

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
22 July 2010 (22.07.2010)

(10) International Publication Number
WO 2010/083464 A2

PCT

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2010/021274

(22) International Filing Date:
15 January 2010 (15.01.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/145,439 16 January 2009 (16.01.2009) US
61/165,835 1 April 2009 (01.04.2009) US

(71) Applicant (for all designated States except US):
CEPHEID [US/US]; 904 Caribbean Drive, Sunnyvale,
CA 94089 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **SVANHOLM
BARRIE, Cecilia** [SE/SE]; Cepheid, P.O. Box 20045,
SE-161 02 Bromma (SE). **DELFOUR, Olivier** [FR/FR];
Cepheid, Vira Solelh, F-81470 Maurens-Scopont (FR).
PERSING, David, H. [US/US]; Cepheid, 904 Caribbean
Drive, Sunnyvale, CA 94089 (US).

(74) Agents: **SCARR, Rebecca, B.** et al.; Michael Best &
Friedrich LLP, 100 East Wisconsin Avenue, Suite 3300,
Milwaukee, WI 53202-4108 (US).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ,
CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO,
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP,
KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD,
ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI,
NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD,
SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR,
TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ,
TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE,
ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,
MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished
upon receipt of that report (Rule 48.2(g))

(54) Title: METHODS OF DETECTING CERVICAL CANCER

(57) Abstract: Methods of detecting cervical dysplasia, such as cervical dysplasia likely to progress to carcinoma in a sample of human cervical cells, are provided. Methods of detecting changes in expression of one or more microRNAs or mRNAs associated with cervical dysplasia or cervical cancer are also provided. Compositions and kits are also provided.



WO 2010/083464 A2

METHODS OF DETECTING CERVICAL CANCER

[001] This application claims priority to U.S. Provisional Application No. 61/145,439, filed January 16, 2009, and U.S. Provisional Application No. 61/165,835, filed April 1, 2009, which are incorporated by reference herein in their entireties for any purpose.

1. BACKGROUND

[002] Cervical cancer is the second most common cause of cancer-related mortality in women worldwide. Epidemiological and laboratory studies suggest a key role for human papillomavirus (HPV) in cervical carcinogenesis (Walboomers, J.M. *et al.* (1999) *J. Pathol.* 189:12-19; Zur, H.H. (2002) *Nat. Rev. Cancer* 2:342-350). Importantly, however, HPV infection alone is not sufficient for cervical carcinogenesis, and additional steps occur over years or decades following initial infection. Most HPV infections resolve spontaneously, but if an oncogenic (high risk) HPV infection persists, there may be progression to a high grade cervical dysplasia or cervical cancer. (Nobbenhuis, M.A. *et al.* (2001) *Lancet* 358:1782-1783). High risk HPVs include HPV-16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, and 68, with HPV-16 and 18 accounting for up to 70% of cervical cancers worldwide.

[003] The Papanicolaou (Pap) smear has become the most commonly used method to screen for cervical dysplasia. It has been a success and the incidence of cervical cancer has been dramatically reduced. However, cytology screening programs have limitations, especially limited sensitivity, estimated at only 51% (Nanda K. *et al.* (2000) *Ann. Intern. Med.* 132:810-819), and repeated tests are therefore necessary. In addition, a high-quality cytology screening program requires highly-trained personnel. Furthermore, although cytological screening programs have reduced the incidence of squamous cervical cancer (SCC), the incidence of cervical adenocarcinoma (AC) has continued to increase. The reason for this is unclear, but it may, in part, be due to difficulties detecting the precursor form of AC using conventional screening methods. (Bray, F. B. *et al.* (2005) *Cancer Epidemiol. Biomarkers Prev.* 14:2191-2199).

[004] HPV DNA testing can be more sensitive than cytologic testing in detecting high-grade cervical dysplasia. However, HPV testing often has lower specificity than cytologic testing since most HPV infections are transient in nature. (Koliopoulous, G. M. *et al.* (2007) *Gynecol. Oncol.* 104:232-246). In order to improve the clinical specificity of

the molecular HPV tests, a number of molecular markers associated with cervical cancer precursor lesions (*i.e.* Cervical Intra-epithelial Neoplasia (“CIN”) grades 1, 2 and 3) have been evaluated. (See *e.g.*, Altieri D.C. (2003) *Nat Rev. Cancer* 3:46-54; Li C. *et al.* (2007) *Mod. Pathol.* 20:242-247; Andersson, S. *et al.* (2006) *Br. J. Cancer* 95:331-338; Martin, C.M. *et al.* (2006) *Expert Rev. Mol. Diagn.* 6:217-229; Branca, M. *et al.* (2006) *Int. J. Gynecol. Pathol.* 25:383-392; Harris C.P. *et al.* (2003) *Genes Chromosomes Cancer* 36:233-241). However, there remains a need for molecular markers in cervical dysplasia which indicate a high risk of progression to cancer.

2. SUMMARY

[005] In some embodiments, the disclosure relates to a method for detecting the presence of cervical dysplasia in a human cervical sample. In some embodiments, the method comprises hybridizing nucleic acids of the sample with a polynucleotide that is complementary to a first target RNA in the sample or to a complement thereof and detecting a complex comprising polynucleotide hybridized to at least one nucleic acid selected from the first target RNA, a DNA amplicon of the first target RNA, and a complement of the first target RNA. In some embodiments, the first target RNA is capable of specifically hybridizing to a sequence selected from SEQ ID NO.:1 to 41 and 133 to 211. In some embodiments, the first target RNA comprises a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NO.:1 to 41 and 133 to 211. In some embodiments, the first target RNA comprises at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388. Detection of a level of the first complex that is greater than a normal level of the first complex indicates the presence of cervical dysplasia in the sample.

[006] In some embodiments, the sample comprises RNA that has been separated from DNA. In some embodiments, the first target RNA, in its mature form, comprises fewer than 30 nucleotides. In some embodiments, the first target RNA is a microRNA.

[007] In some embodiments, the method further comprises hybridizing nucleic acids of the sample with a second polynucleotide that is complementary to a second target RNA in the sample or to a complement thereof, wherein the second RNA is different from the first target RNA, and detecting a second complex comprising the second polynucleotide hybridized to at least one nucleic acid selected from the second target RNA, a DNA amplicon of the second target RNA, and a complement of the second target RNA.

[008] In some embodiments, the second target RNA, in its mature form, comprises fewer than 30 nucleotides. In some embodiments, the second target RNA is a microRNA. In other embodiments, the second target RNA is an mRNA.

[009] In some embodiments, a polynucleotide is complementary to at least 10 contiguous nucleotides of the target RNA or its complement. In some embodiments, a polynucleotide is complementary to at least 15 contiguous nucleotides of the target RNA or its complement. In some embodiments, a polynucleotide comprises a region that is not complementary to the target RNA or its complement.

[0010] In some embodiments, a polynucleotide is provided. In some embodiments, a polynucleotide comprises a first region, wherein the first region comprises a sequence of at least 8 contiguous nucleotides, at least 9 contiguous nucleotides, at least 10 contiguous nucleotides, at least 11 contiguous nucleotides, or at least 12 contiguous nucleotides that are identical to a sequence of at least 8 contiguous nucleotides, at least 9 contiguous nucleotides, at least 10 contiguous nucleotides, at least 11 contiguous nucleotides, or at least 12 contiguous nucleotides of one or more of SEQ ID NOs: 1 to 7, 9 to 37, and 133 to 201.

[0011] In some embodiments, a polynucleotide comprises a first region, wherein the first region comprises a sequence of at least 8 contiguous nucleotides, at least 9 contiguous nucleotides, at least 10 contiguous nucleotides, at least 11 contiguous nucleotides, or at least 12 contiguous nucleotides that are complementary to a sequence of at least 8 contiguous nucleotides, at least 9 contiguous nucleotides, at least 10 contiguous nucleotides, at least 11 contiguous nucleotides, or at least 12 contiguous nucleotides of one or more of SEQ ID NOs: 1 to 7, 9 to 37, and 133 to 201.

[0012] In some embodiments, a polynucleotide comprises a first region, wherein the first region comprises a sequence of at least 8 contiguous nucleotides, at least 9 contiguous nucleotides, at least 10 contiguous nucleotides, at least 11 contiguous nucleotides, or at least 12 contiguous nucleotides that are identical to a sequence of at least 8 contiguous nucleotides, at least 9 contiguous nucleotides, at least 10 contiguous nucleotides, at least 11 contiguous nucleotides, or at least 12 contiguous nucleotides of one or more of SEQ ID NOs: 345 to 388.

[0013] In some embodiments, a polynucleotide comprises a first region, wherein the first region comprises a sequence of at least 8 contiguous nucleotides, at least 9

contiguous nucleotides, at least 10 contiguous nucleotides, at least 11 contiguous nucleotides, or at least 12 contiguous nucleotides that are complementary to a sequence of at least 8 contiguous nucleotides, at least 9 contiguous nucleotides, at least 10 contiguous nucleotides, at least 11 contiguous nucleotides, or at least 12 contiguous nucleotides of one or more of SEQ ID NOs: ~~345 to 388~~.

[0014] In some embodiments, a polynucleotide comprises a detectable label. In some embodiments, a polynucleotide comprises a FRET label. In some embodiments, at least 8, at least 9, at least 10, at least 11, or at least 12 contiguous nucleotides of the polynucleotide are complementary to a target RNA. In some embodiments, a polynucleotide further comprises a second region that is not complementary to the target RNA.

[0015] In some embodiments, a composition is provided that comprises a plurality of polynucleotides, wherein at least one, at least two, at least three, or at least five of the polynucleotides comprise a first region comprising a sequence of at least 8 contiguous nucleotides that is identical or complementary to a sequence of at least 8 contiguous nucleotides of one or more of SEQ ID NOs: 1 to 7, 9 to 37, 133 to 201, and 345 to 388.

[0016] In some embodiments, kits are provided. In some embodiments, a kit comprises a composition described above. In some embodiments, a kit further comprises at least one polymerase. In some embodiments, a kit further comprises dNTPs.

[0017] Further embodiments and details of the inventions are described below

3. BRIEF DESCRIPTION OF THE FIGURES

[0018] FIG. 1 shows an exemplary electropherogram obtained on an Agilent Bioanalyser 2100 to assess the quality of total RNA purified as described in Example 1. Total RNA from cell line CaSki is shown.

[0019] FIG. 2 provides analysis by agarose gel electrophoresis under denaturing conditions of the quality of total RNA purified as described in Example 1 from cell lines CaSki, SW756, ME180, SiHA, C-4I, and C-4II.

[0020] FIG. 3A and 3B show the log₂ fold-changes $\pm$ SD of certain mRNAs in tumor and normal samples relative to the Ambion normal sample, as discussed in Example 4. For each pair of bars, the left bar represents tumor samples and the right bar represents normal samples.

[0021] FIG. 4 shows relative log₂ fold changes of certain mRNAs in liquid PAP samples, as discussed in Example 5.

[0022] FIG. 5 shows relative log₂ fold changes $\pm$ SD of miR-205 in cervical tumor samples and normal tissue, as discussed in Example 6. The left bar represents SCC samples, the middle bar represents ADC samples, and the right bar represents normal samples.

[0023] FIG. 6 shows relative log₂ fold changes $\pm$ SD of miR-1290 in cervical tumor samples and normal tissue, as discussed in Example 6. The left bar represents tumor samples and the right bar represents normal samples.

4. DETAILED DESCRIPTION

4.1. Detecting cervical dysplasia

4.1.1. General methods

[0024] Methods of measuring levels of microRNA species disclosed herein are provided, wherein elevated levels of the microRNA species is indicative of cervical dysplasia. In some embodiments, methods are presented for detecting human cervical dysplasia, such as cervical dysplasia likely to progress to carcinoma. In some embodiments, the method comprises detecting an above-normal level of at least one target RNA that is capable of specifically hybridizing to a sequence selected from SEQ ID NOs: 1 to 41 and 133 to 211. In some embodiments, the method comprises detecting an above-normal level of at least one target RNA, wherein at least one target RNA comprises at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, or at least 24 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388. In some embodiments, the method comprises detecting an above-normal level of at least one target RNA that comprises a sequence that is complementary to at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, or at least 24 contiguous nucleotides of a sequence selected from SEQ ID NO.:1 to 41 and 133 to 211. In some embodiments, the target RNA, in its mature form, comprises fewer than 30 nucleotides. The target RNA, in some embodiments, is a microRNA.

[0025] In the present disclosure, “a sequence selected from” encompasses both “one sequence selected from” and “one or more sequences selected from.” Thus, when “a sequence selected from” is used, it is to be understood that one, or more than one, of the listed sequences may be chosen.

[0026] Detection of a level of target RNA that is greater than a normal level of target RNA indicates the presence of cervical dysplasia in the sample. In some embodiments, the detecting is done quantitatively. In other embodiments, the detecting is done qualitatively. In some embodiments, detecting a target RNA comprises forming a complex comprising a polynucleotide and a nucleic acid selected from a target RNA, a DNA amplicon of a target RNA, and a complement of a target RNA. In some embodiments, the level of the complex is then detected and compared to a normal level of the same complex. The level of the complex, in some embodiments, correlates with the level of the target RNA in the sample.

[0027] “Cervical dysplasia,” which is also known as cervical intraepithelial neoplasia (“CIN”), corresponds to precancerous changes of the cervix that are evidenced by an abnormal growth on the surface of the cervix. Cervical dysplasia is divided into three categories: CIN 1, which is mild dysplasia in which only a few cells are abnormal; CIN 2, which is moderate to marked dysplasia in which the abnormal cells involve about one-half of the thickness of the surface lining of the cervix; and CIN 3, which includes severe dysplasia to carcinoma-in-situ (*i.e.*, precancerous cells limited to the top epithelial layer of the cervix). CIN 3 is unlikely to regress spontaneously, and if left untreated, can penetrate the basement membrane and become an invasive carcinoma.

[0028] Table 1, below, lists 41 hybridization probes that have been found to be complementary to, and hybridize with, target RNAs in cancer cells. These target RNAs were detected at elevated levels in certain human cervical cell lines that were assayed using microarrays (Example 1). Thirty-six of the probes are believed to be complementary to, and hybridize with, target RNA species that are expressed in human cells. The other five probes are complementary to, and hybridize with, publicly known microRNAs that have been deposited by others into miRBase (<http://microrna.sanger.ac.uk/>; see Griffiths-Jones S. *et al.* (2007) *Nucl. Acids Res.* 36:154-158): hsa-miR-423-5p, hsa-miR-765, hsa-miR-92b*, hsa-miR-663, and hsa-miR-936). However, to the knowledge of the inventors, these five known microRNAs have not been disclosed to have utility for detecting cervical dysplasia.

[0029] Table 11, below, lists hybridization probes that have been found to be complementary to, and hybridize with, target RNAs in cancer cells. These target RNAs were detected at elevated levels in certain human clinical cervix samples that were assayed using microarrays (Example 3). Seventy-three of the probes are believed to be

complementary to, and hybridize with, target RNA species that are expressed in human cells. Four of those 73 probes were also detected at elevated levels in certain human cervical cell lines that were assayed using microarrays (Example 1), and are also in Table 1 (836-R4-1, 3371-L4-1, 9053-R3-1, and 9691-L4-1). The remaining 19 probes are complementary to, and hybridize with, publicly known microRNAs that have been deposited by others into miRBase (<http://microrna.sanger.ac.uk/>; *see* Griffiths-Jones S. *et al.* (2007) *Nucl. Acids Res.* 36:154-158). One of those 19 probes was also detected at elevated levels in certain human cervical cell lines that were assayed using microarrays (Example 1), and is also in Table 1 (miR-765). To the knowledge of the inventors, at least 11 of those microRNAs, miR-1246, miR-1290, miR-1308, miR-1826, miR-200c, miR-451, miR-483-5p, miR-491-3p, miR-494, miR-720, and miR-765 have not been disclosed to have utility for detecting cervical dysplasia.

[0030] Table 28, below, lists 44 microRNAs that may be present at elevated levels in certain human cervical cancer cells lines and/or human clinical cervix samples. Some microRNAs in Table 28 are isomirs of one another. In some embodiments, two isomirs have a common core sequence with one or both ends varying by one to three nucleotides. For example, AGCCGCTCTTCTCCCTGCCCACA (SEQ ID NO: 355) and AGCCGCTCTTCTCCCTGCCCACA (SEQ ID NO: 356) are isomirs. Similarly, CCCGGAGAGCGGAGCACAACACA (SEQ ID NO: 346) and CCCGGAGAGCGGAGCACAAC (SEQ ID NO: 347) are isomirs. When multiple isomirs are listed in Table 28, one or more than one of the isomirs may be present at elevated levels in a cervical dysplasia. In some embodiments, a method comprises detecting multiple isomirs with a single probe. Detection of an elevated level of one or multiple isomirs is considered to be indicative of cervical dysplasia.

[0031] For convenience of reference herein, and not by way of limitation, some “target RNA” species are denominated “microRNAs” in the tables set forth herein and Example 1. In some embodiments, the target RNA is a single mature microRNA capable of specifically hybridizing to a hybridization probe set forth in Table 1 or Table 11. In some embodiments, a target RNA is a single mature microRNA that comprises a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NO.:1 to 41 and 133 to 211. In some embodiments, a target RNA is a single mature microRNA that comprises at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388. In some embodiments, target RNA

may include a plurality of target RNAs, all of which are capable of specifically hybridizing to a single complementary probe sequence (for example, when two or more target microRNAs are isomirs). In some embodiments, the so-denominated “microRNA” is one or more RNA species capable of specifically hybridizing to the respective hybridization probe, such that one or more target RNAs do not meet canonical definitions for mature microRNAs. In some embodiments, a target RNA is an mRNA.

[0032] Mature human microRNAs are typically composed of 17-27 contiguous ribonucleotides, and often are 21 or 22 nucleotides in length. The sequences of some target microRNAs that can be detected in accordance with the present disclosure can be found within the pre-microRNA sequences shown in Table 2 (SEQ ID NOs: 42 to 82) and Table 12 (SEQ ID NOs: 226 to 314). The sequences of some microRNAs are shown in Table 28. Further, in some embodiments, a microRNA comprises at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, or at least 26 contiguous nucleotides of a sequence in Table 28 (SEQ ID NOs: 345 to 388). The sequences of the 23 publicly known mature microRNAs, obtained by query of miRBase, are also shown below in Table 3, along with the sequences of other previously known microRNAs that, in some embodiments, can be detected in the methods described herein.

[0033] While not intending to be bound by theory, mammalian microRNAs mature as described herein. A gene coding for a microRNA is transcribed, leading to production of a microRNA precursor known as the “pri-microRNA” or “pri-miRNA.” The pri-miRNA can be part of a polycistronic RNA comprising multiple pri-miRNAs. In some circumstances, the pri-miRNA forms a hairpin with a stem and loop, which may comprise mismatched bases. The hairpin structure of the pri-miRNA is recognized by Drosha, which is an RNase III endonuclease protein. Drosha can recognize terminal loops in the pri-miRNA and cleave approximately two helical turns into the stem to produce a 60-70 nucleotide precursor known as the “pre-microRNA” or “pre-miRNA.” Drosha can cleave the pri-miRNA with a staggered cut typical of RNase III endonucleases yielding a pre-miRNA stem loop with a 5' phosphate and an approximately 2-nucleotide 3' overhang. Approximately one helical turn of the stem (about 10 nucleotides) extending beyond the Drosha cleavage site can be essential for efficient processing. The pre-miRNA is subsequently actively transported from the nucleus to the cytoplasm by Ran-GTP and the export receptor Exportin-5.

[0034] The pre-miRNA can be recognized by Dicer, another RNase III endonuclease. In some circumstances, Dicer recognizes the double-stranded stem of the pre-miRNA. Dicer may also recognize the 5' phosphate and 3' overhang at the base of the stem loop. Dicer may cleave off the terminal loop two helical turns away from the base of the stem loop leaving an additional 5' phosphate and an approximately 2-nucleotide 3' overhang. The resulting siRNA-like duplex, which may comprise mismatches, comprises the mature microRNA and a similar-sized fragment known as the microRNA*. The microRNA and microRNA* may be derived from opposing arms of the pri-miRNA and pre-miRNA. The mature microRNA is then loaded into the RNA-induced silencing complex ("RISC"), a ribonucleoprotein complex. In some cases, the microRNA* also has gene silencing or other activity.

Table 1

Array probe	Array probe sequence (5' to 3', without linker)	probe SEQ ID NO:	fold-changes vs. normal Cervix							C33A HPV neg
			CaSki HPV16	SiHa HPV16	sw756 HPV18	C4-I HPV18	C4-II HPV18	ME180 HPV68 metastasis		
3371-L4-1	TTTCCCTTTCTCCCTCCACACCCCATGACTCCCAACACTTGAG	1	6.14	8.9	6.9	5.43	12.37	4.94	12.18	
4315 D-R4-1	GGAAAGTCAGCCCCCAGCGCCCCCGGAGTTCTTGG	2	6.91	4.11	3.25	7.01	12.19	8.37	14.35	
4988-R4-1	CTCCTCTCCCGCTTTTGGATACCAACACTGGAC	3	3.46	2.72	2.89	4.59	6.32	4.15	10.48	
6647-R2-1	CTCAGCCCCAGCTGGAGAATTTTCCCTCATTA	4	4.49	5.48	4.91	3.05	8.35	2.42	4.16	
9053-R3-1	TTCTGCCCTCCAATCCCCGGGCTCCACCGCC	5	5.40	2.51	3.69	5.76	10.31	5.55	24.91	
6803-R3-1	GCTCCCTCTCTGTTGGACTCACCCAAAGAT	6	2.68	2.00	2.14	3.86	5.38	4.19	19.50	
9691-L4-1	AATCATCCATTTCATCCGCATCTCCCTCTTGGCCCCCTTGC	7	2.83	2.74	2.49	3.82	4.69	5.30	11.57	
miR-423-5p	AAAGTCTCGCTCTCTGCCCCCTCA	8	4.53	3.01	2.89	5.15	7.27	5.86	11.22	
6584-L1-1	TCGGCCCTGCCTCTCTCTCTCT	9	2.16	1.9	1.9	2.69	5.23	2.36	4.03	
7421-R2-1	TAAAGAGACTTCTCCACTGCCAGAGATCT	10	2.46	3.01	2.99	1.5	3.32	2.41	3.5	
8016-L3-1	TCAGCGCAACAAGCCCGCAGTCAACCCCTCT	11	3.31	1.8	1.8	3.41	5.71	3.74	9.14	
8433-L3-1	AAATGGCTCCTTTCCCTTTCCCTCCACCG	12	2.40	2.60	1.6	2.68	4.59	2.61	5.71	
4361-R3-1	CGTCTCCCTCCCTCATGTGC	13	2.21	2.84	2.67	-	4.72	3.52	10.04	
10010 H-L4-1	ACAGGCTACTTTTCAGCAAAATATGTCCATCCT	14	3.67	3.56	1.6	2.92	4.44	-	3.1	
12223-L4-1	CCCAGAGACTACAGACAGAGATGTTCTTCTCCCTCTA	15	-	2.62	2.56	3.73	-	-	24.42	
4610-R3-1	GCCCTCTGGCCCCCTGCCTAATTGGCTGC	16	1.8	-	1.6	2.2	3.56	2.99	6.81	
5192-L3-2	CATTTTCCCTTCTCTCTCTATATCAGCAA	17	5.45	5.05	7.11	-	7.21	2.99	6.59	
5782-L3-1	GATTCAGCCCTTCCCTCC	18	-	2.20	1.5	2.57	-	2.84	5.04	
5836-R3-2	CATTAAAGCCCAATATACAGCAGCGCCCAATTC	19	2.05	7.58	-	-	-	2.61	2.97	
6183-R3-1	GATTCACATTTCTTAATGACTTTCCTCTCT	20	2.68	2.12	2.72	-	-	-	2.74	
6287-L3-2	GCCCCGCCCACTTTCCGGGGCTCACCTGGC	21	2.20	-	1.5	2.70	4.43	4.42	5.90	
6522-L3-1	GGTTTGCCCTCTAATGTGGTAATAGATGTCAAT	22	-	2.58	0.9	-	4.69	2.78	3.76	
6752-R1-1	CCCTCCTTTCCCAACCTCAGT	23	-	3.58	2.99	2.66	5.44	2.26	5.13	
6825-R3-1	CTCAGCTGTTCCCGGTGCCAG	24	-	-	-	2.13	5.17	2.68	2.94	
6930-R3-1	ATTAATCCTTCTCTCCCTCTG	25	-	3.12	2.52	5.01	5.16	5.17	20.54	
7352-R3-2	GCCCCTGCCAGAATCCTCTAACAGCTCTAATTGG	26	-	-	1.6	4.29	5.18	14.54	8.13	
7356-L2-1	ACCGGACATAGCCTCGCCCC	27	2.14	1.9	-	2.46	4.60	2.76	5.57	
7384-R3-1	CTCGCAAGGATCTCCTTCATCCCTCCCCA	28	-	2.26	0.7	1.7	3.27	2.18	3.96	
7764-R3-2	CCCTCTCTGCCTCTCTCATCACCAATAACAGAC	29	-	2.15	2.12	2.41	4.33	3.00	7.07	
8075-L3-1	CCAGCTACACCTCCACGCA	30	2.90	-	2.67	4.02	5.12	-	2.8	
8316-R3-1	ATCAGGGTATCCTCTCCCCA	31	-	-	1.5	2.59	2.84	2.46	9.39	
836-R4-1	AAATAATCATTCGAAATGGTTCTCCCTGCTATGATTCAC	32	2.56	3.24	2.25	-	-	3.60	15.95	
8433 D-R4-1	CCCGAGCCCGCGCCCTGTGTGTGCTCCGCTCTCC- GGGAAATGCCATCACTAAT	33	2.89	2.08	-	3.53	6.19	3.10	8.85	
8724-R3-1	GCCAAAGTTGGAACCTCTCCCTGCCAGCATCAC	34	-	1.6	1.7	3.31	3.92	3.60	11.08	
8832-R4-1	TCTGGAGTACCACCTGTTTTTCCCCCACT	35	-	6.56	2.25	-	4.55	2.21	2.7	
9349-R3-1	GTGATGCAGAGGACTTCTGCTCCAGGTCTC	36	2.10	3.02	1.5	1.9	-	2.80	9.84	
9733-L3-1	AAGGCTGTCCCTCACAGACTTCCCCACCCCT	37	-	2.50	1.5	2.23	4.46	4.20	4.09	
miR-663	GCGGTCCGCGCGGCCCGCCT	38	2.59	-	-	3.09	4.55	3.99	6.49	
miR-765	CATCACCTTCTTCTCTCTCA	39	2.48	-	3.16	3.42	5.46	2.96	21.76	
miR-92b*	CACCTGCACCGGCTCCGCTCCCT	40	2.06	2.17	1.8	-	4.64	3.35	7.45	
miR-936	CTGCGATTCTCCCTCTACTGT	41	-	2.87	2.27	-	3.82	3.03	6.42	

