ismert, hogy csaknem bármely növény különböző sejtjei, szövetei vagy szervei teljes növényekké regenerálhatók. A növények protoplaszt-tenyészetekből való regenerálása Evans és munkatársai (in *„Handbook of Plant Cell Culture”*, pp. 124-176, Macmillan Publ. Co., New York, 1983) és Binding (in *„Plant Protoplasts”*, pp. 21-73, CRC Press, Boca Raton, 1985) leírásaiból ismerhető meg.

A növények regenerálása akár egyetlen protoplasztból vagy

különböző explantátumokból jól ismert a szakterületen [lásd például Weissbach és Weissbach (eds.): „*Methods for Plant Molecular Biology*”, Academic Press, Ind., San Diego, Ca., 1988]. A regenerációs eljárás magában foglalja a transzformált sejtek és hajtások szelekcióját, a transzformált hajtások gyökereztetését és a növénykéek talajban való felnevelését. A kukorica sejttenyészekről és regenerálásukról általában lásd a már idézett, „*The Maize Handbook*” és „*Corn and Corn Improvement*” című kézikönyveket. A kukorica transzformálásáról és regenerálásáról Gordon-Kamm és munkatársai írnak [*The Plant Cell* 2, 603-618 (1990)].

Az *Agrobacterium*-mal bejuttatott találmány szerinti polinukleotidot tartalmazó növények levéldarabokból történő regenerálását Horsch és munkatársai írták le [*Science* 227, 1229-1231 (1985)]. Ebben az eljárásban a transzformátumok egy szelekciós szer jelenlétében és egy olyan tápfolyadékban növekednek, amely hajtásfejlődést indukál [Fraley et al., *Proc. Natl. Acad. Sci. USA* 80, 4803 (1983)]. Ezzel a módszerrel tipikusan 2-4 héten belül hozhatók létre a hajtások, amelyeket ezután egy gyökérfejlődést indukáló, a szelekciós szert és egy antibiotikumot (a baktériumfertőzések megelőzésére) tartalmazó tápközegben tenyésztünk tovább. A találmány szerinti, transzgén növények lehetnek termékenyek vagy terméketlenek.

A szakemberek tudják, hogy miután a rekombináns expressziós keret stabilan beépült a transzgén növényekbe és ott működőképesnek bizonyult, ivaros keresztezéssel tovább vihető más növényekbe. E célra bármilyen szokványos nemesítési technika alkalmazható a keresztezni kívánt fajoktól függően. A vegetatív úton

szaporított növények esetében a kifejlődött transzgén növényekről vett dugványokat szaporítjuk tovább, vagy szövettenyésztési technikákat alkalmazva állítunk elő azonos növényeket. A kívánatos transzgéneket kiválasztjuk és az új változatot vegetatív úton szaporítjuk fel a kereskedelmi forgalom számára. A magról szaporított növények esetében a kifejlődött transzgén növények önbeporzásával hozunk létre homozigóta, beltenyésztett növényeket, amelyek magvai tartalmazzák az újonnan beépített heterológ nukleinsavat. Ezekből a magokból olyan növények nevelhetők, amelyek a kiválasztott fenotípussal rendelkeznek. A regenerált növényekről nyert részek, például virágok, magvak, levelek, hajtások, gyümölcsök és hasonlóak szintén találmányunkhoz tartoznak, feltéve, hogy ezek a részek is tartalmazzanak olyan sejteket, amelyek a találmány szerinti, izolált nukleinsavat tartalmazzák. A regenerált növények leszármazottai, változatai és mutánsai szintén találmányunk oltalmi körébe tartoznak, feltéve, hogy tartalmazzák a bejuttatott nukleinsav-szekvenciákat. A találmány szerinti polinukleotidokat kifejező növények például standard immunológiai vagy DNS-kimutató technikákkal vizsgálhatók a találmány szerinti nukleinsav jelenlétére. Kezdetnek az RNS-szintű expressziót határozhatjuk meg, hogy azonosítsuk és értékeljük az expresszió-pozitív növényeket. E célra az RNS analízisének standard technikáit alkalmazhatjuk; ide tartozik a PCR sokszorozó vizsgálat olyan oligonukleotid primerekkel, amelyek csak a heterológ RNS templátok sokszorozását biztosítják, valamint az oldatban végzett hibridizáció a heterológ nukleinsavra specifikus próbákkal. Az RNS-pozitív növények ezután megvizsgálhatók a

fehérje kifejeződésére például Western-immunblot technikával, a találmány szerinti, specifikus antitesteket használva. Emellett *in situ* hibridizációs és immuncitokémiai vizsgálatokat is végezhetünk szokványos módszerekkel, a heterológ nukleinsavra specifikus polinukleotid próbákkal és antitestekkel, hogy meghatározzuk az expresszió helyét a transzgén szövetben. Általában számos transzgén vonal szűrővizsgálatát kell elvégezni ahhoz, hogy azonosítsuk és kiválasszuk a legmegfelelőbb expressziós profillal rendelkező növényeket.

Egy előnyös megvalósításban a transzgén növény homozigóta a bevitt heterológ nukleinsavra nézve, azaz egy transzgén növény egy-egy bevitt nukleinsav-szekvenciát tartalmaz egy kromoszómapár mindkét tagjának ugyanazon helyén. Egy homozigóta transzgén növény úgy hozható létre, hogy egy heterozigóta transzgén növényről – amely csak egy bevitt heterológ nukleinsavat tartalmaz – önbeporzással magokat nyerünk, amiket kicsíráztatva és a növénykéket analizálva kiválasztjuk azokat, amelyek megváltozott módon fejezik ki a találmány szerinti polinukleotidot a kontroll (natív, nem transzgén) növényekhez képest. Az anyanövénnel való visszakeresztezés, vagy egy nem transzgén növénnel való keresztezés is számításba jöhet.

A polipeptid-mennyiség és/vagy összetétel módosítása

Találmányunk egy eljárást biztosít továbbá a találmány szerinti polipeptidek koncentrációjának vagy arányának módosítására (emelésére vagy csökkentésére) egy növényben vagy annak egy részében. Az eljárás magában foglalja egy növényi sejt transzformálását egy rekombináns expressziós kerettel, amely egy talál-

mány szerinti, fent leírt polinukleotidot tartalmaz, hogy megkapjunk egy transzgén növényi sejtet, majd a transzgén növényi sejt szaporítását a transzgén növény regenerálódását biztosító körülmények között, végül a találmány szerinti polinukleotid kifejeződésének indukálását vagy elnyomását a növényben elegendő ideig ahhoz, hogy a polipeptid koncentrációja és /vagy a polipeptidek aránya módosuljon a transzgén növényben vagy részeiben.

Néhány megvalósításban a találmány szerinti polipeptidek koncentrációja és/vagy aránya úgy befolyásolható egy növényben, hogy *in vivo* vagy *in vitro* megváltoztatjuk egy gén promoterét, és ezzel a gén expresszióját le- vagy felszabályozzuk. Néhány megvalósításban a találmány szerinti, natív gének kódoló területe változtatható meg helyettesítés, hozzáadás vagy kivágás útján, hogy csökkenjen a kódolt enzim aktivitása (US 5,565,350 és WO 93/22443 számú szabadalmi iratok). Még néhány megvalósításban egy promoter-szekvenciát tartalmazó, izolált nukleinsavat (például egy vektort) juttatunk egy növényi sejtbe. Ezt követően a promotert – amihez működőképes módon csatlakozik egy találmány szerinti polinukleotid – tartalmazó növényi sejtet a szakterületen jól ismert módszerek valamelyikével, például Southern-blot-tal, DNS-szekvenálással, PCR analízissel (amihez a promoterra és a génre specifikus primereket használunk) vagy az amplikon kimutatásával kiválasztjuk. A fenti megvalósítások szerint módosított növényt vagy növényi részt a növény kifejlődésére kedvező körülmények között neveljük elegendő ideig ahhoz, hogy a polipeptidek koncentrációja és/vagy aránya megváltozzon a növényben. A növények kifejlődésére kedvező körülmények jól ismertek, és

már fent érintettük azokat.

Általában a polipeptidek koncentrációja vagy aránya legalább 5, 10, 20, 30, 40, 50, 60, 70, 80 vagy 90 %-kal emelkedik vagy csökken egy natív kontroll növényhez, növényi részhez vagy sejthez képest, amiből hiányzik az említett rekombináns expressziós keret. A találmány szerinti módosulás azalatt és/vagy azt követően lép fel, hogy a növény eléri a kívánt fejlődési állapotot. A nukleinsav időszakos és/vagy egy adott szövetben történő expressziójának vezérlésére egy megfelelő promotert alkalmazunk, amely működőképes módon van csatlakoztatva egy találmány szerinti polinukleotidhoz szensz vagy antiszensz orientációban; ezt fentebb részletesen leírtuk. Egy találmány szerinti polinukleotid expressziója egy induktor vegyület hatásos mennyiségének kívülről történő adagolásával is vezérelhető. Az indukálható promoterek és az induktor vegyületek, amelyek aktiválják az ilyen promotereket, jól ismertek a szakterületen. Az előnyös megvalósításokban a találmány szerinti polipeptideket egyszikűekben, különösen kukoricában módosítjuk.

Le nem fordítható területek és kodonpreferencia

A transzláció eredményességét általában specifikus, az RNS 5' nem kódoló vagy le nem fordítható (5'-UTR) területén lévő szekvencia-elemek szabályozzák. A pozitív hatású szekvenciák közé tartoznak a transzlációt indító konszenzus-szekvenciák [Kozak, *Nucl. Acids Res.* 15, 8125 (1987)], és a 7-metil-guanozin „sapka” [Drummond et al., *Nucl. Acids Res.* 13, 7375 (1985)]. Negatív hatású elemek a stabil intramolekuláris 5'-UTR törzs-hurok szerkezetek [Muessing et al., *Cell* 48, 691 (1987)] és az 5'-UTR-

ben lévő AUG szekvenciák vagy azok a rövid leolvasási keretek, amelyeket egy megfelelő AUG előz meg az 5'-UTR-ben [Kozak, *id. közl.*; Rao *et al.*, *Mol. Cell. Biol.* 8, 284 (1988)]. Ennek megfelelően találmányunk 5' és/vagy 3' le nem fordítható területeket biztosít, amelyekkel a heterológ kódoló szekvenciák transzlációja módosítható.

Továbbá, a találmány szerinti polinukleotidok polipeptid-kódoló szegmentumai a kodonhasználat megváltoztatása érdekében is módosíthatók. A kodonhasználat megváltoztatása arra használható, hogy módosítsuk a transzláció hatékonyságát és/vagy optimalizáljuk a kódoló szekvenciát egy kívánt gazdában való kifejeződéshez, vagy optimalizáljunk egy heterológ szekvenciát a kukoricában való kifejeződéshez. A találmány szerinti polinukleotidok kódoló területeinek kodonhasználatát statisztikusan elemezhető kereskedelmi forgalomban lévő szoftvercsomagokkal, mint amilyen a „Codon Preference”, ami az University of Wisconsin Genetics Computer Group-tól kapható [lásd Deveraux *et al.*, *Nucl. Acids Res.* 12, 387-395 (1984)], vagy a MacVector 4.1 (Eastman Kodak Co., New Haven, Cn.). Így tehát találmányunk egy kódhasználat-gyakorisági jellemzőt biztosít a találmány szerinti polinukleotidok legalább egyikének kódoló területe számára. A polinukleotidok száma, amelyeket bevonhatunk a kódhasználat gyakoriságának meghatározásába, 1-től a találmány szerinti polinukleotidok számáig terjedhet. Adott esetben a polinukleotidok teljes hosszúságú szekvenciák lehetnek. A statisztikai analízishez használt szekvenciák példaértékű száma legalább 1, 5, 10, 20, 50 vagy 100 lehet.

Szekvencia-megkeverés

Találmányunk szekvencia-megkeverési eljárásokat is biztosít a találmány szerinti polinukleotidokban vagy az azokból kapott készítményekben. A szekvencia-megkeverés a WO 97/20078 számú szabadalmi iratban, valamint Zhang és munkatársai közleményében [*Proc. Natl. Acad. Sci. USA* 94, 4504-4509 (1997)] van leírva. Általánosságban a szekvencia-megkeverés eszközt biztosít könyvtárak készítésére olyan polinukleotidokból, amelyeknek egy kívánt jellemzőjük van, ami szűrővizsgálattal kiválasztható. A rekombináns polinukleotidok könyvtárai a hasonló szekvenciájú polinukleotidok populációiból készíthetők, amelyek olyan szekvencia-területeket tartalmaznak, amelyek lényeges mértékben azonosak és képesek a homológ rekombinációra *in vitro* és *in vivo*. A szekvencia-rekombinált polinukleotidok populációja tartalmaz egy alpopulációt, amely kívánt vagy előnyös tulajdonságokkal rendelkezik, és amely egy megfelelő szelekciós vagy szűrővizsgálattal kiválogatható. Ez a tulajdonság bármely olyan jellemző vagy sajátosság lehet, amely egy szűrővizsgálattal kiválasztható vagy kimutatható, így például egy kódolt fehérje, egy transzkripciós elem, egy transzkripciót szabályozó szekvencia, RNS-feldolgozás, az RNS stabilitása, kromatin-konformáció, transzláció vagy egy gén vagy transzgén más expressziós sajátossága, egy replikációs elem, egy fehérjekötő elem, vagy hasonló, például bármely olyan tulajdonság, aminek van egy kiválasztható vagy kimutatható sajátossága. Néhány megvalósításban a kiválasztott tulajdonság lehet a vad típusú fehérjéhez képest magasabb K_m - és/vagy k_{cat} -érték. Más megvalósításokban a szekvencia-megkeveréssel létrehozott po-

linukleotid vagy fehérje ligandumkötő kapacitása nagyobb, mint a nem megkevert, vad típusú polinukleotidé. Az ilyen tulajdonságok fokozódásának mértéke legalább 110, 120, 130, 140 vagy 150 % lehet a vad típushoz képest.

Általános és konszenzus-szekvenciák

A találmány szerinti polinukleotidok és polipeptidek magukban foglalják továbbá (a) legalább két, találmány szerinti homológ polinukleotid vagy polipeptid általános szekvenciáját, és (b) legalább három, találmány szerinti homológ polinukleotid vagy polipeptid konszenzus-szekvenciáját. A találmány szerinti általános szekvencia magában foglal minden egyes polipeptidet vagy polinukleotidot, amely tartalmazza az általános polipeptid- vagy polinukleotid-szekvenciát. Az egyes fajtaikat egy polinukleotid fogja össze, aminek aminosav vagy nukleinsav konszenzus-szekvenciája felhasználható antitestek előállítására vagy nukleinsav próbák vagy primerek készítésére, amelyekkel homológok kereshetők más fajokban, nemzetségekben, családokban, rendekben, osztályokban, törzsekben vagy országban. Így például egy polinukleotid, aminek van egy konszenzus-szekvenciája a *Zea mays* egy géncsaládjából, felhasználható olyan antitest, illetve nukleinsav próbák vagy primerek előállítására, amelyek más *Gramineae*-fajok, például búza, rizs vagy cirok szűrővizsgálatára alkalmasak. Másrészt egy olyan polinukleotid, amelynek konszenzus-szekvenciája ortológ génekből származik, felhasználható ortológok azonosítására vagy izolálására más taxonokban. Egy polinukleotid konszenzus-szekvenciája tipikusan legalább 9, 10, 15, 20, 25, 30 vagy 40 aminosav, illetve 20, 30, 40, 50, 100 vagy 150 nukleotid

hosszúságú. A szakemberek tudják, hogy egy konzervatív aminosav-csere végrehajtható az olyan aminosavak helyén, amelyek különböznek az illeszkedő szekvenciákban, de csak a megfelelő konzervatív szubsztitúciós csoportból, amelyeket fentebb ismertettünk. Adott esetben nem több, mint 1 vagy 2 konzervatív aminosav-csere végezhető el a konszenzus-szekvencia minden 10 aminosavnyi hosszán.

Hasonló szekvenciák használhatók egy általános vagy konszenzus-szekvencia elkészítéséhez, amely a találmány szerinti gének allélikus változatait, vagy ortológ vagy paralóg szekvenciáit tartalmazza bármely számban és kombinációban. Adott esetben az általános vagy konszenzus-szekvencia elkészítéséhez felhasznált hasonló szekvenciák a BLAST algoritmus segítségével, a valószínűség legkisebb összegének $[P(N)]$ kiszámításával azonosíthatók. A szekvencia-vizsgáló szoftverek szállítói Ausubel már idézett kézikönyvének 7. fejezetében vannak felsorolva. Egy polinukleotid-szekvencia akkor nevezhető hasonlónak egy viszonyítási szekvenciához, ha az összehasonlítás eredményeként a valószínűség legkisebb összege kisebb, mint körülbelül 0,1, előnyösebben kisebb, mint körülbelül 0,01 vagy 0,001, és legelőnyösebben kisebb, mint körülbelül 0,0001 vagy 0,00001. A hasonló polinukleotidok egymáshoz illeszthetők és egy konszenzus vagy általános szekvencia készíthető el belőlük egy többszörös szekvenciaillesztő szoftverrel, mint amilyen a PILEUP (Genetics Computer Group, Madison, Wi.), az ALIGNX (Vector NTI, North Bethesda, Md.), vagy a SEQUENCHER (Genecode, Ann Arbor, Mi.). A fenti szoftverek alapbeállításai kényelmesen használhatók az általános vagy konszenzus-szekvenciák elkészítésére.