Table 2

Pre-microRNA Candidate	chrom. Location	Pre-microRNA sequences	pre-micro RNA SEQ ID NO:
03371-L	18q21.33	CTCAAGTGTGGGAGTCATGGGGTGTGGAGGGAGGAAAGGATATTTTGTTCTTTGTCTATACATTTCTAGATTTC TATCAGTTGGG	42
12694-R	1q22	GGGACGTGGCCCTCCCCCGGAGCGGGACTCCAAGAACTCCGGGGGGCGCTGGGGGCTGACTTTCC	43
04988-R	14q24.3	CTTTTCTCTCTGCGAAACCTTGCTTGACTTCATGTCAGTGTGGTATCCAAAGACGGGAGGAGGAG	44
06647-R	1q23.3	CTCAGTATCTCAGCTTGGAACTGACCTCGTTAATTTAATGAGGGAAAAATCTCCAGCTGGGCTGAG	45
09053-R	Xq27.3	GGAAGGGCACTGTCTCTGATCCCAAGGGCTGTCAATTTCCCGAGGGCTGGTGAGCCCCGGGATTGAGGGCAAGAA GCCACGCC	46
06803-R	22q12.3	GCCACCTTTTCATGGTGAGGATGCCTGCCACCTTCAGGATCACATCTTTGGGTGAGTCCAAACCAGAGAGGGAGC	47
09691-L	14q24.3	GCAAGGGGCCAAGAGGGAGATGCGGATGAAATGGATGATTAATGGGTCATCTCCTGTAGTTAATTTCTAGATCTCTT GT	48
miR-423-5p	17q11.22	ATAAAGGAAGTTAGGCTGAGGGGCAGAGCGAGACTTTTCTATTTTCCAAAAGCTCGGTCTGAGGCCCTCAGTCTTGCT TCCTAACCCGCGC	49
06584-L	12q24.23	GCTTGGTGAGAGGAGGAGGCGGCGGACCCGCCCTGTCTGCCATCTGGTCCCCCTCCCCCTCCCTCCTCTC ATTGC	50
07421-R	12p13.31	TGAAGAAATCTTCTGGATGACTGACCAAGAGGCTATTCAAGATCTCTGGCAGTGGAGGAATCTCTTTA	51
08016-L	12q21.1	AGAGGGTGACTGCGGGCTGTGCGCTGAAGATTTACAATGTACTTCTTCGAGCGGCTCAGCAACCCCTCT	52
08433-L	17q25.3	CGGTGGAGGAAAGGGAAAGGAGCCATTTCTGCTGCACATCAGTCAGTGCCTGCGCCCTCCCTCCCTCCGCCG	53
04361-R	Xp11.22	TGCTGGAGTAAGGGTTTCTGAAGCTGTGCCATGGCCACATGTGCACATGAGGGAGGGAGAGCTGAGGCTAGCA TGGTCTCTCT	54
12709-L	7q32.1	AGGATGGACATATTTGCTGAAAGTAGCCTGTGCAATTAATGGTATGGAAGTTTAAAAATGGTGTCTCTCT TAGAGGGAGAAGAAACAATCTGTCTGATGCTCTCTGGGATGGCTTAATACAGATAGCATTTGCTCTCCATTTCTG	55
04610-R	8p12	GCCACGTTAATGGTCTCAACCTACATTAAGCTGTGCAATGCAGCCCAATTAGGCAGGGGCCAGAGGGC	57
05192-L	5q34	GTCCTTCTGATATAGAGGAAGGAGGAAATAAGCGCATTAAGTCTCTTTATTAAAGAGTTATTTCAGCATGAC	58
05782-L	5q35.1	GGGGAAGGGCTGGAATCATCGTGGTTGGAACAGTTAAAGAACTCTGTTCAGCCCCAGCCCAAGGCTCCC	59
05836-R	11q23.3	GCCATGGCCCTCCATAGTTTCTGTAGCCCCCTTGGTTCCTCAAGATAGTTTGGAAATGGGCGTGTGTGATAATGGGGG TTAATGGT	60
06183-R	12q21.33	GATTCATCTATCTTTTCTCTCTTCAAAGATACTGTGAAGCACTTAAGGAGGGGAAAGTCATTAAAGAAAATGGAAATC AGCAGCAGGTGAGCCCGAAAGGTGGGCGGGCAGGGCGCTCCAGCCCAACCCCGGATCTGGTGACGCT	61
06287-L	1p34.1	AATGACATCTATTACCACTATAGAGCAACCCATAACAATCCCTTATAGAATGTTTGCTCAATTTGGTTAATTGTCATT CCCCAGTTCCCATAGCAACTGGGCTGTAGCAGCCAGAACTTGATTGAGCCCCAGCACTGGCCCCGACTGAGGTGGGGAA AGAGGGG	62
06522-L	5q23.2	CAAAATACATCTGTTATGCTTCTATTGTTAGACAATCTGGCACCGGGAACAGCTGAGCAGAAGGATTG	63
06752-R	Xq13.1	TGTCATTTGTCCATTTCTCTGACCCAGTGGTATCTGCAAGATCAGAGGGGAGAGGATTAATGTCA	64
06825-R	9q31.1	GCCTCTGTGGCATGATATAATCAGCTTTGATAGGCAGAGGCTGAGGCTGTTTTTCCAAATTAGAGCTGTTAGAGGATTCTG GCAGGGGC	65
06930-R	9p21.3	GGGGCGAGGCTATGTGCGGTGGCAGCCCCGGATGGCCCGGAGTAACGGGACGTCGCCCGGGAGCCTT	66
07352-R	1q25.2	CTTCCCCC	67
07356-L	8q24.3	GGCATTCCTCTTGTTCCCTTCTCCTCTCTGGGAGGGATGAAGAGATCCTTTGCGAGAGGCATGTT	68
07384-R	12q12	TGCTATCTCGCCTCACACATCAACACAGTGCCAGACAGATTCGACTGCAAAGTCTGTTATTGGTGATGAGAGAGGCAGA GAGGGCA	69
07764-R	5q11.2		70

Table 2

Pre-microRNA Candidate	chrom. Location	Pre-microRNA sequences	pre-microRNA SEQ ID NO:
08075-L	10q22.1	CAGCTGGCCTGGTGGCCCTGGTGGAGGTGTAGCTGGGCTCTGACCCAGCTCCTCAAACAGGTTCCATATGGCCCTCCGGCTG	71
08316-R	14q24.3	GTCAGGCTGCTGATTCTCTTACACAGATGCCAGTAAGAACAAAGGCATCACGTGGGAGAGGATACCCCTGATAAATAAGCCATTCCAAACCATCTCTGATTGGCTGTGAGTGGCAGAAATCATTACCCGTGGTGAATCATAGCAGGGAGAACCA	72
00836-R	3q26.2	TTTGGAAATGATTATT	73
12730-R	17q25.3	CCGGCTCGGCCCGCGCTCTCCAGCTCCCGGCTCCTTTTAGTGATATAAATTAGTGATGGCATTTCCCGGAGA-GCGGAGCACACACACAGAGGGCGCGGCTCGGG	74
08724-R	15q23	GGCCAGAGAAGATGAAAAGCTGAAGTCTTCCCTTCCAGCTGAAGCCAGGTGTGATGCTGGCAGGGAGAGGTTCCAAGCTTGGCC	75
08832-R	9q33.2	TTCTGAGATATGATCTGTTGGATTCTCTACTACCAAAGTGGGGGAAAAACAGGTGGTACTCCAGAA	76
09349-R	21q22.11	GGACACCTCTGAACCCCAAGTGGAAATCCAACTGCCAGTTCTTCATCCGAGACCTGGAGCAGGAAGTCTCTGCATCACTGTGTTC	77
09733-L	15q23	AGGGGTGGGGAAGTCTGGTGAAGGACAGCCTTGAGTCAAAGGATGGTCAACCGCTCCATGTGGCTGCCCCACCCCTCCTCCGGCGTCCAGGGCGCGCGGGACCGCCCTCGTGTCTGTGGGGTGGATCCCGGGCGCGTGTTCCTGGTGGCCCGGCCATG	78
miR-663	20p11.1	TTTAGGCGCTGATGAAGTGGAGTTCAGTAGACAGCCCTTTTCAAGCCCTACGAGAACTGGGGTTTCTGGAGG-AGAAAGGAAGGTGATGAAGGATCTGTCTCGTGAGCCTGAA	79
miR-765	1q23.1	CGGGCCCGGGCGGGCGGGAGGACGGGACGCGGTGAGTGTGTTTTTCCCGCCCAATATTGCACCTCGTCCCGGCC	80
miR-92b*	1	TCCGGCCCCCGGCC	81
miR-936	10q25.1	TCAAGGCCCACTGGGACAGTAGAGGGAGGAATCGCAGAAATCACTCCAGGAGCAACTGAGAGACCTTGCTTCTACTTTACCA	82

Table 3

SEQ ID NO	microRNA	Mature microRNA Sequences (5' to 3') sequence
91	miR-423-5p (miR-423)	UGAGGGGCAGAGAGCGAGACUUU
92	miR-663	AGGCGGGGCGCCGCGGGACCGC
93	miR-765	UGGAGGAGAAGGAAGGUGAUG
94	miR-92b*	AGGGACGGGACGCGGUGCAGUG
95	miR-936	ACAGUAGAGGGAGGAAUCGCAG
389	miR-1246	AAUGGAUUUUUGGAGCAGG
390	miR-1290	UGGAUUUUUGGAUCAGGGA
391	miR-1308	GCAUGGGUGGUUCAGUGG
111	miR-142-3p	UGUAGUGUUUCCUACUUUAUGGA
392	miR-1826	AUUGAUCaucGACACUUCGAACGCAU
393	miR-195	UAGCAGCACAGAAUAUUGGC
394	miR-200c	UAAUACUGCCGGGUAAUGAUGGA
395	miR-451	AAACCGUUACCAUUACUGAGUU
396	miR-483-5p	AAGACGGGAGGAAAGAAGGGAG
397	miR-491-3p	CUUAUGCAAGAUUCCCUUCUAC
398	miR-494	UGAAACAUACACGGGAAACCUC
399	miR-720	UCUCGCUGGGGCCUCCA
400	miR-98	UGAGGUAGUAAGUUGUAUUGUU
401	miR-143	UGAGAUGAAGCACUGUAGCUC
100	miR-145	GUCCAGUUUUCCCAGGAUCCCU
402	miR-205	UCCUUCAUUCCACCGGAGUCUG
109	miR-21	UAGCUUAUCAGACUGAUGUUGA
403	miR-31	AGGCAAGAUGCUGGCAUAGCU
96	miR-9	UCUUUGGUUAUCUAGCUGUAUGA
97	miR-199a*	ACAGUAGUCUGCACAUUGGUUA
98	miR-199a	CCCAGUGUUCAGACUACCUGUUC
99	miR-199b	CCCAGUGUUUAGACUAUCUGUUC
101	miR-133a	UUUGGUCCCCUUAACCAGCUG
102	miR-133b	UUUGGUCCCCUUAACCAGCUA
103	miR-214	ACAGCAGGCACAGACAGGCAGU
104	miR-127	CUGAAGCUCAGAGGGCUCUGAU
105	miR-210	CUGUGCGUGUGACAGCGGCUGA
106	miR-182	UUUGGCAAUGGUAGAACUCACACU
107	miR-183	UAUGGCACUGGUAGAAUUCACU
404	miR-155	UAAUAGCUAAUCGUGAUAGGGGU
108	miR-146a	UGAGAACUGAAUCCAUGGGUU
110	miR-301	CAGUGCAAUAGUAUUGUCAAGC
112	miR-142-5p	CAUAAAGUAGAAAGCACUACU

Table 3

SEQ ID NO	microRNA	Mature microRNA Sequences (5' to 3') sequence
113	miR-194	UGU AACAGCAACUCCAUGUGGA
114	miR-215	AUGACCUAUGAAUUGACAGAC
115	miR-32	UAUUGCACAUUACUAAGUUGCA
116	miR-374b	AUAUAAUACAACCUGCUAAGUG
117	miR-933	UGUGCGCAGGGAGACCUCUCCC
118	miR-769-3p	CUGGGAUUCUCCGGGUCUUGGUU
119	miR-671	AGGAAGCCCUGGAGGGGCUGGAG
120	miR-934	UGUCUACUACUGGAGACACUGG
121	miR-935	CCAGUUACCGCUUCCGCUACCGC
122	miR-937	AUCCGCGCUCUGACUCUCUGCC
123	miR-938	UGCCCUUAAAGGUGAACCCAGU
124	miR-939	UGGGGAGCUGAGGCUCUGGGGGUG
125	miR-940	AAGGCAGGGCCCCCGCUCCCC
126	miR-941	CACCCGGCUGUGUGCACAUGUGC
127	miR-942	UCUUCUCUGUUUUGGCCAUGUG
128	miR-943	CUGACUGUUGCCGUCCUCCAG
129	miR-944	AAAUUAUUGUACAUCGGAUGAG
130	miR-708	AAGGAGCUUACAAUCUAGCUGGG
131	miR-874-5p	CGGCCCCACGCACCAGGGUAAG
132	miR-874-3p	CUGCCCUGGCCCGAGGGACCGA

[0035] In Table 1, the expression levels of target RNAs measured for each of the identified sample cell lines are expressed as fold-changes in expression relative to expression levels measured in normal human cervix total RNA (see Example 1). The expression levels of the target RNAs detected by the probes in Table 11, expressed as fold-changes for each of the clinical cervix samples, are shown in Table 10 (Example 3).

[0036] In some embodiments, target RNAs can be measured in samples collected at one or more times from a patient to monitor the status or progress of cervical dysplasia in the patient.

[0037] In some embodiments, a sample to be tested is obtained using one or more techniques commonly used for preparing Pap smears, *e.g.*, (i) endocervical swab, using a cotton applicator stick (or wire brush for endocervical specimens) advanced into the os of the cervix, with the stick gently rolled between the thumb and index finger; (ii) cervical

scrape, in which the longer end of a cervical spatula is inserted into the os of the cervix and pressed gently, with turning and scraping. In some embodiments, the sample to be tested is a cervical biopsy, such as a punch biopsy or cone biopsy. In some embodiments, the sample to be tested is from a loop excision, or LEEP, procedure.

[0038] The clinical sample to be tested is, in some embodiments, freshly obtained. In other embodiments, the sample is a fresh frozen specimen. In some embodiments, the sample is a tissue sample, such as a formalin-fixed paraffin embedded sample. In some embodiments, the sample is a liquid cytology sample.

[0039] Exemplary liquid cytology preservative solutions include, but are not limited to, ThinPrep™ PreservCyt™ solution (Hologic, Bedford, MA) and SurePath™ preservative solution (BD Diagnostics, NJ). Additional exemplary preservative solutions include, but are not limited to, RNAlater® (Ambion), formalin (e.g., 10% aqueous formalin), Universal Viral Transport Media (BD Diagnostics, NJ), M4, M4RT, PVA (polyvinyl-alcohol), PolyCyte (American Mastertech Scientific), Spray-Cyte cytological fixative (Becton-Dickinson), formaldehyde (e.g., 10% in phosphate buffer), NuFix Complete Collection Solution (QC Sciences), CarboFix (StatLab Medical Products), Cyto Jar (Surgipath Medical Industries), SED Fix (Surgipath Medical Industries), SprayFix (Surgipath Medical Industries), cytology fixative 50% alcohol solution (U.S. Biotex), Cyto-Prep (Wakefield), Cyto-Fix (Wakefield), PVA with zinc or copper, merthiolate-iodine-formaldehyde (MIF), sodium acetate-acetic acid-formalin (SAF), mercuric chloride-based Schaudinn's, zinc-based Schaudinn's preservative (Meridian Diagnostics, Inc.), EcoFix® (Meridian Bioscience), Parasafe®, Unifix, Proto-fix™, and STF.

[0040] In some embodiments, the clinical sample to be tested is obtained in conjunction with routine cytologic screening (e.g., by Pap smear), currently recommended for all women between the ages of 21 and 65, and women who are under 21 years old who have been sexually active for three years or more. In some embodiments, the sample to be tested is obtained from a woman who has a predisposition to develop cervical cancer, e.g., a woman who has tested positive for HPV infection, and especially positive for a high risk HPV type. In some embodiments, the clinical sample to be tested is obtained from women who have one or more of the following risk factors: multiparous, many sexual partners, first sexual intercourse at a young age, smoke

cigarettes, use of oral contraceptives, and a weakened immune system. In some embodiments, the clinical sample is obtained from women who have diagnostic signs or clinical symptoms that may be associated with cervical cancer, such as abnormal Pap tests, abnormal bleeding or visible cervical lesions.

[0041] In some embodiments, the methods described herein are used for early detection of cervical dysplasia in a sample of cervical cells, such as those obtained by routine Pap smear. In some embodiments, methods described herein can be used for early detection of cervical dysplasia in a sample of cervical cells, and to determine a likelihood that the detected cervical dysplasia will progress to cervical cancer.

[0042] Thus, in some embodiments, methods of the present disclosure can be used for routine screening of healthy women with no risk factors. In some embodiments, methods herein are used to (1) screen women who have a history of abnormal Pap smears and/or of assays showing infection by one or more HPV strains associated with the development of cervical cancer, (2) screen women with one or more of the above-described risk factors, (3) confirm a diagnosis made by cytology, histology or HPV assay, and/or further characterize a diagnosis made by cytology or histology.

[0043] In some embodiments, the methods described herein can be used to assess the effectiveness of a treatment for cervical cancer in a patient. In some embodiments, the target RNA expression levels are determined at various times during the treatment, and are compared to target RNA expression levels from an archival sample taken from the patient, *e.g.*, by Pap smear, before the manifestation of any signs of cervical dysplasia or cervical cancer or before beginning treatment. Ideally, target RNA expression levels in the normal Pap smear sample evidence no aberrant changes in target RNA expression levels. Thus, in such embodiments, the progress of treatment of an individual with cervical dysplasia or cervical cancer can be assessed by comparison to a sample of cervical cells from the same individual when she was healthy or prior to beginning treatment.

[0044] In some embodiments, a target RNA is capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7 and 8. In some embodiments, a target RNA is capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11

and 12. In some embodiments, a target RNA is selected from miR-1246, miR-1308, miR-491-3p, miR-1826, and miR-1290 (SEQ ID NOs: 208, 210, 205, 211, and 209), and target RNAs that are capable of specifically hybridizing to probes 13254-R5-1, 13252-L5-3, 13532-L5-2, 4440-L3-2, 6216-L1-1, and 6235-R5-2 (SEQ ID NOs: 194, 193, 172, 142, 151, and 153). In some embodiments, a target RNA is capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1, 5, 7, and 32. In some embodiments, a target RNA is capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1 to 41 set forth in Table 1 and SEQ ID NOs: 133 to 211 in Table 11. In some embodiments, a target RNA comprises at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388. In some embodiments, a target RNA comprises a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 1 to 41 and 133 to 211. In some embodiments, a target RNA, in its mature form, comprises fewer than 30 nucleotides. In some embodiments, a target RNA is a microRNA.

[0045] In embodiments in which the method comprises detecting expression of more than one target RNA, the expression levels of the plurality of target RNAs may be detected concurrently or simultaneously in the same assay reaction. In some embodiments, expression levels are detected concurrently or simultaneously in separate assay reactions. In some embodiments, expression levels are detected at different times, *e.g.*, in serial assay reactions.

[0046] The common, or coordinate, expression of target RNAs that are physically proximal to one another in the genome permits the informative use of such chromosome-proximal target RNAs in methods herein.

[0047] Table 2 identifies the chromosomal location of each of the 41 target RNAs capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1 to 41 in Table 1. Table 12 identifies the chromosomal location of each of the target RNAs capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 133 to 211 in Table 11. Thus, in some embodiments, the level of expression of one or more target RNAs located within about 1 kilobase (kb), within about 2 kb, within about 5 kb, within about 10 kb, within about 20 kb, within about 30 kb, within about 40 kb, and even within about 50 kb of the

chromosomal locations in Table 2 and Table 12 is detected in lieu of, or in addition to, measurement of expression of the respective tabulated target RNA in the methods described herein. See Baskerville, S. and Bartel D.P. (2005) RNA 11:241-247.

[0048] In some embodiments, in combination with detecting one or more target RNAs capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs:1 to 41 and 133 to 211 and/or detecting one or more target RNAs comprising at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388 and/or detecting one or more target RNAs that comprise a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs:1 to 41 and 133 to 211, methods herein further comprise detecting the level(s) of expression of at least one microRNA selected from miR-21, miR-31, miR-182, miR-183, miR-155, miR-9, miR-199a*, miR-199a, miR-199b, miR-205, miR-145, miR-133a, miR-133b, miR-214, miR-127, miR-210, miR-146a, miR-301, miR-142-5p, miR-194, miR-215, miR-32, miR-374b, miR-933, miR-769-3p, miR-671, miR-934, miR-935, miR-937, miR-938, miR-939, miR-940, miR-941, miR-942, miR-943, miR-944, miR-708, miR-874-5p, and miR-874-3p. In some embodiments, an increase in expression of one or more of these microRNAs, in combination with an elevated level of one or more target RNAs capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs:1 to 41 and 133 to 211 and/or an elevated level of one or more target RNAs comprising at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388 and/or an elevated level of one or more target RNAs that comprise a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs:1 to 41 and 133 to 211, is indicative of the presence of cervical dysplasia in a sample of human cervical cells.

[0049] In some embodiments, in combination with detecting one or more target RNAs capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs:1 to 41 and 133 to 211 and/or detecting one or more target RNAs comprising at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388 and/or detecting one or more target RNAs that comprise a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs:1 to 41 and 133 to 211, methods herein further comprise detecting in a sample of

human cervical cells the expression of at least one microRNA selected from miR-9, miR-199a*, miR-199a, miR-199b, miR-145, miR-133a, miR-133b, miR-214 and miR-127 where invasive squamous cell cervical carcinoma is implicated. In some embodiments, an increase in expression of one or more microRNAs selected from miR-9, miR-199a*, miR-199a, miR-199b, miR-145, miR-133a, miR-133b, miR-214 and miR-127, in combination with an elevated level of one or more target RNAs capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs:1 to 41 and 133 to 211 and/or an elevated level of one or more target RNAs comprising at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388 and/or an elevated level of one or more target RNAs that comprise a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs:1 to 41 and 133 to 211, is indicative of the presence of cervical carcinoma in a sample of human cervical cells.

[0050] In some embodiments, in combination with detecting one or more target RNAs capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs:1 to 41 and 133 to 211 and/or detecting one or more target RNAs comprising at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388 and/or detecting one or more target RNAs that comprise a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs:1 to 41 and 133 to 211, methods herein further comprise detecting in a sample of human cervical cells the expression of at least one microRNA selected from miR-210, miR-182 and miR-183 where human papilloma virus 16 ("HPV-16") is implicated. In some embodiments, an increase in expression of one or more of miR-210, miR-182 and miR-183, in combination with an elevated level of one or more target RNAs capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs:1 to 41 and 133 to 211 and/or an elevated level of one or more target RNAs comprising at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388 and/or an elevated level of one or more target RNAs that comprise a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs:1 to 41 and 133 to 211, is indicative of HPV16 infection in a sample of cervical cells.

[0051] In some embodiments, in combination with detecting one or more target RNAs capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NO.:1 to 41 and 133 to 211 and/or detecting one or more target RNAs comprising at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388 and/or detecting one or more target RNAs that comprise a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs:1 to 41 and 133 to 211, methods herein further comprise detecting in a sample of human cervical cells the expression of miR-146a in order to distinguish cervical cancer from pre-neoplastic lesions, e.g., HPV-infected cervical cells.

[0052] In some embodiments, a target RNA, in its mature form, comprises fewer than 30 nucleotides. In some embodiments, a target RNA is a microRNA.

[0053] In some embodiments, the methods further comprise detecting in a sample of human cervical cells the expression of at least one target RNA gene located in close proximity to chromosomal features, such as cancer-associated genomic regions, fragile sites, and human papilloma virus integration sites.

[0054] In some embodiments, in combination with detecting one or more target RNAs capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs:1 to 41 and 133 to 211 and/or detecting one or more target RNAs comprising at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388 and/or detecting one or more target RNAs that comprise a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs:1 to 41 and 133 to 211, methods herein further comprise detecting in a sample of human cervical cells the expression of at least one mRNA species. In some embodiments, the at least one mRNA is selected from the mRNAs for the genes set forth in Table 4, below. In some embodiments, at least one mRNA is selected from mRNAs for CDKN2A, MKI67, TOP2A, and MCM5. In some embodiments, at least one mRNA is selected from mRNAs for CDKN2A, MKI67, TOP2A, MCM5, BIRC5, MMP9, and MCM2.

Table 4

gene	name	alias
BIRC5	survivin	survivin
IGF2BP3	insulin-like growth factor 2 mRNA	L523S, IMP-3, KOC1

Table 4

gene	name	alias
	binding protein 3	
TERC	telomerase RNA component	hTR
CDKN2A	cyclin-dependent kinase inhibitor 2A	P16 ^{ink4}
MCM5	minichromosome maintenance complex component 5	-
TOP2A	topoisomerase II- α	
MYBL2	v-myb myeloblastosis viral oncogene homolog (avian)-like 2	B-myb
PIK3CA	phosphoinositide-3-kinase, catalytic, alpha polypeptide	PI3K
DROSHA	class 2 RNase III enzyme that initiates processing of microRNA	Drosha, Rnasen
MKI67	antigen identified by monoclonal antibody Ki-67	Ki-67
MMP9	matrix metalloproteinase 9	gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase
MCM2	minichromosome maintenance complex component 2	

[0055] In some embodiments, an increase in expression of one or more mRNAs listed in the table above is indicative of the presence of cervical dysplasia or cervical cancer in a sample of human cervical cells.

[0056] In some embodiments, more than one target RNA is detected simultaneously in a single reaction. In some embodiments, at least 2, at least 3, at least 5, or at least 10 target RNAs are detected simultaneously in a single reaction. In some embodiments, all target RNAs are detected simultaneously in a single reaction.

4.1.2. Exemplary controls

[0057] In some embodiments, a normal level (a “control”) for each target RNA can be determined as an average level or range that is characteristic of normal cervical cells or other reference material, against which the level measured in the sample can be compared. The determined average or range of target RNA in normal subjects can be used as a benchmark for detecting above-normal levels of target RNA indicative of cervical dysplasia. In some embodiments, normal levels of target RNA can be determined using individual or pooled RNA-containing samples from one or more individuals, such as from patients undergoing hysterectomy for benign gynecologic disease.

[0058] In some embodiments, determining a normal level of expression of a target RNA comprises detecting a complex comprising a probe hybridized to a nucleic acid selected from a target RNA, a DNA amplicon of the target RNA, and a complement of the target RNA. That is, in some embodiments, a normal level of expression can be determined by detecting a DNA amplicon of the target RNA, or a complement of the target RNA rather than the target RNA itself. In some embodiments, a normal level of such a complex is determined and used as a control. The normal level of the complex, in some embodiments, correlates to the normal level of the target RNA. Thus, when a normal level of a target is discussed herein, that level can, in some embodiments, be determined by detecting such a complex.

[0059] In some embodiments, a control comprises RNA from cells of a single individual, *e.g.*, a patient undergoing hysterectomy for benign gynecologic disease. In some embodiments, a control comprises RNA from a pool of cells from multiple individuals. In some embodiments, a control is drawn from anatomically and/or cytologically normal areas of the cervix of the individual from whom the test sample was obtained. In some embodiments, a control comprises commercially-available human RNA, such as, for example, human cervix total RNA (Ambion; AM6992). In some embodiments, a normal level or normal range has already been predetermined prior to testing a sample for an elevated level.

[0060] In some embodiments, the normal level of target RNA can be determined from one or more continuous cell lines, typically cell lines previously shown to have expression levels of the at least one target RNA that approximate the level of expression in normal cervical cells.

[0061] In some embodiments, a method comprises detecting the level of expression of at least one target RNA. In some embodiments, a method further comprises comparing the level of expression of at least one target RNA to a normal level of expression of the at least one target RNA. In some embodiments, a method further comprises comparing the level of expression of at least one target RNA to a control level of expression of the at least one target RNA. A control level of expression of the at least one target RNA is, in some embodiments, the level of expression of the at least one target RNA in a normal cell. In some such embodiments, a control level may be referred to as a

normal level. In some embodiments, a greater level of expression of the at least one target RNA relative to the level of expression of the at least one target RNA in a normal cell indicates cervical dysplasia.

[0062] In some embodiments, the level of expression of the at least one target RNA is compared to a reference level of expression, e.g., from a confirmed cervical dysplasia. In some such embodiments, a similar level of expression of the at least one target RNA relative to the reference sample indicates cervical dysplasia.

[0063] In some embodiments, a level of expression of at least one target RNA that is at least about two-fold greater than a normal level of expression of the respective at least one target RNA indicates the presence of cervical dysplasia. In some embodiments, a level of expression of at least one target RNA that is at least about two-fold greater than the level of the respective at least one target RNA in a control sample comprised of normal cells indicates the presence of a cervical dysplasia. In various embodiments, a level of expression of at least one target RNA that is at least about 3-fold, at least about 4-fold, at least about 5-fold, at least about 6-fold, at least about 7-fold, at least about 8-fold, at least about 9-fold, or at least about 10-fold greater than the level of expression of the respective at least one target RNA in a control sample comprised of normal cells indicates the presence of cervical dysplasia. In various embodiments, a level of expression of at least one target RNA that is at least about 3-fold, at least about 4-fold, at least about 5-fold, at least about 6-fold, at least about 7-fold, at least about 8-fold, at least about 9-fold, or at least about 10-fold greater than a normal level of expression of the at least one target RNA indicates the presence of cervical dysplasia.

[0064] In some embodiments, an increase in expression of one or more target RNAs capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1 to 41 in Table 1 and SEQ ID NOs: 133 to 211 in Table 11 is indicative of the presence of cervical dysplasia or cervical cancer in a sample of human cervical cells. In some embodiments, an increase in expression of one or more target RNAs comprising at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388 is indicative of the presence of cervical dysplasia or cervical cancer in a sample of human cervical cells. In some embodiments, an increase in expression of one or more target RNAs comprising a sequence that is complementary to at least 15

contiguous nucleotides of a sequence selected from SEQ ID NOs: 1 to 41 and 133 to 211 is indicative of the presence of cervical dysplasia or cervical cancer in a sample of human cervical cells.

[0065] In some embodiments, an increase in expression of one or more target RNAs capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1 to 41 in Table 1 and SEQ ID NOs: 133 to 211 in Table 11 is indicative of the presence of cervical dysplasia in a sample of human cervical cells that is likely to proceed to cervical cancer. In some embodiments, an increase in expression of one or more target RNAs comprising at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388 is indicative of the presence of cervical dysplasia in a sample of human cervical cells that is likely to proceed to cervical cancer. In some embodiments, an increase in expression of one or more target RNAs comprising a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 1 to 41 and 133 to 211 is indicative of the presence of cervical dysplasia in a sample of human cervical cells that is likely to proceed to cervical cancer.

[0066] In some embodiments, a control level of expression of a target RNA is determined contemporaneously, such as in the same assay or batch of assays, as the level of expression of the target RNA in a sample. In some embodiments, a control level of expression of a target RNA is not determined contemporaneously as the level of expression of the target RNA in a sample. In some such embodiments, the control level of expression has been determined previously.

[0067] In some embodiments, the level of expression of a target RNA is not compared to a control level of expression, for example, when it is known that the target RNA is expressed at very low levels, or not at all, in normal cells. In such embodiments, detection of a high level of the target RNA in a sample is indicative of cervical dysplasia.

4.1.3. Exemplary methods of preparing RNAs

[0068] Target RNA can be prepared by any appropriate method. Total RNA can be isolated by any method, including, but not limited to, the protocols set forth in Wilkinson, M. (1988) Nucl. Acids Res. 16(22):10,933; and Wilkinson, M. (1988) Nucl. Acids Res. 16(22): 10934, or by using commercially-available kits or reagents, such as

the TRIzol® reagent (Invitrogen™), Total RNA Extraction Kit (iNtRON Biotechnology), Total RNA Purification Kit (Norgen Biotek Corp.), RNAqueous™ (Ambion), MagMAX™ (Ambion), RecoverAll™ (Ambion), RNeasy (Qiagen), etc.

[0069] In some embodiments, small RNAs are isolated or enriched. In some embodiments “small RNA” refers to RNA molecules smaller than about 200 nucleotides (nt) in length. In some embodiments, “small RNA” refers to RNA molecules smaller than about 100 nt, smaller than about 90 nt, smaller than about 80 nt, smaller than about 70 nt, smaller than about 60 nt, smaller than about 50 nt, or smaller than about 40 nt.

[0070] Enrichment of small RNAs can be accomplished by method. Such methods include, but are not limited to, methods involving organic extraction followed by adsorption of nucleic acid molecules on a glass fiber filter using specialized binding and wash solutions, and methods using spin column purification. Enrichment of small RNAs may be accomplished using commercially-available kits, such as mirVana™ Isolation Kit (Applied Biosystems), mirPremier™ microRNA Isolation Kit (Sigma-Aldrich), PureLink™ miRNA Isolation Kit (Invitrogen), miRCURY™ RNA isolation kit (Exiqon), microRNA Purification Kit (Norgen Biotek Corp.), miRNeasy kit (Qiagen), etc. In some embodiments, purification can be accomplished by the TRIzol® (Invitrogen) method, which employs a phenol/isothiocyanate solution to which chloroform is added to separate the RNA-containing aqueous phase. Small RNAs are subsequently recovered from the aqueous by precipitation with isopropyl alcohol. In some embodiments, small RNAs can be purified using chromatographic methods, such as gel electrophoresis using the flashPAGE™ Fractionator available from Applied Biosystems.

[0071] In some embodiments, small RNA is isolated from other RNA molecules to enrich for target RNAs, such that the small RNA fraction (e.g., containing RNA molecules that are 200 nucleotides or less in length, such as less than 100 nucleotides in length, such as less than 50 nucleotides in length, such as from about 10 to about 40 nucleotides in length) is substantially pure, meaning it is at least about 80%, 85%, 90%, 95% pure or more, but less than 100% pure, with respect to larger RNA molecules. Alternatively, enrichment of small RNA can be expressed in terms of fold-enrichment. In some embodiments, small RNA is enriched by about, at least about, or at most about 5X, 10X, 20X, 30X, 40X, 50X, 60X, 70X, 80X, 90X, 100X, 110X, 120X, 130X, 140X, 150X,

160X, 170X, 180X, 190X, 200X, 210X, 220X, 230X, 240X, 250X, 260X, 270X, 280X, 290X, 300X, 310X, 320X, 330X, 340X, 350X, 360X, 370X, 380X, 390X, 400X, 410X, 420X, 430X, 440X, 450X, 460X, 470X, 480X, 490X, 500X, 600X, 700X, 800X, 900X, 1000X, 1100X, 1200X, 1300X, 1400X, 1500X, 1600X, 1700X, 1800X, 1900X, 2000X, 3000X, 4000X, 5000X, 6000X, 7000X, 8000X, 9000X, 10,000X or more, or any range derivable therein, with respect to the concentration of larger RNAs in an RNA isolate or total RNA in a sample.

[0072] In yet other embodiments, expression is measured in a sample in which RNA has not first been purified from the cells.

[0073] In some embodiments, RNA is modified before target RNAs are detected. In some embodiments, the modified RNA is total RNA. In other embodiments, the modified RNA is small RNA that has been purified from total RNA or from cell lysates, such as RNA less than 200 nucleotides in length, such as less than 100 nucleotides in length, such as less than 50 nucleotides in length, such as from about 10 to about 40 nucleotides in length. RNA modifications that can be utilized in the methods described herein include, but are not limited to, the addition of a poly-dA or a poly-dT tail, which can be accomplished chemically or enzymatically, and/or the addition of a small molecule, such as biotin.

[0074] In some embodiments, one or more target RNAs are reverse transcribed. In some embodiments, where present, RNA is modified when it is reverse transcribed, such as when a poly-dA or a poly-dT tail is added to the cDNA during reverse transcription. In other embodiments, RNA is modified before it is reverse transcribed. In some embodiments, total RNA is reverse transcribed. In other embodiments, small RNAs are isolated or enriched before the RNA is reverse transcribed.

[0075] When a target RNA is reverse transcribed, a complement of the target RNA is formed. In some embodiments, the complement of the target RNA is detected rather than the target RNA itself (or a DNA copy thereof). Thus, when the methods discussed herein indicate that a target RNA is detected, or the level of a target RNA is determined, such detection or determination may be carried out on a complement of the target RNA instead of, or in addition to, the target RNA itself. In some embodiments, when the complement of the target RNA is detected rather than the target RNA, a probe is

used that is complementary to the complement of the target RNA. In such embodiments, the probe comprises at least a portion that is identical in sequence to the target RNA, although it may contain thymidine in place of uridine, and/or comprise other modified nucleotides.

[0076] In some embodiments, the method of detecting one or more target RNAs comprises amplifying cDNA complementary to said target RNA. Such amplification can be accomplished by any method. Exemplary methods include, but are not limited to, real time PCR, endpoint PCR, and amplification using T7 polymerase from a T7 promoter annealed to a cDNA, such as provided by the SenseAmp Plus™ Kit available at Implen, Germany.

[0077] When a target RNA or a cDNA complementary to a target RNA is amplified, in some embodiments, a DNA amplicon of a target RNA is formed. A DNA amplicon may be single stranded or double-stranded. In some embodiments, when a DNA amplicon is single-stranded, the sequence of the DNA amplicon is related to the target RNA in either the sense or antisense orientation. In some embodiments, the DNA amplicon of the target RNA is detected rather than the target RNA itself. Thus, when the methods discussed herein indicate that a target RNA is detected, or the level of a target RNA is determined, such detection or determination may be carried out on a DNA amplicon of the target RNA instead of, or in addition to, the target RNA itself. In some embodiments, when the DNA amplicon of the target RNA is detected rather than the target RNA, a probe is used that is complementary to the complement of the target RNA. In some embodiments, when the DNA amplicon of the target RNA is detected rather than the target RNA, a probe is used that is complementary to the target RNA. Further, in some embodiments, multiple probes may be used, and some probes may be complementary to the target RNA and some probes may be complementary to the complement of the target RNA.

[0078] In some embodiments, the method of detecting one or more target RNAs comprises RT-PCR, as described below. In some embodiments, detecting one or more target RNAs comprises real-time monitoring of an RT-PCR reaction, which can be accomplished by any method. Such methods include, but are not limited to, the use of TaqMan®, Molecular beacon, or Scorpion probes (i.e., FRET probes) and the use of

intercalating dyes, such as SYBR green, EvaGreen, thiazole orange, YO-PRO, TO-PRO, etc.

4.1.4. Exemplary analytical methods

[0079] As described above, methods are presented for detecting cervical dysplasia, including cervical dysplasia likely to progress to carcinoma, in a sample of human cervical cells. In some embodiments, the method comprises detecting a level of expression of at least one target RNA capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1 to 41 set forth in Table 1 and SEQ ID NOs: 133 to 211 set forth in Table 11 that is greater in the sample than a normal level of expression of the at least one target RNA in a control sample, such as a sample derived from normal cervical cells. In some embodiments, a method comprises detecting a level of one or more target RNAs comprising at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388 that is greater in the sample than a normal level of expression of the at least one target RNA in a control sample. In some embodiments, a method comprises detecting a level of one or more target RNAs that comprise a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 1 to 41 and 133 to 211 that is greater in the sample than a normal level of expression of the at least one target RNA in a control sample. In some embodiments, a target RNA, in its mature form, comprises fewer than 30 nucleotides. In some embodiments, a target RNA is a microRNA.

[0080] In some embodiments, such as those described above, the method further comprises detecting a level of expression of at least one target RNA of the human miRNome that does not specifically hybridize to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1 to 41 set forth in Table 1 and SEQ ID NOs: 133 to 211 set forth in Table 11 that is greater in the sample than a normal level of expression of the at least one target RNA in a control sample. As used herein, the term “human miRNome” refers to all microRNA genes in a human cell and the mature microRNAs produced therefrom.

[0081] Any analytical procedure capable of permitting specific and quantifiable (or semi-quantifiable) detection of the desired at least one target RNA may be used in the methods herein presented. Such analytical procedures include, but are not limited to, the

microarray methods set forth in Example 1 and the RT-PCR methods set forth in Example 6, and methods known to those skilled in the art.

[0082] In some embodiments, detection of a target RNA comprises forming a complex comprising a polynucleotide that is complementary to a target RNA or to a complement thereof, and a nucleic acid selected from the target RNA, a DNA amplicon of the target RNA, and a complement of the target RNA. Thus, in some embodiments, the polynucleotide forms a complex with a target RNA. In some embodiments, the polynucleotide forms a complex with a complement of the target RNA, such as a cDNA that has been reverse transcribed from the target RNA. In some embodiments, the polynucleotide forms a complex with a DNA amplicon of the target RNA. When a double-stranded DNA amplicon is part of a complex, as used herein, the complex may comprise one or both strands of the DNA amplicon. Thus, in some embodiments, a complex comprises only one strand of the DNA amplicon. In some embodiments, a complex is a triplex and comprises the polynucleotide and both strands of the DNA amplicon. In some embodiments, the complex is formed by hybridization between the polynucleotide and the target RNA, complement of the target RNA, or DNA amplicon of the target RNA. The polynucleotide, in some embodiments, is a primer or probe.

[0083] In some embodiments, a method comprises detecting the complex. In some embodiments, the complex does not have to be associated at the time of detection. That is, in some embodiments, a complex is formed, the complex is then dissociated or destroyed in some manner, and components from the complex are detected. An example of such a system is a TaqMan® assay. In some embodiments, when the polynucleotide is a primer, detection of the complex may comprise amplification of the target RNA, a complement of the target RNA, or a DNA amplicon of a target RNA.

[0084] In some embodiments the analytical method used for detecting at least one target RNA in the methods set forth herein includes real-time quantitative RT-PCR. See Chen, C. et al. (2005) *Nucl. Acids Res.* 33:e179 and PCT Publication No. WO 2007/117256, which are incorporated herein by reference in its entirety. In some embodiments, the analytical method used for detecting at least one target RNA includes the method described in U.S. Publication No. US2009/0123912 A1, which is incorporated herein by reference in its entirety. In an exemplary method described in that publication,

an extension primer comprising a first portion and second portion, wherein the first portion selectively hybridizes to the 3' end of a particular microRNA and the second portion comprises a sequence for universal primer, is used to reverse transcribe the microRNA to make a cDNA. A reverse primer that selectively hybridizes to the 5' end of the microRNA and a universal primer are then used to amplify the cDNA in a quantitative PCR reaction.

[0085] In some embodiments, the analytical method used for detecting at least one target RNA includes the use of a TaqMan® probe. In some embodiments, the analytical method used for detecting at least one target RNA includes a TaqMan® assay, such as the TaqMan® MicroRNA Assays sold by Applied Biosystems, Inc. In an exemplary TaqMan® assay, total RNA is isolated from the sample. In some embodiments, the assay can be used to analyze about 10 ng of total RNA input sample, such as about 9 ng of input sample, such as about 8 ng of input sample, such as about 7 ng of input sample, such as about 6 ng of input sample, such as about 5 ng of input sample, such as about 4 ng of input sample, such as about 3 ng of input sample, such as about 2 ng of input sample, and even as little as about 1 ng of input sample containing microRNAs.

[0086] The TaqMan® assay utilizes a stem-loop primer that is specifically complementary to the 3'-end of a target RNA. In an exemplary TaqMan® assay, hybridizing the stem-loop primer to the target RNA is followed by reverse transcription of the target RNA template, resulting in extension of the 3' end of the primer. The result of the reverse transcription is a chimeric (DNA) amplicon with the stem-loop primer sequence at the 5' end of the amplicon and the cDNA of the target RNA at the 3' end. Quantitation of the target RNA is achieved by real time RT-PCR using a universal reverse primer having a sequence that is complementary to a sequence at the 5' end of all stem-loop target RNA primers, a target RNA-specific forward primer, and a target RNA sequence-specific TaqMan® probe.

[0087] The assay uses fluorescence resonance energy transfer ("FRET") to detect and quantitate the synthesized PCR product. Typically, the TaqMan® probe comprises a fluorescent dye molecule coupled to the 5'-end and a quencher molecule coupled to the 3'-end, such that the dye and the quencher are in close proximity, allowing the quencher to suppress the fluorescence signal of the dye via FRET. When the polymerase replicates

the chimeric amplicon template to which the TaqMan® probe is bound, the 5'-nuclease of the polymerase cleaves the probe, decoupling the dye and the quencher so that FRET is abolished and a fluorescence signal is generated. Fluorescence increases with each RT-PCR cycle proportionally to the amount of probe that is cleaved.

[0088] Additional exemplary methods for RNA detection and/or quantification are described, e.g., in U.S. Publication No. US 2007/0077570 (Lao et al.), PCT Publication No. WO 2007/025281 (Tan et al.), U.S. Publication No. US2007/0054287 (Bloch), PCT Publication No. WO2006/0130761 (Bloch), and PCT Publication No. WO 2007/011903 (Lao et al.), which are incorporated by reference herein in their entireties for any purpose.

[0089] In some embodiments, quantitation of the results of real-time RT-PCR assays is done by constructing a standard curve from a nucleic acid of known concentration and then extrapolating quantitative information for target RNAs of unknown concentration. In some embodiments, the nucleic acid used for generating a standard curve is an RNA (e.g., microRNA) of known concentration. In some embodiments, the nucleic acid used for generating a standard curve is a purified double-stranded plasmid DNA or a single-stranded DNA generated in vitro.

[0090] In some embodiments, where the amplification efficiencies of the target nucleic acids and the endogenous reference are approximately equal, quantitation is accomplished by the comparative Ct (cycle threshold, e.g., the number of PCR cycles required for the fluorescence signal to rise above background) method. Ct values are inversely proportional to the amount of nucleic acid target in a sample. In some embodiments, Ct values of the target RNA of interest can be compared with a control or calibrator, such as RNA (e.g., microRNA) from normal tissue. In some embodiments, the Ct values of the calibrator and the target RNA samples of interest are normalized to an appropriate endogenous housekeeping gene.

[0091] In addition to the TaqMan® assays, other real-time RT-PCR chemistries useful for detecting and quantitating PCR products in the methods presented herein include, but are not limited to, Molecular Beacons, Scorpion probes and intercalating dyes, such as SYBR Green, EvaGreen, thiazole orange, YO-PRO, TO-PRO, etc., which are discussed below.

[0092] In some embodiments, real-time RT-PCR detection is performed specifically to detect and quantify the expression of a single target RNA. The target RNA, in some embodiments, is selected from a target RNA capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1 to 41 set forth in Table 1 and SEQ ID NOs: 133 to 211 set forth in Table 11. In some embodiments, the target RNA specifically hybridizes to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. In some embodiments, the target RNA specifically hybridizes to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12. In some embodiments, the target RNA specifically hybridizes to a nucleic acid comprising a sequence selected from SEQ ID NOs: 142, 151, 153, 193, 194, 205, 172, 208, 210, and 211. In some embodiments, the target RNA specifically hybridizes to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1, 5, 7, and 32. In some embodiments, the target RNA comprises at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388. In some embodiments, the target RNA comprises a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 1 to 41 and 133 to 211. In some embodiments, a target RNA, in its mature form, comprises fewer than 30 nucleotides. In some embodiments, a target RNA is a microRNA.

[0093] As described above, in some embodiments, in addition to detecting expression of one or more target RNAs capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1 to 41 and SEQ ID NOs: 133 to 211, and/or detecting expression of at least one target RNA comprising at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388 and/or detecting expression of at least one target RNA comprising a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 1 to 41 and 133 to 211, the methods comprise detection of expression of one or more microRNAs selected from miR-21, miR-31, miR-182, miR-183, miR-155, miR-9, miR-199a*, miR-199a, miR-199b, miR-145, miR-133a, miR-133b, miR-214, miR-127, miR-205, miR-210, miR-146a, miR-301, miR-142-5p, miR-194, miR-215, miR-32, miR-374b, miR-933, miR-769-3p, miR-671, miR-934, miR-935, miR-937, miR-938, miR-939, miR-940, miR-941, miR-942, miR-943, miR-944, miR-708, miR-874-5p, and miR-874-3p.

[0094] In various other embodiments, real-time RT-PCR detection is utilized to detect, in a single multiplex reaction, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, or at least 8 target RNAs. At least one target RNA, in some embodiments, is capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1 to 41 and 133 to 211. In some embodiments, at least one target RNA comprises at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388. In some embodiments, at least one target RNA comprises a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 1 to 41 and 133 to 211. In some embodiments, a target RNA, in its mature form, comprises fewer than 30 nucleotides. In some embodiments, a target RNA is a microRNA.

[0095] In some embodiments, the method comprises detecting expression in a multiplex RT-PCR reaction of at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, or at least 8 target RNAs, wherein each target RNA is capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, and 8. In some embodiments, the method comprises detecting greater than normal expression, using a single multiplex RT-PCR reaction, of at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, or at least 12 target RNAs, wherein each target RNA is capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12. In some embodiments, the method comprises detecting greater than normal expression, using a single multiplex RT-PCR reaction, of at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, or at least 12 target RNAs, wherein each target RNA is capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 142, 151, 153, 193, 194, 205, 172, 208, 210, and 211. In some embodiments, the method comprises detecting expression in a multiplex RT-PCR reaction of at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, or at least 9 target RNAs, wherein each target RNA is capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1, 4, 5, 7, 12, 17, 25, 26, 32.

[0096] In some multiplex embodiments, a plurality of probes, such as TaqMan® probes, each specific for a different RNA target, is used. In some embodiments, each target RNA-specific probe is spectrally distinguishable from the other probes used in the same multiplex reaction.

[0097] In some embodiments, quantitation of real-time RT PCR products is accomplished using a dye that binds to double-stranded DNA products, such as SYBR Green, EvaGreen, thiazole orange, YO-PRO, TO-PRO, etc. In some embodiments, the assay is the QuantiTect SYBR Green PCR assay from Qiagen. In this assay, total RNA is first isolated from a sample. Total RNA is subsequently poly-adenylated at the 3'-end and reverse transcribed using a universal primer with poly-dT at the 5'-end. In some embodiments, a single reverse transcription reaction is sufficient to assay multiple target RNAs. Real-time RT-PCR is then accomplished using target RNA-specific primers and an miScript Universal Primer, which comprises a poly-dT sequence at the 5'-end. SYBR Green dye binds non-specifically to double-stranded DNA and upon excitation, emits light. In some embodiments, buffer conditions that promote highly-specific annealing of primers to the PCR template (e.g., available in the QuantiTect SYBR Green PCR Kit from Qiagen) can be used to avoid the formation of non-specific DNA duplexes and primer dimers that will bind SYBR Green and negatively affect quantitation. Thus, as PCR product accumulates, the signal from SYBR Green increases, allowing quantitation of specific products.

[0098] Real-time RT-PCR is performed using any RT-PCR instrumentation available in the art. Typically, instrumentation used in real-time RT-PCR data collection and analysis comprises a thermal cycler, optics for fluorescence excitation and emission collection, and optionally a computer and data acquisition and analysis software.

[0099] In some embodiments, the analytical method used in the methods described herein is a DASL® (cDNA-mediated Annealing, Selection, Extension, and Ligation) Assay, such as the MicroRNA Expression Profiling Assay available from Illumina, Inc. (See <http://www.illumina.com/downloads/MicroRNAAssayWorkflow.pdf>). In some embodiments, total RNA is isolated from a sample to be analyzed by any method. Additionally, in some embodiments, small RNAs are isolated from a sample to be analyzed by any method. Total RNA or isolated small RNAs may then be

polyadenylated (> 18 A residues are added to the 3'-ends of the RNAs in the reaction mixture). The RNA is reverse transcribed using a biotin-labeled DNA primer that comprises from the 5' to the 3' end, a sequence that includes a PCR primer site and a poly-dT region that binds to the poly-dA tail of the sample RNA. The resulting biotinylated cDNA transcripts are then hybridized to a solid support via a biotin-streptavidin interaction and contacted with one or more target RNA-specific polynucleotides. The target RNA-specific polynucleotides comprise, from the 5'-end to the 3'-end, a region comprising a PCR primer site, region comprising an address sequence, and a target RNA-specific sequence.

[00100] In some DASL® embodiments, the target RNA-specific sequence comprises at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, or at least 24 contiguous nucleotides having a sequence identically present in one of SEQ ID NOs: 1 to 41 and 133 to 211. In some embodiments, the target RNA-specific sequence comprises a probe sequence that is complementary to at least a portion of a microRNA of the human miRNome, such as miR-21, miR-31, miR-182, miR-183, miR-155, miR-9, miR-199a*, miR-199a, miR-199b, miR-145, miR-133a, miR-133b, miR-205, miR-214, miR-127, miR-210, miR-146a, miR-301, miR-142-5p, miR-194, miR-215, miR-32, miR-374b, miR-933, miR-769-3p, miR-671, miR-934, miR-935, miR-937, miR-938, miR-939, miR-940, miR-941, miR-942, miR-943, miR-944, miR-708, miR-874-5p, and miR-874-3p.

[00101] After hybridization, the target RNA-specific polynucleotide is extended, and the extended products are then eluted from the immobilized cDNA array. A second PCR reaction using a fluorescently-labeled universal primer generates a fluorescently-labeled DNA comprising the target RNA-specific sequence. The labeled PCR products are then hybridized to a microbead array for detection and quantitation.

[00102] In some embodiments, the analytical method used for detecting and quantifying the expression of the at least one target RNA in the methods described herein is a bead-based flow cytometric assay. See Lu J. et al. (2005) Nature 435:834-838, which is incorporated herein by reference in its entirety. An example of a bead-based flow cytometric assay is the xMAP® technology of Luminex, Inc. (See

<http://www.luminexcorp.com/technology/index.html>). In some embodiments, total RNA is isolated from a sample and is then labeled with biotin. The labeled RNA is then hybridized to target RNA-specific capture probes (e.g., FlexmiR™ products sold by Luminex, Inc. at <http://www.luminexcorp.com/products/assays/index.html>) that are covalently bound to microbeads, each of which is labeled with 2 dyes having different fluorescence intensities. A streptavidin-bound reporter molecule (e.g., streptavidin-phycoerythrin, also known as “SAPE”) is attached to the captured target RNA and the unique signal of each bead is read using flow cytometry. In some embodiments, the RNA sample (total RNA or enriched small RNAs) is first polyadenylated, and is subsequently labeled with a biotinylated 3DNA™ dendrimer (i.e., a multiple-arm DNA with numerous biotin molecules bound thereto), such as those sold by Marligen Biosciences as the Vantage™ microRNA Labeling Kit, using a bridging polynucleotide that is complementary to the 3'-end of the poly-dA tail of the sample RNA and to the 5'-end of the polynucleotide attached to the biotinylated dendrimer. The streptavidin-bound reporter molecule is then attached to the biotinylated dendrimer before analysis by flow cytometry. See <http://www.marligen.com/vantage-microrna-labeling-kit.html>. In some embodiments, biotin-labeled RNA is first exposed to SAPE, and the RNA/SAPE complex is subsequently exposed to an anti-phycoerythrin antibody attached to a DNA dendrimer, which can be bound to as many as 900 biotin molecules. This allows multiple SAPE molecules to bind to the biotinylated dendrimer through the biotin-streptavidin interaction, thus increasing the signal from the assay.

[00103] In some embodiments, the analytical method used for detecting and quantifying the expression of the at least one target RNA in the methods described herein is by gel electrophoresis and detection with labeled probes (e.g., probes labeled with a radioactive or chemiluminescent label), such as by Northern blotting. In some embodiments, total RNA is isolated from the sample, and then is size-separated by SDS polyacrylamide gel electrophoresis. The separated RNA is then blotted onto a membrane and hybridized to radiolabeled complementary probes. In some embodiments, exemplary probes contain one or more affinity-enhancing nucleotide analogs as discussed below, such as locked nucleic acid (“LNA”) analogs, which contain a bicyclic sugar moiety instead of deoxyribose or ribose sugars. See, e.g., Várallyay, E. et al. (2008) *Nature Protocols* 3(2):190-196, which is incorporated herein by reference in its entirety. In some

embodiments, the total RNA sample can be further purified to enrich for small RNAs. In some embodiments, target RNAs can be amplified by, e.g., rolling circle amplification using a long probe that is complementary to both ends of a target RNA (“padlocked probes”), ligation to circularize the probe followed by rolling circle replication using the target RNA hybridized to the circularized probe as a primer. See, e.g., Jonstrup, S.P. et al. (2006) RNA 12:1-6, which is incorporated herein by reference in its entirety. The amplified product can then be detected and quantified using, e.g., gel electrophoresis and Northern blotting.