Gépi alkalmazások

Találmányunk eljárásokat is biztosít a találmány szerinti polinukleotidok és polipeptidek modellezéséhez vagy analizálásához.

Találmányunk biztosít egy eljárást egy találmány szerinti polinukleotid vagy polipeptid homológ-jelöltjének (azaz egy ortológjának vagy paralógjának) azonosítására. Az eljárás magában foglalja egy találmány szerinti polinukleotid vagy polipeptid szekvencia-adatainak betáplálását egy olyan gépbe, amelynek egy hardver vagy szoftver alapú szekvencia-analizáló rendszere van, egy adatstruktúra felépítését, amely meggyorsítja a szekvencia-adatok elérését, az adatok feldolgozását a polinukleotid vagy polipeptid szerkezetének analizálása céljából, és az analízis eredményének kijelzését. Egy homológ-jelölt esetében statisztikailag szignifikáns annak valószínűsége, hogy ugyanaz a biológiai funkciója (ugyanazt a reakciót katalizálja, kötődik a homológ fehérjékhez és/vagy nukleinsavakhoz, hasonló szerepet játszik egy szerkezetben), mint az összehasonlítás alapjául szolgáló referencia-szekvenciának. Ennek megfelelően a találmány szerinti polinukleotidok és polipeptidek felhasználhatók homológok azonosítására állatokban vagy más növényfajokban, különösen a *Gramineae* család fajaiban, például – nem kizárólag – cirokban, búzában vagy rizsben.

A találmány szerinti eljárás magában foglalja egy polinukleotid vagy polipeptid teszt-szekvenciát képviselő adatok megszerzését. A teszt-szekvenciák egy állat vagy egy növény nukleinsavából nyerhetők. Teszt-szekvenciák kaphatók közvetlenül vagy közvetve adatbázisokból is, mint amilyenek – nem kizárólag – a

következők: GenBank, EMBL, GenSeq, SWISS-PROT vagy a GenomeWeb. Néhány megvalósításban a teszt-szekvenciát egy kukoricától eltérő növényfajból vesszük, és bár funkciója bizonytalan, mégis összehasonlítható a vizsgált szekvenciával a szekvencia-hasonlóság vagy -azonosság meghatározása céljából. A teszt-szekvencia adatait egy gépbe, például egy számítógépbe tápláljuk, amely tartalmazza (i) egy referencia-szekvencia adatait, és (ii) egy hardver vagy szoftver szekvencia-összehasonlító eszközt a referencia- és a teszt-szekvencia közötti szekvencia-hasonlóság vagy -azonosság feltárásához.

A szekvencia-analízis céljára példaszerű szekvencia-összehasonlító szoftvert kínál a Genetics Computer Group (Madison, Wi.), az InforMax (Bethesda, Md.) vagy az Intelligenetics (Mountain View, Ca.). Adott esetben a szekvencia-összehasonlítás megoldható a BLAST vagy GAP programcsaládokkal is. A BLAST 2.0 programot alapbeállítási paraméterekkel használva a teszt-szekvenciát akkor tekintjük a referencia-szekvencia homológjának (azaz alléljének, ortológjának vagy paralógjának), ha a valószínűség legkisebb összege $[P(N)]$ kisebb, mint 0,1, vagy kisebb, mint 0,01, 0,001, 0,0001 vagy 0,00001. A szakemberek felismerik, hogy úgy növekszik annak statisztikai valószínűsége, hogy egy homológ azonos vagy hasonló funkcióval rendelkezik, amint a hasonlóság foka növekszik a vizsgált és a referencia-szekvencia között.

A referencia-szekvencia egy találmány szerinti polipeptid vagy polinukleotid szekvenciája lehet. Mind a referencia-, mind a teszt-szekvencia adott esetben legalább 25 aminosav vagy legalább 100 nukleotid hosszúságú, de hosszuk elérheti a „Nuklein-

savak" (különösen a (g) pontban), illetve a „Fehérjék” fejezetekben leírt polinukleotidok illetve polipeptidek teljes hosszúságát is. A szakemberek figyelmét felhívjuk arra a tényre, hogy minél nagyobb fokú a szekvencia-azonosság vagy hasonlóság egy ismert működésű referencia-szekvencia és egy teszt-szekvencia között, annál nagyobb annak valószínűsége, hogy a teszt-szekvencia működése is azonos vagy hasonló lehet. A szekvenciák összehasonlításának eredményét kijelezzük (képernyőn megjelenítjük, kinyomtatjuk, rögzítjük) a kijelző berendezések és/vagy médiumok (például képernyő, nyomtatvány vagy számítógép által olvasható médiumok) egyikén.

Nukleinsavak kimutatása

Találmányunk egy eljárást biztosít továbbá egy találmány szerinti polinukleotid kimutatására egy nukleinsav-mintában, amiről feltételezhető, hogy egy találmány szerinti polinukleotidot tartalmaz, így egy növényi sejt-lizátumban, különösen kukorica-lizátumban. Néhány megvalósításban egy találmány szerinti polinukleotiddal rokon gént vagy annak egy részletét sokszorozhatjuk, mielőtt a nukleinsav-mintát egy találmány szerinti polinukleotiddal érintkeztetjük. A polinukleotiddal érintkeztetett nukleinsav-minta egy hibridizációs komplexet képez. A polinukleotid szigorú körülmények között hibridizálódik egy találmány szerinti polipeptidet kódoló génnel. A hibridizációs komplex kialakulását használjuk egy találmány szerinti polipeptidet kódoló gén kimutatására a nukleinsav-mintában. A szakemberek megértik, hogy egy izolált nukleinsavból, amely egy találmány szerinti polinukleotidot tartalmaz, hiányozniok kell a megcélzott génektől

eltérő génekben közönséges, kereszt-hibridizálódó szekvenciáknak, hogy ne okozzanak ál-pozitív eredményt. A hibridizációs komplex kimutatása bármely ismert módon történhet. Így például a nukleinsav-minta vagy annak egy része oldat-fázisú, szilárd fázisú, vegyes fázisú vagy *in situ* hibridizációs vizsgálattal elemezhető.

A találmány szerinti használatra alkalmas, kimutatható jelölés bármely olyan kompozíció lehet, amely spektroszkópos, radioizotópos, fotokémiai, biokémiai, immunkémiai, elektromos, optikai vagy kémiai mérésekkel kimutatható. Megfelelő jelölés például a biotin a jelölt sztreptavidin-konjugátummal való festés céljára, a mágneses gyöngyök, a fluoreszcens festékek, a radioizotópok, az enzimek vagy a kolorimetriás jelölések. A másféle jelölések közé tartoznak azok a ligandumok, amelyek fluorofórokkal, kemilumineszkáló reagensekkel vagy enzimekkel kapcsolt antitestekhez tudnak kötődni. A találmány szerinti nukleinsavak jelölése könnyen megvalósítható például jelölt PCR primerek használatával.

Bár találmányunkat a bemutatás érdekében részletesen leírtuk és az érthetőség kedvéért példákkal is illusztráljuk, nyilvánvaló, hogy bizonyos változtatások és módosítások végezhetők azon a mellékelt igénypontok oltalmi körén belül.

1. példa:

cDNS-könyvtár szerkesztése

Kukorica-szövetek teljes RNS-tartalmát izoláljuk „TRIZOL Reagent” segítségével (Life Technology, Inc., Gaithersburg, Md.), ami a guanidin-izotiocianát/savas-fenol módszer módosított változata [Chomczynski és Sacchi, *Anal. Biochem.* 162, 156 (1987)]. Röviden, a növényi szövetmintákat folyékony nitrogénben elporít-

juk, majd hozzáadjuk a TRIzol reagenst és tovább homogenizáljuk dörzsmozsárban. A mintához kloroformot adunk, majd a vizes és szerves fázisokat centrifugálással elválasztjuk. A teljes RNS-t a vizes fázisból izopropanollal csapjuk ki.

A poli(A)+ RNS-t az összes RNS-ből választjuk el a „Poly-ATtract” rendszerrel (Promega Corp; Madison, Wi.). Röviden, biotinezett oligo(dT) primereket használunk arra, hogy hibridizálódjanak a mRNS 3' poli(A) farkaival. A hibrideket paramágneses részecskékhez kapcsolt sztreptavidinnel kötjük meg és mágneses térrel elválasztjuk. A mRNS-t erősen szigorú feltételek mellett mossuk és RNáz-mentes desztillált vízzel eluáljuk.

A cDNS-szintézist és az egyirányú cDNS-könyvtárak létrehozását a „SuperScript Plasmid System” segítségével végezzük (Life Technologies, Inc., Gaithersburg, Md.). A cDNS első szálát egy oligo(dT) primerrel szintetizáljuk, ami egy *NotI* restrikciós helyet tartalmaz. A reakciót a „SuperScript Reverse Transcriptase II” katalizálja 45 °C hőmérsékleten. A cDNS második szálát α -³²P-dCTP-vel jelöljük, és a reakcióelegy egy részét agarózgél-elektroforézissel analizáljuk, hogy meghatározzuk a cDNS méretét. Az 500 bázispárnál rövidebb cDNS-molekulákat és a nem kapcsolódott adaptereket Sephacryl-S400 kromatográfiával eltávolítjuk. A kiválasztott cDNS-molekulákat a pSPORT1 vektorba ligáljuk a *NotI* és *SalI* restrikciós helyeken.

Másrészt a cDNS-könyvtárak a számos elérhető módszer bármelyikével előállíthatók. Így például a cDNS-ek plazmid-vektorokba vihetők be úgy, hogy a cDNS-könyvtárat először Uni-ZAP™ XR vektorokban készítjük el a gyártó receptje szerint (Stratagene Clo-

ning Systems, La Jolla, Ca.). Az Uni-ZAP™ XR könyvtárakat plazmid-könyvtárakká alakíthatjuk át szintén a Stratagene receptje szerint. Az átalakítás nyomán a cDNS-inszertumokat a pBluescript plazmid-vektor fogja tartalmazni. Emellett a cDNS-ek közvetlenül is beépíthetők az előre felnyitott Bluescript II SK(+) vektorokba (Stratagene) T4 DNS-ligáz használatával (New England Biolabs), amit a DH10B sejtekbe való transzfektálás követ a gyártó receptje szerint (GIBCO BRL Products). Amikor a cDNS-inszertumok már a plazmid-vektorokban vannak, a plazmid-DNS-t véletlenszerűen felszedett, a rekombináns pBluescript plazmidokat tartalmazó baktériumtelepekből nyerjük ki, vagy az inszertált cDNS-szekvenciákat polimeráz láncreakcióval sokszorozzuk az inszertált cDNS-szekvenciákat határoló vektor-szekvenciákra specifikus primerek segítségével. A sokszorozott inszertum-DNS-eket vagy plazmid-DNS-eket a festék-primer szekvenáló reakcióval szekvenáljuk, hogy részleges cDNS-szekvenciákat („kifejezett szekvencia-végek” vagy „EST”-k) kapjunk [Adams et al., *Science* 252, 1651-1656 (1991)]. Az így keletkezett EST-ket egy Perkin-Elmer Model 377 fluoreszcenciás szekvenátorral analizáljuk.

2. példa:

Teljes hosszúságú, dúsított cDNS-könyvtár elkészítése

Egy teljes hosszúságú, dúsított cDNS-könyvtár Carninci és munkatársai módszere két változatának egyikével készíthetők el [Genomics 37, 327-336 (1996)]. Ezek a változatok azon alapszanak, hogy kémiai úton egy biotincsoportot viszünk be az eukarióta mRNS 5' „sapkájának” diolcsoportjára, hogy kiválasszuk a teljes hosszúságú, első cDNS-szálat. A szelekció úgy valósul meg,

hogy a sapkán lévő biotincsoportot sztreptavidinnel bevont mágneses gyöngyökkel megfogjuk, majd egy RNáz I-gyel végzett kezelés következik az inkompletten szintetizálódott cDNS-ek eltávolítása céljából. A második cDNS-szálak jól bevált eljárásokkal szintetizálhatók, mint amilyen például a Life Technologies (Rockville, Md.) „SuperScript Plasmid System for cDNA Synthesis and Plasmid Cloning” készlete. Az így készített könyvtárakról kimutatták, hogy 50-70 % teljes hosszúságú cDNS-t tartalmaznak.

Az első szál szintézisének módját az alábbiakban részletesen ismertetjük. A * azt jelzi, hogy a reagenst a Life Technologies, Inc.-től szereztük be.

A. Az első cDNS-szál szintézisének eljárása (trehalózzal)

mRNS (10 µg)	25 µl
*NotI primer (5 µg)	10 µl
*5X első szál puffer	43 µl
*0,1 M DTT	20 µl
*dNTP keverék, 10 mM	10 µl
trehalóz (telített oldat)	59,2 µl
RNáz inhibitor (Promega)	1,8 µl
*SuperScript II RT 200 U/µl	20 µl
glicerin (100 %-os)	18 µl
víz	7 µl

A mRNS-t és a NotI primert összekeverjük és 65 °C-on denatúráljuk 10 percen át. Ezután az elegyet jégben hűtjük és hozzáadjuk a többi összetevőt. A reakcióelegyet 45 °C-on inkubáljuk 2 percen át. Ezután 20 µl RT-t (reverz transzkriptáz) adunk hozzá és elindítjuk a termociklus-programot (MJ Research, Waltham, Ma.):

1. lépés: 45 °C, 10 perc
2. lépés: 45 °C, -0,3 °C/ciklus, 2 s/ciklus
3. lépés: a 2. lépés 33-szoros ismétlése
4. lépés: 35 °C, 5 perc
5. lépés: 45 °C, 5 perc
6. lépés: 45 °C, 0,2 °C/ciklus, 1 s/ciklus
7. lépés: a 6. lépés 49-szeres ismétlése
8. lépés: 55 °C, 0,1 °C/ciklus, 12 s/ciklus
9. lépés: a 8. lépés 49-szeres ismétlése
10. lépés: 55 °C, 2 perc
11. lépés: 60 °C, 2 perc
12. lépés: a 11. lépés 9-szeres ismétlése
13. lépés: 4 °C a végéig
14. lépés: befejezés

B. Az első cDNS-szál szintézisének második módszere

mRNS (10 µg) 25 µl

víz 30 µl

*NotI adapter primer (5 µg) 10 µl

Inkubálás 65 °C-on, 10 percen át, utána jégben lehűtjük és hozzáadjuk a további reagenseket:

*5X első szál puffer 20 µl

*0,1 M DTT 10 µl

*dNTP keverék, 10 mM 5 µl

A reakcióelegyet 45 °C-on inkubáljuk 2 percen át, majd hozzáadunk 10 µl *SuperScript II RT plazmidot (200 U/µl) és elindítjuk az alábbi programot

1. lépés: 45 °C, 6 s, majd -0,1 °C/ciklus

2. lépés: a hőmérséklet csökkentése 99 ciklusban
3. lépés: 35 °C, 5 perc
4. lépés: 45 °C, 60 perc
5. lépés: 50 °C, 10 perc
6. lépés: 4 °C a végéig
7. lépés: befejezés

Az első cDNS-szál szintézise után a DNS-t fenollal, standard módon extraháljuk, majd nátrium-acetáttal és etanollal kicsapjuk és -20 °C-on tároljuk.

C. A mRNS diolcsoportjának oxidálása a biotines jelöléshez

Az első cDNS-szálat „kigombolyítjuk” és 70 %-os etanollal egyszer mossuk. Az üledéket 23,2 µl, dietil-pirokarbonáttal (DPEC) kezelt vízben felfuszpendáljuk és jégre tesszük. Készítünk friss, 100 mmólos nátrium-perjodát-oldatot, majd összemérjük a következő reagenseket:

mRNS : első cDNS-szál (20 µg mRNS-sel kezelve)	46,4 µl
nátrium-perjodát-oldat (100 mM)	2,5 µl
nátrium-acetát-oldat (3 M, pH=4,5)	1,1 µl

A 100 mmólos oldat készítéséhez 21,39 µg nátrium-perjodátot adunk 1 µl vízhez.

A kémcsövet fóliába burkoljuk és jégen inkubáljuk 45 percen át. Az inkubálás után a reakcióelegyet kicsapjuk 10 µl 5 mólos nátrium-klorid-oldat, 0,5 µl 20 %-os SDS-oldat és 61 µl izopropanol elegyében. Jégen inkubáljuk legalább 30 percen át, majd maximális fordulattal centrifugáljuk 4 °C-on, 30 percen át, ami után az üledéket előbb 70, majd 80 %-os etanollal mossuk.