[00104] In alternative embodiments, labeled probes are hybridized to isolated total RNA in solution, after which the RNA is subjected to rapid ribonuclease digestion of single-stranded RNA, e.g., unhybridized portions of the probes or unhybridized target RNAs. In these embodiments, the ribonuclease treated sample is then analyzed by SDS-PAGE and detection of the radiolabeled probes by, e.g., Northern blotting. See mirVana™ miRNA Detection Kit sold by Applied Biosystems, Inc. product literature at <http://www.ambion.com/catalog/CatNum.php?1552>.

[00105] In some embodiments, the analytical method used for detecting and quantifying the at least one target RNA in the methods described herein is by hybridization to a microarray. See, e.g., Liu, C.G. et al. (2004) Proc. Nat’l Acad. Sci. USA 101:9740-9744; Lim, L.P. et al. (2005) Nature 433:769-773, each of which is incorporated herein by reference in its entirety, and Example 1.

[00106] In some embodiments, detection and quantification of a target RNA using a microarray is accomplished by surface plasmon resonance. See, e.g., Nanotech News (2006), available at http://nano.cancer.gov/news_center/nanotech_news_2006-10-30b.asp. In these embodiments, total RNA is isolated from a sample being tested. Optionally, the RNA sample is further purified to enrich the population of small RNAs. After purification, the RNA sample is bound to an addressable microarray containing probes at defined locations on the microarray. Nonlimiting exemplary probes include probes comprising sequences set forth in SEQ ID NOs: 1 to 41 and 133 to 211. Exemplary probes also include, but are not limited to, probes comprising a region that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388. Exemplary probes

also include, but are not limited to, probes comprising at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 1 to 41 and 133 to 211. In some embodiments, the probes contain one or more affinity-enhancing nucleotide analogs as discussed below, such as locked nucleic acid (“LNA”) nucleotide analogs. After hybridization to the microarray, the RNA that is hybridized to the array is first polyadenylated, and the array is then exposed to gold particles having poly-dT bound to them. The amount of bound target RNA is quantitated using surface plasmon resonance.

[00107] In some embodiments, microarrays are utilized in a RNA-primed, Array-based Klenow Enzyme (“RAKE”) assay. See Nelson, P.T. et al. (2004) *Nature Methods* 1(2):1-7; Nelson, P.T. et al. (2006) *RNA* 12(2):1-5, each of which is incorporated herein by reference in its entirety. In some embodiments, total RNA is isolated from a sample. In some embodiments, small RNAs are isolated from a sample. The RNA sample is then hybridized to DNA probes immobilized at the 5’-end on an addressable array. The DNA probes comprise, in some embodiments, from the 5’-end to the 3’-end, a first region comprising a “spacer” sequence which is the same for all probes, a second region comprising three thymidine-containing nucleosides, and a third region comprising a sequence that is complementary to a target RNA of interest.

[00108] Exemplary target RNAs of interest include, but are not limited to, target RNAs capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1 to 41 and 133 to 211, and target RNAs comprising at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388, and target RNAs comprising a region that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 1 to 41 and 133 to 211. Target RNAs also include target RNAs in the miRNome that do not specifically hybridize to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1 to 41 and 133 to 211. In some embodiments, a target RNA, in its mature form, comprises fewer than 30 nucleotides. In some embodiments, a target RNA is a microRNA.

[00109] After the sample is hybridized to the array, it is exposed to exonuclease I to digest any unhybridized probes. The Klenow fragment of DNA polymerase I is then applied along with biotinylated dATP, allowing the hybridized target RNAs to act as primers for the enzyme with the DNA probe as template. The slide is

then washed and a streptavidin-conjugated fluorophore is applied to detect and quantitate the spots on the array containing hybridized and Klenow-extended target RNAs from the sample.

[00110] In some embodiments, the RNA sample is reverse transcribed. In some embodiments, the RNA sample is reverse transcribed using a biotin/poly-dA random octamer primer. When than primer is used, the RNA template is digested and the biotin-containing cDNA is hybridized to an addressable microarray with bound probes that permit specific detection of target RNAs. In typical embodiments, the microarray includes at least one probe comprising at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, or at least 24 contiguous nucleotides identically present in, or complementary to a region of, a sequence selected from SEQ ID NOs: 1 to 41, 133 to 211, and 345 to 388. After hybridization of the cDNA to the microarray, the microarray is exposed to a streptavidin-bound detectable marker, such as a fluorescent dye, and the bound cDNA is detected. See Liu C.G. et al. (2008) *Methods* 44:22-30, which is incorporated herein by reference in its entirety.

[00111] In some embodiments, target RNAs are detected and quantified in an ELISA-like assay using probes bound in the wells of microtiter plates. See Mora J.R. and Getts R.C. (2006) *BioTechniques* 41:420-424 and supplementary material in *BioTechniques* 41(4):1-5; U.S. Patent Publication No. 2006/0094025 to Getts et al., each of which is incorporated by reference herein in its entirety. In these embodiments, a sample of RNA that is enriched in small RNAs is either polyadenylated, or is reverse transcribed and the cDNA is polyadenylated. The RNA or cDNA is hybridized to probes immobilized in the wells of a microtiter plates, wherein each of the probes comprises a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1 to 41, 133 to 211, or 345 to 388, or a sequence such as one or more sequences of target RNAs (or the reverse complement thereof) of the human miRNome, depending on whether RNA or cDNA is hybridized to the array. In some embodiments, the hybridized RNAs are labeled using a capture sequence, such as a DNA dendrimer (such as those available from Genisphere, Inc., http://www.genisphere.com/about_3dna.html) that is labeled with a plurality of biotin molecules or with a plurality of horseradish peroxidase

molecules, and a bridging polynucleotide that contains a poly-dT sequence at the 5'-end that binds to the poly-dA tail of the captured nucleic acid, and a sequence at the 3'-end that is complementary to a region of the capture sequence. If the capture sequence is biotinylated, the microarray is then exposed to streptavidin-bound horseradish peroxidase. Hybridization of target RNAs is detected by the addition of a horseradish peroxidase substrate such as tetramethylbenzidine (TMB) and measurement of the absorbance of the solution at 450nm.

[00112] In still other embodiments, an addressable microarray is used to detect a target RNA using quantum dots. See Liang, R.Q. et al. (2005) Nucl. Acids Res. 33(2):e17, available at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=548377>, which is incorporated herein by reference in its entirety. In some embodiments, total RNA is isolated from a sample. In some embodiments, small RNAs are isolated from the sample. The 3'-ends of the target RNAs are biotinylated using biotin-X-hydrazide. The biotinylated target RNAs are captured on a microarray comprising immobilized probes comprising sequences that are identically present in, or complementary to a region of, one or more of SEQ ID NOs: 1 to 41, 133 to 211, and 345 to 388 and/or probes comprising sequences other than those that are complementary to one or more microRNAs of the human miRNome. The hybridized target RNAs are then labeled with quantum dots via a biotin-streptavidin binding. A confocal laser causes the quantum dots to fluoresce and the signal can be quantified. In alternative embodiments, small RNAs can be detected using a colorimetric assay. In these embodiments, small RNAs are labeled with streptavidin-conjugated gold followed by silver enhancement. The gold nanoparticles bound to the hybridized target RNAs catalyze the reduction of silver ions to metallic silver, which can then be detected colorimetrically with a CCD camera

[00113] In some embodiments, detection and quantification of one or more target RNAs is accomplished using microfluidic devices and single-molecule detection. In some embodiments, target RNAs in a sample of isolated total RNA are hybridized to two probes, one which is complementary to nucleic acids at the 5'-end of the target RNA and the second which is complementary to the 3'-end of the target RNA. Each probe comprises, in some embodiments, one or more affinity-enhancing nucleotide analogs,

such as LNA nucleotide analogs and each is labeled with a different fluorescent dye having different fluorescence emission spectra. The sample is then flowed through a microfluidic capillary in which multiple lasers excite the fluorescent probes, such that a unique coincident burst of photons identifies a particular target RNA, and the number of particular unique coincident bursts of photons can be counted to quantify the amount of the target RNA in the sample. See U.S. Patent Publication No. 2006/0292616 to Neely et al., which is hereby incorporated by reference in its entirety. In some alternative embodiments, a target RNA-specific probe can be labeled with 3 or more distinct labels selected from, e.g., fluorophores, electron spin labels, etc., and then hybridized to an RNA sample, such as total RNA, or a sample that is enriched in small RNAs.

Nonlimiting exemplary target RNA-specific probes include probes comprising sequences selected from of SEQ ID NOs: 1 to 41 and 133 to 211. Nonlimiting exemplary target RNA-specific probes include probes comprising sequences that are complementary to sequences selected from of SEQ ID NOs: 1 to 41 and 133 to 211. Nonlimiting exemplary target RNA-specific probes also include probes comprising at least 15 contiguous nucleotides of, or the complement of at least 15 contiguous nucleotides of, a sequence selected from SEQ ID NOs: 1 to 41, 133 to 211, and 345 to 388.

[00114] Optionally, the sample RNA is modified before hybridization. The target RNA/probe duplex is then passed through channels in a microfluidic device and that comprise detectors that record the unique signal of the 3 labels. In this way, individual molecules are detected by their unique signal and counted. See U.S. Patent Nos. 7,402,422 and 7,351,538 to Fuchs et al., U.S. Genomics, Inc., each of which is incorporated herein by reference in its entirety.

[00115] In some embodiments, the detection and quantification of one or more target RNAs is accomplished by a solution-based assay, such as a modified Invader assay. See Allawi H.T. et al. (2004) RNA 10:1153-1161, which is incorporated herein by reference in its entirety. In some embodiments, the modified invader assay can be performed on unfractionated detergent lysates of cervical cells. In other embodiments, the modified invader assay can be performed on total RNA isolated from cells or on a sample enriched in small RNAs. The target RNAs in a sample are annealed to two probes which form hairpin structures. A first probe has a hairpin structure at the 5' end and a

region at the 3'-end that has a sequence that is complementary to the sequence of a region at the 5'-end of a target RNA. The 3'-end of the first probe is the "invasive polynucleotide". A second probe has, from the 5' end to the 3'-end a first "flap" region that is not complementary to the target RNA, a second region that has a sequence that is complementary to the 3'-end of the target RNA, and a third region that forms a hairpin structure. When the two probes are bound to a target RNA target, they create an overlapping configuration of the probes on the target RNA template, which is recognized by the Cleavase enzyme, which releases the flap of the second probe into solution. The flap region then binds to a complementary region at the 3'-end of a secondary reaction template ("SRT"). A FRET polynucleotide (having a fluorescent dye bound to the 5'-end and a quencher that quenches the dye bound closer to the 3' end) binds to a complementary region at the 5'-end of the SRT, with the result that an overlapping configuration of the 3'-end of the flap and the 5'-end of the FRET polynucleotide is created. Cleavase recognizes the overlapping configuration and cleaves the 5'-end of the FRET polynucleotide, generates a fluorescent signal when the dye is released into solution.

4.1.5. Exemplary polynucleotides

[00116] In some embodiments, polynucleotides are provided. In some embodiments, synthetic polynucleotides are provided. Synthetic polynucleotides, as used herein, refer to polynucleotides that have been synthesized in vitro either chemically or enzymatically. Chemical synthesis of polynucleotides includes, but is not limited to, synthesis using polynucleotide synthesizers, such as OligoPilot (GE Healthcare), ABI 3900 DNA Synthesizer (Applied Biosystems), and the like. Enzymatic synthesis includes, but is not limited, to producing polynucleotides by enzymatic amplification, e.g., PCR.

[00117] In some embodiments, a polynucleotide is provided that comprises at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, or at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 1 to 41, 133 to 211, and 345 to 388, and sequences complementary to SEQ ID NOs: 1 to 41, 133 to 211, and 345 to 388. In some embodiments, the polynucleotide further comprises a region having a sequence that is not found in, or complementary to, any of SEQ ID NOs: 1 to 41, 133 to

211, and 345 to 388. In some embodiments, a polynucleotide is provided that comprises at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, or at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 1 to 7, 9 to 37, 133 to 201, and 345 to 388, and sequences complementary to SEQ ID NOs: 1 to 7, 9 to 37, 133 to 201, and 345 to 388. In some embodiments, the polynucleotide further comprises a region having a sequence that is not found in, or complementary to, any of SEQ ID NOs: 1 to 7, 9 to 37, 133 to 201, and 345 to 388.

[00118] In various embodiments, a polynucleotide comprises fewer than 500, fewer than 300, fewer than 200, fewer than 150, fewer than 100, fewer than 75, fewer than 50, fewer than 40, or fewer than 30 nucleotides. In various embodiments, a polynucleotide is between 8 and 200, between 8 and 150, between 8 and 100, between 8 and 75, between 8 and 50, between 8 and 40, or between 8 and 30 nucleotides long.

[00119] In some embodiments, the polynucleotide is a primer. In some embodiments, the primer is labeled with a detectable moiety. In some embodiments, a primer is not labeled. A primer, as used herein, is a polynucleotide that is capable of specifically hybridizing to a target RNA or to a cDNA reverse transcribed from the target RNA or to an amplicon that has been amplified from a target RNA or a cDNA (collectively referred to as “template”), and, in the presence of the template, a polymerase and suitable buffers and reagents, can be extended to form a primer extension product.

[00120] In some embodiments, the polynucleotide is a probe. In some embodiments, the probe is labeled with a detectable moiety. A detectable moiety, as used herein, includes both directly detectable moieties, such as fluorescent dyes, and indirectly detectable moieties, such as members of binding pairs. When the detectable moiety is a member of a binding pair, in some embodiments, the probe can be detectable by incubating the probe with a detectable label bound to the second member of the binding pair. In some embodiments, a probe is not extendable, e.g., by a polymerase. In other embodiments, a probe is extendable.

[00121] In some embodiments, the polynucleotide is a FRET probe that in some embodiments is labeled at the 5'-end with a fluorescent dye (donor) and at the 3'-end with a quencher (acceptor), a chemical group that absorbs (i.e., suppresses) fluorescence emission from the dye when the groups are in close proximity (i.e., attached

to the same probe). In other embodiments, the donor and acceptor are not at the ends of the FRET probe. Thus, in some embodiments, the emission spectrum of the donor moiety should overlap considerably with the absorption spectrum of the acceptor moiety.

4.1.5.1. Exemplary polynucleotide modifications

[00122] In some embodiments, the methods of detecting at least one target RNA described herein employ one or more polynucleotides that have been modified, such as polynucleotides comprising one or more affinity-enhancing nucleotide analogs. Modified polynucleotides useful in the methods described herein include primers for reverse transcription, PCR amplification primers, and probes. In some embodiments, the incorporation of affinity-enhancing nucleotides increases the binding affinity and specificity of a polynucleotide for its target nucleic acid as compared to polynucleotides that contain only deoxyribonucleotides, and allows for the use of shorter polynucleotides or for shorter regions of complementarity between the polynucleotide and the target nucleic acid.

[00123] In some embodiments, affinity-enhancing nucleotide analogs include nucleotides comprising one or more base modifications, sugar modifications and/or backbone modifications.

[00124] In some embodiments, modified bases for use in affinity-enhancing nucleotide analogs include 5-methylcytosine, isocytosine, pseudoisocytosine, 5-bromouracil, 5-propynyluracil, 6-aminopurine, 2-aminopurine, inosine, diaminopurine, 2-chloro-6-aminopurine, xanthine and hypoxanthine.

[00125] In some embodiments, affinity-enhancing nucleotide analogs include nucleotides having modified sugars such as 2'-substituted sugars, such as 2'-O-alkyl-ribose sugars, 2'-amino-deoxyribose sugars, 2'-fluoro-deoxyribose sugars, 2'-fluoro-arabinose sugars, and 2'-O-methoxyethyl-ribose (2'MOE) sugars. In some embodiments, modified sugars are arabinose sugars, or d-arabino-hexitol sugars.

[00126] In some embodiments, affinity-enhancing nucleotide analogs include backbone modifications such as the use of peptide nucleic acids (PNA; *e.g.*, an oligomer including nucleobases linked together by an amino acid backbone). Other backbone modifications include phosphorothioate linkages, phosphodiester modified

nucleic acids, combinations of phosphodiester and phosphorothioate nucleic acid, methylphosphonate, alkylphosphonates, phosphate esters, alkylphosphonothioates, phosphoramidates, carbamates, carbonates, phosphate triesters, acetamides, carboxymethyl esters, methylphosphorothioate, phosphorodithioate, p-ethoxy, and combinations thereof.

[00127] In some embodiments, a polynucleotide includes at least one affinity-enhancing nucleotide analog that has a modified base, at least nucleotide (which may be the same nucleotide) that has a modified sugar, and/or at least one internucleotide linkage that is non-naturally occurring.

[00128] In some embodiments, an affinity-enhancing nucleotide analog contains a locked nucleic acid ("LNA") sugar, which is a bicyclic sugar. In some embodiments, a polynucleotide for use in the methods described herein comprises one or more nucleotides having an LNA sugar. In some embodiments, a polynucleotide contains one or more regions consisting of nucleotides with LNA sugars. In other embodiments, a polynucleotide contains nucleotides with LNA sugars interspersed with deoxyribonucleotides. *See, e.g., Frieden, M. et al. (2008) Curr. Pharm. Des. 14(11):1138-1142.*

4.1.5.2. Exemplary primers

[00129] In some embodiments, a primer is provided. In some embodiments, a primer is identical or complementary to at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, or at least 24 contiguous nucleotides of a target RNA. In some embodiments, a primer may also comprise portions or regions that are not identical or complementary to the target RNA. In some embodiments, a region of a primer that is identical or complementary to a target RNA is contiguous, such that any region of a primer that is not identical or complementary to the target RNA does not disrupt the identical or complementary region.

[00130] In some embodiments, a primer comprises a portion that is identically present in a target RNA. In some such embodiments, a primer that comprises a region that is identically present in the target RNA is capable of selectively hybridizing to a cDNA that has been reverse transcribed from the RNA, or to an amplicon that has

been produced by amplification of the target RNA or cDNA. In some embodiments, the primer is complementary to a sufficient portion of the cDNA or amplicon such that it selectively hybridizes to the cDNA or amplicon under the conditions of the particular assay being used.

[00131] As used herein, “selectively hybridize” means that a polynucleotide, such as a primer or probe, will hybridize to a particular nucleic acid in a sample with at least 5-fold greater affinity than it will hybridize to another nucleic acid present in the same sample that has a different nucleotide sequence in the hybridizing region. Exemplary hybridization conditions are discussed in Example 1. In some embodiments, a polynucleotide will hybridize to a particular nucleic acid in a sample with at least 10-fold greater affinity than it will hybridize to another nucleic acid present in the same sample that has a different nucleotide sequence in the hybridizing region.

[00132] Nonlimiting exemplary primers include primers comprising sequences that are identically present in, or complementary to a region of, sequences selected from SEQ ID NOs: 1 to 41, 133 to 211, and 345 to 388. Exemplary primers also include, but are not limited to, primers comprising regions that are identical or complementary to at least 15 contiguous nucleotides of sequences selected from SEQ ID NOs: 1 to 41, 133 to 211, and 345 to 388.

[00133] In some embodiments, a primer is used to reverse transcribe a target RNA, for example, as discussed herein. In some embodiments, a primer is used to amplify a target RNA or a cDNA reverse transcribed therefrom. Such amplification, in some embodiments, is quantitative PCR, for example, as discussed herein. In some embodiments, a primer comprises a detectable moiety.

4.1.5.3. Exemplary probes

[00134] In various embodiments, methods of detecting the presence of a cervical dysplasia comprise hybridizing nucleic acids of a human cervical sample with a probe. In some embodiments, the probe comprises a portion that is complementary to a target RNA. In some embodiments, the probe comprises a portion that is identically present in the target RNA. In some such embodiments, a probe that is complementary to a target RNA is complementary to a sufficient portion of the target RNA such that it selectively hybridizes to the target RNA under the conditions of the particular assay being

used. In some embodiments, a probe that is complementary to a target RNA is complementary to at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, or at least 24 contiguous nucleotides of the target RNA. In some embodiments, a probe that is complementary to a target RNA comprises a region that is complementary to at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, or at least 24 contiguous nucleotides of the target RNA. That is, a probe that is complementary to a target RNA may also comprise portions or regions that are not complementary to the target RNA. In some embodiments, a region of a probe that is complementary to a target RNA is contiguous, such that any region of a probe that is not complementary to the target RNA does not disrupt the complementary region.

[00135] In some embodiments, the probe comprises a portion that is identically present in the target RNA. In some such embodiments, a probe that comprises a region that is identically present in the target RNA is capable of selectively hybridizing to a cDNA that has been reverse transcribed from the RNA, or to an amplicon that has been produced by amplification of the target RNA or cDNA. In some embodiments, the probe is complementary to a sufficient portion of the cDNA or amplicon such that it selectively hybridizes to the cDNA or amplicon under the conditions of the particular assay being used. In some embodiments, a probe that is complementary to a cDNA or amplicon is complementary to at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, or at least 24 contiguous nucleotides of the cDNA or amplicon. In some embodiments, a probe that is complementary to a target RNA comprises a region that is complementary to at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, or at least 24 contiguous nucleotides of the cDNA or amplicon. That is, a probe that is complementary to a cDNA or amplicon may also comprise portions or regions that are not complementary to the cDNA or amplicon. In some embodiments, a region of a probe that is complementary to a cDNA or amplicon is contiguous, such that any region of a probe that is not complementary to the cDNA or amplicon does not disrupt the complementary region.

[00136] Nonlimiting exemplary probes include probes comprising sequences set forth in SEQ ID NOs: 1 to 41 and 133 to 211. Nonlimiting exemplary probes include probes comprising sequences that are identically present in, or complementary to a region of, sequences selected from SEQ ID NOs: 1 to 41, 133 to 211, and 345 to 388. Exemplary probes also include, but are not limited to, probes comprising regions that are identical or complementary to at least 15 contiguous nucleotides of sequences selected from SEQ ID NOs: 1 to 41, 133 to 211, and 345 to 388.

[00137] In some embodiments, the method of detectably quantifying one or more target RNAs comprises: (a) isolating total RNA; (b) reverse transcribing a target RNA to produce a cDNA that is complementary to the target RNA; (c) amplifying the cDNA from (b); and (d) detecting the amount of a target RNA using real time RT-PCR and a detection probe.

[00138] As described above, in some embodiments, the real time RT-PCR detection is performed using a FRET probe, which includes, but is not limited to, a TaqMan® probe, a Molecular beacon probe and a Scorpion probe. In some embodiments, the real time RT-PCR detection and quantification is performed with a TaqMan® probe, *i.e.*, a linear probe that typically has a fluorescent dye covalently bound at one end of the DNA and a quencher molecule covalently bound at the other end of the DNA. The FRET probe comprises a sequence that is complementary to a region of the cDNA such that, when the FRET probe is hybridized to the cDNA, the dye fluorescence is quenched, and when the probe is digested during amplification of the cDNA, the dye is released from the probe and produces a fluorescence signal. In such embodiments, the amount of target RNA in the sample is proportional to the amount of fluorescence measured during cDNA amplification.

[00139] The TaqMan® probe typically comprises a region of contiguous nucleotides having a sequence that is complementary to a region of a target RNA or its complementary cDNA that is reverse transcribed from the target RNA template (*i.e.*, the sequence of the probe region is complementary to or identically present in the target RNA to be detected) such that the probe is specifically hybridizable to the resulting PCR amplicon. In some embodiments, the probe comprises a region of at least 6 contiguous nucleotides having a sequence that is fully complementary to or identically present in a

region of a cDNA that has been reverse transcribed from a target RNA template, such as comprising a region of at least 8 contiguous nucleotides, at least 10 contiguous nucleotides, at least 12 contiguous nucleotides, at least 14 contiguous nucleotides, or at least 16 contiguous nucleotides having a sequence that is complementary to or identically present in a region of a cDNA reverse transcribed from a target RNA to be detected.

[00140] In some embodiments, the region of the cDNA that has a sequence that is complementary to the TaqMan® probe sequence is at or near the center of the cDNA molecule. In some embodiments, there are independently at least 2 nucleotides, such as at least 3 nucleotides, such as at least 4 nucleotides, such as at least 5 nucleotides of the cDNA at the 5'-end and at the 3'-end of the region of complementarity.

[00141] In some embodiments, Molecular Beacons can be used to detect and quantitate PCR products. Like TaqMan® probes, Molecular Beacons use FRET to detect and quantitate a PCR product via a probe having a fluorescent dye and a quencher attached at the ends of the probe. Unlike TaqMan® probes, Molecular Beacons remain intact during the PCR cycles. Molecular Beacon probes form a stem-loop structure when free in solution, thereby allowing the dye and quencher to be in close enough proximity to cause fluorescence quenching. When the Molecular Beacon hybridizes to a target, the stem-loop structure is abolished so that the dye and the quencher become separated in space and the dye fluoresces. Molecular Beacons are available, e.g., from Gene Link™ (see <http://www.genelink.com/newsite/products/mbintro.asp>).

[00142] In some embodiments, Scorpion probes can be used as both sequence-specific primers and for PCR product detection and quantitation. Like Molecular Beacons, Scorpion probes form a stem-loop structure when not hybridized to a target nucleic acid. However, unlike Molecular Beacons, a Scorpion probe achieves both sequence-specific priming and PCR product detection. A fluorescent dye molecule is attached to the 5'-end of the Scorpion probe, and a quencher is attached to the 3'-end. The 3' portion of the probe is complementary to the extension product of the PCR primer, and this complementary portion is linked to the 5'-end of the probe by a non-amplifiable moiety. After the Scorpion primer is extended, the target-specific sequence of the probe binds to its complement within the extended amplicon, thus opening up the stem-loop structure and allowing the dye on the 5'-end to fluoresce and generate a signal. Scorpion

probes are available from, e.g, Premier Biosoft International (see http://www.premierbiosoft.com/tech_notes/Scorpion.html).

[00143] In some embodiments, labels that can be used on the FRET probes include colorimetric and fluorescent labels such as Alexa Fluor dyes, BODIPY dyes, such as BODIPY FL; Cascade Blue; Cascade Yellow; coumarin and its derivatives, such as 7-amino-4-methylcoumarin, aminocoumarin and hydroxycoumarin; cyanine dyes, such as Cy3 and Cy5; eosins and erythrosins; fluorescein and its derivatives, such as fluorescein isothiocyanate; macrocyclic chelates of lanthanide ions, such as Quantum DyeTM; Marina Blue; Oregon Green; rhodamine dyes, such as rhodamine red, tetramethylrhodamine and rhodamine 6G; Texas Red; fluorescent energy transfer dyes, such as thiazole orange-ethidium heterodimer; and, TOTAB.

[00144] Specific examples of dyes include, but are not limited to, those identified above and the following: Alexa Fluor 350, Alexa Fluor 405, Alexa Fluor 430, Alexa Fluor 488, Alexa Fluor 500, Alexa Fluor 514, Alexa Fluor 532, Alexa Fluor 546, Alexa Fluor 555, Alexa Fluor 568, Alexa Fluor 594, Alexa Fluor 610, Alexa Fluor 633, Alexa Fluor 647, Alexa Fluor 660, Alexa Fluor 680, Alexa Fluor 700, and, Alexa Fluor 750; amine-reactive BODIPY dyes, such as BODIPY 493/503, BODIPY 530/550, BODIPY 558/568, BODIPY 564/570, BODIPY 576/589, BODIPY 581/591, BODIPY 630/650, BODIPY 650/655, BODIPY FL, BODIPY R6G, BODIPY TMR, and, BODIPY-TR; Cy3, Cy5, 6-FAM, Fluorescein Isothiocyanate, HEX, 6-JOE, Oregon Green 488, Oregon Green 500, Oregon Green 514, Pacific Blue, REG, Rhodamine Green, Rhodamine Red, Renographin, ROX, SYPRO, TAMRA, 2', 4', 5', 7'-Tetrabromosulfonefluorescein, and TET.

[00145] Specific examples of fluorescently labeled ribonucleotides useful in the preparation of RT-PCR probes for use in some embodiments of the methods described herein are available from Molecular Probes (Invitrogen), and these include, Alexa Fluor 488-5-UTP, Fluorescein-12-UTP, BODIPY FL-14-UTP, BODIPY TMR-14-UTP, Tetramethylrhodamine-6-UTP, Alexa Fluor 546-14-UTP, Texas Red-5-UTP, and BODIPY TR-14-UTP. Other fluorescent ribonucleotides are available from Amersham Biosciences (GE Healthcare), such as Cy3-UTP and Cy5-UTP.

[00146] Examples of fluorescently labeled deoxyribonucleotides useful in the preparation of RT-PCR probes for use in the methods described herein include Dinitrophenyl (DNP)-1'-dUTP, Cascade Blue-7-dUTP, Alexa Fluor 488-5-dUTP, Fluorescein-12-dUTP, Oregon Green 488-5-dUTP, BODIPY FL-14-dUTP, Rhodamine Green-5-dUTP, Alexa Fluor 532-5-dUTP, BODIPY TMR-14-dUTP, Tetramethylrhodamine-6-dUTP, Alexa Fluor 546-14-dUTP, Alexa Fluor 568-5-dUTP, Texas Red-12-dUTP, Texas Red-5-dUTP, BODIPY TR-14-dUTP, Alexa Fluor 594-5-dUTP, BODIPY 630/650-14-dUTP, BODIPY 650/665-14-dUTP; Alexa Fluor 488-7-OBEA-dCTP, Alexa Fluor 546-16-OBEA-dCTP, Alexa Fluor 594-7-OBEA-dCTP, Alexa Fluor 647-12-OBEA-dCTP. Fluorescently labeled nucleotides are commercially available and can be purchased from, e.g., Invitrogen.

[00147] In some embodiments, dyes and other moieties, such as quenchers, are introduced into polynucleotide used in the methods described herein, such as FRET probes, via modified nucleotides. A "modified nucleotide" refers to a nucleotide that has been chemically modified, but still functions as a nucleotide. In some embodiments, the modified nucleotide has a chemical moiety, such as a dye or quencher, covalently attached, and can be introduced into a polynucleotide, for example, by way of solid phase synthesis of the polynucleotide. In other embodiments, the modified nucleotide includes one or more reactive groups that can react with a dye or quencher before, during, or after incorporation of the modified nucleotide into the nucleic acid. In specific embodiments, the modified nucleotide is an amine-modified nucleotide, *i.e.*, a nucleotide that has been modified to have a reactive amine group. In some embodiments, the modified nucleotide comprises a modified base moiety, such as uridine, adenosine, guanosine, and/or cytosine. In specific embodiments, the amine-modified nucleotide is selected from 5-(3-aminoallyl)-UTP; 8-[(4-amino)butyl]-amino-ATP and 8-[(6-amino)butyl]-amino-ATP; N6-(4-amino)butyl-ATP, N6-(6-amino)butyl-ATP, N4-[2,2-oxy-bis-(ethylamine)]-CTP; N6-(6-Amino)hexyl-ATP; 8-[(6-Amino)hexyl]-amino-ATP; 5-propargylamino-CTP, 5-propargylamino-UTP. In some embodiments, nucleotides with different nucleobase moieties are similarly modified, for example, 5-(3-aminoallyl)-GTP instead of 5-(3-aminoallyl)-UTP. Many amine modified nucleotides are commercially available from, e.g., Applied Biosystems, Sigma, Jena Bioscience and TriLink.

[00148] Exemplary detectable moieties also include, but are not limited to, members of binding pairs. In some such embodiments, a first member of a binding pair is linked to a polynucleotide. The second member of the binding pair is linked to a detectable label, such as a fluorescent label. When the polynucleotide linked to the first member of the binding pair is incubated with the second member of the binding pair linked to the detectable label, the first and second members of the binding pair associate and the polynucleotide can be detected. Exemplary binding pairs include, but are not limited to, biotin and streptavidin, antibodies and antigens, etc.

[00149] In some embodiments, multiple target RNAs are detected in a single multiplex reaction. In some such embodiments, each probe that is targeted to a unique cDNA is spectrally distinguishable when released from the probe. Thus, each target RNA is detected by a unique fluorescence signal.

[00150] One skilled in the art can select a suitable detection method for a selected assay, e.g., a real-time RT-PCR assay. The selected detection method need not be a method described above, and may be any method.

4.2. Exemplary compositions and kits

[00151] In another aspect, compositions are provided. In some embodiments, compositions are provided for use in the methods described herein.

[00152] In some embodiments, a composition comprises at least one polynucleotide. In some embodiments, a composition comprises at least one primer. In some embodiments, a composition comprises at least one probe. In some embodiments, a composition comprises at least one primer and at least one probe.

[00153] In some embodiments, compositions are provided that comprise at least one target RNA-specific primer. The term “target RNA-specific primer” encompasses primers that have a region of contiguous nucleotides having a sequence that is (i) identically present in one of SEQ ID NOs: 1 to 41 or 133 to 211, (ii) complementary to the sequence of a region of contiguous nucleotides found in one of SEQ ID NOs: 1 to 41 or 133 to 211; (iii) complementary to the sequence of a region of contiguous nucleotides found in one of SEQ ID NOs: 345 to 388; or (iv) identically present in one of SEQ ID NOs: 345 to 388.

[00154] In some embodiments, compositions are provided that comprise at least one target RNA-specific probe. The term “target RNA-specific probe” encompasses probes that have a region of contiguous nucleotides having a sequence that is (i) identically present in one of SEQ ID NOs: 1 to 41 or 133 to 211, (ii) complementary to the sequence of a region of contiguous nucleotides found in one of SEQ ID NOs: 1 to 41 or 133 to 211; (iii) complementary to the sequence of a region of contiguous nucleotides found in one of SEQ ID NOs: 345 to 388; or (iv) identically present in one of SEQ ID NOs: 345 to 388.

[00155] In some embodiments, target RNA-specific primers and probes comprise deoxyribonucleotides. In other embodiments, target RNA-specific primers and probes comprise at least one nucleotide analog. Nonlimiting exemplary nucleotide analogs include, but are not limited to, analogs described herein, including LNA analogs and peptide nucleic acid (PNA) analogs. In some embodiments, target RNA-specific primers and probes comprise at least one nucleotide analog which increases the hybridization binding energy (e.g., an affinity-enhancing nucleotide analog, discussed above). In some embodiments, a target RNA-specific primer or probe in the compositions described herein binds to one target RNA in the sample. In some embodiments, a single primer or probe binds to multiple target RNAs, such as multiple isomers.

[00156] In some embodiments, more than one primer or probe specific for a single target RNA is present in the compositions, the primers or probes capable of binding to overlapping or spatially separated regions of the target RNA.

[00157] It will be understood, even if not explicitly stated hereinafter, that in some embodiments in which the compositions described herein are designed to hybridize to cDNAs reverse transcribed from target RNAs, the composition comprises at least one target RNA-specific primer or probe (or region thereof) having a sequence that is identically present in a target RNA (or region thereof).

[00158] In some embodiments, a target RNA is capable of specifically hybridizing to at least one probe comprising a sequence selected from SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7 and 8. In some embodiments, a target RNA is capable of specifically hybridizing to at least one nucleic acid probe comprising a sequence selected from SEQ

ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12. In some embodiments, a target RNA is capable of specifically hybridizing to at least one nucleic acid probe comprising a sequence selected from SEQ ID NOs: 142, 151, 153, 193, 194, 205, 172, 208, 210, and 211. In some embodiments, a target RNA is capable of specifically hybridizing to at least one nucleic acid probe comprising a sequence selected from SEQ ID NOs: 1, 5, 7, and 32. In some embodiments, a target RNA is capable of specifically hybridizing to at least one probe comprising a sequence selected from SEQ ID NOs: 1 to 41 and 133 to 211. In some embodiments, a target RNA comprises at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388. In some embodiments, a target RNA comprises a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 1 to 41 and 133 to 211. In some embodiments, a target RNA, in its mature form, comprises fewer than 30 nucleotides. In some embodiments, a target RNA is a microRNA.

[00159] In some embodiments, the composition comprises a plurality of target RNA-specific primers and/or probes for each of at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, or at least 8 target RNAs, the target RNAs comprising a region of contiguous nucleotides having a sequence that is identically present in one of SEQ ID NOs: 42, 43, 44, 45, 46, 47, 48, or 49. In some embodiments, the plurality includes a target RNA-specific primer and/or probe specific for each of at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, or at least 12 target RNAs, the target RNAs comprising a region of contiguous nucleotides having a sequence that is identically present in one of SEQ ID NOs: 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, and 53. In some embodiments, the plurality includes a target RNA-specific primer and/or probe specific for each of at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, at least 75, or at least 100 target RNAs comprising a region of contiguous nucleotides having a sequence that is identically present in one of SEQ ID NOs: 42 to 82 or 226 to 317. It will be understood that, in some embodiments, target RNAs described herein comprise a sequence identically present in a sequence set forth in Table 2 or Table 12, except that thymine (T) bases in the sequences shown in Table 2 or Table 12 are replaced by uracil (U) bases in the target RNAs.

[00160] In some embodiments, a composition is an aqueous composition. In some embodiments, the aqueous composition comprises a buffering component, such as phosphate, tris, HEPES, etc., and/or additional components, as discussed below. In some embodiments, a composition is dry, for example, lyophilized, and suitable for reconstitution by addition of fluid. A dry composition may include a buffering component and/or additional components.

[00161] In some embodiments, a composition comprises one or more additional components. Additional components include, but are not limited to, salts, such as NaCl, KCl, and MgCl₂; polymerases, including thermostable polymerases; dNTPs; RNase inhibitors; bovine serum albumin (BSA) and the like; reducing agents, such as β -mercaptoethanol; EDTA and the like; etc. One skilled in the art can select suitable composition components depending on the intended use of the composition.

[00162] In some embodiments, an addressable microarray component is provided that comprises target RNA-specific probes attached to a substrate.

[00163] Microarrays for use in the methods described herein comprise a solid substrate onto which the probes are covalently or non-covalently attached. In some embodiments, probes capable of hybridizing to one or more target RNAs or cDNAs are attached to the substrate at a defined location ("addressable array"). Probes can be attached to the substrate in a wide variety of ways, as will be appreciated by those in the art. In some embodiments, the probes are synthesized first and subsequently attached to the substrate. In other embodiments, the probes are synthesized on the substrate. In some embodiments, probes are synthesized on the substrate surface using techniques such as photopolymerization and photolithography.

[00164] In some embodiments, the solid substrate is a material that is modified to contain discrete individual sites appropriate for the attachment or association of the probes and is amenable to at least one detection method. Representative examples of substrates include glass and modified or functionalized glass, plastics (including acrylics, polystyrene and copolymers of styrene and other materials, polypropylene, polyethylene, polybutylene, polyurethanes, Teflon[®], etc.), polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified

silicon, carbon, metals, inorganic glasses and plastics. In some embodiments, the substrates allow optical detection without appreciably fluorescing.

[00165] In some embodiments, the substrate is planar. In other embodiments, probes are placed on the inside surface of a tube, such as for flow-through sample analysis to minimize sample volume. In other embodiments, probes can be in the wells of multi-well plates. In still other embodiments, probes can be attached to an addressable microbead array. In yet other embodiments, the probes can be attached to a flexible substrate, such as a flexible foam, including closed cell foams made of particular plastics.

[00166] The substrate and the probe can each be derivatized with functional groups for subsequent attachment of the two. For example, in some embodiments, the substrate is derivatized with one or more chemical functional groups including, but not limited to, amino groups, carboxyl groups, oxo groups and thiol groups. In some embodiments, probes are attached directly to the substrate through one or more functional groups. In some embodiments, probes are attached to the substrate indirectly through a linker (*i.e.*, a region of contiguous nucleotides that space the probe regions involved in hybridization and detection away from the substrate surface). In some embodiments, probes are attached to the solid support through the 5' terminus. In other embodiments, probes are attached through the 3' terminus. In still other embodiments, probes are attached to the substrate through an internal nucleotide. In some embodiments the probe is attached to the solid support non-covalently, e.g., via a biotin-streptavidin interaction, wherein the probe is biotinylated and the substrate surface is covalently coated with streptavidin.

[00167] In some embodiments, the compositions comprise a microarray having probes attached to a substrate, wherein at least one of the probes (or a region thereof) comprises a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1 to 41, 133 to 211, or SEQ ID NOs: 345 to 388. In some embodiments, at least 2, at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, or at least 100 of the probes comprise a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1 to 41, 133 to 211, or 345 to 388. In some embodiments, the microarray comprises at least one target

RNA-specific probe comprising a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1 to 41, 133 to 211, or 345 to 388 and at least one target RNA-specific probe comprising a sequence that is identically present in, or complementary to a region of, a target RNA set forth in Table 3. In some embodiments, the microarray comprises each target RNA-specific probe at only one location on the microarray. In some embodiments, the microarray comprises at least one target RNA-specific probe at multiple locations on the microarray.

[00168] As used herein, the terms “complementary” or “partially complementary” to a target RNA (or target region thereof), and the percentage of “complementarity” of the probe sequence to that of the target RNA sequence is the percentage “identity” to the reverse complement of the sequence of the target RNA. In determining the degree of “complementarity” between probes used in the compositions described herein (or regions thereof) and a target RNA, such as those disclosed herein, the degree of “complementarity” is expressed as the percentage identity between the sequence of the probe (or region thereof) and the reverse complement of the sequence of the target RNA that best aligns therewith. The percentage is calculated by counting the number of aligned bases that are identical as between the 2 sequences, dividing by the total number of contiguous nucleotides in the probe, and multiplying by 100.

[00169] In some embodiments, the microarray comprises at least one probe having a region with a sequence that is fully complementary to a target region of a target RNA. In other embodiments, the microarray comprises at least one probe having a region with a sequence that comprises one or more base mismatches when compared to the sequence of the best-aligned target region of a target RNA.

[00170] In some embodiments, the microarray comprises at least one probe having a region of at least 10, at least 11, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, or at least 25 contiguous nucleotides with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1 to 41 or 133 to 211. In some embodiments, the microarray comprises at least one probe having a region of at least 10, at least 11, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, or at least 25 contiguous

nucleotides with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1 to 41 or 133 to 211, and at least one probe comprising a region of at least 10, at least 11, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, or at least 25 contiguous nucleotides having a sequence that is identically present in, or complementary to a region of, a target RNA set forth in Table 3.

[00171] In some embodiments, the microarray comprises at least one probe having a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. In some embodiments, the microarray comprises at least one, at least two, at least three, at least five, or eight probes that each comprise a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7 or 8. In some embodiments, the microarray further comprises additional probes that do not have a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7 and 8.

[00172] In some embodiments, the microarray comprises at least one probe having a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12. In some embodiments, the microarray comprises at least one, at least two, at least three, at least five, at least eight, at least 10, or at least 12 probes that each comprise a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12. In some embodiments, the microarray further comprises additional probes that do not have a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12.

[00173] In some embodiments, the microarray comprises at least one probe having a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 142, 151, 153, 193, 194, 205, 172, 208, 210, and 211. In some embodiments, the microarray comprises at least one, at least two, at least three, at least five, at least eight, or at least 10 probes that each comprise a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 142,

151, 153, 193, 194, 205, 172, 208, 210, and 211. In some embodiments, the microarray further comprises additional probes that do not have a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 142, 151, 153, 193, 194, 205, 172, 208, 210, and 211.

[00174] In some embodiments, the microarray comprises at least one probe having a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1, 5, 7, or 32. In some embodiments, the microarray comprises at least one, at least two, at least three, or at least four probes that each comprise a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1, 5, 7, or 32. In some embodiments, the microarray further comprises additional probes that do not have a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1, 5, 7, or 32.

[00175] In some embodiments, the microarrays comprise probes having a region with a sequence that is complementary to target RNAs that comprise a substantial portion of the human miRNome (*i.e.*, the publicly known microRNAs that have been accessioned by others into miRBase (<http://microrna.sanger.ac.uk/> at the time the microarray is fabricated), such as at least about 60%, at least about 70%, at least about 80%, at least about 90%, even at least about 95% of the human miRNome. In some embodiments, the microarrays comprise probes that have a region with a sequence that is identically present in target RNAs that comprise a substantial portion of the human miRNome, such as at least about 60%, at least about 70%, at least about 80%, at least about 90%, even at least about 95% of the human miRNome.

[00176] In some embodiments, components are provided that comprise probes attached to microbeads, such as those sold by Luminex, each of which is internally dyed with red and infrared fluorophores at different intensities to create a unique signal for each bead. In some embodiments, the compositions useful for carrying out the methods described herein include a plurality of microbeads, each with a unique spectral signature. Each uniquely labeled microbead is attached to a unique target RNA-specific probe such that the unique spectral signature from the dyes in the bead is associated with a particular probe sequence. Nonlimiting exemplary probe sequences include SEQ ID

NOs: 1 to 41 and 133 to 211. Nonlimiting exemplary probe sequences also include probes comprising a region that is identically present in, or complementary to, a sequence selected from SEQ ID NOs: 1 to 41, 133 to 211, and SEQ ID NOs: 345 to 388. In some embodiments, a probe sequence comprises at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, or at least 24 contiguous nucleotides that are identically present in, or complementary to a region of, SEQ ID NOs: 1 to 41, 133 to 211, and SEQ ID NOs: 345 to 388.

[00177] In some embodiments, a uniquely labeled microbead has attached thereto a probe having a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1 to 41, 133 to 211, or 345 to 388. In other embodiments, the uniquely labeled microbead has attached thereto a probe having a region with a sequence that comprises one or more base mismatches when compared to the most similar sequence selected from SEQ ID NOs: 1 to 41, 133 to 211, and 345 to 388, and sequences complementary to SEQ ID NOs: 1 to 41, 133 to 211.

[00178] In some embodiments, a composition is provided that comprises a plurality of uniquely labeled microbeads, wherein at least one microbead has attached thereto a probe having a region of at least 10, at least 11, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, or at least 25 contiguous nucleotides with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1 to 41, 133 to 211 or 345 to 388. In some embodiments, a composition comprises a plurality of uniquely labeled microbeads, wherein at least one of the microbeads has attached thereto a probe having a region of at least 10, at least 11, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, or at least 25 contiguous nucleotides with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1 to 41, 133 to 211, or 345 to 388, and at least a second microbead having attached thereto a probe comprising a region of at least 10, at least 11, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, or at least 25

contiguous nucleotides having a sequence that is identically present in, or complementary to a region of, a target RNA set forth in Table 3.

[00179] In some embodiments, the compositions comprise a plurality of uniquely labeled microbeads, at least one of which has attached thereto a target RNA-specific probe having a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. In some embodiments, the compositions comprise at least two, at least three, at least five, or at least 8 uniquely labeled microbeads that each have attached thereto a unique target RNA-specific probe having a region with a sequence that is identically present in, or complementary to a region of, a different one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7 or 8. In some embodiments, the composition comprises at least one uniquely labeled microbead having attached thereto a target RNA-specific probe having a region with a sequence that is not present in, or complementary to a region of, any of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8.

[00180] In some embodiments, the compositions comprise plurality of uniquely labeled microbeads, wherein at least one microbead has attached thereto a probe having a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12. In some embodiments, the compositions comprise at least two, at least three, at least five, at least eight, at least 10, or at least 12 uniquely labeled microbeads that each have attached thereto a unique target RNA-specific probe having a region with a sequence that is identically present in, or complementary to a region of, a different one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12. In some embodiments, the composition comprises at least one uniquely labeled microbead having attached thereto a target RNA-specific probe having a region with a sequence that is not present in, or complementary to a region of, any of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12.

[00181] In some embodiments, the compositions comprise plurality of uniquely labeled microbeads, wherein at least one microbead has attached thereto a probe having a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 142, 151, 153, 193, 194, 205, 172, 208, 210, and 211. In some embodiments, the compositions comprise at least two, at least three, at least five, at

least eight, or at least 10 uniquely labeled microbeads that each have attached thereto a unique target RNA-specific probe having a region with a sequence that is identically present in, or complementary to a region of, a different one of SEQ ID NOs: 142, 151, 153, 193, 194, 205, 172, 208, 210, and 211. In some embodiments, the composition comprises at least one uniquely labeled microbead having attached thereto a target RNA-specific probe having a region with a sequence that is not present in, or complementary to a region of, any of SEQ ID NOs: 142, 151, 153, 193, 194, 205, 172, 208, 210, and 211.

[00182] In some embodiments, the compositions comprise plurality of uniquely labeled microbeads, wherein at least one microbead has attached thereto a probe having a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1, 5, 7, or 32. In some embodiments, the compositions comprise at least two, at least three, or at least four uniquely labeled microbeads that each have attached thereto a unique target RNA-specific probe having a region with a sequence that is identically present in, or complementary to a region of, a different one of SEQ ID NOs: 1, 5, 7, or 32. In some embodiments, the composition comprises at least one uniquely labeled microbead having attached thereto a target RNA-specific probe having a region with a sequence that is not present in, or complementary to a region of, any of SEQ ID NOs: 1, 5, 7, or 32.

[00183] In some embodiments, the compositions comprise a plurality of uniquely labeled microbeads, wherein the plurality comprises at least one microbead having attached thereto a probe having a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1 to 41, 133 to 211, or SEQ ID NOs: 345 to 388. In some embodiments, the plurality comprises at least two, at least five, at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, at least 60, at least 75, or at least 100 microbeads each of which having attached thereto a probe having a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1 to 41, 133 to 211, or 345 to 388. In some embodiments, a composition comprises at least one uniquely labeled microbead having attached thereto a target RNA-specific probe having a region with a sequence that is not present in, or complementary to a region of, any of SEQ ID NOs: 1 to 41 or 133 to 211.

[00184] In some embodiments, the compositions comprise a plurality of uniquely labeled microbeads, at least one of which has attached thereto a probe having a region with a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1 to 41, 133 to 211, or 345 to 388, and at least a second bead that has attached thereto a probe having a region with a sequence that is identically present in, or complementary to a region of, a target RNA set forth in Table 3.

[00185] In some embodiments, the compositions comprise a plurality of uniquely labeled microbeads, each of which has attached thereto a unique probe having a region that is complementary to target RNAs that comprise a substantial portion of the human miRNome, such as at least about 60%, at least about 70%, at least about 80%, at least about 90%, or at least about 95% of the human miRNome. In some embodiments, the compositions comprise a plurality of uniquely labeled microbeads having attached thereto a unique probe having a region with a sequence that is identically present in target RNAs that comprise a substantial portion of the human miRNome, such as at least about 60%, at least about 70%, at least about 80%, at least about 90%, or at least about 95% of the human miRNome.

[00186] In some embodiments, compositions are provided that comprise at least one polynucleotide for detecting at least one target RNA. In some embodiments, the polynucleotide is used as a primer for a reverse transcriptase reaction. In some embodiments, the polynucleotide is used as a primer for amplification. In some embodiments, the polynucleotide is used as a primer for RT-PCR. In some embodiments, the polynucleotide is used as a probe for detecting at least one target RNA. In some embodiments, the polynucleotide is detectably labeled. In some embodiments, the polynucleotide is a FRET probe. In some embodiments, the polynucleotide is a TaqMan® probe, a Molecular Beacon, or a Scorpion probe.

[00187] In some embodiments, a composition comprises at least one FRET probe having a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1 to 41, 133 to 211, or 345 to 388. In some embodiments, a composition comprises at least two, at least five, at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, at least 60, at least 75, or at least 100 FRET probes, each of which has a sequence that is identically present in, or complementary to a

region of, a different one of SEQ ID NOs: 1 to 41, 133 to 211, or 345 to 388. In some embodiments, a FRET probe is labeled with a donor/acceptor pair such that when the probe is digested during the PCR reaction, it produces a unique fluorescence emission that is associated with a specific target RNA. In some embodiments, when a composition comprises multiple FRET probes, each probe is labeled with a different donor/acceptor pair such that when the probe is digested during the PCR reaction, each one produces a unique fluorescence emission that is associated with a specific probe sequence and/or target RNA. In some embodiments, the sequence of the FRET probe is complementary to a target region of a target RNA. In other embodiments, the FRET probe has a sequence that comprises one or more base mismatches when compared to the sequence of the best-aligned target region of a target RNA.

[00188] In some embodiments, a composition comprises a FRET probe consisting of at least 8, at least 9, at least 10, at least 11, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, or at least 25 nucleotides, wherein at least a portion of the sequence is identically present in, or complementary to a region of, one of SEQ ID NOs: 1 to 41, 133 to 211, and 345 to 388. In some embodiments, at least 8, at least 9, at least 10, at least 11, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, or at least 25 nucleotides of the FRET probe are identically present in, or complementary to a region of, one of SEQ ID NOs: 1 to 41, 133 to 211, and 345 to 388. In some embodiments, the FRET probe has a sequence with one, two or three base mismatches when compared to the sequence or complement of one of SEQ ID NOs: 1 to 41 or 133 to 211.

[00189] In some embodiments, the compositions further comprise a FRET probe consisting of at least 10, at least 11, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, or at least 25 contiguous nucleotides, wherein the FRET probe comprises a sequence that is identically present in, or complementary to a region of, a region of a target RNA set forth in Table 3. In some embodiments, the FRET probe is identically present in, or complementary to a region of, at least at least 10, at least 11, at least 13, at least 14, at

least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, or at least 24 contiguous nucleotides of a target RNA set forth in Table 3.

[00190] In some embodiments, the compositions comprise at least one target RNA-specific FRET probe comprising a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. In some embodiments, the compositions comprise at least two, at least three, at least five, or at least 8 uniquely labeled target RNA-specific FRET probes, each comprising a sequence that is identically present in, or complementary to a region of, a different one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7 and 8.

[00191] In some embodiments, the compositions comprise at least one target RNA-specific FRET probe comprising a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12. In some embodiments, the compositions comprise at least two, at least three, at least five, at least eight, at least 10, or at least 12 uniquely labeled target RNA-specific FRET probes, each of which comprises a sequence that is identically present in, or complementary to a region of, a different one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12.

[00192] In some embodiments, the compositions comprise at least one target RNA-specific FRET probe comprising a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 142, 151, 153, 193, 194, 205, 172, 208, 210, and 211. In some embodiments, the compositions comprise at least two, at least three, at least five, at least eight, at least nine, or at least 10 uniquely labeled target RNA-specific FRET probes, each of which comprises a sequence that is identically present in, or complementary to a region of, a different one of SEQ ID NOs: 142, 151, 153, 193, 194, 205, 172, 208, 210, and 211.

[00193] In some embodiments, the compositions comprise at least one target RNA-specific FRET probe comprising a sequence that is identically present in, or complementary to a region of, one of SEQ ID NOs: 1, 5, 7, or 32. In some embodiments, the compositions comprise at least two, at least three, or at least four uniquely labeled target RNA-specific FRET probes, each of which comprises a sequence that is identically

present in, or complementary to a region of, a different one of SEQ ID NOs: 1, 5, 7, or 32.

[00194] In some embodiments, a kit comprises a polynucleotide discussed above. In some embodiments, a kit comprises at least one primer and/or probe discussed above. In some embodiments, a kit comprises at least one polymerase, such as a thermostable polymerase. In some embodiments, a kit comprises dNTPs. In some embodiments, kits for use in the real time RT-PCR methods described herein comprise one or more target RNA-specific FRET probes and/or one or more primers for reverse transcription of target RNAs and/or one or more primers for amplification of target RNAs or cDNAs reverse transcribed therefrom.

[00195] In some embodiments, one or more of the primers and/or probes is “linear”. A “linear” primer refers to a polynucleotide that is a single stranded molecule, and typically does not comprise a short region of, for example, at least 3, 4 or 5 contiguous nucleotides, which are complementary to another region within the same polynucleotide such that the primer forms an internal duplex. In some embodiments, the primers for use in reverse transcription comprise a region of at least 4, such as at least 5, such as at least 6, such as at least 7 or more contiguous nucleotides at the 3'-end that has a sequence that is complementary to region of at least 4, such as at least 5, such as at least 6, such as at least 7 or more contiguous nucleotides at the 5'-end of a target RNA.