D. A mRNS diolcsoportjának biotinezése

A DNS-t felszuszpendáljuk 110 μ l, DEPC-tal kezelt vízben, majd hozzáadunk 5 μ l 20 %-os SDS-oldatot, 5 μ l 2 mólos nátrium-acetát-oldatot (pH = 6,1) és 300 μ l frissen készített, 10 mmólos biotin-hidrazid-oldatot. A kémcsövet fóliába burkoljuk és szobahőmérsékleten inkubáljuk egy éjszakán át.

E. Kezelés RNáz I-gyel

Kicsapjuk a DNS-t a következő elegyben: 10 μ l 5 M nátrium-klorid-oldat, 75 μ l 2 M nátrium-acetát-oldat (pH = 6,1), 420 μ l biotinezett mRNS:cDNS és 1262,5 μ l (2,5 térfogat) 100 %-os etanol. A kicsapást két csőben hajtjuk végre a fenti térfogatok megfelezésével. A reakcióelegyet -20 °C-on tartjuk legalább 30 percen át. Ezután a DNS-t maximális fordulatszámmal lecentrifugáljuk 4 °C-on, 30 perc alatt, az üledéket 80 %-os etanollal kétszer mossuk, majd feloldjuk a DNS-t 70 μ l RNáz-mentes vízben. A két csövet összeöntjük és kiegészítjük 140 μ l-re. Ezután hozzáadunk 40 μ l (10 U/ μ l) RNáz I-et, 140 μ l első szál cDNS:RNS oldatot és 20 μ l 10X puffert. A reakcióelegyet 37 °C-on inkubáljuk 15 percen át, majd hozzáadunk 5 μ l élesztő tRNS-t (40 μ g/ μ l) minden mintához a cDNS „megfogása” céljából.

F. A teljes hosszúságú első cDNS szál megfogása

A gyöngyöket élesztő tRNS-sel blokkoljuk: 1 ml gyöngyhez 5 μ l élesztő tRNS oldatot (40 μ g/ μ l) adunk. Az elegyet 30 percen át jégen keverjük, majd háromszor mossuk a következő oldattal: 1-1 ml 2 M nátrium-klorid 50 mM EDTA-val, pH = 8,0.

Ezután a gyöngyöket felszuszpendáljuk az utóbbi oldat 800 μ l-ében, hozzáadjuk a 200 μ l RNáz I-gyel kezelt mintát, és az elegyet szobahőmérsékleten inkubáljuk 30 percen át. A gyöngyöket

mágneses állvánnyal megfogjuk, a felülúszót eltesszük és elkezdjük az alábbi mosást:

kétszer 1 ml 2 M nátrium-klorid-oldat, 50 mM EDTA, pH = 8,0;
egyszer 1 ml 0,4 %-os SDS-oldat 50 µg/ml tRNS-sel;
egyszer 1 ml 10 mM Tris-HCl (pH = 7,5), 0,2 mM EDTA, 10 mM nátrium-klorid és 20 %-os glicerin elegye;
egyszer 1 ml 50 µg/ml-es tRNS oldat; és
egyszer 1 ml első szál puffer.

G. A második cDNS szál szintézise

A gyöngyöket felszuszpendáljuk az alábbi elegyben:

*5X első puffer	8 µl
*0,1 M DTT-oldat	4 µl
*10 mM dNTP keverék	8 µl
*5X második puffer	60 µl
* <i>E. coli</i> ligáz, 10 U/µl	2 µl
* <i>E. coli</i> DNS-polimeráz, 10 U/µl	8 µl
* <i>E. coli</i> RNáz H, 2 U/µl	2 µl
³² P-dCTP, 10 µCi/µl vagy víz	2 µl
víz	208 µl

A reakcióelegyet 16 °C-on inkubáljuk 2 órán át, miközben minden 30. percben megkeverjük. Ezután hozzáadunk 4 µl T4 DNS-polimeráz oldatot és még 5 percen át inkubáljuk 16 °C-on.

A második cDNS szálak elúciójához a gyöngyöket mágneses állványon immobilizáljuk, majd 200 µl vízben reszuszpendáljuk, újra elválasztjuk és egyesítjük a mintákat (körülbelül 500 µl). A gyöngyökhöz 200 µl vizet adunk, utána 200 µl fenol/kloroform elegyet, erélyesen összekeverjük, majd centrifugáljuk, hogy elvá-

lasszuk a mintát a fenoltól.

Egyesítjük a DNS-mintákat (körülbelül 700 μl) és fenollal újra tisztítjuk a DNS-t, majd kicsapjuk a következők hozzáadásával: 2 μg glikogén, 0,5 térfogat 7,5 M ammónium-acetát-oldat és 2 térfogat 100 %-os etanol. A kicsapódáshoz állni hagyjuk egy éjszakán át. Az üledéket lecentrifugáljuk, mossuk 70 %-os etanollal és levegőn megszárítjuk.

H. A *SalI* adapter ligálása

Az üledéket 26 μl vízben felfuszpendáljuk és 1 μl -t a Tris-acetát-EDTA (TAE) gélhez használunk.

Összeállítjuk az alábbi reakcióelegyet:

második cDNS szál	25 μl
*5X T4 DNS-ligáz puffer	10 μl
*SalI adapterek	10 μl
*T4 DNS-ligáz	5 μl

Finoman összekeverjük és 16 °C-on inkubáljuk egy éjszakán át. Adott esetben a második napon hozzáadunk még 2 μl ligázt és szobahőmérsékleten inkubáljuk további 2 órán át.

A reakcióelegyhez 50 μl vizet adunk, a DNS-t 100 μl fenollal megtisztítjuk, a felső fázisból 90 μl -t átviszünk egy új csőbe és kicsapjuk az alábbi elegyben.

glikogén-oldat 1 $\mu\text{g}/\mu\text{l}$	2 μl
felső fázis DNS	90 μl
7,5 M ammónium-acetát-oldat	50 μl
100 %-os etanol	300 μl

A kicsapás -20 °C-on, egy éjszakán át történik. Az üledéket 4 °C-on lecentrifugáljuk, 70 %-os etanollal mossuk és megszárítjuk.

I. **Emésztés NotI-gyel**

Összemérjük az alábbi elegyet:

második DNS szál	41 µl
*3. reakciópuffer	5 µl
*NotI, 15 U/µl	4 µl

Finoman összekeverjük és 37 °C-on inkubáljuk 2 órán át. Ezután hozzáadunk 50 µl vizet és 100 µl fenolt, keverjük, a felső fázisból 90 µl-t mérünk egy új csőbe, majd hozzáadunk 50 µl ammónium-acetát-oldatot és 300 µl etanol. A kicsapást -20 °C-on, egy éjszakán át folytatjuk.

A klónozást, ligálást és transzformálást a SuperScript cDNS szintézis készlettel végezzük.

3. példa:

A cDNS szekvenálása és a könyvtár szubsztrahálása

Egyedi telepeket veszünk fel és a DNS-t vagy PCR-val (M13 „előre” és „fordított” primerekkel), vagy plazmid-izolálással kinyerjük. A cDNS-klónokat M13 fordított primerekkel szekvenálhatjuk.

A cDNS-könyvtárakat 22 x 22 cm-es agarlemezekre szélesztjük körülbelül 3000 telep/lemez sűrűségben. A lemezeket 37 °C-on inkubáljuk 12-24 órán keresztül. A telepeket 384 lyukas lemezekbe szedjük fel egy telepfelszedő robottal (Q-bot; GENETIX Ltd.). Ezeket a lemezeket 37 °C-on inkubáljuk egy éjszakán át. Amikor a kellő számú telepet felszedtük, a Q-bot 22 x 22 cm-es nejlon membránokra „tűzi le” azokat. Egy-egy membrán 9216 vagy 36864 telepet tartalmaz. Ezeket a membránokat a megfelelő antibiotikumot tartalmazó agarlemezekre simítjuk és 37 °C-on inkubáljuk egy éjszakán át.

Miután a telepek kinőttek, a membránokat a második napon levesszük, denaturáló oldattal átitatott szűrőpapírra fektetjük 4 percre, majd egy forró vízfürdő fölött tartjuk újabb 4 pecen át. Ezután közömbösítő oldattal átitatott szűrőpapírra tesszük a membránokat 4 percre, az oldat feleslegét száraz szűrőpapírral (1 perc) eltávolítjuk, majd a membrán telepekkel borított oldalát „Proteinase K” oldatba mártjuk és 37 °C-on inkubáljuk 40-50 percen át. Ezután a membránokat száraz szűrőpapírra fektetjük és egy éjszakán át szárítjuk, majd a DNS-t UV fénnel végzett kezeléssel a nejlon membránhoz rögzítjük.

A telepek hibridizálását Sambrook, Fritsch és Maniatis leírása szerint végezzük (in „*Molecular Cloning: A Laboratory Manual*”, 2. kiadás). Ehhez az alábbi próbákat használjuk:

1. első szál cDNS ugyanabból a szövetből, amelyből a könyvtár készült, a legnagyobb mennyiségben jelen lévő klónok eltávolítására;
2. az első szekvenálási adatok alapján leggyakoribb 48-192 cDNS-klón ugyanabból a könyvtárból;
3. a leggyakoribb 192 cDNS-klón a teljes kukorica szekvenancia-adatbázisból;
4. egy Sal-A20 oligonukleotid (TCG ACC CAC GCG TCC GAA AAA AAA AAA AAA AAA AAA), ami eltávolítja a poliA farokkal rendelkező, de nem cDNS molekulákat; és
5. az rRNS-ből származó cDNS-klónok.

Az autoradiográfiás képet számítógépbe olvassuk, majd kiértékeljük a jelek intenzitását és megállapítjuk a hideg telepek címét. A hideg telepeket a 384 lyukas lemezekből 96 lyukas leme-

zekbe oltjuk át a Q-bot segítségével.

4. példa:

Gén azonosítása számítógépes homológia-kereséssel

Gének azonosságát a BLAST programmal vizsgáljuk [Basic Local Alignment Search Tool; Altschul et al., *J. Mol. Biol.* 215, 403-410 (1993)] az alapbeállításokat használva. A hasonló szekvenciákat a BLAST „nr” adatbázisban keressük, ami tartalmazza az összes nem redundáns CDS-transzlátumot, a Brookhaven Protein Data Bank háromdimenziós szerkezeteiből levezetett szekvenciákat, a SWISS-PROT fehérje-szekvencia adatbázis legújabb kiadását, valamint az EMBL és DDBJ adatbázisok adatait. A cDNS-szekvenciákat összehasonlítottuk az összes hozzáférhető DNS-szekvenciával, ami az „nr” adatbázisban fellelhető; ehhez a BLASTN algoritmust használtuk. A DNS-szekvenciákat minden leolvasási keretben lefordítottuk és összehasonlítottuk az összes hozzáférhető fehérje-szekvenciával, amihez a BLASTX algoritmust használtuk [Gish és States, *Nature Genetics* 3, 266-272 (1993)]. Néhány esetben két vagy több klón szekvenálásából származó adatokat – amelyek átfedő DNS-szakaszokat tartalmaznak – használhatunk fel a folytonos DNS-szekvenciák megszerkesztéséhez.

A szekvenciák egyeztetését és a százalékos hasonlóság kiszámítását a LASERGENE programcsalád Megalign programjával (DNA-STAR, In., Madison, Wi.) végezzük. Szekvenciák többszörös összehasonlítása a CLUSTAL illesztőprogrammal (Higgins és Sharp, *id. közl.*), alapbeállítási paraméterekkel (hézag-büntetőpont = 10, hézaghossz-büntetőpont = 10) is elvégezhető. Páros összehasonlítás esetében a CLUSTAL alapbeállítási paraméterei a következők:

KTUPLE = 1, hézag-büntetőpont = 3, ablak = 5 és megőrzött átlók = 5.

5. példa:

A transzgének expressziója egyszikűek sejtjeiben

Elkészítünk egy transzgén szerkezetet, ami egy találmány szerinti polipeptid cDNS-ét tartalmazza szensz orientációban a kukorica 27 kD zein promoteréhez képest – ami a cDNS-fragmentum 5' végénél helyezkedik el –, és a 10 kD zein 3' végét, ami a cDNS-fragmentum 3' végénél helyezkedik el. A gént képviselő cDNS-fragmentumot polimeráz láncreakcióval (PCR) készítjük el a cDNS-klónból megfelelő oligonukleotid primerek alkalmazásával. Klónozó helyek (*NcoI* vagy *SmaI*) építhetők be az oligonukleotidokba, hogy biztosítsuk a DNS-fragmentum megfelelő irányultságát, amikor beépítjük az emésztett pML103 vektorba az alább leírt módon. A sokszorozást ezután végezzük le szokványos PCR-val. A sokszorozott DNS-t a *NcoI* és *SmaI* restriktációs enzimekkel emésztjük és agarózgélen frakcionáljuk. A megfelelő csíkot izoláljuk a gélből és egyesítjük a pML103 plazmid 4,9 kbp *NcoI*-*SmaI* fragmentumával. A pML103 plazmidot a Budapesti Egyezménynek megfelelően letétbe helyeztük az American Type Culture Collection-nál, ahol nyilvántartási száma ATCC 97366. A pML103 DNS-szegmentuma tartalmaz egy 1,05 kbp *SalI*/*NcoI* promoter-fragmentumot a kukorica 27 kD zein génjéből, és egy 0,96 kbp *SmaI*-*SalI* fragmentumot a kukorica 10 kD zein génje 3' végéből (utóbbi a Promega pGem9Zf(+)) vektorából származik). A vektort és a DNS-inszertumot 15 °C-on ligáljuk egy éjszakán át, lényegében Maniatis leírása szerint. Ezután a ligált DNS-sel az *E. coli* XL1-Blue törzset (Epicurian Coli XL-1 Blue; Stratagene) transzformáljuk. A transz-

formált baktériumok szűrővizsgálatát a plazmid-DNS restrikciós enzim emésztésével és korlátozott nukleotid-szekvencia analízissel végezzük (Sequenase DNA Sequencing Kit; U.S. Biochemical; didezoxi-láncterminációs eljárás). Az így kapott plazmid-szerkezet egy transzgén, amely 5' -> 3' irányban haladva a kukorica 27 kD zeint promoterét, a találmány szerinti polipeptidet kódoló cDNS-fragmentumot és a 10 kD zein 3'-területét tartalmazza.

A fent leírt transzgén kukoricasejtekbe juttatható az alábbi eljárással. A H99 és LH132 beltenyésztett kukorica-vonalak keresztezéséből származó, éretlen embriókat kimetsszük a magkezdeményből a megporzás utáni 10-11. napon, amikor 1,0-1,5 mm hosszúak. Az embriókat a tengelyi oldalukkal lefelé agarózzal szilárdított N6 táptalajra fektetjük [Chu et al., *Sci. Sin. Peking* 18, 659-668 (1975)] és sötétben, 27 °C-on tartjuk. Az éretlen embriók szkutellumából kisarjadó szuszpenzor-struktúrákon morzsalékony embriogén kallusz jelenik meg, amiben szomatikus proembrioid és embrioid (embrió-képző) sejtek vannak a differenciálatlan sejtek halmazai között. Az izolált embriogén kalluszokat N6 táptalajon tenyésztjük és minden 2-3. héten osztjuk tovább.

A transzkripciós kísérletekhez a p35S/Ac (Hoechst Ag, Frankfurt, NSZK) vagy más, ezzel egyenértékű plazmidot használhatunk, hogy biztosítsuk a szelektálható markert. Ez a plazmid a *Pat* gént tartalmazza, ami a foszfinotricin-acetil-transzferázt (PAT) kódolja (EP 0 242 236). A PAT enzim a glutamin-szintetáz inhibitor herbicidek, például a foszfinotricin ellen biztosít rezisztenciát. A p35S/Ac plazmidban a *Pat* gén a karfiolmozaik vírus 35S promoterének [Odell et al., *Nature* 313, 810-812 (1985)] és

az *Agrobacterium tumefaciens* Ti plazmidjának T-DNS-éből eredő nopalin-szintetáz gén 3'-régiójának szabályozó hatása alatt áll.

Gének bejuttatására a kallusztényészet sejtjeibe a részecske-bombázást használhatjuk [Klein et al., *Nature* 317, 70-73 (1987)]. Ehhez arany szemcséket (1 μm átmérő) vonunk be DNS-sel az alábbi módon. A plazmid-DNS-ből 10 μg -ot adunk 50 μl arany szemcse-szuszpenzióhoz (60 mg/ml). Az elegyhez 50 μl 2,5 M kalcium-klorid-oldatot és 20 μl 1,0 M spermidin-oldatot (szabad bázis) adunk. Az oldatok hozzáadása közben a szuszpenziót erélyesen keverjük. Tíz perc múlva a csöveket röviden centrifugáljuk (15000 f/perc, 5 s) és a felülúszót eltávolítjuk. A szemcséket 200 μl vízmentes etanolban felfuszpendáljuk, újra centrifugáljuk és a felülúszót eltávolítjuk. Az etanolos öblítést megismételjük, majd a szemcséket 30 μl etanolban felfuszpendáljuk. A DNS-sel bevont arany szemcsék szuszpenziójából 5 μl -t teszünk egy Kapton repülő korong (Bio-Rad Labs.) közepére. Ezután a szemcséket a kukoricaszövetbe lőjük egy Biolistic PDS-1000/He készülékkel (Bio-Rad Instruments, Hercules, Ca.), 6880 kPa hélium-nyomást, 0,5 cm réstávolságot és 1,0 cm repülési távolságot használva.

A belövéshez az embriogén szövetet szűrőpapírra tesszük agarózzal szilárdított N6 táptalajon. A szövetet vékony gyeppe rendezzük el ami egy körülbelül 5 cm átmérőjű, kör alakú területet fed be. A szövetet tartalmazó Petri-csészét a készülék kamrájába tesszük körülbelül 8 cm-re a fogóhálótól. A kamra légnyomását 94,9 kPa-ra csökkentjük. A makrohordozót a hélium lökeshulláma gyorsítja fel, ami akkor lép fel, amikor a gáz nyomása a kilövő-csőben meghaladja a 6880 kPa-t és a szakadómembrán átszakad.

Hét nappal a belövés után a szövetet átrakjuk olyan N6 táptalajra, amely 2 mg/l glufozinátot tartalmaz, és nincs benne ka-zein vagy prolin. A szövet lassan tovább nő ezen a táptalajon. További 2 hét múlva a szövetet friss, glufozinátot tartalmazó N6 táptalajra oltjuk át. Hat hét múlva körülbelül 1 cm átmérőjű, aktívan növekvő kalluszkok találhatók a glufozináttal kiegészített táptalajt tartalmazó lemezek néhányán. Ezek a kalluszkok képesek tovább növekedni a szelektív táptalajon.

A transzgén kalluszkokból úgy regenerálhatók növények, hogy a szövetcsomókat először 0,2 mg/l 2,4D-vel kiegészített N6 táptalajra, majd két hét múlva regeneráló táptalajra tesszük át [Fromm et al., *Bio/Technology* 8, 833-839 (1990)].

6. példa:

A transzgének expressziója kétszikűek sejtjeiben

Egy magspecifikus expressziós keretet – amely a *Phaseolus vulgaris* fazeolin raktározott magfehérjéjének β alegységét kódoló gén promoterét és transzkripciós terminátorát tartalmazza [Doyle et al., *J. Biol. Chem.* 261, 9228-9238 (1986)] – használhatunk a találmány szerinti polipeptidek kifejeztetéséhez szójababban. A fazeolin-keret körülbelül 500 nukleotidot tartalmaz a transzlációs kezdő kodontól felfelé (5' irányban) és körülbelül 1650 nukleotidot a fazeolin stop-kodonjától lefelé (3' irányban). Az 5'- és 3'-régiók között vannak az egyedi restrikciós endonukleáz hasítási helyek: az *NcoI* (amely magában foglalja az ATG transzlációs kezdő kodont), az *SmaI*, a *KpnI* és az *XbaI*. A teljes keretet *HindIII* restrikciós helyek határolják.

A gén cDNS-fragmentumát PCR-val állítjuk elő a cDNS-klónból

megfelelő oligonukleotid primerek használatával. Az oligonukleotidokba klónozó helyeket építhetünk be, hogy biztosítsuk a DNS-fragmentum helyes orientációját, amikor beépítjük az expressziós vektorba. A sokszorozást ezután végezzük el a fent leírt módon, majd az izolált fragmentumot egy pUC18 vektorba inszertáljuk, ami a magspecifikus expressziós keretet hordozza.

A találmány szerinti polipeptidet kódoló szekvenciát tartalmazó expressziós vektort használhatjuk szójaembriók transzformálására. Szomatikus embriók indukálása céljából 3-5 mm hosszú sziklevél-darabokat metszünk ki az A2872 szójafajta sterilizált felszínű, érett magvaiból, és világosban vagy sötétben, 26 °C-on tenyésztjük azokat egy megfelelő agar-táptalajon, 6-10 héten át. A másodlagos embriókat leválasztjuk és egy alkalmas tápfolyadékba tesszük át. A szomatikus embrió-csomókból ismételt szelektálással kiválasztjuk azokat, amelyek korai, globuláris stádiumú embriókként szaporodnak, és szuszpenzióban tartjuk fenn azokat az alább leírt módon.

A szója embriogén szuszpenzió-tenyészeteket 35 ml tápfolyadékban, körkörös rázógépen (150 f/perc), 16 °C-on, fluoreszcens lámpával megvilágítva, 16:8 órás világos/sötét periódus mellett tartjuk fenn. A tenyészetekből minden második héten körülbelül 35 mg szövetet oltunk át 35 ml friss tápfolyadékba.

A szója embriogén szuszpenzió-tenyészetek részecske-bombázásos módszerrel transzformálhatók (Klein et al., *id. közl.*; US 4,945,050). Egy hélium meghajtású Biolistic PDS1000/HE készüléket használhatunk erre a célra.

A szója transzformálásának elősegítésére használható, sze-

lektálható markergén egy transzgénben van, ami a CaMV 35S promoterét, a pJR225 plazmidból vett higromcin-foszfotranszferáz gént [*E. coli*-ból; Gritz et al., *Gene* 25, 179-188 (1983)], és az *Agrobacterium tumefaciens* Ti plazmidja nopalin-szintetáz génjének 3' régióját tartalmazza. A magspecifikus expressziós keret tartalmazza a fazeolin-gén 5'-régióját, a találmány szerinti polipeptidet kódoló fragmentumot és a restrikciós fragmentumként izolált fazeolin 3'-régiót. Ezt a keretet inszertálhatjuk a markergént hordozó vektor egy egyedi restrikciós helyére.

Egy aranyszemcse-szuszpenzió (1 μm átmérőjű szemcsék, 60 mg/ml) 50 μl -hez sorban hozzáadunk 5 μl DNS-t (1 $\mu\text{g}/\mu\text{l}$), 20 μl 0,1 M spermidin-oldatot és 50 μl 2,5 M kalcium-klorid-oldatot. Az elegyet 3 percen át keverjük, majd 10 s-ig centrifugáljuk és a felülúszót eltávolítjuk. A DNS-sel bevont szemcséket ezután egyszerűen mossuk 400 μl 70 %-os etanollal, majd 40 μl vízmentes etanolban reszuszpendáljuk és háromszor 1 s-ig szonikáljuk. A DNS-sel bevont aranyszemcse-szuszpenzióból 5 μl -t teszünk egy makrohordozó-korongra.

Egy kéthetes szuszpenzió-tenyészetből körülbelül 300-400 mg szövetet teszünk egy üres, 60x15 mm-es Petri-csészébe és a felesleges folyadékot pipettával eltávolítjuk. Az egyes transzformációs kísérletekben rendszeren körülbelül 5-10 lemez szövetet bombázunk. A membránt átszakító nyomást 7568 kPa-ra állítjuk be, a kamrában a nyomást 94,9 kPa-ra csökkentjük. A szövetet a fogóhálótól körülbelül 9 cm-re tesszük és minden mintát háromszor bombázunk. A belővéseket követően a szövetet kétfelé osztjuk, friss tápfolyadékba tesszük és a fent leírt módon tenyésztjük tovább.

A kezelést követő 5-7. napon a tápfolyadékot frissre cseréljük, majd a 11-12. napon újra friss, de 50 mg/ml higromicint tartalmazó tápfolyadékra. Ezt a szelektív tápfolyadékot hetente frissítjük. A kezelés után 7-8 héttel zöld, transzformált szövet növekedését figyelhetjük meg a transzformálatlan, elhaló szövet-csomók között. Az elkülönült, zöld szövetcsomókat kiemeljük és egyedi palackokba oltjuk át, ahol új, klónként fenntartott, transzformált embriogén szuszpenzió-tenyészetekké lesznek. Minden új vonalat független transzkripciós eseményként kezelünk. Ezeket a szuszpenziókat a továbbiakban éretlen embriók tenyészeteként tartjuk fenn, vagy teljes növényeket regenerálunk belőlük.

7. példa:

A transzgének expressziója mikroorganizmusokban

A találmány szerinti polipeptideket kódoló cDNS-ek az *E. coli* T7 pBT430 expressziós vektorába inszertálhatók. Ez a vektor a pET-3a [Rosenberg et al., *Gene* 56, 125-135 (1987)] származéka, amiben a T7 bakteriofág RNS-polimeráz/T7 promoter rendszerét alkalmazzák. A pBT430 plazmidot úgy készítjük el, hogy először tönkretesszük az EcoRI és HindIII helyeket a pET-31 plazmidban, az eredeti pozíciókban. Ezután EcoRI és HindIII helyeket tartalmazó oligonukleotid-adaptert építünk be a pET-31 BamHI helyére. Ezzel létrejött a pET-3aM vektor, amely járulékos egyedi klónozó helyeket tartalmaz a gének beillesztése céljára. Ekkor a transzláció kezdőhelyénél lévő NdeI helyet irányított nutagenezissel átalakítjuk egy NcoI helyé, azaz a pET-3aM DNS-szekvenciáját ezen a területen (5'-CATATGG) átalakítjuk a pBT430 5'-CCCATGG szekvenciájává.

A cDNS-t tartalmazó plazmid-DNS-t megfelelő módon emésztjük, hogy szabaddá tegyük a fehérjét kódoló nukleinsav-fragmentumot. A fragmentumot 1 %-os NuSieve GTG alacsony olvadáspontú agarózgélben (FMC) tisztíthatjuk. A puffer és az agarózgél 10 µg/ml etídiu-bromidot tartalmaz a DNS-fragmentum láthatóvá tételére. A fragmentumot az agarózgéltől GELase (Epicentre Technologies) enzim es emésztéssel tisztítjuk meg a gyártó receptje szerint, majd etanollal kicsapjuk, megszáritjuk és 20 µl vízben felszuszpendáljuk. Megfelelő oligonukleotid adaptert csatolhatunk a fragmentumhoz T4 DNS-ligázzal (New-England Biolabs, Beverly, Ma.). A csatolt adaptereket tartalmazó fragmentumot megtisztítjuk az adapterek feleslegétől alacsony olvadáspontú agarózgélben, a fent leírt módon. A pBT430 vektort emésztjük, alkalikus foszfátázzal (NEB) defoszforiláljuk és fenol/kloroform eleggyel fehérjementesítjük a fent leírt módon. Az előkészített pBT430 vektort és a fragmentumot összekapcsoljuk (16 °C, 15 óra), majd DH5 elektrokompetens sejteket (Gibco BRL) transzformálunk vele. A transzformánsokat 100 µg/ml ampicillint tartalmazó LB táptalajon szelektáljuk. A találmány szerinti polipeptideket kódoló géneket tartalmazó transzformánsokban restrikciós enzim es analízissel ellenőrizzuk, hogy a gén megfelelő irányultságú-e a T7 promoterhez képest.

A magas szintű expresszió eléréséhez egy megfelelő orientációjú plazmid-klónnal transzformáljuk az *E. coli* BL 21(DE3) törzsét [Studier et al., *J. Mol. Biol.* 189, 113-130 (1986)]. A baktériumot 100 mg/l ampicillint tartalmazó LB tápfolyadékban, 25 °C-on tenyésztjük. Amikor a tenyészet 600 nm-nél mért optikai



denzitása körülbelül 1, izopropil-tio- β -galaktozidot (IPTG, az induktor) adunk hozzá 0,4 mM végkoncentrációban, és az inkubálást még 3 órán át folytatjuk 25 °C-on. Ezután a sejteket centrifugálással kinyerjük és 50 μ l 50 mM Tris-HCl pufferben (pH = 8,0; 0,1 mM DTT-vel és 0,2 mM fenil-metilszulfonil-fluoriddal kiegészítve) felfuszpendáljuk. Kis mennyiségű, 1 mm-es üveggyöngyöt adunk hozzá és az elegyet háromszor körülbelül 5 s-ig szonikáljuk egy mikrofejes ultrahang-generátorral. Ezután az elegyet lecentrifugáljuk és a felülúszó fehérje-tartalmát meghatározzuk. A tenyészet oldható frakciójából 1 μ g fehérjét SDS-poliakrilamid gélen, elektroforézissel elválasztunk, és a gélen megfigyeljük az elvárt molekulatömegű fehérjecsík mozgását.

8. példa:

A *Sus1* génbe inszertált *Mutator* (*Mu*) hatásának meghatározása

Ez a példa a szacharóz-szintetáz génbe inszertált *Mu* elemet tartalmazó növények azonosításának, valamint a mutáns növények fenotípusos és biokémiai vizsgálatának módszerét írja le.

A „Trait Utility System for Corn” (TUSC; US 5,962,764 számú szabadalmi irat) egy eljárás, amely genetikai és molekuláris technikákat alkalmaz a gének funkcióinak vizsgálatára kukoricában. A génfunkció vizsgálata feltételezi, hogy a gén szekvenciája már ismert, így az eljárás fordítva működik: a szekvenciától a fenotípus felé. Ezt a megközelítést „fordított genetiká”-nak nevezik szemben az „egyenes” eljárásokkal, amelyek célja az adott tulajdonságért (fenotípusért) felelős gén azonosítása és izolálása.

A Pioneer Hi-Breed International, Inc.-nek van egy bejegyzett kukorica genom-DNS gyűjteménye, amely körülbelül 42.000

egyedi F_1 növényből áll [Meeley és Briggs, *Maize Genet. Coop. Newslett.* 69, 82 (1995)]. A növényegyedek mindegyikének genomja több másolatban tartalmazza a *Mutator* (*Mu*) transzpozabilis elem-családot. A *Mu* család erősen mutagén; az *Mu*-DR aktív elem jelenlétében az elemek genomszerte áthelyeződnek, génterületekbe inszertálódnak és gyakran tönkreteszik a gén működését. A genom-DNS-t nagy számú egyedből begyűjtve a Pioneer összeállította mutagenizált kukorica-genom könyvtárát. A *Mu* inszertálódásának eseménye a legtöbb esetben heterozigóta, és mivel a legtöbb inszerciós mutáció recesszív természetű, az F_1 növények vad típusúnak látszanak. Az összes F_1 növényről önbeporzással hozzuk létre az F_2 magvakat. Az F_2 nemzedékben az inszerciós mutációk a mendeli szabályokat követve jelennek meg, és ez lehetőséget ad a mutáns allél fenotípusra gyakorolt hatásának felismerésére. A TUSC rendszert számos laboratóriumban eredményesen használják egy sor gén funkciójának azonosítására [a kukorica *An1* génjének klónozása és jellemzése: Bensen et al., *Plant Cell* 7, 75-84 (1995); a C-funkció aktivitásának szétterjedése a kukorica virágfejlődése során: Mena et al., *Science* 274, 1537-1540 (1996); a kémiai védekező mechanizmusok vizsgálata fűfélékben: Frey et al., *Science* 277, 696-699 (1997); a kukorica-címer merisztémája fejlődésének szabályozása az *APETALA2*-szerű gén által: Chuck et al., *Genes Develop.* 12, 1145-1154 (1998); egy *SecY* homológra van szükség a kloroplaszt tilakoid-membrán kialakulásához és a normális kloroplaszt-gén expresszióhoz: Roy és Barkan, *J. Cell. Biol.* 141, 1-11 (1998)].

A *Mu*-inszertumok szűrővizsgálata PCR-val a *Sus1*-ben a követ-

kező.

Két primert terveztünk a *Sus1* cDNS-en belül és génspecifikus primereknek (GSP) neveztük azokat:

„egyenes” primer (GSP1): 5'-ACGGAATCGTTCGCAAGTGGATCTC-3' (8. számú szekvenciavázlat);

„fordított” primer (GSP2): 5'-GATGATTGGCTTGTTCTGTCGTTTC-3' (9. számú szekvenciavázlat).

Ezek a primerek körülbelül 1 kbp távolságra vannak a *Sus1* genom-szekvenciájától.

Mu TIR primer: 5'-AGA GAA GCC AAC GCC AWC GCC TCY ATT TCG TC-3'

A „Pickoligo” programot használjuk a primerek kiválasztásához. Ez a program az alábbi egyenlettel határozza meg a T_m hőmérsékletet:

$$T_m = [\{ (GC*3 + AT*2) * 37 - 562 \} / \text{hosszúság}] - 5$$

A PCR 62 °C összeolvasztási hőmérsékleten és az alábbi termociklussal fut:

94 °C, 2 perc (kezdeti denaturálás)
/ 94 °C, 30 s - 1 perc
35 ciklus 62 °C, 30 s - 2 perc
 \ 72 °C, 1-3 perc
 72 °C, 5 perc (végső meghosszabbítás)

A PCR termékeinek gélelektroforézisével kimutattuk, hogy nem volt hibás párosodás a primerek reakcióiban, és csak egyetlen fragmentum sokszorozódott a párosított GSP-reakciókban. A 42.000 növény genom-DNS-ét csoportokban egyesítjük, amelyek mindegyikében 48 növény van. A csoportokat a GSP1- vagy GSP2-vel és a Mu

TIR-rel sokszorozzuk. Azokat a csoportokat, amelyek pozitív eredményt adtak a *Sus1* cDNS-sel, mint próbával, tovább vizsgáltuk, hogy meghatározzuk a sokszorozódó fragmentum méretét. Azokat a csoportokat, amelyekben tiszta fragmentumok voltak azonosíthatók, még tovább vizsgáltuk, hogy azonosítsuk azt a növényegyet, amelyből az *Mu*-inszertum(ok) ered(nek).