[00196] In some embodiments, a kit comprises one or more pairs of linear primers (a “forward primer” and a “reverse primer”) for amplification of a cDNA reverse transcribed from a target RNA. Accordingly, in some embodiments, a first primer comprises a region of at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, or at least 10 contiguous nucleotides having a sequence that is identical to the sequence of a region of at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, or at least 10 contiguous nucleotides at the 5'-end of a target RNA. Furthermore, in some embodiments, a second primer comprises a region of at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, or at least 10 contiguous nucleotides having a sequence that is complementary to the sequence of a region of at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, or at least 10 contiguous nucleotides at the 3'-end of a target RNA. In some embodiments, the kit comprises at least a first set of primers for amplification of a

cDNA that is reverse transcribed from a target RNA capable of specifically hybridizing to a nucleic acid comprising a sequence identically present in one of SEQ ID NOs: 1 to 41 and 133 to 211 and/or a cDNA that is reverse transcribed from a target RNA that comprises at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388. In some embodiments, the kit further comprises at least a second set of primers for amplification of a cDNA that is reverse transcribed from a target RNA set forth in Table 3.

[00197] In some embodiments, the kit comprises at least two, at least five, at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, at least 60, at least 75, or at least 100 sets of primers, each of which is for amplification of a cDNA that is reverse transcribed from a different target RNA capable of specifically hybridizing to a sequence selected from SEQ ID NOs: 1 to 41 and 133 to 211 and/or a cDNA that is reverse transcribed from a target RNA that comprises at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388. In some embodiments, the kit comprises at least one set of primers that is capable of amplifying more than one cDNA reverse transcribed from a target RNA in a sample.

[00198] In some embodiments, probes and/or primers for use in the compositions described herein comprise deoxyribonucleotides. In some embodiments, probes and/or primers for use in the compositions described herein comprise deoxyribonucleotides and one or more nucleotide analogs, such as LNA analogs or other duplex-stabilizing nucleotide analogs described above. In some embodiments, probes and/or primers for use in the compositions described herein comprise all nucleotide analogs. In some embodiments, the probes and/or primers comprise one or more duplex-stabilizing nucleotide analogs, such as LNA analogs, in the region of complementarity.

[00199] In some embodiments, the compositions described herein also comprise probes, and in the case of RT-PCR, primers, that are specific to one or more housekeeping genes for use in normalizing the quantities of target RNAs. Such probes (and primers) include those that are specific for one or more products of housekeeping genes selected from U6 snRNA, ACTB, B2M, GAPDH, GUSB, HPRT1, PPIA, RPLP, RRN18S, TBP, TUBB, UBC, YWHA (TATAA), PGK1, and RPL4.

[00200] In some embodiments, the kits for use in real time RT-PCR methods described herein further comprise reagents for use in the reverse transcription and amplification reactions. In some embodiments, the kits comprise enzymes such as reverse transcriptase, and a heat stable DNA polymerase, such as Taq polymerase. In some embodiments, the kits further comprise deoxyribonucleotide triphosphates (dNTP) for use in reverse transcription and amplification. In further embodiments, the kits comprise buffers optimized for specific hybridization of the probes and primers.

4.2.1. Exemplary normalization of RNA levels

[00201] In some embodiments, quantitation of target RNA expression levels requires assumptions to be made about the total RNA per cell and the extent of sample loss during sample preparation. In order to correct for differences between different samples or between samples that are prepared under different conditions, the quantities of target RNAs in some embodiments are normalized to the expression of at least one endogenous housekeeping gene.

[00202] Appropriate genes for use as reference genes in the methods described herein include those as to which the quantity of the product does not vary between normal and dysplastic or cancerous cervical cells, or between different cell lines or under different growth and sample preparation conditions. In some embodiments, endogenous housekeeping genes useful as normalization controls in the methods described herein include, but are not limited to, U6 snRNA, RNU44, RNU 48, and U47. In typical embodiments, the at least one endogenous housekeeping gene for use in normalizing the measured quantity of microRNAs is selected from U6 snRNA, U6 snRNA, RNU44, RNU 48, and U47. In some embodiments, one housekeeping gene is used for normalization. In some embodiments, more than one housekeeping gene is used for normalization.

4.2.2. Exemplary qualitative methods

[00203] In some embodiments, methods comprise detecting a qualitative change in a target RNA profile generated from a clinical sample of human cervical cells as compared to a normal target RNA profile (in some exemplary embodiments, a target RNA profile of a control sample). Some qualitative changes in the expression profile are indicative of the presence of cervical dysplasia in a sample of human cervical cells.

Various qualitative changes in the expression profile are indicative of the propensity to proceed to cervical cancer. The term "target RNA profile" refers to a set of data regarding the concurrent expression of a plurality of target RNAs in the same sample.

[00204] In some embodiments, at least one, at least two, at least three, at least four, at least five, at least six, at least seven, or at least eight of the target RNAs of the plurality of target RNAs are capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7 and 8. In some embodiments, at least one, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least 10, at least 11, or at least 12 of the target RNAs of the plurality of target RNAs is capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12. In some embodiments, at least one, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, or at least 10 of the target RNAs of the plurality of target RNAs is capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 142, 151, 153, 193, 194, 205, 172, 208, 210, and 211. In some embodiments, at least one, at least two, at least three, or at least four of the target RNAs of the plurality of target RNAs is capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1, 5, 7, and 32.

[00205] In some embodiments, at least one, at least two, at least five, at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, at least 60, at least 75, or at least 100 of the plurality of target RNAs is capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs: 1 to 41 and 133 to 211. In some embodiments, at least one, at least two, at least five, at least 10, at least 15, at least 20, at least 25, at least 30, at least 40 of the plurality of target RNAs comprises at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388. In some embodiments, at least one, at least two, at least five, at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, at least 60, at least 75, or at least 100 of the plurality of target RNAs comprises a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 1 to 41 and 133 to 211. In some embodiments, a target RNA, in its mature form,

comprises fewer than 30 nucleotides. In some embodiments, a target RNA is a microRNA.

[00206] Qualitative expression data for use in preparing target RNA expression profiles is obtained using any suitable analytical method, including the analytical methods presented herein.

[00207] In some embodiments, for example, concurrent expression data are obtained using, e.g., a microarray, as described above. Thus, in addition to use for quantitative expression level assays of specific target RNAs as described above, a microarray comprising probes having sequences that are complementary to a substantial portion of the miRNome may be employed to carry out target RNA gene expression profiling, for analysis of target RNA expression patterns.

[00208] In some embodiments, distinct target RNA signatures are associated with established markers for cervical dysplasia, or directly with the presence of cervical dysplasia. In some embodiments, distinct target RNA signatures are associated with established markers for CIN-1, CIN-2 or CIN-3 cervical dysplasia, or directly with the level of severity of cervical dysplasia. In some embodiments, distinct target RNA signatures are associated with established markers for cervical dysplasia likely to progress to carcinoma, or directly with cervical dysplasia that is likely to progress to carcinoma. In some embodiments, distinct target RNA signatures are associated with HPV infection and/or integration into the genome of the host cell. In some embodiments, distinct target RNA signatures are associated with established markers for cervical cancer, or directly with the presence of cervical cancer.

[00209] According to the expression profiling method, in some embodiments, total RNA from a sample from a subject suspected of having cervical dysplasia is quantitatively reverse transcribed to provide a set of labeled polynucleotides complementary to the RNA in the sample. The polynucleotides are then hybridized to a microarray comprising target RNA-specific probes to provide a hybridization profile for the sample. The result is a hybridization profile for the sample representing the expression pattern of target RNAs in the sample. The hybridization profile comprises the signal from the binding of the polynucleotides reverse transcribed from the sample to the target RNA-specific probes in the microarray. In some embodiments, the profile is recorded as the

presence or absence of binding (signal vs. zero signal). In some embodiments, the profile recorded includes the intensity of the signal from each hybridization. The profile is compared to the hybridization profile generated from a normal, *i.e.*, noncancerous, or in some embodiments, a control sample. An alteration in the signal is indicative of the presence of cervical dysplasia or cervical cancer in the subject.

4.3. Exemplary additional target RNAs

[00210] In some embodiments, in combination with detecting one or more target RNAs that are capable of specifically hybridizing to a nucleic acid comprising a sequence selected from SEQ ID NOs:1 to 41 and 133 to 211 and/or detecting one or more target RNAs comprising at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388 and/or detecting one or more target RNAs that comprise a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs:1 to 41 and 133 to 211, methods herein further comprise detecting the level(s) of expression of at least one other marker associated with cervical dysplasia or HPV integration.

[00211] Accordingly, in some embodiments, the methods described herein further comprise detecting increased expression of any one or more of miR-21, miR-31, miR-182, miR-183, miR-146a, miR-155, and miR-205. In some embodiments, the methods described herein further comprise detecting increased expression of any one or more of miR-663, miR-765, miR-92b*, miR-936, miR-9, miR-199a*, miR-199a, miR-199b, miR-145, miR-133a, miR-133b, miR-214, miR-127, miR-210, miR-301, miR-142-3p, miR-142-5p, miR-194, miR-215 and miR-32.

[00212] In some embodiments, the methods described herein further comprise detecting altered expression of target RNAs associated with HPV integration sites. As used herein, the term “associated with” a given HPV integration site means that the target RNA gene is located in close proximity to the HPV integration site; *i.e.*, when the target RNA is located within the same chromosomal band or within 3 megabases (3 Mb), preferably within 2.5 Mb, of the HPV integration site. Thus, in some embodiments, the methods further comprise detecting increased expression of target RNAs associated with HPV integration sites, such as fragile sites which are preferential targets for HPV16 associated with cervical tumors. Such target RNAs include: miR-186, miR-101

(associated with FRA1A on chromosome 1p36 and FRA1C on chromosome 1p31); miR-194 and miR-215 (associated with FRA1F on chromosome 1q21 and FRA1H on chromosome 1q42.1); miR-106b, miR-25 and miR93 (associated with FRA7F on chromosome 7q22); miR-29b, miR-29a, miR-96, miR-182-5p, miR-182-3p, miR-183, and miR-129-1 (associated with FRA7G on chromosome 7q31.2 and FRA7H on chromosome 7q32.3); let7-1a, let7-d, let-7f-1, miR-23b, miR-24-1, and miR-27b (associated with FRA9D on chromosome 9q22.1); miR-32 (associated with FRA9E on chromosome 9q32-33.1); miR159-1 and miR-192 (associated with FRA11A on chromosome 11q13.3); miR-125b-1, let-7a-2, and miR-100 (associated with FRA11B on chromosome 11q23.3); miR-196-2 and miR-148b (associated with FRA12A on chromosome 12q13.1); miR-190 (associated with FRA15A on chromosome 15q22); miR-21, miR-301, miR-142-5p, and miR-142-3p (associated with FRA17B on chromosome 17q23.1); and miR-105-1 and miR-175 (associated with FRAXF on chromosome Xq28).

[00213] In other embodiments, the methods described herein further comprise detecting altered expression of cervical cancer-associated small RNAs with non-canonical hairpins.

[00214] In other embodiments, the methods described herein further comprise detecting increased expression of the mRNA of one or more of the following genes: BIRC5, IGF2BP3, TERC, CDKN2A, MCM5, TOP2A, MYBL2, PIK3CA, DROSHA, MKI67, MMP9, and MCM2. In some embodiments, the methods described herein further comprise detecting increased expression of the mRNA of one or more of the following genes: CDKN2A, MKI67, TOP2A, and MCM5. In some embodiments, the methods described herein further comprise detecting increased expression of the mRNA of one or more of the following genes: CDKN2A, MKI67, TOP2A, MCM5, BIRC5, MMP9, and MCM2. Appropriate genes for use as reference genes when detecting mRNA expression include those as to which the quantity of the product does not vary between normal and cancerous cervical cells, or between different cell lines or under different growth and sample preparation conditions. In some embodiments, endogenous housekeeping genes useful as normalization controls in the methods described herein include, but are not limited to, ACTB, B2M, GAPDH, GUSB, HPRT1, PPIA, RPLP, TBP, TUBB, UBC, PGK1 and RPL4. In typical embodiments, the at least one endogenous

housekeeping gene for use in normalizing the measured quantity of mRNAs is selected from GAPDH, TBP and ACTB. In some embodiments, one housekeeping gene is used for normalization. In some embodiments, more than one housekeeping gene is used for normalization.

[00215] In alternative embodiments, the methods described herein further comprise detecting chromosomal codefendants, *i.e.*, target RNAs clustered near each other in the human genome which tend to be regulated together. Accordingly, in further embodiments, the methods comprise detecting the expression of one or more target microRNAs, each situated within the chromosome no more than 50,000 bp from the chromosomal location of the pre-microRNA sequences in Table 2.

[00216] The following examples are for illustration purposes only, and are not meant to be limiting in any way.

5. EXAMPLES

5.1 Example 1: MicroRNAs from Cervical Cancer Cell Lines

[00217] Using microarray analysis, 41 distinct microRNAs were demonstrated to be overexpressed in cervical cell lines.

Cell lines

[00218] Total RNA was prepared from eight different cell lines of cervical origin that are commonly used in studies of cervical dysplasia and/or carcinoma. The RNA was used for both microRNA array profiling, further described below, and mRNA expression studies.

[00219] As set forth in Table 5 below, cell lines were selected for diversity, deriving from various squamous cervical cancers (SCC) and adenocarcinomas (AC) and, in most cases, chronically infected with HPV. Cell line C-33A appears to be HPV negative, but likely originally contained HPV. In order to identify early molecular markers that indicate a high progression rate from cervical dysplasia to cancer, seven of the eight cell lines chosen were derived from primary lesions. One cell line, ME-180, was derived from a metastatic source. All cell lines were purchased from LGC Promochem (ATCC) and cultured according to ATCC's guidelines.

Table 5

Cell line	ATCC accession no.	Cancer type	HPV-type
C4-I	CRL-1594	carcinoma	HPV18
C4-II	CRL-1595	carcinoma	HPV18
HELA S3	CCL-2.2	adenocarcinoma	HPV18
Ca Ski	CRL-1550	epidermoid carcinoma	HPV16
SIHA	HTB-35	squamous cell carcinoma	HPV16
SW756	CRL-10302	squamous cell carcinoma	HPV18
C-33A	HTB-31	carcinoma	HPV negative
ME-180	HTB-33	epidermoid carcinoma metastatic site: omentum	HPV 68

[00220] All cell lines except for HeLa S3 grew normally. Growth of HeLa S3 was very slow in the beginning of culturing, taking about two weeks before the first passage was done.

Total RNA preparation and analysis

[00221] Cells from two confluent 75cm² flasks were harvested (totaling approximately 10⁷ cells). Total RNA was prepared using TRIzol® (Invitrogen™) according to the manufacturer's protocol. All RNA samples were diluted in RNase-free water and stored in -80°C (-112°F). OD260/280 was measured on a spectrophotometer.

[00222] The quantity of RNA obtained is set forth in Table 6, below.

Table 6

	[µg/ml]	Volume µl	Total µg	Ratio 28S/18S
CaSki (CRL-1550)	1000	300	300	1.8
sw756 (CRL-10302)	2716	150	407.4	1.5
C33A (HTB-31)	2236	150	335.4	1.6
ME-180 (HTB-33)	1628	150	244.2	1.6
SiHa (HTB-35)	1508	150	226.2	1.4
C4-I (CRL-1594)	1452	150	217.8	1.7
C4-II (CRL-1595)	1656	150	248.4	1.4
HeLa S3 (CCL-2.2)	3954	150	593.1	1.6

[00223] RNA quality was assessed by calculating OD 260/280 ratios, and by electrophoresis on agarose gels under denaturing conditions. The quality of all RNA samples was high as assessed using an Agilent Bioanalyser 2100, as exemplified by the electropherogram shown in Figure 1 obtained for total RNA from cell line CaSki. Figure

2 shows denaturing gel electrophoresis of total RNA from the cell lines. The quantity was sufficient for microRNA array profiling and quantitative RT-PCR of both microRNA and mRNA.

[00224] Total RNA from normal cervix was purchased for use as a control from Ambion (Applied Biosystems).

MicroRNA Enrichment

[00225] MicroRNA enrichment was performed using a Flash PAGE Fractionator (Ambion). The gel purification protocol enriches for small RNAs less than about 40 nucleotides (nt) long, including microRNAs. Briefly, a total RNA sample (prepared as above) was loaded onto a pre-cast gel using the Flash PAGE Fractionator. The total RNA fraction smaller than 40 nt (the “microRNA fraction”) was recovered after gel migration and resuspended into nuclease free water.

Microarray Analysis

Probe design and spotting

[00226] The polynucleotide probes used for microarray preparation had the configuration 5'-NH₂-(C)₆-(spacer)-(oligomer probe sequence)-3'. The 5'-amino group allowed chemical bonding onto the array support. Each also included an identical spacer sequence of 15 nt, as shown below, to prevent non-specific interactions of the polynucleotide probes with the array support:

5'AminoC6-TTGTAATACGACTCA – Oligo probe sequence. (SEQ ID NO: 90)

Probe sequences given in Table 1 omit the linker.

[00227] The probes were synthesized according to standard protocols by Eurofins MWG Operon (Ebersberg, Germany). Nexterion (Schott) microarray glass slides were used as the solid support for the microarray.

[00228] The polynucleotide probe concentration used for the spotting was 25 µmol. The probes were spotted in duplicate using the Nexterion spotting buffer provided with the array glass support by Schott with 1 % SDS (sodium dodecyl sulfate) added to allow larger spot sizes (*e.g.*, 100-150 microns compared to 70-100 microns without SDS). The spotter used was the QArray mini (Genetix) equipped with Stealth

SMP3 pins (Telechem). After deposition of one series of spots, the spotting needle was washed 5 times with 60mM NaOH before spotting the next series of probes. Each slide is designed with 32 blocks of spotted probes, with each block being a 20x20 square of spotted probes. Each probe was spotted in duplicate. Spotted glass slides were stored at 4°C until use.

MicroRNA labelling

[00229] The labelling of the microRNA fraction was adapted from a published protocol developed at EMBL (Heidelberg, Germany) by the European Molecular Biology Group (Castoldi *et al.*, “A sensitive array for microRNA expression profiling (miChip) based on locked nucleic acids (LNA),” *RNA* 2006 May;12(5):913-20. Epub 2006 Mar 15, incorporated herein by reference in its entirety). Briefly, the microRNA fraction was incubated for 6 hours at 4°C with a mixture containing 10 µM of dye-labelled tetra-nucleotide (5'-rUrUrUrU-Cy5-3') (or alternatively, 5'-rUrUrUrU-Cy3-3') (Biospring, Germany) in Ambion buffer diluted to 1X with RNase free water, 8% polyethylene glycol (PEG), 2 mM adenosine triphosphate (ATP), and T4 RNA ligase (0.7U/µl). The labelling reaction was run by heating the mixture for 15 minutes at 65°C. This procedure ligated the poly-U dye-labelled tail to the 3' end of all the microRNAs. Labelled samples were stored at 4°C before hybridization.

Array Hybridization

[00230] The labelled microRNA fraction was hybridized to the spotted arrays using a Discovery hybridization station (Ventana, Tucson, AZ). Briefly, 2 mL of a mixture of 1% BSA, 2X SSC, and 0.2 % SDS was incubated with the chips for 30 min at 42°C. Then the chips were washed once using EZ Prep buffer (Ventana) and then three more times with Ribowash (Ventana). Next, 20 µl of the labelled microRNA mixture and 180µl of ChipHybe Reagent (Ventana) were added to the array. The arrays were heated for 6 minutes at 37°C, then were incubated at 42°C for 8 hours, after which the heating was stopped. The chips were washed once with Ribowash (Ventana) and then heated for 2 minutes at 37°C. The chips were washed again with Ribowash (Ventana) with one drop of CheapClean (Ventana) added, and incubated for 2 minutes at 37°C. The chips were washed two more times using Ribowash (Ventana). The chips were then stored dry at room temperature overnight. On the following day, the final washes were done according

to Ventana's instructions for the Discovery hybridization station. The slides were washed twice with 2X SSC + 0.2X SDS buffer and then one more time with 0.1X SSC. All the slides were dried using a speed centrifuge from Arrayit (TeleChem International, Sunnyvale, CA) at room temperature and kept in the dark before scanning.

[00231] As an alternative to the ChipHybe Reagent solution (solution 1), the following solution may be used for array hybridization (solution 2) to form probe:target RNA hybrids by mixing 2 parts of 1.5X TMAC Hybridization Solution to 1 part (v:v) sample, so that the final component concentrations are 3M TMAC, 0.10% Sarkosyl, 50 mM Tris, and 4 mM EDTA, and incubating on the array at 42°C for 8 h:

1.5X TMAC Hybridization Solution

Reagent	Catalog Number	Final Conc	Amount/ 250 mL
5 M TMAC*	Sigma T3411	4.5 M	225 mL
20% Sarkosyl	-----	0.15%	1.88 mL
1 M Tris-HCl, pH 8.0	Sigma T3038	75mM	18.75 mL
0.5 M EDTA, pH 8.0	Invitrogen 15575-020	6mM	3.0 mL
H ₂ O	-----	-----	1.37 mL

*TMAC is tetramethyl ammonium chloride

Array Image Acquisition

[00232] The arrays were scanned using an Axon™ scanner (Molecular Devices, Sunnyvale, CA) and their Genepix™ software. The image was formatted in tif format, defined by an image color depth of 16 bits/pixel (1600*1600). At such setting, pixels can assume intensity values ranging from 0 to 65,535. Pixels exhibiting the maximum intensity value are "saturated" and were assigned the value of 65,535. The resolution of the array scan was set at 10 μm/pixel. For hybridization experiments using different fluorescent dyes (e.g., Cy5 and Cy3) the photomultiplier tube (PMT) was adjusted to the higher intensity spot (Cy3 is scanned at lower PMT settings than Cy5).

Array Image Analysis

[00233] The PMT of the laser scanner digitized the captured fluorescence intensity for each given “point” of a slide and stored the numerical value as a pixel corresponding to that point. A picture composed of such pixels was then analyzed.

[00234] The first task for image analysis was to detect the spot position, using a process called segmentation. Spots were segmented by circles of adaptable or fixed radius. To be reliably segmented and quantified, the spot diameter was required to be more than 5 - 6 pixels. Before segmentation an indexing grid was provided giving the approximate positions of the spots. The segmentation itself detected the limits of spots near the grid circles. Briefly, the Genepix software assigns a circle to each spot on the array (segmentation). The segmentation had to be conducted in a somewhat flexible way due to spotting imperfections and/or support deformation, as the spots were almost never on a perfectly rectangular grid.

[00235] After segmentation by the software, the circles were modified manually and adjusted onto the spots until all the spots on the array were clearly identified. At this stage, if the array presented high background noise preventing real spots from being distinguished from the background, the array was rejected for further analysis.

[00236] The second task of image analysis was to quantify spots and export the data into a result file. This was a relatively easy and well-defined task once the spots were located on the image. The statistical approach used most frequently to quantify spot intensity was the mean or median of pixels belonging to a spot. The median approach was more robust than the mean value in the presence of outlier pixels. In practice, however, there was little difference in the results obtained using mean or median.

Array Data Analysis

[00237] All the array data were analysed using the R bioconductor package (“Bioconductor: open software development for computational biology and bioinformatics,” *Genome Biol.* 2004;5(10):R80. Epub 2004 Sep 15, which is incorporated herein by reference in its entirety).

[00238] Array data were first tested for quality by comparing the spot intensities for the internal controls. One internal control (SEQ ID NO: 83; Table 7) was used as a labelling control (this synthetic RNA is added to the purified microRNA fraction before labelling), and 7 other internal controls (SEQ ID NOs: 84-89 and 405; Table 7) were used for the normalization of the data (these synthetic RNA controls are added to the total RNA fraction before hybridization at 520 fmol each/array). The probe sequences that bind to the synthetic RNAs, and certain mutant probe sequences, are also shown in Table 7 (SEQ ID NOs: 406 to 409 and 212 to 217).

Table 7	
Control Sequences used in microarray experiments	
Sequence (5'-3')	Sequence identification number
CGCGCGUCGCUUUUAUCUACUGU	SEQ ID NO: 83; CTL30_COMP
UUAUCGUUCGAUAAGUCGCGUU	SEQ ID NO: 84; CTL11_COMP
GAAGUUACUAUGUAGGCAACCU	SEQ ID NO: 85; CTL23_COMP
CGCGGGACUAAUUGUUACCGGG	SEQ ID NO: 86; CTL26_COMP
UCGCGUCGAACUCCGCAACCGA	SEQ ID NO: 87; CTL29_COMP
ACCGAACGCCGUACCCAUCGGG	SEQ ID NO: 88; CTL31_COMP
CGAGGGUAAACGACUCUCGUGUC	SEQ ID NO: 89; CTL36_COMP
GCGUACCGACGCGUAGACGGAC	SEQ ID NO: 405; CTL13_COMP
TTGTAATACGACTCAACAGTAGATAAAGCGACGCGCG	SEQ ID NO: 406; CTL30
TTGTAATACGACTCAAACGCGACTTATCGAACGATAA	SEQ ID NO: 407; CTL11
TTGTAATACGACTCAAGGTTGCCTACATAGTAACTTC	SEQ ID NO: 408; CTL23
TTGTAATACGACTCACCCGGTAACAATTAGTCCCGCG	SEQ ID NO: 409; CTL26
TTGTAATACGACTCATCGGTTGCGGAGTTCGACGCGA	SEQ ID NO: 212; CTL29
TTGTAATACGACTCACCCGATGGGTACGGCGTTCGGT	SEQ ID NO: 213; CTL31
TTGTAATACGACTCAGACACGAGAGTCGTTACCCTCG	SEQ ID NO: 214; CTL36
TTGTAATACGACTCAGTCCGTCTACGCGTCGGTACGC	SEQ ID NO: 215; CTL13
TTGTAATACGACTCAGGCCGTCTACGCGTCGGTACGC	SEQ ID NO: 216; CTL13_MUT
TTGTAATACGACTCACCCGGTAACAATTAGACCCGCG	SEQ ID NO: 217; CTL26_MUT

[00239] All sequences for which the intensity of the spot was higher than the mean local background intensity plus 1.5 times its standard deviation were categorized as expressed microRNAs. The following criteria were required to be met:

1. Specificity of the hybridization controls had to be within acceptance criteria (e.g. CTL26) vs. its corresponding single base mutant,

CTL26_MUT, or CTL13 vs. its corresponding single base mutant, CTL13_mut).

2. Approximate equality of the signal intensity of the replicates of the positive controls
3. Approximate equality between median block signal intensities based on the positive controls for each block
4. Approximate equality between median array signals based on all sequences detected
5. Signal intensity for the purification and labelling control (CTL30).

[00240] Statistical normalization of the data was done by computing the Log2ratio where the Log2ratio equals average intensity signal of the duplicated spots/median intensity of all positives controls for the block. The normalization was done per block to avoid non-homogenous labelling of all blocks of the array. This block-by-block normalization has been shown to be more efficient then using overall normalization of the slide. The obtained values are Log2 values.

[00241] The intensities of the spots for each polynucleotide probe were compared in the sample from the cervical cancer cell line versus normal cervical tissue, resulting in an evaluation of the relative expression for each microRNA.

[00242] The expression fold-change corresponds to $2(\text{Log2ratio})$. The Log2ratio is the ratio between the two conditions compared, or $\log_2(\text{Xcell-line}/\text{Xnormal})$, which is the same as $(\log_2\text{Xcell-line} - \log_2\text{Xnormal})$, where X is the measured intensity value. In cases where there was no signal from the “normal” condition, the lowest measured intensity value in the experiment was used as the baseline from which a fold-change expression value was calculated. A fold-change value of less than zero corresponds to a down-regulation of (1/fold-change) times.

[00243] Data are tabulated in Table 1, and include all microRNAs over-expressed in more than 50% of tested cell lines. Expression in HeLa was not used to qualify microRNAs for inclusion in Table 1, because of an observed odd expression and signal pattern.