Az ilyen módon azonosított F_1 növények magvait szabadföldön elvetjük. Minden F_2 sorban 20 növényről levélkorongokat szedünk és izoláljuk a genom-DNS-t. Ugyanezt a 20 növényt önbeporozzuk és F_3 magvakat nyerünk róluk. A 12 soronként egyesített DNS-t (20-20 növényből) PCR-val vizsgáljuk a GSP1 vagy GSP2 és a *Mu* TIR primerek alkalmazásával, a fent leírt módon. Három csoportban, amelyek *Mu* inszertumokat tartalmaznak, tovább vizsgáltuk az egyes növényeket és azonosítottuk a homozigótákat. A PCR-val sokszorozott fragmentumokat a TOPO vektorba (Invitrogen) klónoztuk és szekvenáltuk. A *Mu* inszerciós helyeket úgy határoztuk meg, hogy összevetettük a *Sus1* megkapott szekvenciáját a *Mu* szekvenciákkal; az eredményt a 6. ábrán mutatjuk be. A két inszertum 3 nukleotid távolságra van egymástól a 12. exonnak megfelelő nyílt leolvasási keretben, ami arra utal, hogy a génnek ez a területe egy „forró pontot” képviselhet a *Mu* inszerció számára.

A homozigóta mutáns növények és vad típusú leszármazottaik szárának a cső alatti két izközét begyűjtöttük a betakarítás előtt két héttel. A megszárított szárdarabokat megőröltük, majd megmértük azok cellulóz- és lignin-koncentrációját. A mutáns növények cellulóz-koncentrációja 30 %-kal kisebb volt a vad típusú leszármazottakénál, ha az eredményt a teljes szárazanyagra vonat-

koztatva adtuk meg, és 6 %-kal volt kisebb, ha a szerkezeti szárazanyagot vettük alapul (7. ábra). A szerkezeti szárazanyag-tartalom jelentős csökkenését is megfigyeltük a mutáns növényekben. Ez összhangban van azzal a feltevással, hogy a szacharózsztetetáz működéséből származó UDP-glükóz jelentős szerepet játszik a cellulóz bioszintézisében. Felúnt az is, hogy a cellulóztermelés csökkenése hátrányosan befolyásolja a sejtfal képződését.

9. példa:

A szacharóz-sztetetáz aminosav-szekvenciák összehasonlítása

Az összehasonlítást az AlignX programmal (Vector NTI) végeztük el. Az SH1 és a SUS1 körülbelül 80 %-ban azonosak és körülbelül 90 %-ban hasonlók egymáshoz. A SUS3 körülbelül 70 %-ban azonos és körülbelül 80 %-ban hasonló mind az SH1-hez, mind a SUS1-hez (8. ábra). Elég különbség létezik mind az SH1-hez, mind a SUS1-hez képest ahhoz, hogy a SUS3-at más fehérjének tekintsük. Mivel az 1. kromoszóma rövid karját a 9. kromoszóma hosszú karja duplikátumának tartják, a *Sus3* elhelyezkedése (bin 1.04) alapján feltételezhető, hogy a *Sus1* (térképpontja 9.05) leszármazottja. A homológia vizsgálata alapján azonban úgy tűnik, hogy az *Sh1*-től és a *Sus1*-től függetlenül fejlődött ki. A szacharózsztetetáz gének evolúciója kromoszóma-függőnek tűnik, mert az *Sh1* és a *Sus1* – mindkettő a 9. kromoszómán van – lényegesen nagyobb fokban hasonlítanak egymáshoz, mint akármelyikük a *Sus3*-ra, holott a *Sus3* látszólag a *Sus1* duplikátuma.

Az 1. számú szekvenciavázlatból úgy látszik, hogy a *Sus3*-ból 5-10 aminosav hiányzik az N-terminális végből. A kissé megcsonkított *Sus3*, ami 802 aminosav hosszúságú, számított molekulatömege

91 kDa. A molekulatömeg 0,5-1 kDa-nal emelkedhet, ha egyszer a teljes hosszúságú cDNS-t sikerül izolálni. Az SH1 és a SUS1 számított molekulatömege rendre 91,7 és 92,8 kDa. A SUS3, SH1 és SUS1 enzimek izoelektromos pontjai (pI) rendre 6,07, 5,96 és 6,04.

10. példa:

A kukorica szacharóz-szintetáz polinukleotidok összehasonlítása

Az összehasonlítást az AlignX programmal végeztük el. Az *Sh1* és a *Sus1* 67 %-ban hasonlók; mindkét gén 60 %-ban hasonlít a *Sus3*-ra (9. ábra). A hasonlóság terén ugyanolyan viszonyok figyelhetők meg, mint az aminosav-szekvenciák esetében (8. ábra és 9. példa).

11. példa:

A levezetett teljes hosszúságú *Sus3* polinukleotid és az általa kódolt polipeptid: az 1. számú szekvenciavázlat és a 13. számú szekvenciavázlat szerinti cirok-szekvencia felhasználása az N-terminális vég kiegészítésére

Egy natív, a teljes hosszúságú *Sus3*-at kódoló, teljes hosszúságú *Zea mays* cDNS szekvenálására még nem volt lehetőségünk, mert a *Sus3* igen alacsony szinten fejeződik ki a kukoricában, és ennek megfelelően igen kis arányban fordul elő a kukorica cDNS-könyvtárakban. Azonban egy körülbelül 345 nukleotidból álló cirok-EST (13. számú szekvenciavázlat, génbanki azonosítója BF481989) 3' vége igen magas fokú homológiát mutat az 1. számú szekvenciavázlat szerinti szekvencia 5' végével. A két szekvencia összehasonlításával (10. és 11. ábrák) lehetségessé vált az első metionint kódoló ATG kodon azonosítása a 13. számú szekven-

ciavázlat szerinti szekvencián, és meghatározni azt a nyílt leolvasási keretet, ami átmegy az illesztett, 1. számú szekvenciavázlat szerinti szekvenciába. A cirok-szekvencia e kis szegmensét a *Sus3* cDNS-ének csaknem teljes hosszúságú szekvenciájába inszertálva kapjuk meg a *Sus3* levezetett teljes hosszúságú polinukleotid-szekvenciáját (11. számú szekvenciavázlat), ami a teljes hosszúságú *Sus3* polipeptidet (12. számú szekvenciavázlat) kódolja. Így tehát ezt a cirok-kukorica hibrid szekvenciát használtuk a levezetett *Sus3* fehérje (12. számú szekvenciavázlat) N-terminális végének kiegészítésére.

12. példa:

A teljes hosszúságú *Zea mays* *Sus3* polinukleotid meghatározása genom-DNS-ből

A *Zea mays* egy teljes hosszúságú, natív *Sus3* polinukleotid szekvenálása a 11. vagy 13. számú szekvenciavázlat felhasználásával valósítható meg, hogy az izolálás céljára azonosítani tudjuk a *Zea mays* genom-DNS-ének azt a fragmentumát, ami a *Sus3* gén le nem fordítható területét és 5' végét kódolja. Az így izolált genom-fragmentumot szekvenáljuk és azonosítjuk benne az exonokat, hogy ezzel igazoljuk a kukorica *Sus3* cDNS-szekvenciáját.

A fenti példák csak illusztrálják találmányunkat, de nem korlátozzák oltalmi körét. A szakemberek könnyen felismerik a találmány másféle változatait, amelyek azonban az igénypontokon belül maradnak. A leírásban idézett összes közleményt, szabadalmi iratot és számítógépes programot referenciaként tartjuk számon.

S Z A B A D A L M I I G É N Y P O N T O K

1. Izolált polinukleotid, amely egy szacharóz-szintetáz aktivitással bíró polipeptidet kódol és az alábbi csoport egyik eleme:

(a) polinukleotid, ami legalább 80 %-os szekvencia-azonosságot mutat – a GAP algoritmussal, alapbeállítási paraméterek mellett vizsgálva – egy 1. vagy 11. számú szekvenciavázlat szerinti polinukleotiddal;

(b) polinukleotid, ami egy 2. vagy 12. számú szekvenciavázlat szerinti polipeptidet kódol;

(c) polinukleotid, amit *Zea mays* nukleinsavakból sokszoroztunk olyan primerek használatával, amelyek szigorú körülmények között szelektíven hibridizálódnak egy 1. vagy 11. számú szekvenciavázlat szerinti polinukleotidon belül lévő lokuszokkal;

(d) polinukleotid, ami szigorú körülmények között és 0,1X SSC-vel 65 °C-on végzett mosás mellett szelektíven hibridizálódik egy 1. vagy 11. számú szekvenciavázlat szerinti polinukleotiddal;

(e) egy 1. vagy egy 12. számú szekvenciavázlat szerinti polinukleotid;

(f) polinukleotid, ami egy (a), (b), (c) vagy (e) pont szerinti polinukleotiddal komplementer; és

(g) polinukleotid, ami legalább 50 összefüggő nukleotidot tartalmaz egy (a), (b), (c), (d), (e) vagy (f) pont szerinti polinukleotidból.

2. Rekombináns expressziós keret, amely egy 1. igénypont sze-

rinti elemet tartalmaz működőképes módon, szensz vagy antiszensz irányultságban egy promoterhoz kapcsolva.

3. Gazdasejt, amely egy 2. igénypont szerinti rekombináns expressziós keretet tartalmaz.

4. Transzgén növény, amely egy 2. igénypont szerinti rekombináns expressziós keretet tartalmaz.

5. A 4. igénypont szerinti transzgén növény, amely növény egy-szikű.

6. A 4. igénypont szerinti transzgén növény, amely növény két-szikű.

7. A 4. igénypont szerinti transzgén növény, amely növény a kukorica, szója, napraforgó, cirok, olajrepce, búza, lucerna, gyapot, rizs, árpa, köles, földimogyoró és kakaó közül van kiválasztva.

8. A 4. igénypont szerinti transzgén növény magja.

9. Eljárás a szacharóz-szintetáz szintjének módosítására egy transzgén növényben, azzal jellemezve, hogy az említett eljárás az alábbi lépéseket foglalja magában:

(a) egy rekombináns expressziós keret – ami egy 1. igénypont szerinti polinukleotidot tartalmaz működőképes módon összekapcsolva egy promoterral – bejuttatása egy növényi sejtbe;

(b) a növényi sejt tenyésztése a növényi sejtek szaporítására alkalmas körülmények között;

(c) az említett transzgén növény regenerálása; és

(d) az említett polinukleotid kifejeztetése – aminek eredményeként egy általa kódolt fehérje termelődik – elegendő ideig ahhoz, hogy módosuljon a szacharóz-szintetáz szintje az emlí-

tett növényben.

10. A 9. igénypont szerinti eljárás, azzal jellemezve, hogy az említett növény a kukorica, szója, napraforgó, cirok, olajrepce, búza, lucerna, gyapot, rizs, árpa, köles, földimogyoró és kakaó közül van kiválasztva.

11. A 9. igénypont szerinti eljárás, azzal jellemezve, hogy a kódolt fehérje az alábbi csoport egyik eleme:

(a) egy 2. vagy 12. számú szekvenciavázlat szerinti polipeptid;

(b) polipeptid, ami legalább 80 %-ban azonos és legalább egy közös epitópja van egy 2. vagy 12. számú szekvenciavázlat szerinti polipeptiddel, ahol az említett szekvencia-azonosságot a GAP algoritmussal, alapbeállítási paraméterek használatával határozzuk meg; és

(c) legalább egy polipeptid, amit egy 1. igénypont szerinti elem kódol.

12. Izolált fehérje, amely az alábbi csoport egyik eleme:

(a) egy 2. vagy 12. számú szekvenciavázlat szerinti polipeptid;

(b) polipeptid, ami legalább 80 %-ban azonos és legalább egy közös epitópja van egy 2. vagy 12. számú szekvenciavázlat szerinti polipeptiddel, ahol az említett szekvencia-azonosságot a GAP algoritmussal, alapbeállítási paraméterek használatával határozzuk meg; és

(c) legalább egy polipeptid, amit egy 1. igénypont szerinti elem kódol.

13. Eljárás a cellulóz-termelés fokozására egy transzgén növény

szárában, azzal jellemezve, hogy az említett eljárás az alábbi lépéseket foglalja magában:

(a) egy rekombináns expressziós keret – ami egy szacharóz-szintetáz polinukleotidot tartalmaz működőképes módon összekapcsolva egy promoterral – bejuttatása egy növényi sejtbe;

(b) a növényi sejt tenyésztése a növényi sejtek szaporítására alkalmas körülmények között;

(c) az említett transzgén növény regenerálása; és

(d) az említett polinukleotid kifejeztetése elegendő ideig ahhoz, hogy emelkedjen a szacharóz-szintetáz szintje az említett növényben.

14. A 13. igénypont szerinti eljárás, azzal jellemezve, hogy az említett növény a kukorica, szója, napraforgó, cirok, olajrepce, búza, lucerna, gyapot, rizs, árpa, köles, földimogyoró és kakaó közül van kiválasztva.

15. A 13. igénypont szerinti eljárás, azzal jellemezve, hogy az említett szacharóz-szintetáz polinukleotid a *Sus1*, *Sh1* vagy a *Sus3* a kukoricából.

16. A 13. igénypont szerinti eljárás, azzal jellemezve, hogy az említett promoter előnyösen a szár szöveteiben vezérli a kifejeződést.

17. Eljárás a cellulóz-koncentráció emelésére egy transzgén növény magvának szöveteiben, azzal jellemezve, hogy az említett eljárás az alábbi lépéseket foglalja magában:

(a) egy rekombináns expressziós keret – ami egy szacharóz-szintetáz polinukleotidot tartalmaz működőképes módon összekapcsolva egy promoterral – bejuttatása egy növényi sejtbe;

(b) a növényi sejt tenyésztése a növényi sejtek szaporítására alkalmas körülmények között;

(c) az említett transzgén növény regenerálása; és

(d) az említett polinukleotid kifejeztetése elegendő ideig ahhoz, hogy emelkedjen a szacharóz-szintetáz szintje az említett transzgén növény említett magvaiban.

18. A 17. igénypont szerinti eljárás, azzal jellemezve, hogy az említett növény a kukorica, szója, napraforgó, cirok, olajrepce, búza, lucerna, gyapot, rizs, árpa, köles, földimogyoró és kakaó közül van kiválasztva.

19. A 17. igénypont szerinti eljárás, azzal jellemezve, hogy az említett szacharóz-szintetáz polinukleotid a *Sus1*, *Sh1* vagy a *Sus3* a kukoricából.

20. A 17. igénypont szerinti eljárás, azzal jellemezve, hogy az említett promoter előnyösen a magban vezérli a kifejeződést.

21. A 17. igénypont szerinti eljárás, azzal jellemezve, hogy az említett promoter előnyösen a maghéjban (perikarpium) vezérli a kifejeződést.

115
20
16

151 sz.

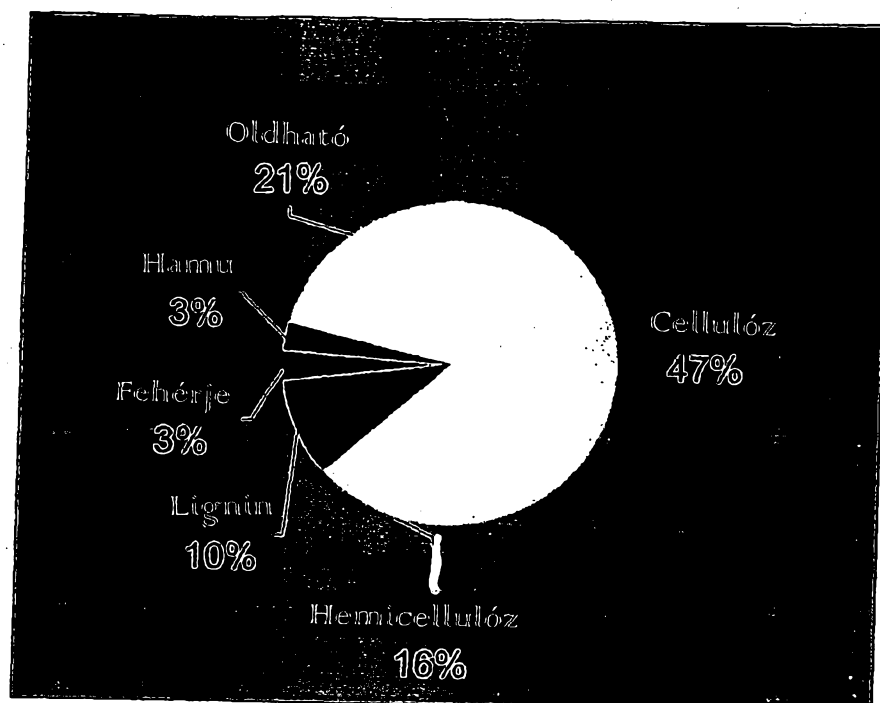
A meghatalmazott:


S. B. G. & K.
Szabadalmi Ügyvivői Iroda
H-1062 Budapest, Andrássy út 113.
Telefon: 461-1000, Fax: 461-1099

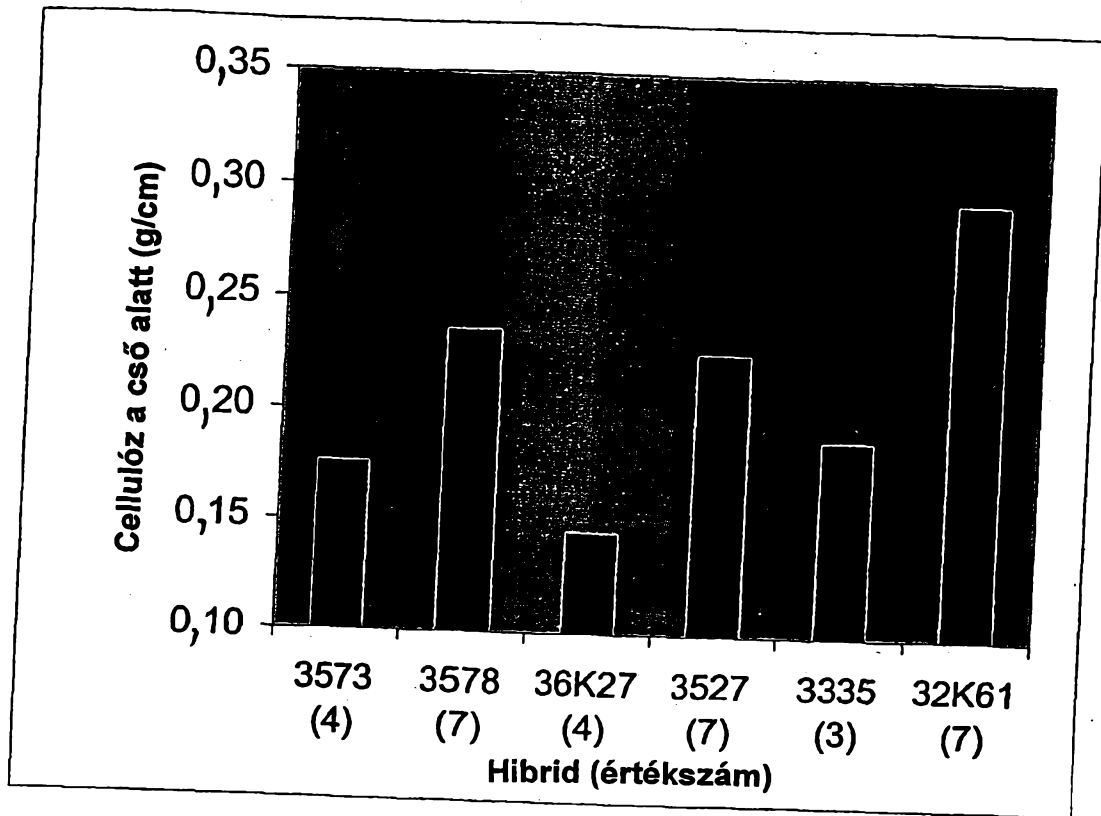
"H2"

1/16

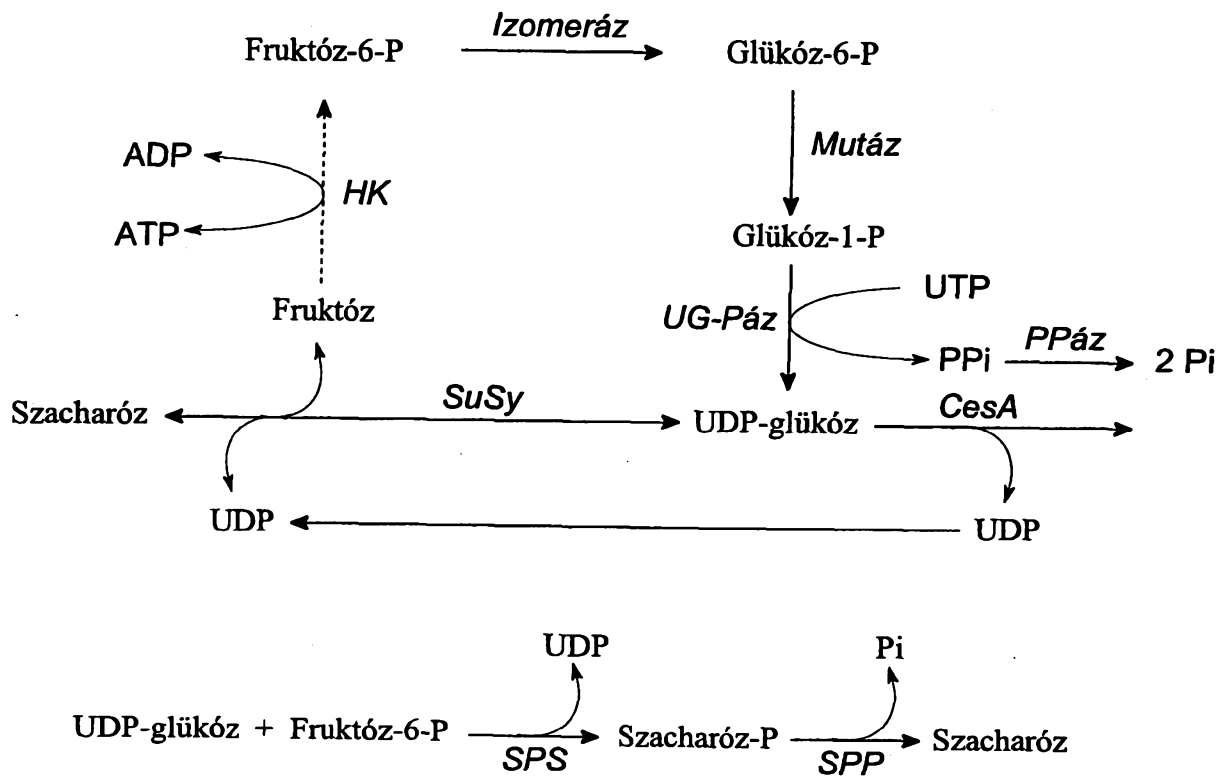
77:382/BE



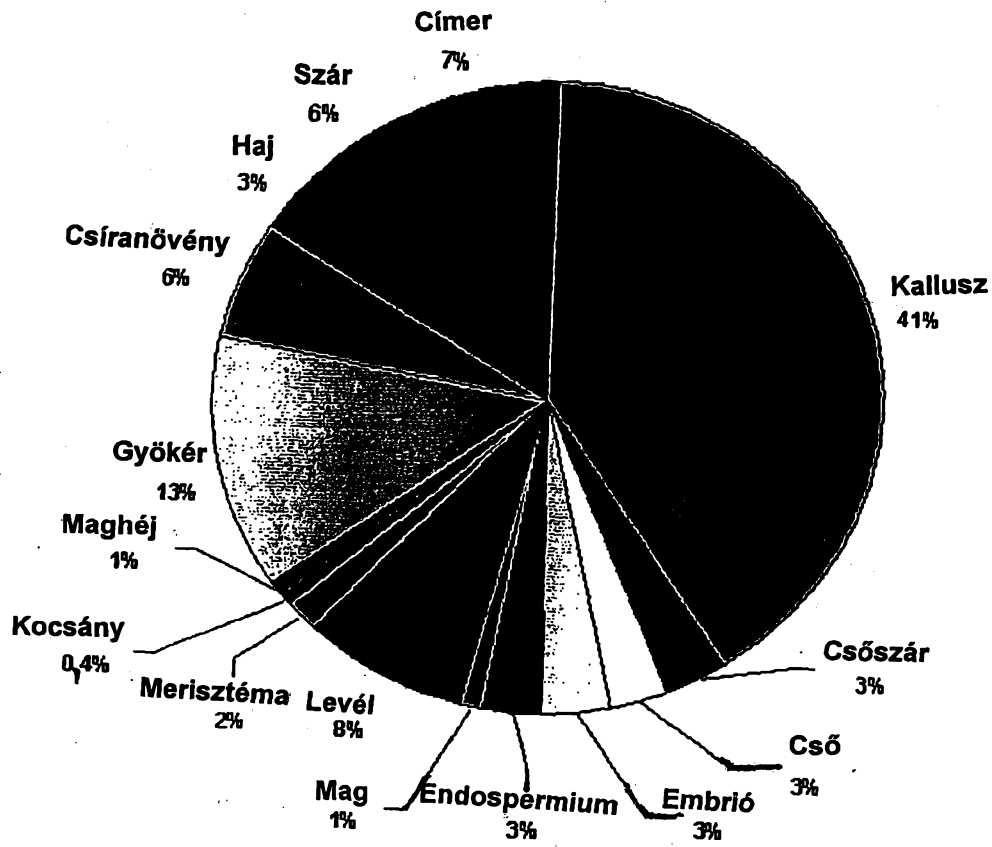
1. ábra



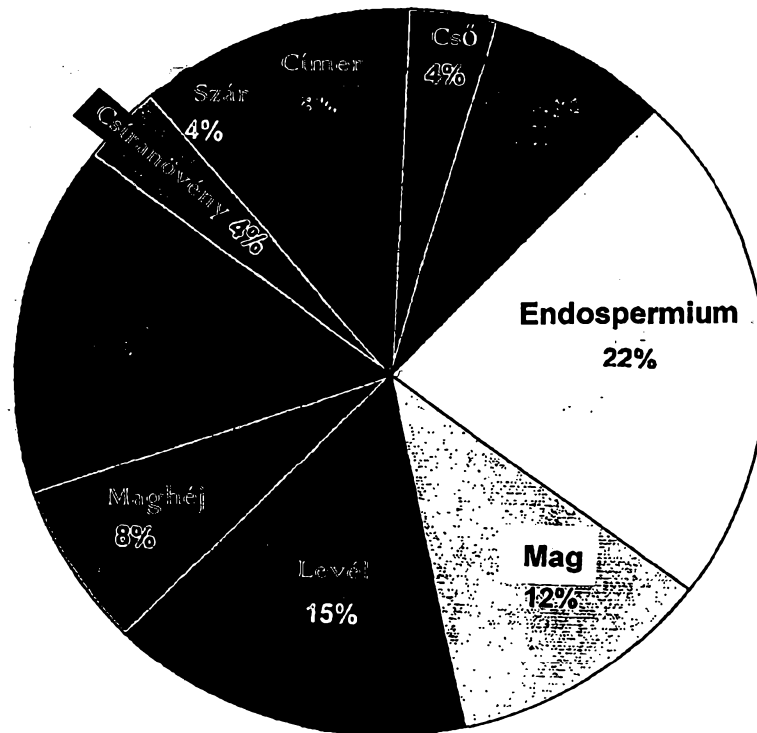
2. ábra



3. ábra



4. ábra

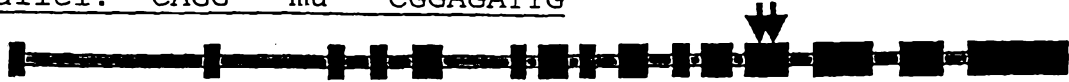


5. ábra

6/16

1. allél: CACCCGG - mu - AGATTG

2. allél: CAGG - mu - CGGAGATTG



6. ábra

Genotípus	Szerkezeti szárazanyag (az összes szárazanyag %-ában)	Cellulóz (az összes szárazanyag %-ában)	Cellulóz (a szerkezeti szárazanyag %-ában)
<i>Sus-I</i> (vad típus)	63,2	25,2 +/- 0,38	39,9
<i>Sus-I</i> (mutáns)	47,3	17,7 +/- 0,34	37,4

7. ábra

		1		50
Sh1	(1)	-----MAAKLRLSREHIGATFSSEPNEELIALFSRYVHOGKMLQREHQ		
Sus1	(1)	MGEAGDRVLRLSREHIGDLSAHPNELVAVFIRLKNLKGKMLQPHQ		
Sus3	(1)	-----STHSCDRVEDTLHARRMIVALLSKYVNKGKGLQPHH		
Consensus	(1)	LSRLHSLRERIGDTLSAHPNELVALFSRYVN GKGMLOPHQ		
		51		100
Sh1	(46)	LLAEFD-AIFDSDE--KYAPFEDILRAAQEAIVLPFWVALAIRPRPGVT		
Sus1	(51)	ILAEYNNALPEAREKCLKGAEEDVLRAAQEAIVIPPFWVALAIRPRPGVT		
Sus3	(40)	ILDALDEVQSGGGR-LAEGPFLDVLRSAAQEAIVLPPFWAIAVPRPGVT		
Consensus	(51)	ILAEFD AI DADRE LKDGPFEDVLRAAQEAIVLPPFWVALAIRPRPGVT		
		101		150
Sh1	(93)	DYIRVNVSELAVEELSVSEYLAFKEQLVDGQSN SNFVLELDFEPPFNASFP		
Sus1	(101)	EYVRVNVSELAVEELRPPEYLOPKFOLVVEGPN NNFVLELDFEPPFNASFP		
Sus3	(89)	EYVRVNVHISVLOLTVSEYLRFKELVDGQHNDPYVLELDFEPPFNVSVE		
Consensus	(101)	EYVRVNVSELAVEELSVSEYL FKEQLVDGQ N NFVLELDFEPPFNASFP		
		151		200
Sh1	(143)	RESMSKSIQNGVOFLNRHLSSKLFQDKESLYELINELKARNYKCTTMMLN		
Sus1	(151)	RESLSKSIQNGVOFLNRHLSSKLFHDKESLYELINELKARNYKCTTMMLN		
Sus3	(139)	RNRRESIISQNGVOFLNRHLSSIMERNRDCLEPILDFIRGRHMHCHVMGN		
Consensus	(151)	RPSLSKSIQNGVOFLNRHLSSKLF DKESLYPLLNFLRAHNYKG TMMLN		
		201		250
Sh1	(193)	DRIQSLRGLSSSRKAEEYLSVPQDTPYSEFNHRFQELGLEKGGWTAK		
Sus1	(201)	DRIQSLRGLSALGAKRKAEEHLSLQADTPYSEFHHRFQELGLEKGGWTAK		
Sus3	(189)	DRIQSLRGLGSLTKAEEHLSKLPADTPYSEFAYKGFQWGLEKGGWTAK		
Consensus	(201)	DRIQSL ALQSALRKAEEHLSLQADTPYSEF HRFQELGLEKGGWTAK		
		251		300
Sh1	(243)	RVLETIHLLELLEAPDPANLEKFLGTIPMIFNVVILSPHGYFAQANVLG		
Sus1	(251)	RAQETIHLLELLEAPDSTLEKFLGTIPMIFNVVILSPHGYFAQANVLG		
Sus3	(239)	HVLEMIHLLELLEAPDSTLEKFLGTIPMIFNVVILSPHGYFAQANVLG		
Consensus	(251)	RVLETIHLLELLEAPDPSTLEKFLGTIPMIFNVVILSPHGYFAQANVLG		
		301		350
Sh1	(293)	YPDITGGQVYILQVRALENEMLLRIKQGLDITPKILIVTRLLPDAQT		
Sus1	(301)	YPDITGGQVYILQVRALENEMLLRIKQGLDITPKILIVTRLLPDAQT		
Sus3	(289)	LPDITGGQVYILQVRALENEMVLRLLKQGLDITPKILIVTRLLPDAQT		
Consensus	(301)	YPDITGGQVYILQVRALENEMLLRIKQGLDITPKILIVTRLLPDA GT		
		351		400
Sh1	(343)	TCGQRLEKVICTEHTDIRVPPFRNENGILRKWISRFDWPYLETYTEDVA		
Sus1	(351)	TCGQRLEKVICTEHTDIRVPPFRNENGILRKWISRFDWPYLETYTEDVA		
Sus3	(339)	SCNQRLEKVICTEHTDIRVPPFRNENGILRKWISRFDWPYLETYTEDVA		
Consensus	(351)	TCGQRLEKVICTEHTDIRVPPFRNENGILRKWISRFDWPYLETYTEDVA		
		401		450
Sh1	(393)	SEIMKEMOAKPELLIGNYSDGNLVATLLAHKLGVTOSTIAHALEKTKYPN		
Sus1	(401)	HPIAGELIANPELLIGNYSDGNLVATLLAHKMGVHTIAHALEKTKYPN		
Sus3	(389)	GPIAELIGTETPELLIGNYSDGNLVATLLAHKMGVHTIAHALEKTKYPN		
Consensus	(401)	EIAAELQA PDLIIGNYSDGNLVATLLAHKMGVTQCTIAHALEKTKYPN		
		451		500
Sh1	(443)	SEIYLDKEDSOYHFSCQFTADLIAMNHTDFIITSTFOBIAGSKDITVGOYE		
Sus1	(451)	SDIYWKKEDEYHFSCQFTADLIAMNHTDFIITSTFOBIAGSKDITVGOYE		
Sus3	(439)	SEIYWKNEDEYHFSCQFTADLIAMNHTDFIITSTFOBIAGSKDITVGOYE		
Consensus	(451)	SDIYWKKEDEYHFSCQFTADLIAMNHTDFIITSTFOBIAGSKDITVGOYE		
		501		550
Sh1	(493)	SHIAFTLPGLYRVVHGIDVFDPKFNIVSPGADMSIYFPYTES KRLTSLH		
Sus1	(501)	SHIAFTLPGLYRVVHGIDVFDPKFNIVSPGADMSIYFPYTES KRLTSLH		
Sus3	(489)	SHIAFTLPGLYRVVHGIDVFDPKFNIVSPGADMSIYFPYTES KRLTSLH		
Consensus	(501)	SHIAFTLPGLYRVVHGIDVFDPKFNIVSPGADMSIYFPYTES KRLTSLH		

```

551                                     600
Sh1 (543) PEIEELIYSDVENSRAKFLVKKKKPIIFSMARLDRVKNTGLVLYGKN
Sus1 (551) PEIEELIYSQTEENRAKFLVNDNRNPIIFSMARLDRVKNTGLVLYGKN
Sus3 (539) GSIEENIIDPEQNDRIIGHIDORSKPIIFSMARLDRVKNTGLVLYGKN
Consensus (551) PEIEELIYS ENSEHKFVL DR KPIIFSMARLDRVKNTGLVLYGKN

601                                     650
Sh1 (593) ARRLRELANLVIVAGDHGK-EKERNHQAIEFKQVSLIDEYKLGKCHIRWIS
Sus1 (601) KRLOELVNLVIVAGDHGN-PKDKSEQAIEFKQVFDLIEQYNLNGHIRWIS
Sus3 (589) AKLRELVNLVIVAGYNDVNKSKREHIEIEKQVHLLIKTENVLFQPFQWIS
Consensus (601) ARRLRELVNLVIVAGDHG SKDREEQAIEFKKMHDLLID YNL GHIRWIS

651                                     700
Sh1 (642) AQMNRVRNGELYRYICDTKGAFWQPAFYEAAGLTVVIEAMTCGLPTFATYCH
Sus1 (650) AQMNRVRNGELYRYICDTKGAFWQPAFYEAAGLTVVIEAMTCGLPTFATAY
Sus3 (639) ACTNRVRNGELYRYIADTHGAFWQPALYEAAGLTVVIEAMTCGLPTFATLH
Consensus (651) AQMNRVRNGELYRYICDTKGAFWQPAFYEAAGLTVVIEAMTCGLPTFAT H

701                                     750
Sh1 (692) GGPAAEIIIVDGVSGFLIDDPYHSDKAADILVNFDDKCKADESYNDEILSQGGL
Sus1 (700) GGPAAEIIIVHGVSGFLIDDPYQGDKASALLVDFFDKCQAEESHWSKLSQGGI
Sus3 (689) GGPAAEIIIEHGVSGFLIDDPYHPQAVNLIADFFDRCKQDEEDHVTNLSGAGL
Consensus (701) GGPAAEIIIVHGVSGFLIDDPYH DKAALLVDFDKCKADPSHW ISQGGI

751                                     800
Sh1 (742) QRIYEKYTWKLYSERLMTLTGVYGFWKYVSNLERRRTRYLEMFPYALKYR
Sus1 (750) QRIYEKYTWKLYSERLMTLTGVYGFWKYVSNLERRRTRYLEMFPYALKYR
Sus3 (739) QRIYEKYTWKLYSERLMTLTGVYGFWKYVSNLERRRTRYLEMFPYALKYR
Consensus (751) QRIYEKYTWKLYSERLMTLTGVYGFWKYVSNLERRRTRYLEMFPYALKYR

801                                     817
Sh1 (792) SLASTVPLSFD-----
Sus1 (800) SLASTVPLAVEGEPSSK
Sus3 (789) ELAKTVPLAID-QPQ--
Consensus (801) SLASTVPLAID P