5.2 Example 2: Analysis of microRNA on Luminex Platform

[00244] The Luminex technology (Luminex Corp., Austin, TX) is based on liquid phase hybridization to probe-labelled beads, followed by flow cytometry detection of beads with differing ratios of fluorescent dyes. Beads with up to 100 different dye ratios are available, making it possible to interrogate a single sample for up to 100 analytes simultaneously.

Coupling of Probes to Luminex Beads

[00245] Aliquots of each 5'-amino-modified probe having sequences as set forth in Example 1 and Table 1 are prepared at a concentration of 0.1nmol/ μ L in molecular biology grade water. The probes are coupled to the beads using carbodiimide chemistry according to the manufacturer's protocol (Luminex bead coupling protocol). The probe-coupled beads are stored at 4°C.

Total RNA Preparation for Luminex Analysis

[00246] Fifty fmoles of each of 7 internal controls (the same synthetic RNAs used for the array controls) are added to the total RNA fraction isolated from the biological samples. Prior to hybridization with Luminex beads, the total RNA preparation is treated to avoid the formation of dendrimers, which result from the circularization of a single RNA molecule, or concatenation to another RNA molecule. To avoid the formation of dendrimers, the RNA is pre-treated with calf intestinal phosphatase (CIP) to remove the 5'-phosphate groups. The CIP reagent can be obtained from Invitrogen (Carlsbad, CA) and the CIP reaction is run according to the manufacturer's protocol.

Bead Labelling and Hybridization

[00247] After CIP treatment, the total RNA fraction is then labelled with biotin using the Vantage microRNA Labelling Kit (Marligen). The labelled fraction is hybridized to the Luminex beads using the Marligen protocol. Briefly, the polynucleotide beads are mixed with the Marligen hybridization solution (1.5 X TMAC) and the labelled total RNA. The hybridization is performed at 60°C for an hour in the dark. After hybridization, the beads are washed using the Luminex standard 6X SSPET wash buffer (sodium phosphate, sodium chloride, EDTA, Triton X-100, pH 7.4).

Detection of Bead Hybridization

[00248] The detection of the Luminex beads is done using streptavidin phycoerythrin (SAPE) (Europa Bioproducts, Cambridge, UK). The SAPE is added to the washed beads according to the Luminex protocol. The beads are then read using the Luminex IS-200 instrument using the high gain setting for better resolution.

Data Acquisition and Analysis

[00249] The Luminex IS-200 reads at least 25 beads of each dye-ratio in the reaction mix. Each dye-ratio bead corresponds to a particular probe sequence, and the intensity value is returned as an average value of all read beads. The mean fluorescence intensity (MFI) data is normalized using synthetic RNA controls, and fold changes between normal and diseased samples are computed using the Bioplex software (Bio-Rad, Hercules, CA) and the R bioconductor package (Bioconductor: open software development for computational biology and bioinformatics, Genome Biol. 2004;5(10):R80. Epub 2004 Sep 15).

5.3 Example 3: MicroRNAs from Clinical Cervix Samples

Tissue samples

[00250] Archived formalin-fixed, paraffin-embedded (FFPE) blocks from cervical tumors were cut into 10 to 20 μm sections. Three to four sections per sample were extracted using RecoverAll™ Total Nucleic Acid Isolation Kit (Applied Biosystems, Inc.; Foster City, CA) according to the manufacturer's protocol. RNA samples were diluted in RNase-free water and stored in -80°C (-112°F).

[00251] Archived or freshly snap-frozen specimens from cervical tumors were also used. Tissue samples were homogenized by mortar and pestle in TRIzol® Reagent (Invitrogen; Carlsbad, CA) and RNA was extracted according to manufacturer's protocol. RNA samples were diluted in RNase-free water and stored in -80°C (-112°F).

[00252] The cervical samples used in this experiment are shown in Table 8:

Table 8: Clinical samples

<i>Sample name</i>	<i>Sample type</i>	<i>Description</i>
ASCC-1a	FFPE	AdenoSquamous Cervical Carcinoma (ASCC)

SCC-1a	FFPE	Squamous Cervical Carcinoma (SCC)
SCC-2	FFPE	SCC
ADC-1a	FFPE	AdenoCarcinoma (ADC)
SCC-3a	FFPE	SCC
SCC-1b	FFPE	SCC
SCC-3b	Frozen	SCC
ADC-1b	Frozen	ADC
SCC-5	Frozen	SCC
ASCC-1b	Frozen	ASCC
SCC-7	Frozen	SCC
SCC-8	Frozen	SCC
cx-normal-4	Frozen	
cx-normal-7	Frozen	
cx-normal-11	Frozen	

MicroRNA Preparation:

[00253] All samples were enriched for the microRNA fraction using a Flash PAGE Fractionator (Ambion). Briefly, a total RNA sample was loaded onto a pre-cast gel using the Flash PAGE Fractionator. The total RNA fraction smaller than 40 nt (the “microRNA fraction”) was recovered after gel migration and resuspended into nuclease free water.

Microarray Analysis

Probe design and spotting

[00254] The polynucleotide probes used for microarray preparation had the configuration 5'-NH₂-(C)₆-(spacer)-(oligomer probe sequence)-3'. The 5'-amino group

allowed chemical bonding onto the array support. Each also included an identical spacer sequence of 15 nt, as shown below, to prevent non-specific interactions of the polynucleotide probes with the array support:

5'AminoC6-TTGTAAATACGACTCA – Oligo probe sequence. (SEQ ID NO: 90)

Probe sequences given in Table 11 omit the linker.

[00255] The probes were synthesized according to standard protocols by Eurofins MWG Operon (Ebersberg, Germany). Nexterion (Schott) microarray glass slides were used as the solid support for the microarray.

[00256] The polynucleotide probe concentration used for the spotting was 25 µmol. The probes were spotted in duplicate using the Nexterion spotting buffer provided with the array glass support by Schott with 1 % SDS (sodium dodecyl sulfate) added to allow larger spot sizes (*e.g.*, 100-150 microns compared to 70-100 microns without SDS). The spotter used was the QArray mini (Genetix) equipped with Stealth SMP3 pins (Telechem). After deposition of one series of spots, the spotting needle was washed 5 times with 60mM NaOH before spotting the next series of probes. Each slide is designed with 48 blocks of spotted probes, with each block being a 20x18 square of spotted probes. Each probe was spotted in duplicate. Spotted glass slides were stored at 4°C until use.

MicroRNA labelling

[00257] The labelling of the microRNA fraction was adapted from a published protocol developed at EMBL (Heidelberg, Germany) by the European Molecular Biology Group (Castoldi *et al.*, “A sensitive array for microRNA expression profiling (miChip) based on locked nucleic acids (LNA),” *RNA* 2006 May;12(5):913-20. Epub 2006 Mar 15, incorporated herein by reference in its entirety). Briefly, the microRNA fraction was incubated for 6 hours at 4°C with a mixture containing 10 µM of dye-labelled tetra-nucleotide (5'-rUrUrUrU – Cy5- 3') (or alternatively, 5'-rUrUrUrU – Cy3-3') (Biospring, Germany) in Ambion buffer diluted to 1X with RNase free water, 8% polyethylene glycol (PEG), 2 mM adenosine triphosphate (ATP), and T4 RNA ligase (0.7U/µl). The labelling reaction was run by heating the mixture for 15 minutes at 65°C.

This procedure ligated the poly-U dye-labelled tail to the 3' end of all the microRNAs. Labelled samples were stored at 4°C before hybridization.

Array Hybridization

[00258] The labelled microRNA fraction was hybridized to the spotted arrays using a Discovery hybridization station (Ventana, Tucson, AZ). Briefly, 2 mL of a mixture of 1% BSA, 2X SSC, and 0.2 % SDS was incubated with the chips for 30 min at 42°C. Then the chips were washed once using EZ Prep buffer (Ventana) and then three more times with Ribowash (Ventana). Next, 20 µl of the labelled microRNA mixture and 180µl of ChipHybe Reagent (Ventana) were added to the array. The arrays were heated for 6 minutes at 37°C, then were incubated at 42°C for 8 hours, after which the heating was stopped. The chips were washed once with Ribowash (Ventana) and then heated for 2 minutes at 37°C. The chips were washed again with Ribowash (Ventana) with one drop of CheapClean (Ventana) added, and incubated for 2 minutes at 37°C. The chips were washed two more times using Ribowash (Ventana). On the following day, the final washes were done according to Ventana's instructions for the Discovery hybridization station. The slides were washed twice with 2X SSC + 0.2X SDS buffer and then one more time with 0.1X SSC. All the slides were dried using a speed centrifuge from Arrayit (TeleChem International, Sunnyvale, CA) at room temperature and kept in the dark before scanning.

Array Image Acquisition

[00259] The arrays were scanned using an Axon™ scanner (Molecular Devices, Sunnyvale, CA) and their Genepix™ software. The image was formatted in tif format, defined by an image color depth of 16 bits/pixel (1600*1600). At such setting, pixels can assume intensity values ranging from 0 to 65,535. Pixels exhibiting the maximum intensity value are "saturated" and were assigned the value of 65,535. The resolution of the array scan was set at 10 µm/pixel. For hybridization experiments using different fluorescent dyes (e.g., Cy5 and Cy3) the photomultiplier tube (PMT) was adjusted to the higher intensity spot (Cy3 is scanned at lower PMT settings than Cy5).

Array Image Analysis

[00260] The PMT of the laser scanner digitized the captured fluorescence intensity for each given “point” of a slide and stored the numerical value as a pixel corresponding to that point. A picture composed of such pixels was then analyzed.

The first task for image analysis was to detect the spot position, using a process called segmentation. Spots were segmented by circles of adaptable or fixed radius. To be reliably segmented and quantified, the spot diameter was required to be more than 5 - 6 pixels. Before segmentation an indexing grid was provided giving the approximate positions of the spots. The segmentation itself detected the limits of spots near the grid circles. Briefly, the Genepix software assigns a circle to each spot on the array (segmentation). The segmentation had to be conducted in a somewhat flexible way due to spotting imperfections and/or support deformation, as the spots were almost never on a perfectly rectangular grid.

[00261] After segmentation by the software, the circles were modified manually and adjusted onto the spots until all the spots on the array were clearly identified. At this stage, if the array presented high background noise preventing real spots from being distinguished from the background, the array was rejected for further analysis.

[00262] The second task of image analysis was to quantify spots and export the data into a result file. This was a relatively easy and well-defined task once the spots were located on the image. The statistical approach used most frequently to quantify spot intensity was the mean or median of pixels belonging to a spot. The median approach was more robust than the mean value in the presence of outlier pixels. In practice, however, there was little difference in the results obtained using mean or median.

Array Data Analysis

[00263] All the array data were analysed using the R bioconductor package (“Bioconductor: open software development for computational biology and bioinformatics,” *Genome Biol.* 2004;5(10):R80. Epub 2004 Sep 15, which is incorporated herein by reference in its entirety).

[00264] Array data were first tested for quality by comparing the spot intensities for the internal controls. One internal control (SEQ ID NO: 83; Table 9) was used as a labelling control (this synthetic RNA is added to the purified microRNA fraction before labelling), and 6 other internal controls (SEQ ID NOs: 84-89; Table 9) were used for the normalization of the data (these synthetic RNA controls are added to the total RNA fraction before hybridization at 520 fmol each/array). The probe sequences that bind to the synthetic RNAs, and a mutant probe sequence, are also shown in Table 9 (SEQ ID NOs: 406 to 409, 212 to 214, and 217).

Table 9
Control Sequences used in microarray experiments

Sequence (5'-3')	Sequence identification number
CGCGCGUCGCUUUUAUCUACUGU	SEQ ID NO: 83; CTL30_COMP
UUAUCGUUCGAUAAGUCGCGUU	SEQ ID NO: 84; CTL11_COMP
GAAGUUACUAUGUAGGCAACCU	SEQ ID NO: 85; CTL23_COMP
CGCGGGACUAAUUGUUACCGGG	SEQ ID NO: 86; CTL26_COMP
UCGCGUCGAACUCCGCAACCGA	SEQ ID NO: 87; CTL29_COMP
ACCGAACGCCGUACCCAUCGGG	SEQ ID NO: 88; CTL31_COMP
CGAGGGUAAACGACUCUCGUGUC	SEQ ID NO: 89; CTL36_COMP
TTGTAATACGACTCAACAGTAGATAAAGCGACGCGCG	SEQ ID NO: 406; CTL30
TTGTAATACGACTCAAACGCGACTTATCGAACGATAA	SEQ ID NO: 407; CTL11
TTGTAATACGACTCAAGGTTGCCTACATAGTAACTTC	SEQ ID NO: 408; CTL23
TTGTAATACGACTCACCCGGTAACAATTAGTCCCGCG	SEQ ID NO: 409; CTL26
TTGTAATACGACTCATCGGTTGCGGAGTTCGACGCGA	SEQ ID NO: 212; CTL29
TTGTAATACGACTCACCCGATGGGTACGGCGTTCGGT	SEQ ID NO: 213; CTL31
TTGTAATACGACTCAGACACGAGAGTCGTTACCCTCG	SEQ ID NO: 214; CTL36
TTGTAATACGACTCACCCGGTAACAATTAGACCCGCG	SEQ ID NO: 217; CTL26_MUT

[00265] All sequences for which the intensity of the spot was higher than the mean local background intensity plus 1.5 times its standard deviation were categorized as expressed microRNAs. The following criteria were required to be met in order consider the array intensity data valid for further analysis:

1. Specificity of the hybridization controls had to be within acceptance criteria (e.g. CTL26 vs. its corresponding single base mutant, CTL26_MUT).
2. Approximate equality of the signal intensity of the replicates of the positive controls
3. Approximate equality between median block signal intensities based on the positive controls for each block
4. Approximate equality between median array signals based on all sequences detected
5. Signal intensity for the purification and labelling control (CTL30).

[00266] Statistical normalization of the data was done by computing the Log2ratio where the Log2ratio equals average intensity signal of the duplicated spots/median intensity of all positives controls for the block. The normalization was done per block to avoid non-homogenous labelling of all blocks of the array. This block-by-block normalization has been shown to be more efficient then using overall normalization of the slide. The obtained values are Log2 values.

[00267] The intensities of the spots for each polynucleotide probe were compared in the sample from the cervical cancer cell line versus normal cervical tissue, resulting in an evaluation of the relative expression for each microRNA.

[00268] The expression fold-change corresponds to $2^{(\text{Log2ratio})}$. The Log2ratio is the ratio between the two conditions compared, or $\log_2(\text{Xcell-line}/\text{Xnormal})$, which is the same as $(\log_2\text{Xcell-line} - \log_2\text{Xnormal})$, where X is the measured intensity value. In cases where there was no signal from the “normal” condition, the lowest measured intensity value in the experiment was used as the baseline from which a fold-change expression value was calculated. A fold-change value of less than zero corresponds to a down-regulation of $(1/\text{fold-change})$ times.

Results

[00269] All of the samples generated low signals, possibly due to degradation of the samples. For three of the normal samples (normal-4, -7, and -11), a

reliable number of sequences were detected, so those samples were used as controls for the analysis.

[00270] In this experiment, miR-21 was up-regulated in all of the tumors tested compared to the normal samples. In certain tumor samples (for example, ADC-1a, ADC-1b, ASCC-1b, SCC-4a, and SCC-7), a particularly high or low number of sequences was detected. Two microRNAs, miR-145 and miR-143, which have previously been reported to be down-regulated in certain cancer tissues (Wang et al., *PLoS One* (2008) 3: e2557), were also found to be downregulated in this experiment. Table 10 shows a list of the microRNAs that were found to be upregulated in at least one of the tumor samples tested. Table 11 shows the probe sequences that were used to detect the microRNAs listed in Table 10. Table 12 shows the microRNA precursor sequences and their chromosomal location.

Table 10: Fold-change in expression relative to normal controls

Sequence	SCC1a	SCC1b	SCC2	SCC3a	SCC3b	SCC5	SCC7	SCC8	ADC1a	ADC1b	ASCC1a	ASCC1b
10030-R5-1	0.48	0.62996	0.42	0.68	0.45	5.13	1.09	0.83	0.48	6.55	0.34	7.08
10435-R4-1	0.55	0.62996	0.42	0.68	0.45	4.01	0.59	0.83	0.48	NA	0.37	4.24
12730-R5-2	0.30	0.56850	0.38	0.89	0.40	0.69	0.76	0.75	0.43	NA	0.30	3.09
12917-R5-1	0.29	0.54270	0.36	0.58	0.38	0.47	2.93	0.72	0.41	NA	0.29	0.51
12917-R5-2	0.29	0.53431	0.35	0.57	0.38	0.47	2.79	0.71	0.40	NA	0.29	0.50
13075-L5-1	0.34	0.62996	0.42	1.17	0.45	0.55	1.71	0.83	0.48	NA	0.34	0.94
13108-L5-2	1.23	1.34983	0.87	0.68	0.60	10.57	0.81	0.83	0.48	11.37	0.34	16.00
13111-L5-3	1.50	1.52386	0.79	1.69	0.99	1.30	3.85	0.85	0.58	NA	0.55	1.26
13122-L5-1	1.79	1.68168	0.69	0.68	0.54	2.80	1.68	0.83	0.48	4.43	0.34	3.47
13124-L5-2	0.48	0.43358	0.40	0.82	0.23	0.62	1.22	0.43	0.25	NA	0.18	6.48
13129-L5-3	0.80	0.72991	0.77	0.82	0.40	1.14	4.07	0.25	0.34	NA	0.10	42.27
13168-L5-1	0.92	2.67482	1.46	1.65	1.02	2.21	0.59	0.83	0.48	NA	6.61	0.59
13181-L5-2	0.61	0.50095	0.33	0.54	0.35	1.04	1.13	0.66	0.38	NA	0.27	48.83
13195-L5-1	0.40	1.75224	0.42	0.68	0.45	0.55	0.59	0.83	0.48	5.52	0.52	0.59
13207-R5-4	1.50	1.36744	1.01	0.68	0.45	2.13	2.09	0.83	0.48	NA	0.34	2.06
13209-L5-2	0.47	0.39869	0.38	0.26	0.17	1.77	1.59	1.08	0.26	NA	0.41	3.61
13219-L5-1	1.01	0.85828	0.64	0.68	0.45	1.79	1.42	0.83	0.48	NA	0.34	2.00
13227-L5-2	0.78	0.73130	0.47	0.78	0.47	1.30	3.21	1.02	0.31	NA	0.22	2.28
13229-R5-3	0.68	0.81104	0.57	2.52	1.40	1.07	0.70	1.14	0.31	NA	0.69	1.55
13231-L5-3	0.19	0.35780	0.24	0.38	0.25	0.31	1.08	0.47	0.27	NA	0.19	5.11
13247-L5-3	0.48	0.51572	0.20	0.92	0.62	0.83	5.13	1.41	0.27	NA	0.41	2.28
13252-L5-3	1.77	2.03164	0.42	0.68	0.45	3.37	0.59	0.83	0.48	17.08	0.34	75.31
13254-R5-1	3.95	3.51150	0.94	0.55	0.69	16.00	0.48	1.86	0.39	26.31	0.28	9.67
13260-L5-2	0.31	0.49907	0.33	0.53	0.35	0.43	0.47	0.66	0.38	NA	0.27	0.47
13267-L5-1	0.66	0.51565	0.56	1.62	0.55	0.83	2.21	1.01	0.19	NA	0.14	31.04
13274-L5-3	0.90	1.05855	0.54	0.67	0.55	1.71	3.59	0.93	0.33	NA	0.38	2.23
13283-L5-3	0.75	0.84669	0.39	0.58	0.38	5.35	1.84	0.72	0.41	6.57	0.29	9.44
13291-L5-1	0.49	0.58218	0.31	0.50	0.33	0.64	2.38	0.83	0.35	NA	0.25	1.82
13296-L5-3	0.60	0.50124	0.33	0.54	0.35	1.04	1.14	0.66	0.38	NA	0.27	11.21
13312-L5-1	0.29	0.24998	0.26	0.92	0.30	0.36	1.29	0.23	0.13	NA	1.25	4.81

Sequence	SCC1a	SCC1b	SCC2	SCC3a	SCC3b	SCC5	SCC7	SCC8	ADC1a	ADC1b	ASCC1a	ASCC1b
13325-R5-2	0.56	0.70384	0.63	1.81	0.99	1.22	5.31	0.67	0.38	NA	0.56	0.47
13335-L5-2	0.49	0.43467	0.34	0.32	0.21	0.97	1.51	0.39	0.22	NA	0.16	3.91
13335-L5-3	0.43	0.39498	0.30	0.59	0.44	1.00	2.18	0.89	0.17	NA	0.24	2.14
13339-L5-1	0.68	0.62012	0.43	0.46	0.32	1.17	2.49	1.05	0.26	NA	0.28	3.69
13504-R5-3	1.55	1.49604	1.52	0.68	0.45	1.57	0.83	0.83	0.48	NA	0.65	0.59
13532-L5-2	1.98	2.37574	0.73	0.64	0.42	5.40	0.56	0.88	0.46	12.41	0.32	2.08
227-L5-1	0.64	0.74093	0.56	0.81	0.51	0.96	2.97	0.95	0.32	NA	0.54	1.31
25-R5-2	0.90	1.05584	0.29	0.47	0.31	1.62	0.78	0.58	0.33	10.98	0.23	1.40
2819-L5-2	0.34	0.62996	7.35	0.68	0.45	0.55	0.59	0.83	0.48	NA	0.34	0.59
2819-R5-4	0.47	0.38389	0.32	0.57	0.40	0.67	1.32	0.64	0.23	0.87	0.32	1.86
3371-L4-1	0.21	0.38626	0.25	1.04	0.45	0.52	1.00	0.51	0.29	NA	0.21	0.91
3744-R5-1	1.33	1.09513	0.78	0.67	0.54	1.93	2.78	1.04	0.25	NA	0.18	3.05
3911-R5-1	0.34	0.62996	0.42	2.28	0.90	0.55	0.59	0.83	0.48	NA	0.34	1.79
4417-R5-2	1.29	2.32318	0.96	0.68	0.45	3.07	0.59	0.83	0.48	NA	1.14	1.56
4440-L3-2	4.04	7.35978	3.34	8.03	6.55	3.01	2.06	0.53	1.66	3.29	2.85	1.51
4440-R3-2	0.50	0.50446	0.38	0.67	0.39	3.32	1.03	0.36	0.16	2.82	0.28	4.80
4498-L3-2	0.84	0.78263	0.50	0.68	0.45	5.48	0.91	0.83	0.48	5.05	0.46	8.39
4719-R5-1	2.00	2.68058	0.84	0.68	0.45	3.15	0.59	0.83	0.48	8.15	0.34	0.59
4765-L5-1	0.54	0.62996	0.42	3.01	1.55	1.22	0.59	0.83	0.48	NA	0.34	2.20
4829-R2-1	0.53	0.61657	0.43	1.12	0.62	1.73	5.34	0.87	0.25	NA	0.56	2.07
4855-R5-1	1.84	2.01422	1.31	1.88	1.00	1.82	2.61	0.33	0.65	3.60	3.08	0.24
4988-R5-2	0.57	0.58473	0.57	0.85	0.46	0.32	1.47	0.48	0.28	NA	2.25	0.34
6216-L1-1	3.25	5.58042	2.97	8.16	5.47	2.04	2.31	0.67	1.39	NA	2.38	1.62
6216-R5-2	0.57	0.55598	0.45	0.80	0.46	4.04	1.04	0.33	0.20	3.13	0.30	5.54
6235-R5-2	4.37	4.36450	1.88	0.68	1.20	6.64	2.24	0.83	0.48	7.33	0.34	8.03
6803-R5-2	0.57	0.53791	0.39	2.83	1.52	1.31	1.27	0.57	0.33	NA	0.33	3.89
7067-L5-1	0.44	0.62996	0.42	0.68	0.45	2.73	0.59	0.83	0.48	NA	0.34	4.27
7126-L3-1	0.56	0.49062	0.37	0.74	0.50	0.64	2.54	1.08	0.23	NA	0.39	1.37
7182-L5-1	0.48	0.50155	0.43	0.82	0.51	0.78	4.34	0.82	0.18	NA	0.32	1.29
7292-R3-2	0.37	0.63424	0.35	0.57	0.38	1.26	0.50	0.71	0.41	18.66	0.29	3.23
7578-L3-1	4.60	0.54709	1.08	0.59	0.85	0.48	0.51	1.78	0.85	39.02	0.57	7.03
7781-R5-2	0.43	0.51223	0.34	0.55	0.36	0.84	1.58	0.68	0.39	NA	0.27	2.40

Sequence	SCC1a	SCC1b	SCC2	SCC3a	SCC3b	SCC5	SCC7	SCC8	ADC1a	ADC1b	ASCC1a	ASCC1b
7887-L5-3	0.47	0.48451	0.32	0.52	0.34	1.27	0.45	0.64	0.37	NA	0.26	3.37
8004-R3-2	1.02	0.91481	0.42	0.68	0.45	1.94	1.27	0.83	0.48	NA	0.34	1.95
8298-R5-1	0.38	0.38316	0.34	0.50	0.34	0.63	2.32	1.05	0.29	NA	0.30	1.85
8339-R5-1	0.70	1.10768	0.91	0.68	0.45	4.60	0.59	0.83	0.48	NA	0.57	3.15
836-R4-1	1.33	1.43068	0.42	0.68	0.45	1.85	1.53	0.83	0.48	NA	0.34	0.59
9053-R3-1	0.68	0.80482	0.64	0.87	0.48	3.04	1.01	0.89	0.33	NA	0.23	4.22
9164-R5-1	0.34	0.62996	0.42	2.42	1.14	1.06	0.72	0.83	0.48	NA	0.34	1.09
9485-R5-1	2.15	2.24932	2.77	0.68	0.82	1.32	0.59	0.83	1.14	NA	0.34	0.59
9691-L4-1	0.58	0.70634	0.38	0.62	0.41	0.98	2.22	0.76	0.44	NA	0.31	1.55
9816-R5-1	0.46	0.58377	0.39	0.63	0.41	0.51	1.55	0.82	0.44	NA	0.31	1.76
99996-L4-1	0.25	0.46163	0.30	0.49	0.33	0.40	0.43	0.61	0.35	5.53	3.56	0.43
miR-1246	3.51	4.24691	1.92	0.68	1.16	5.40	2.72	0.96	0.55	5.82	0.55	7.41
miR-1290	1.46	1.29141	0.77	0.68	0.45	1.94	1.74	0.83	0.48	NA	0.34	1.84
miR-1308	5.41	5.16798	1.26	0.50	0.97	18.09	0.50	2.59	1.12	39.12	0.79	15.99
miR-142-3p	0.79	0.68772	1.20	0.68	1.44	4.46	0.59	1.37	0.48	NA	0.55	0.97
miR-143	0.21	0.39339	0.26	0.42	0.28	0.42	0.37	0.52	0.30	5.20	0.60	0.37
miR-145	0.31	0.40117	0.43	0.43	0.28	0.48	0.37	0.53	0.65	5.99	0.81	0.60
miR-1826	2.93	4.04445	2.24	0.57	1.23	4.97	1.14	1.44	1.02	11.52	1.00	5.45
miR-195	0.37	0.41556	0.27	0.45	0.55	1.21	0.39	0.55	0.31	5.64	0.53	1.19
miR-200c	1.00	1.22187	0.95	0.68	0.99	1.34	0.59	0.83	0.48	NA	1.11	2.05
miR-205	6.23	5.69362	2.07	0.61	1.47	4.24	0.53	1.13	0.43	NA	1.45	2.54
miR-21	3.01	3.70795	10.92	0.56	7.46	24.36	1.01	7.32	3.70	21.96	11.87	24.14
miR-31	0.34	0.62996	0.42	0.68	0.45	0.55	0.59	0.83	0.48	NA	0.58	0.92
miR-451	0.65	0.52572	0.33	0.47	0.65	0.38	5.85	0.58	0.33	NA	0.70	0.41
miR-483-5p	0.54	0.37151	0.25	1.24	0.26	0.32	1.72	0.49	0.28	NA	0.20	0.35
miR-491-3p	0.50	0.62996	0.42	0.68	0.81	12.59	0.59	0.83	0.48	58.76	0.34	18.82
miR-494	0.63	0.81477	0.54	0.68	0.45	1.56	0.59	0.83	0.48	NA	0.49	2.70
miR-720	1.15	1.54266	1.73	1.02	0.92	1.78	0.59	0.83	0.85	NA	1.27	1.45
miR-765	0.28	0.42837	0.28	0.84	0.46	0.51	1.56	0.55	0.32	NA	0.28	0.84
miR-98	1.11	0.57757	0.61	0.41	1.01	1.00	0.36	0.51	0.59	7.06	1.04	2.14