```

8b. ábra

9a. ábra

		551	600
Sh1	(540)	CTTACCGGACACCTGTCTTCAAGTGTGTCAGGAGACAAAGGAGAGCTTTCGT	
Sus1	(520)	CTCAACAGCGACCTGTCTCAAGCTCTTCCATBACAATACAGACAGCTA	
Sus3	(458)	CTCAACCGACAGTGTCTCTCAATCATGTTCCGCAACAGGATTECTCTGGA	
Consensus	(551)	CTCAACCGACACCTGTCTCAAGTGTTCCTA GACAAGGAGAGCTTGTGA	
		601	650
Sh1	(590)	CCCCTTGGTGAAGTTCTGCAAGGCTCA TAACTACAAAGGCTCAGGAGATCA	
Sus1	(570)	CCCCTTGGTCAAGTTCTGCTGCECCACAACTACAAGGAGATGACCATCA	
Sus3	(508)	CCCCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	
Consensus	(601)	CCCCTTGGTGAAGTTCTCTCCG GCCCACAACCTACAAGGGGA GAC ATGA	
		651	700
Sh1	(640)	TGTTGAATGACAGAATCCAAAGCCTTCCTGGCTTCAATCATCTCTGAG	
Sus1	(620)	TGTTGAATGACAGAATCCGAGCTCTCAAGCTCTCTGCTGCTGCTGCTGCT	
Sus3	(558)	TGCTTATGATGAGAATCAAGCTTGGGAGGCTTCAAGCTCTCTGCTGCTGCT	
Consensus	(651)	TGTTGAATGACAGAATCCAAAGCCT GTGGTCT CAATCTGCGCTGAG	
		701	750
Sh1	(690)	AAGGCTGAGGAGCTTCTACTGAGTGTTCCTCAAGGAGCTCTCTACTCTGA	
Sus1	(670)	AAGGCTGAGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Sus3	(608)	AAGGCTGAGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Consensus	(701)	AAGGCTGAGGAGCACCCTGTCT A CTCTGCTGACAC CCATCTCTCTGA	
		751	800
Sh1	(740)	GTTCAAGGCTAGCTTCCAAAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCT	
Sus1	(720)	ATTTCAAGGCTAGCTTCCAAAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCT	
Sus3	(658)	ATTTGCTTATTAATTTCAAGGCTTGGGCTCTCTCTCTCTCTCTCTCTCTCT	
Consensus	(751)	ATTT ACCATAGGTTCCAAGAGCTTGGCCTGGAGAAGGCTTGGGGTGATA	
		801	850
Sh1	(790)	CTCTGAGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Sus1	(770)	GCCTGAGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Sus3	(708)	CAGCAGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Consensus	(801)	C GC AAGCGTGTCTCTGGA AC ATCCACCTCTCTCTCT GACCTCTCTTGA	
		851	900
Sh1	(840)	CCCCCTATCTCTGCTCAAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Sus1	(820)	CCCCCAGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Sus3	(758)	CCCCCAGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Consensus	(851)	GCCCCAGATCC TCCACCCTGGAGAAGTTCCTTGGAAACGATCCCCATGAT	
		901	950
Sh1	(890)	GTTCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Sus1	(870)	GTTCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Sus3	(808)	TCTTACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Consensus	(901)	GTTCAACGTTGTTATCCT TCCCCTCATGG TACTTCTGCTCAAGCTAATG	
		951	1000
Sh1	(940)	TGCTTCTCTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Sus1	(920)	TCTTCTCTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Sus3	(858)	TCTTCTCTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Consensus	(951)	T TT GG TACCCTGACAC GGAGG CAGGTTGTCTACAT CTGGATCAA	
		1001	1050
Sh1	(990)	GTCCCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Sus1	(970)	GTCCCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Sus3	(908)	GTCCCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Consensus	(1001)	GTCCGTGCT TGGAGAATGAGATGCTTCTGAGGAT AAGCAGCAAGG CT	
		1051	1100
Sh1	(1040)	TGATATCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Sus1	(1020)	TGATATCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Sus3	(958)	TGATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
Consensus	(1051)	TGATATCAC CCGAAGATCCTCATTTGTTACCAGGCTG T CCTGATGCAA	

		1101		1150
Sh1	(1090)	CTGGGACTAGGTCGGCTCAGCGGCTGAGAAAGGTCATTGGTAC	TCACGAC	
Sus1	(1070)	CTGGGACTAGGTCGGCTCAGCGGCTGAGAAAGGTCATTGGTAC	TCACGAC	
Sus3	(1008)	AAAGAAATATCAATGCAATCAGCGGCTGAGAAAGGTCATTGGTAC	TCACGAC	
Consensus	(1101)	CTGG AC AC TCGGTCAGCGGCTGAGAAAGGTCATTGG AC GAGCAC		
		1151		1200
Sh1	(1140)	ACAGACATCATTCGGCTTCGCTTCAGAAATGCAATGGCATCC	CEGCA	
Sus1	(1120)	TGCCATATCCCTTCGGCTTCGCTTCAGAAATGCAATGGCATCC	CEGCA	
Sus3	(1058)	AGTTACATATTCAGCTTCGCTTCAGAAATGCAATGGCATCC	CEGCA	
Consensus	(1151)	AC ACATC TTCGCTTCCTTCAGAAATGCAATGG ATCCTTCGCA		
		1201		1250
Sh1	(1190)	GTGGATCTC CGATTGATGTCTGGCCATACCTGGAGACATACACTGAGG		
Sus1	(1170)	GTGGATCTC CGATTGATGTCTGGCCATACCTGGAGACATACACTGAGG		
Sus3	(1108)	ATGGATATCAATGATTCGCTTCGCTTCAGAAATGCAATGGCATCC	CEGCA	
Consensus	(1201)	GTGGATCTC CGATTGATGTCTGGCCATACCTGGAGACATACACTGAGG		
		1251		1300
Sh1	(1240)	ATGTTGC GTGAAATTGCTG AGAA T CAGGCCAATCTGACCT ATC		
Sus1	(1220)	ATGTTGC GTGAAATTGCTG AGAA T CAGGCCAATCTGACCT ATC		
Sus3	(1158)	ATGTTGC GTGAAATTGCTG AGAA T CAGGCCAATCTGACCT ATC		
Consensus	(1251)	ATGTTGC GTGAAATTGCTG AGAA T CAGGCCAATCTGACCT ATC		
		1301		1350
Sh1	(1290)	ATTGGAAACTACAGTGATGGAAACCTTGT GCGTCTTGCTCGC CACAA		
Sus1	(1270)	ATTGGAAACTACAGTGATGGAAACCTTGT GCGTCTTGCTCGC CACAA		
Sus3	(1208)	ATTGGAAACTACAGTGATGGAAACCTTGT GCGTCTTGCTCGC CACAA		
Consensus	(1301)	ATTGGAAACTACAGTGATGGAAACCTTGT GCGTCTTGCTCGC CACAA		
		1351		1400
Sh1	(1340)	GATGGGAGTTACTCAGTGATGACCTGCTCATGC CTGGAGAAAACCTAAGT		
Sus1	(1320)	GATGGGAGTTACTCAGTGATGACCTGCTCATGC CTGGAGAAAACCTAAGT		
Sus3	(1258)	GATGGGAGTTACTCAGTGATGACCTGCTCATGC CTGGAGAAAACCTAAGT		
Consensus	(1351)	GATGGGAGTTACTCAGTGATGACCTGCTCATGC CTGGAGAAAACCTAAGT		
		1401		1450
Sh1	(1390)	ACCC AACTC GACATCTACTGGAAGAA TTCGA GA CAGTACCCTTC		
Sus1	(1370)	ACCC AACTC GACATCTACTGGAAGAA TTCGA GA CAGTACCCTTC		
Sus3	(1308)	ACCC AACTC GACATCTACTGGAAGAA TTCGA GA CAGTACCCTTC		
Consensus	(1401)	ACCC AACTC GACATCTACTGGAAGAA TTCGA GA CAGTACCCTTC		
		1451		1500
Sh1	(1440)	TC TCCAGTTCAC GCTGAC T ATTGC ATGAACCATGCCGATTCAT		
Sus1	(1420)	TC TCCAGTTCAC GCTGAC T ATTGC ATGAACCATGCCGATTCAT		
Sus3	(1358)	TC TCCAGTTCAC GCTGAC T ATTGC ATGAACCATGCCGATTCAT		
Consensus	(1451)	TC TCCAGTTCAC GCTGAC T ATTGC ATGAACCATGCCGATTCAT		
		1501		1550
Sh1	(1490)	CATCACCAGCACATTCCAAGAAATCGC GGAAGCAAGGACACCGT GG C		
Sus1	(1470)	CATCACCAGCACATTCCAAGAAATCGC GGAAGCAAGGACACCGT GG C		
Sus3	(1408)	CATCACCAGCACATTCCAAGAAATCGC GGAAGCAAGGACACCGT GG C		
Consensus	(1501)	CATCACCAGCACATTCCAAGAAATCGC GGAAGCAAGGACACCGT GG C		
		1551		1600
Sh1	(1540)	AGTACGAGTC CATAT GCGTTCACTCTGCCTGG CTGTACCG GTTGTC		
Sus1	(1520)	AGTACGAGTC CATAT GCGTTCACTCTGCCTGG CTGTACCG GTTGTC		
Sus3	(1458)	AGTACGAGTC CATAT GCGTTCACTCTGCCTGG CTGTACCG GTTGTC		
Consensus	(1551)	AGTACGAGTC CATAT GCGTTCACTCTGCCTGG CTGTACCG GTTGTC		
		1601		1650
Sh1	(1590)	CATGGCATCGATGT TTCGATCCCAAGTTCAACAT GTCTCTCCTGGAGC		
Sus1	(1570)	CATGGCATCGATGT TTCGATCCCAAGTTCAACAT GTCTCTCCTGGAGC		
Sus3	(1508)	CATGGCATCGATGT TTCGATCCCAAGTTCAACAT GTCTCTCCTGGAGC		
Consensus	(1601)	CATGGCATCGATGT TTCGATCCCAAGTTCAACAT GTCTCTCCTGGAGC		

		1651		1700
Sh1	(1640)	AGACATGAGTGT	TTACTACCGT	TATACGAAAGCCAGAACAGACATCAGTG
Sus1	(1620)	GGATCTGTCCAT	CACCTCCGGA	CACCGAGTCCAGAACAGGGTCAGCT
Sus3	(1558)	TGACATGTCCAT	TACTTTTCACAT	ACCGAGAGAGCCATAGCAATCAGCT
Consensus	(1651)	GACATGTCCAT	TACTTCCC	TATACCGAGACGGACAAGAGACTCACCT
		1701		1750
Sh1	(1690)	CCTTC	CATCCTGAAATCGAGGAGCTCAT	CACAGCGACCTCGAGAACTCC
Sus1	(1670)	CCCTTC	ACCGGACAT	TCAGGAGCTCTGACAGGCAACCGAGAAAG
Sus3	(1608)	CTCTTCAT	GGTTCAACGGAATTTGATTTATGAGCCG	AGCTAAAGGAT
Consensus	(1701)	CCCTTCAT	CCTGAAATCGAGGAGCTCAT	TACAGCCA G CGAGAAC C
		1751		1800
Sh1	(1740)	GAGCACAAGTTC	GTTCTGAAACGACAGAGG	GATCATCTTCTCCAT
Sus1	(1720)	GAGCACAAGTTC	GTTCTGAAACGACAGAGG	CAATCATCTTCTCCAT
Sus3	(1658)	GAGCACAAGTTC	GTTCTGAAACGACAGAGG	CAATCATCTTCTCCAT
Consensus	(1751)	GAGCACAAGTTC	GTTCTGAAACGACAGAGG	AAGCC ATCATCTTCTCCAT
		1801		1850
Sh1	(1790)	GGGCGGTC	CGACCGCTGAGGAAACATGACAGC	CTGCTCGAGATGATAGC
Sus1	(1770)	GGGCGGTC	CGACCGCTGAGGAAACATGACAGC	CTGCTCGAGATGATAGC
Sus3	(1708)	GGGCGGTC	CGACCGCTGAGGAAACATGACAGC	CTGCTCGAGATGATAGC
Consensus	(1801)	GGC CGTCTCGACCG	GTGAAGAACATGACAGGGCTGGTCGAG	TGTACG
		1851		1900
Sh1	(1840)	GCAAGAACCGCG	GCTGAGGGAGCTGGTGAACCTCGTGGT	CGTTGCCGG
Sus1	(1820)	GCAAGAACCGCG	GCTGAGGGAGCTGGTGAACCTCGTGGT	CGTTGCCGG
Sus3	(1758)	GCAAGAACCGCG	GCTGAGGGAGCTGGTGAACCTCGTGGT	CGTTGCCGG
Consensus	(1851)	GCAAGAACCGCG	GCTGAGGGAGCTGGTGAACCTCGTGGT	CGTTGCCGG
		1901		1950
Sh1	(1890)	GACCACTG	CAAGGAGTCCAAAGGACAGGAGGAGCAGGCGGAGTTCAA	
Sus1	(1870)	GACCACTG	CAAGGAGTCCAAAGGACAGGAGGAGCAGGCGGAGTTCAA	
Sus3	(1808)	GACCACTG	CAAGGAGTCCAAAGGACAGGAGGAGCAGGCGGAGTTCAA	
Consensus	(1901)	GACCACTG	CAAC AGTCCAAGGACAGGAGGAGCAGGCGGAGTTCAA	
		1951		2000
Sh1	(1937)	GAAGATGTATG	ACCTCATCGAGAGTACAA GTTCAAGG	CCATATGCGGT
Sus1	(1917)	GAAGATGTATG	ACCTCATCGAGAGTACAA GTTCAAGG	CCATATGCGGT
Sus3	(1858)	GAAGATGTATG	ACCTCATCGAGAGTACAA GTTCAAGG	CCATATGCGGT
Consensus	(1951)	GAAGATGTATG	ACCTCATCGAGAGTACAA GTTCAAGG	CCATATGCGGT
		2001		2050
Sh1	(1987)	GGATCTC	GCCAGATGAACCG GTCCGCAACGGCGAGCTGTACCGCTAC	
Sus1	(1967)	GGATCTC	GCCAGATGAACCG GTCCGCAACGGCGAGCTGTACCGCTAC	
Sus3	(1908)	GGATCTC	GCCAGATGAACCG GTCCGCAACGGCGAGCTGTACCGCTAC	
Consensus	(2001)	GGATCTC	GCCAGATGAACCG GTCCGCAACGGCGAGCTGTACCGCTAC	
		2051		2100
Sh1	(2037)	ATCTGCGATACCA	AGGGCGCATTCTGTCAGCCTGCTTCTACGAAGCGTT	
Sus1	(2017)	ATCTGCGATACCA	AGGGCGCATTCTGTCAGCCTGCTTCTACGAAGCGTT	
Sus3	(1958)	ATCTGCGATACCA	AGGGCGCATTCTGTCAGCCTGCTTCTACGAAGCGTT	
Consensus	(2051)	ATCTGCGATACCA	AGGGCGCATTCTGTCAGCCTGCTTCTACGAAGCGTT	
		2101		2150
Sh1	(2087)	CGGCTGAC	GTGATCGACTCCATGACCTCGGCTTGTGCAACGATCGGCA	
Sus1	(2067)	CGGCTGAC	GTGATCGACTCCATGACCTCGGCTTGTGCAACGATCGGCA	
Sus3	(2008)	CGGCTGAC	GTGATCGACTCCATGACCTCGGCTTGTGCAACGATCGGCA	
Consensus	(2101)	CGG CTGAC	GTGGTTGAGGCCATGACCTGCGG CTGCC ACGTTCGCGA	
		2151		2200
Sh1	(2137)	CTGCGCATGCGCG	CGCTGCTGACATCATCTGCGCGAATCTGCGCTG	
Sus1	(2117)	CTGCGCATGCGCG	CGCTGCTGACATCATCTGCGCGAATCTGCGCTG	
Sus3	(2058)	CTGCGCATGCGCG	CGCTGCTGACATCATCTGCGCGAATCTGCGCTG	
Consensus	(2151)	CC CCATGGCGGTCC	GCTGAGATCATCGTGCACGGCGT TCTGGCTTC	

		2201		2250
Sh1	(2187)	CACATTGACCCCTTACCACAGCGACAAGGC GCGGATATCTCGTGCACCTT		
Sus1	(2167)	CACATCAGCCCTTACCAGGCGGACAAGGCCTGCGCCCTGCTCCTTGACTT		
Sus3	(2108)	CACATTGACCCGTTACCAGCCGTAACAGGCTTTTAATCTGATCGCGACTT		
Consensus	(2201)	CACATTGACCCCTTACCAC GCGACAAGGC GCGGATCTGCTGGTGCACCTT		
		2251		2300
Sh1	(2237)	CTTTTACAAAATGCAAGGCAGATCCGAGCTACTGCGACGAGAATCTACAGG		
Sus1	(2217)	CTTCGACAAGTCCGAGCGGAGCGGAGCCACTGAGCAAGATCTCCGAGG		
Sus3	(2158)	CTTCGACCGGTGCAAGCAGACCGGAGATCACTGGGTGAATATACTGGAG		
Consensus	(2251)	CTTCGACAAGTGCAAGGCAGA CCGAGCCACTGGG CAAGATCTC CAGG		
		2301		2350
Sh1	(2287)	CGGCCTTGCAAGAAATTTATGAGAAGTACACCTGCAAGCTCTACTCCAG		
Sus1	(2267)	CGCGGCTCCAGGCTATCGAGGAGAAGTACACCTGCAAGCTGTACTCGAG		
Sus3	(2208)	CACGGCTGCGAGGCATATACGAGAAGTACACATGCAAGCAATACTGAG		
Consensus	(2301)	GCGGGCTGCAGCG AT TA GAGAAGTACACCTGGAAGCT TACTC GAG		
		2351		2400
Sh1	(2337)	AGGCTGATGACCCCTGACCGCGGTGTACGGGTTCTGGAAGTACGTGTCCAA		
Sus1	(2317)	AGGCTGATGACCCCTGACCGCGGTGTACGGGTTCTGGAAGTACGTGTCCAA		
Sus3	(2258)	AGGCTGATGACCCCTGACCGCGGTGTACGGGTTCTGGAAGTACGTGTCCAA		
Consensus	(2351)	AGGCTGATGACCCCTGACCGCGGTGTACGGGTTCTGGAAGTACGTGTCCAA		
		2401		2450
Sh1	(2387)	CCTGGAGAGGCGCGAGACCCGCGCTACCTGAGATGTTCTACGCCTGA		
Sus1	(2367)	CCTGGAGAGGCGCGAGACCCGCGCTACCTGAGATGTTCTACGCCTGA		
Sus3	(2308)	GCTCAGAGGCTGAGAGGAGACCCCTACCTGAGATGTTCTACGATCTGA		
Consensus	(2401)	CCTGGAGAGGCGCGAGACCCGCGCTACCT GAGATGTTCTACGC CTGA		
		2451		2500
Sh1	(2437)	AGTACCTGAGCTGCGCAAGGCAGGTTCCGGTCTCTTCTCA-----TTAG		
Sus1	(2417)	AGTACCTGAGCTGCGCAAGGCAGGTTCCGGTCTCTTCTCA-----GGGA		
Sus3	(2358)	AGTACCTGAGCTGCGCAAGGCAGGTTCCGGTCTCTTCTCA-----GAG		
Consensus	(2451)	AGTACCGCA CTTGGCGAGCACCGTGCCGCTGGCC T GA G AG		
		2501		2550
Sh1	(2481)	TACGGGCAAGCAAGGAGT-AGAAGAGAGAGAGCCAGGCGGCA-----G		
Sus1	(2461)	GAGGCCTGCGCAAGTGT-TCGGTACGGCGGCAAGAGAGCTGATC---G		
Sus3	(2408)	TAGGTTGCGCAACTGCGA-TCGGTACGGGCAAGTGTGAGAGAGCTG		
Consensus	(2501)	TAGC GC AGAA G GA TGCCTAACA GGCACAGGCCTGA G		
		2551		2600
Sh1	(2525)	AACGATCGCTGCTTTCGATCT-----GT-TTCAACCGCAATTTCG		
Sus1	(2507)	ATCGATGAGGAGAGGGAGCACTCGGA-----GT-CTCGTGTCTTTTC		
Sus3	(2458)	AGGATCCTTAAATTAATTAGGCGCGGAGACCGTATCCGATAAAATGTG		
Consensus	(2551)	AACGATC C G A TT G CTCGG GT GTCA CAATTTCG		
		2601		2650
Sh1	(2565)	ATTGTTAGTTCGCTTANGGAGTATGTTG--TACTTGGTTTCAGAACTT		
Sus1	(2551)	TTCCTATTCTTTCTTTCTTTTTC--TTCCCGAGCGGAGAAATAA		
Sus3	(2508)	CGTAGCTCAACTCGTTTATTATATGATACATAATGCTAGTATATCAAAAT		
Consensus	(2601)	TG TGTC TG TTT TT TTATGT TACT GGAGTC AA AAAAT		
		2651		2700
Sh1	(2613)	TGGTTCCTTCTCGTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT		
Sus1	(2599)	AGAGTC-TGCTT-TTCTTACGGCGCGCGCTTCTTCTTCTTCTTCTTCTT		
Sus3	(2558)	TACTGAAGGAGAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT		
Consensus	(2651)	TG TTC TGCT GTTG TTG CGTTGTGTTCGTT CTG C GCTT		
		2701		2750
Sh1	(2663)	GCCTGCTTCCATGATGCTGGGAATTCGCTGCACTTTTCTTCTTCTTCTT		
Sus1	(2647)	CAGAGTTAAATTTACCTACG--TTCTAAGGTCTTCTTCTTCTTCTTCTT		
Sus3	(2600)	TACTGTA-TATGTTCAAGCTGTC-----GCTGCAATTTCTTCTTCTTCTT		
Consensus	(2701)	ACTGTT ATATTAAGCTG C TTGGCTGCA CTT TTC TGA T A		

```

                2751                                2800
Sh1 (2713) AAATGCCTGCTGCTCACTGCTTCCAGAGTC-----
Sus1 (2695) TCCGGGTGTGCTTCTAGTAGTCTCAGGACCTTAGTAGTTTCGTTGC
Sus3 (2644) AGCCGACCGAATGCTGAAGTGCTCATTAATACATCATTTCTTTGACC
Consensus (2751) A C GC GGC CTGTGA GTCTGATAGA TG T TA T TG C
                2801                                2850
Sh1 (2747) -----
Sus1 (2745) GTCGTTGAGAGGGAACGTTGGTGGTGGTGGTGTCTCTCAGTCAGGCGT
Sus3 (2694) TGTAAAAAATAAAAAAATAAAAAAAAAAAAAAAAAAGGCCGCCG-----
Consensus (2801) G A A AA G G G C

                2851                                2900
Sh1 (2747) -----
Sus1 (2795) GGTGCTCCCTTTGTTTCCTGGATGGGATGTTGCTCCTTGAATAATAATCG
Sus3 (2738) -----
Consensus (2851)

                2901                                2950
Sh1 (2747) -----
Sus1 (2845) TAGTGGCCTTGGAGCCCTTTTCCTGAAATAAGAGCAGCATCCTAGTGCTT
Sus3 (2738) -----
Consensus (2901)

                2951                                2964
Sh1 (2747) -----
Sus1 (2895) ACCTTTGCAGCTGT
Sus3 (2738) -----
Consensus (2951)

```

9f. ábra

CGCCAGTCGCCAGTCGCCACAGCCACACCACACCACACTAGCCGCGGCCGCGGGTAGGAG
CGCGCGCGGCGCGGCGGAACGACCCACCGGTGGCGGCAGCCATGCTCTGCCCCGAAGCTGA
ACCGCAACGCGAGCATCCGGGACCGCGTCGAGGACACCCTCCACGCGCACCGCAACGAGC
TCGTGCCCCCTCCTCTCCAAGTACGTGAACAAGGGGAAGGGCATCCTGCAGCCGCACCACA
TCCTCGACGCGCTCGACGAGGTCCAGGGCTCCGGGGTCCGCGCGCTCGCCGAGGGACCCT
TCCTCGACGTCTCCGCTCCGCGCAGGAGGCGATCGTGCTGCCGCCG

10. ábra

Círok-szekvencia a 13. sz. szekvenciavázlatból a 11. sz. szekvenciavázlatban

ATGTCTGCCCCGAAGCTGAACCGCAACGCGAGCATCCGG

Kukorica-szekvencia az 1. sz. szekvenciavázlatból a 11. sz. szekvenciavázlatban

accgcgctcga	ggacaccctc	cacgcgcacc	gcaacgagct	60		
cgctcgccctc	ctgtccaagt	acgtgaacaa	ggggaagggc	atcctgcagc	cgcaccacat	120
cctcgacgcg	ctcgacgagg	tccagggctc	cggggggccg	gcgctagccg	agggaccctt	180
cctcgacgtc	ctccgctccg	cgcaggaggc	gatcgtgctg	ccgcccgttc	tggccatcgc	240
ggtgcgcccc	cgcccgggag	tttgggagta	cgtccgctgc	aacgttcacg	agctcagcgt	300
cgagcagctc	acagtctcgg	agtacctccg	cttcaaggag	gagcttgctg	acggccagca	360
caatgatccc	tacgttctcg	agcttgactt	cgagccgttc	aatgtctcag	tcccacgccc	420
aaatcgggtc	tcatctattg	gaaacgggtg	gcagttccct	aaccgacact	tgtcctcaat	480
catgttccgc	aacagggatt	gcttggagcc	cctgttggat	ttcctccgtg	gccaccggca	540
caaggggcat	gttatgatgc	ttaatgatag	aatacaaagc	ttggggaggc	ttcagtcctg	600
gctgacaaaa	gctgaggagc	acttgtcaaa	gctccctgct	gacacaccat	actcacaatt	660
tgtttataaa	tttcaagagt	ggggcctgga	gaaagggttg	ggtgatacag	caggacatgt	720
tttgaaatg	atccatctcc	ttctagacat	cattcaggcg	ccagacccat	ctaccctaga	780
gaaattcttg	gggaggatcc	ccatgatatt	taacgttgtt	gtggtatccc	ctcatggata	840
ctttggtcaa	gctaattgat	taggcttgcc	agacacagga	ggacagatcg	tctatatact	900
ggaccaagtc	cgtgcactag	aaaatgagat	ggttctccgt	ttaaagaaac	aagggttga	960
tgtttcccca	aagattctca	ttgttactcg	gctgatacca	gatgcaaaag	gaacatcatg	1020
caatcagcgg	cttgagagaa	ttagtggaac	acagcatact	tacatattac	gagttccctt	1080
cagaaatgaa	aatgggatac	ttaagaaatg	gatatcaaga	tttgatgtgt	ggccatatct	1140
ggaaacattt	gctgaggatg	ctgctggtga	aattgctgct	gaattacaag	gtactccaga	1200
cttcataatt	ggaaactaca	gtgatggaaa	tcttggtggc	tcattgctat	cttacaagat	1260
gggaattacc	cagtgaaca	ttgctcatgc	tctggaaaag	actaagtatc	cagattcaga	1320
catatttttg	aagaatttgc	atgagaagta	ccatttctcc	tgccagttca	ctgctgatat	1380
aattgctatg	aacaatgctg	attttatcat	caccagcaca	taccaagaaa	ttgctggaag	1440
caaaaatact	gttgagacgt	atgagagtca	tactgccttt	actctgcctg	gtctgtaccg	1500
agttgtccat	gggatcgatg	tcttcgatcc	aaagttcaat	atagtctctc	ctggagctga	1560
catgtccata	tactttccac	ataccgagaa	ggccaagcga	ctcacctctc	ttcatggttc	1620
aatcgaaaat	ttgatattat	acccggagca	aaacgatgaa	cacattgggc	atctggatga	1680
ccggtcaaa	cccatcctct	tctccatggc	aagactcgac	aggggtgaaga	acataacagg	1740
gctgggtcgaa	gcttttgcta	agtgcgctaa	gctgaggga	ctggtaaacc	ttgtcgtcgt	1800
tgccgggtac	aatgatgtca	acaagttccaa	ggacagggaa	gagatcgccg	agatagagaa	1860
gatgcatgaa	ctcatcaaga	cccacaactt	gttcgggcag	ttccgctgga	tctctgccca	1920
gacaaacagg	gcccgtaacg	gcgagctcta	tcgctacatc	gctgataccc	atgggtgctt	1980
cgtagacgcg	gccttgatg	aagcgcttcg	tctcaccgtc	gttgaggcca	tgacctgtgg	2040
gcttcctact	ttcgcgacgc	tccatggagg	tccagctgag	atcatagagc	atggcgctct	2100
gggcttccac	attgaccctg	accacccgga	acaggctgtt	aatctgatgg	ccgacttctt	2160
cgaccggtgc	aagcaagacc	cagatcactg	ggtgaatata	tctggagcag	ggctgcagcg	2220
catatacag	aagtacacat	ggaagatata	ctcagagagg	ttgatgacac	tggccggggg	2280
ctacgggttc	tggaaagtac	tgtcgaagct	cgagaggctg	gagacgaggc	gctaccttga	2340
gatgttctac	atactgaagt	tccgcgagct	ggcgaagacc	gtgccgcttg	caattgacca	2400
accgcagtag	cttgcgcaac	tgcgactgcg	tagcacttgg	tacaagactg	aaacctgaag	2460
gaccttcagt	aatttaggag	cggcagacgg	tagccaataa	aatgtgccgg	agctgaactg	2520
gttttttatt	atgtacataa	tggcagatata	acaaaattac	tgaaggcagg	tgggttgcag	2580
ttgtgtgttc	gttactgttt	actgtattat	gtcaagctgt	cggctgcaat	ttctttgctg	2640
gcaagccgca	ggcactgggt	aagtgtctgt	aaatacatca	tattctgttg	acctgtgaaa	2700
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	cggccgc			