Table 11: Probe sequences

probe	probe sequence 5' -> 3'	SEQ ID
25-R5-2	TTCTGCTTTCCCAGAGCCTCACCCCTCTTTT	133
227-L5-1	ACACCTGTCTCTCCCCAGTGCTTCCGCCCCCTCA	134
836-R4-1	AAATAATCATTCCAAATGGTTCTCCCTGCTATGATTCAC	32
2819-L5-2	CCACACTTCTAATTGGACAAAAGTGCCTTTCAAAC	136
2819-R5-4	CAGCCTGCCACCGCCGCTTTTGAAGAAGCACTTCA	137
3371-L4-1	TTTCCTTTCCCTCCCCCTCCACACCCCATGACTCCCCACACTTGAG	1
3744-R5-1	CTTCTCCTTCCCTCCCTGCTCCCTCCCACTAATGCCAAAT	138
3911-R5-1	GGCTCCCTAGTGAAAAAATGCAAAATTTGTATAAT	139
4417-R5-2	ACTCGGCGCTCATCAAAAAGTTCCTGTCCG	141
4440-L3-2	TTTGACATTGAGAGCACTGGGCAGAAATCACA	142
4440-R3-2	GTCATAGTTACTCCCGCCGTTTACCCGCATTTC	143
4498-L3-2	GAGATCCAGACGGCCGTGCGCCTGCTGCTGCCT	144
4719-R5-1	ACAGCATCACATGGATTCTGTGTCCAGTGGCCTTAGCA	145
4765-L5-1	ACATGCTCCTGACACTTTCTCTTAGTTTCTCGGGCTCC	146
4829-R2-1	TCCCTTTGTGCTGCCCCAGTGCCTTCCCCCTG	147
4855-R5-1	CGGGTCTCCCGCTTCCCCCTCCTGCTCCAAGG	148
4988-R5-2	CTCCTCCTCCCGCTCTTTGGATACCAAACAC	149
6216-L1-1	GACATTCAGAGCACTGGGCAGAAATCACATG	151
6216-R5-2	CATAGTTACTCCCGCCGTTTACCCGTGCTTC	152
6235-R5-2	TCTGCTCCAAAAATCCATTTAATATATTGT	153
6803-R5-2	GCTCCCTCTCTGGTTGGACCTCACCCAAA	154
7067-L5-1	GGAGATCCAGACGGCCGAGCGCCTGCTGCTGCC	155
7126-L3-1	GCACACCCGCTCTCCGGCCCCGCGCCCCTG	156
7182-L5-1	AACTAGCCGTTTCCGTACCTTCCCCTGCCCCC	157
7292-R3-2	ACAATATTTATCCAGGGATGGGAGTCAGATGCA	158
7578-L3-1	CGCAGTGCACACCCTGAGCTACAGCCCCTC	159
7781-R5-2	AGCCTGTGCCTGCCGCTGTCTAGTACTGGT	160
7887-L5-3	CAAGAGCCAGCCTGCACTACCAGTCCCATGCCA	161
8004-R3-2	GGAAGTGTCTCTCCTTGCTCCAGTCATTGAAG	162
8298-R5-1	GATGCTGGCGTCCGCCGAGCCTCTCGCCCCATCCCGG	163
8339-R5-1	AAAAGCCAATACATTTTCACTGTACCGGCCAC	164
9053-R3-1	TTCTTGCCCTCCAATCCCCGGGCTCCACCAGCC	5
9164-R5-1	TGCTTCCATCCCGCCAGTTTGGTTTCATTGTACTGACAACC	166
9485-R5-1	CTGGGTGAGGTCCCACCGTGGTGCGCTTGGCTGTGCCAGC	167
9691-L4-1	AATCATCCATTTTCATCCGCATCTCCCTCTTGGCCCCTTGC	7
9816-R5-1	CCCTTTAAGAGCCTCTCCGCGCGCTGCCG	169
10030-R5-1	CCGTGGATGTCAACTCAGCTGCCTTCCGCC	170
10435-R4-1	GCATGCTAATTGTGCCCTGTTGTCTTTCTTAACT	171
999996-L4-1	GGGAGGAGTCAGGTGTGTGCTGTGGGTGGGGGAAGAC	173
12730-R5-2	GCGCCCTGTGTTGTGCTCCGCTCTCCGGGAAATGC	174
12917-R5-1	GGGCCCTTCCCTTCCCCCAACATTGAGCCTTG	175
12917-R5-2	GGACCTATGGGCCCTTCCCTTCCCCCAACATTG	176
13075-L5-1	TGAAAGCTGAAGTCCAGCCCAGCCCTCT	177
13108-L5-2	CTGCTGCCTTCCCTTGGTTGAGGGGCTGAGCACG	178
13111-L5-3	TCTCCGCCGGGCCTTACCCTGCCCTGCTCTTCT	179
13122-L5-1	TTAGGAAATTCATCTCACCTGCTCCAGTCC	180
13124-L5-2	GCTCCATGTCTCCTCCCCTCCGCGAAAGCCTAAAC	181
13129-L5-3	AGCCTTCCCTGTCCCCTGGCCCCGACCTGCTCCA	182
13168-L5-1	CGCTTCCTTAACCATTTTTTTTTTTTTTAACCAC	183

13181-L5-2	TGGACGTCTGAACAGTCACTGCCTGCCCCAACCT	184
13195-L5-1	ATGACCATTGTATTAGTATCTTTTTTTTTTTTT	185
13207-R5-4	CTGCGGCAAGTGCTTCTACATCCCTGCTCCAACAA	186
13209-L5-2	TAACTCGCCTGCTGCCCCGGCGGCTGCCCCGCG	187
13219-L5-1	CTCTGACTCCCTCACTCAGTCTCTCTGCTCCAGC	188
13227-L5-2	GGGCCCAGTCCTCCTCGTCCCCCTTCCCACCTCGG	189
13229-R5-3	GCAGCTCCGCCAGTCTCTGTGGGCAGGGAGAAG	190
13231-L5-3	GGCCCACCCGGGGGCGGCTCCCCAGCACCGACGCC	191
13247-L5-3	TCCTGAGCCGCCTTCCCCCTCCCGACCTCAGAGCCCT	192
13252-L5-3	ACGTGCCTTCCTGACTGTGAGCTCCTTGAGAGC	193
13254-R5-1	CAATGAACCACTGAACCACTCATGCACTGAACC	194
13260-L5-2	CTGTAGACCCACACTCAGTCTCTATAGCTA	195
13267-L5-1	CACTCCCTGCTGGCCCCACCTCACCTATGGTG	196
13274-L5-3	CCTTCTCTTCTCCCGTGCTCCACCCCTCCCTCAGGG	197
13283-L5-3	GGACCCCTGCCTTCCTTGCTGCCACCCCTTGACACA	198
13291-L5-1	CCCAAGCGCCCTTCCCTCCCTCCTTCCCTCCCG	199
13296-L5-3	CAGTCACCTCAGATTCCCTGTGCCCTCTGCCCTGG	200
13312-L5-1	CCACCCCTCCCCACAGCCCAGCCCCACTCAC	201
13325-R5-2	TCCAACACTGCCTGGCGCTGGGCTCTTCCCCA	134
13335-L5-2	CCACTGCCCTCCTGCCGCATCCTATGCTCCTCT	140
13335-L5-3	ACCTCAGCCTCCACTGCCCTCCTGCCGCATCCTAT	168
13339-L5-1	GACTGAGGGTTTAAAGAAGATGGTGTCCGCCGC	150
13504-R5-3	AGACTGCTGTAAATGCGGACAAAGCGTCCCTGC	165
13532-L5-2	TGCTCTACCGGCTATGACATTAGGTGTGACCG	172
miR-1246	CCTGCTCCAAAAATCCATT	208
miR-1290	TCCCTGATCCAAAAATCCA	209
miR-1308	CCACTGAACCACCCATGC	210
miR-1826	ATTGCGTTTGAAGTGTGATGATCAAT	211
miR-200c	TCCATCATTACCCGGCAGTATTA	203
miR-451	AACTCAGTAATGGTAACGGTTT	204
miR-483-5p	CTCCCTTCTTTCCCTCCCGTCTT	202
miR-491-3p	GTAGAAGGGAATCTTGATAAG	205
miR-494	GAGGTTTCCCGTGTATGTTTCA	206
miR-720	TGGAGGCCCCAGCGAGA	207
miR-765	CATCACCTTCCTTCTCCTCCA	39
miR-143	GAGCTACAGTGCTTCATCTCA	218
miR-145	AGGGATTCCCTGGGAAAACCTGGAC	219
miR-205	CAGACTCCGGTGGAATGAAGGA	220
miR-21	TCAACATCAGTCTGATAAGCTA	221
miR-31	AGCTATGCCAGCATCTTGCCCT	222
miR-142-3p	TCCATAAAGTAGGAAACACTACA	223
miR-195	GCCAATATTTCTGTGCTGCTA	224
miR-98	AACAATACAACCTTACTACCTCA	225

Table 12: microRNA precursor sequences and chromosomal locations

probe	microRNA precursor sequence 5' -> 3'	SEQ ID	chr	start	end	strand
25-R5-2	TCCCGCAGCCGGTGACTGGAGCCCACTCTGCAGAGACAAAAGGTTAGAAAAAG AGGGGTGAGGCTCTGGGAAAGCAGAAATGCGGGG	226	2	176709550	176709636	-1
227-L5-1	TGAGGGCGGAAAGCACTGGGAGAGACAGGTGTGAGCTTCCCACGTTGGTGATC AGCTCACACCTGTCTTGTGTTCTTGGTATTCACAGACTCTCA	227	3	187350863	187350957	1
836-R4-1	AAATAAGCCATTCCAAACCACTTCTCTGATTGCTGTGAGTGGCAGAATCATTC ACCGTGGTGAATCATAGCAGGGAGAACCAATTTGGAATGATTATTT AATGCCAGTGAGTTTGAAAGGCACCTTTGTCCAATTAGAAAGTGTGGAGAAATAT TCATCCTGTCCATGACAAAAGATGAAGTGCTTCTTTCAAAAGCGCGGTGGCAG GCTG	73	3	170758581	170758678	-1
2819-L5-2	AATGCCAGTGAGTTTGAAGGCACCTTTGTCCAATTAGAAAGTGTGGAGAAATAT TCATCCTGTCCATGACAAAAGATGAAGTGCTTCTTTCAAAAGCGCGGTGGCAG GCTG	228	15	59266509	59266618	1
2819-R5-4	CTCAAGTGTGGGAGTCATGGGGTGTGGAGGGGAGGAAAGGAAAGGATATTTTG TTTCTTTTGTCTATACATTTCTAGATTTCTATGAGTTGGG	229	15	59266509	59266618	1
3371-L4-1	CTTCTCTTATTCTCCCTGTTTTCATCTACTTTTAAAGTAATAAATTTGGCATT AGTGGAGGGGAGCAGGAGGAGGAGAAAG	42	18	58394821	58394914	1
3744-R5-1	GGCCCTTAGGAAATTAGAGTGTGTTTGAATTTCAAAAGTATAATTTTAATTAT ACAAATTTTGCAATTTTTCACTAGGGAGCC	230	19	14176021	14176103	1
3911-R5-1	TGGCCTGACGTGAGGAGGAGGACTTTTCGAAAGTTTATAGGAAAGTTTCCGC TTTCCAGTCCCCCTCCCCGTCCCA	231	17	52705568	52705650	1
3995-L2-1	GCTGGGGTTTCATCGAGAAACTCCCTGCGATGAGCCACTAGGGTCACGGACAG GGAACTTTTGATGAGCGCCGAGT	232	7	19123856	19123933	1
4417-R5-2	GTGATGTGATTTCTGCCAGTGCTCTGAATGTCAAACTGAAGAAATTCAGTGA AATGCGGGTAAACGGCGGAGTAACATATGAC	233	14	34943859	34943935	1
4440-L3-2	GTGATGTGATTTCTGCCAGTGCTCTGAATGTCAAACTGAAGAAATTCAGTGA AATGCGGGTAAACGGCGGAGTAACATATGAC	234	7	68165348	68165431	1
4440-R3-2	TTCCCCCAGGCAGCAGCGGCACGGCCGTCTGGATCTCCCTGGAGGTGATGG TCGAGCGCTTGTCAATAATGCGCCAGGCGGGA	235	7	68165348	68165431	1
4498-L3-2	ACAGCGGCATGGTTTCATGCCAAATTCGGAAGCAATCTTCTGTCTAAGGCCACT GGACACAGAAATCCATGTGATGCTGT	236	6	25840053	25840136	1
4719-R5-1	GGAGCCCGAGAAACTAAGAGAAAGTGTCAAGAGCATGTTAATCAGACTCGTTTA CACTGTAAACAATAACGTCCTCTCTCGGGTCTCC	237	13	71318356	71318433	1
4765-L5-1	GGTGTGCTGCCTCTCTTTCTGCCCCCTTATACCCCTTGACCCCGAGGGGAAG GCACTCGGGCAGCACAAAGGAGCAGATGCC	238	13	99346548	99346632	1
4829-R2-1		239	1	149949355	149949439	1

8433-L3-1	CGGTGGAGGAAAGGGAAAGGAGCCATTTTCTGCTGCACATCAGTCAGTGCC	53	17	75427043	75427117	1
	TGCGCCCTCCCTCCCTCCGCCG					
7887-L5-3	ACAGTAGGTATAGCTGGCATGGACTGGTAGTGCAGGCTGGCTCTTGGAAGG	257	3	29199527	29199609	-1
	AGTATGTATCCAGGCTGGTTGGCTGCTGT					
9053-R3-1	GGAAGGGCACTGTCTCTGTGATTCCAGGGCTGTCAATTTCCCGAGGGCTGGT	46	X	144618949	144619035	1
	GGAGCCCGGGGATTGGAGGGCAAGAAGCCCGCC					
9164-R5-1	TGTTTTCATCTTGCTTTCTTTCATGGTCCATGATGCCAGCTGAGGTTGTTCAGTAC	258	1	218383957	218384038	1
	AATGAAACCAAACTGGCGGGATGGAAGCA					
9485-R5-1	CTGGGAACAATGGGGCCATTGTGGGAGGATGGAGTGCAGCAGACTGCTGGCAC	259	11	118290213	118290297	1
	AGCCAAGCGCACACCGTGGGACCTCACCCAG					
9691-L4-1	GCAAGGGGCAAGAGGGAGATGCGGATGAAATGGATGATTTAATGGGTCACTCT	48	14	77897549	77897632	1
	CTCCTGTAGTTAATTTCTCTAGATCTCTTGT					
9774-R2-1	GCTTGTCTTAAAGATCTTCTTCTGTTCCTTGGTTTATCCACTTGGTTGG	260	13	35312134	35312220	1
	CCTGATGGGAGCAGGCGGTGAGGGGCGGGC					
9816-R5-1	CTGGCCCATTTTCATTCGTGATAAAATTTTAATGGTCTCTCTGGCTGATCCGG	261	17	35028328	35028415	1
	GACGGCAGCGCGGAGAGGCTCTTAAAGGGCCAG					
10030-R5-1	GGATGCAACCGTGGAAAGCCGGTGCCGTTGAGGATCTGCCACAGCGGGAAGGCA	262	10	98752662	98752739	1
	GCTGAGTTGACATCCACGGGCATCC					
10435-R4-1	GAGGCTGCTTAATGAGGTGCCCTTTTCAAAATGTCACTTAAATCTTTTATTAG	263	5	168043163	168043253	-1
	TTTAAAGAAAGACAAACAGGGCACAAATAGCATGCAACTC					
4315_D-R4-1	GGGACGTGGCCCTCCCCCGGAGCGGGACTCCAAGAACTCCGGGGGGCGC	264	1	153319527	153319595	1
	TGGGGGTGACTTTCC					
99996-L4-1	GTCTTCCCCCAACCCACAGCACACACCTGACTCCTCCTTCCAGGGAAGAC	265	17	35759252	35759338	-1
	CTCAGGGCTGTGGTGAGTCAGATAAGAAAGAC					
	CCCGGCTCGGCCCGGTCTCTCCAGCTCCTCCGGCTCCTTTTAGTGATATAA					
12730-R5-2	TTAGTGATGGCATTTCCCGAGAGCGGAGCACAAACACAGGGCGCCGGCTCGG	266	17	75427123	75427229	1
	G					
12917-R5-1	GGACCTGGGGGCTTCTCTGACCCCTTGAACAGCTTATACTATGAGACCTTGGGA	267	1	45246668	45246780	1
	ACCTCCTCCATGCAGACACACAAGGCTCAATGTTGGGGAAGGGAAGGCCCA					
	TAGGTCC					
12917-R5-2	GGACCTGGGGGCTTCTCTGACCCCTTGAACAGCTTATACTATGAGACCTTGGGA	268	1	45246668	45246780	1
	ACCTCCTCCATGCAGACACACAAGGCTCAATGTTGGGGAAGGGAAGGCCCA					
	TAGGTCC					
13075-L5-1	AGAGGGCTGGGCTGGACTTCAGCTTTCACTAGGAAATGAGTCTTGCTGCCCT	269	2	42137734	42137788	1
	TT					
13108-L5-2	TTCCACACGTGCTCAGGCCCTCAACCAAGGAAGCAGCAGGCCCTTGCC	270	2	31479610	31479698	1
	TCCTTATTACAGAGGGCTGACTGACCCCTAGGGAG					
13111-L5-3	AGCCTGTGGGAAGAGAAGAGAGAGGAGGGTGAAGGCCCGCGGAGACACTC	271	16	3475382	3475470	-1
	TGCCCCACCCACACCCCTGCCCTATGGGCCACACAGCT					

13122-L5-1	272	2	85447047	85447127	1
13124-L5-2	273	1	154700544	154700634	-1
13129-L5-3	274	20	61388629	61388733	1
13168-L5-1	275	11	78788919	78789022	-1
13181-L5-2	276	1	98283397	98283491	-1
13195-L5-1	277	3	54069253	54069352	1
13207-R5-4	278	10	677614	677734	-1
13209-L5-2	279	10	74122118	74122238	1
13219-L5-1	280	11	100895740	100895860	-1
13227-L5-2	281	11	133373044	133373159	1
13229-R5-3	282	11	199324	199422	1
13231-L5-3	283	11	34919943	34920052	1
13247-L5-3	284	1	165748638	165748746	1
13252-L5-3	285	1	176734861	176734981	-1
13254-R5-1	286	1	181542244	181542361	-1

13260-L5-2	GTGCTATAGCTATAGAGACTGAGTGTGGGTCTACAGAAAAATGTGGCCATGCC CTCCACCCAGTGGCTGGCAGCCTTTGGCACAG	287	1	210949402	210949488	-1
13267-L5-1	CACCATAGGTGAGGTGGGGCCAGCAGGAGTGGGCTGGGCTGGGCTGGGCCA AGGTACAAAGCCTCACCTCACCCTGCCACCCAGGCTTCAACGTGG	288	1	226351579	226351677	1
13274-L5-3	AGGTGGTGGTGGGAGGACCCCTGAGGAGGGTGGGAGCACGGGAGAGAAGAAG GCATACCCAACTGACCTACTTACCTGTCCCCCTACCCACACAGAGGGCTTCCCT GGAGGCCGCCATTGC	289	12	51578925	51579045	-1
13283-L5-3	GGGACGGGGTTGGGTGTCAAAAGGTGGCAGCAAGGAAGGAGGGGTCCCTA AGGTGTCTCCTGCCCTCCTTGTGTAGACTTTGGCCTGAGCAAAAGAGGCC CGGAGGGAAGGAGGAGGAGGGCGCTTGGCAGAAACCAAGGTTGCCAGAT TATCCTAGGGACTCTTGGGSCAGAAACAGACGCCCTCTGCGTCTCCCTCTCC	290	1	26753591	26753696	1
13291-L5-1	CC	291	1	36545595	36545702	1
13296-L5-3	CAGGAATTCCACTGGCAGCCAGGGCAGAGGGCACAGGAATCTGAGGTGACTGG CACAGAAAGACTCAGGCCCTGTGGCTCCTCCCTCAGGACTGCTTCCTA	292	1	43686783	43686881	1
13312-L5-1	GTGAGTGGGGTGGGCTGTGGGGAGGGTGGGCTGGCAGGGAACAGGCAGAC CATCCCTTCTACCCACAGGATCCTGCTGCTGCAGACAG	293	15	72689606	72689696	-1
13325-R5-2	ACTCAGGCACTGCCCTCTGACGATGCTTCCAGATCTGGTAGCTCATGGGGA AGAGCCACAGCCAGGCAGTGTGGGA	294	16	88088866	88088944	1
13335-L5-2	TGGCTGGGAGAGGAGATAGATGGGAGGAGGAGGAGCTGAGGCTGAGGTAC GGATTTCTAGCCCCGCCCTACCTCCTCTGCCCCCTAGTGCCCCGTGGCCAA	295	17	4803649	4803753	-1
13335-L5-3	TGGCTGGGAGAGGAGATAGATGGGAGGAGGAGGAGCTGAGGCTGAGGTAC GGATTTCTAGCCCCGCCCTACCTCCTCTGCCCCCTAGTGCCCCGTGGCCAA CGGCGGACACCATCTTCTTTAAACCCCTCAGTCCGTATTGGTCTCTATGGCAT CCATAGAGGCCATTCTGGCTCTGAGGTCTCTCAGTAAAGAACTTAGATGGTATT ACTGTGT	296	17	4803649	4803753	-1
13339-L5-1	TCACATGTCCTCAGCTGTTGTCTGGGTGAGGCATCCCTGTCTGGGAGCAGCC ACAGCTCTGCCCTGGTCTCCAGAGCAGGAGCGCTTTGTCCGCATTTACAGCAG TCTACACAGATG	297	17	7150842	7150954	-1
13504-R5-3	AGCCACACGGTCAACACCTAATGTATAGCCGGTAGAGCAGGAGCCCTCTGGA TGGAAGCACTGTGAGGCTC	298	7	150369385	150369502	1
13532-L5-2	TGTATCCTTGAATGGATTTTGGAGCAGGAGTGGACACCTGACCCAAAGGAA TCAATCCATAGGCTAGCAAT	299	9	131989877	131989948	-1
miR-1246	GAGCGTCACGTTGACACTCAAAAAGTTTTCAGATTTTGGAAACATTTTCGGATTTT GGATTTTGGATCAGGGATGCTCAA	300	2	177173954	177174026	-1
miR-1290	CCCCGCATGGGTGGTTTCAGTGGCAGAAATCTCAAATTTGTAATCCCCATAATCC C	301	1	19096152	19096229	-1
miR-1308	GACAGTGCAGTCACCCATAAAGTAGAAAGCACTACTAACAGCACTGGAGGGTG TAGTGTTCCTACTTTATGGATGAGTGTACTGTG	302	X	21990180	21990233	-1
miR-142-3p		303	17	53763592	53763678	-1

miR-1826	ATTGATCATCGACACTTCGAACGCAATTGCAGCCCGGTTCTCCAGGGCTT TGCTGTCTGAGCGTCGCTTGCCGATCAGTAG	304	16	33873009	33873093	1
miR-195	AGCTTCCCTGGCTCTAGCAGCACAGAAAATATTGGCACAGGGAAGCGAGTCTGC CAATATTGGCTGTGCTGCCAGGCAGGGTGGTG	305	17	6861658	6861744	-1
miR-200c	CCCTCGTCTTACCCAGCAGTGTTTGGGTGCGGTTGGGAGTCTCTAATACTGCC GGTAATGATGGAGG	306	12	6943123	6943190	1
miR-451	CTTGGGAATGGCAAGGAAAACCGTTACCAATTACTGAGTTTAGTAATGGTAATGG TTCTCTTGCTATACCCAGA	307	17	24212513	24212584	-1
miR-483-5p	GAGGGGAAGACGGGAGGAAAAGAGGGAGTGGTTCCATCAGCCCTCCTCACTC CTCTCTCCCGTCTTCTCCTCTC	308	11	2111940	2112015	-1
miR-491-3p	TTGACTTAGCTGGGTAGTGGGAAACCCCTCCATGAGGAGTAGAACACTCCTTA TGCAAGATTCCCTTCTACCTGGCTGGGTGG	309	9	20706104	20706187	1
miR-494	GATACTCGAAGGAGAGGTGTCCGTGTGTCTTCTCTTTATTATGATGAAAC ATACACGGGAAACCTCTTTTATGATC	310	14	100565724	100565804	1
miR-720	CCGGATCTCACACGGTGGTGTAAATATCTCGCTGGGGCCTCCAAAATGTTGTG CCCAGGGGTGTTAGAGAAAAACACACACTTTGAGATGAATTAAGAGTCCCTTA TTAG	311	3	165541823	165541932	1
miR-765	TTTAGGGCGTGATGAAAAGTGGAGTTCACTAGACAGCCCTTTTCAAGCCCTACG AGAAACTGGGGTTTCTGGAGGAGAAAGGAGTGATGAAGGATCTGTTCTCGTG AGCCTGAA	80	1	155172547	155172660	-1
miR-98	AGGATTCGTCTATGCCAGGTGAGGTAGTAAGTTGTATTGTTGTGGGTAGG GATATTAGGCCCAATTAGAAGATACTATACAACCTACTACTTCCCTGGTG TGTGGCAATTTCA	312	X	53599909	53600027	-1
miR-143	GCGAGGCCCTGTCTCCAGCCTGAGGTGCAGTGGTGCATCTCTGGTCAGTT GGGAGTCTGAGATGAAGCACTGTAGCTCAGGAAGAGAGAGTTGTTCTGCAGC CACCTTGTCTCACGGTCCAGTTTTCCAGGAATCCCTTAGATGCTAAGATGG GGATTCCCTGGAATACTGTTCTTGAGGTCAATGGTT	313	5	148788674	148788779	1
miR-145	AAAGATCCTCAGACAAATCCATGTGCTTCTCTCTCTTCAATCCACCGGAGTC TGCTCTCATACCCCAACCAGATTTTCAGTGGAGTGAAGTTCAGGAGGCATGGAGCT GACA	314	5	148790402	148790489	1
miR-205	TGTCGGGTAGCTTATCAGACTGATGTTGACTGTTGAATCTCATGGCAACACCA GTCGATGGGCTGCTGACA	315	1	207672101	207672210	1
miR-21	GGAGAGGAGGCAAGATGCTGGCATAGCTGTTGAACCTGGGAACCTGCTATGCCA ACATATTGCCATCTTTCC	316	17	55273409	55273480	1
miR-31		317	9	21502114	21502184	-1

Microarray validation

[00271] One or two microarrays were used for each of the samples in Table 8.

[00272] All sequences for which the intensity of the spot was higher than the local background mean intensity plus 1.5 times its standard deviation were declared “expressed.” The normalization was based on the signal obtained for positive controls.

[00273] The following parameters were used to check the quality of the microarray data for this experiment.

Control *HeatMap*: allows verification, after normalization, of:

- a. the low signal of the negative control,
- b. the specificity (*CTL26* versus *CTL26_MUT*),
- c. the approximate equality between positive control signals (without *CTL30*, which is a purification and labelling control),
- d. the approximate equality between block signal medians (based only on positive controls),
- e. the approximate equality between array signal medians (based on all sequences detected)
- f. the purification and labelling control (*CTL30*).

[00274] *Correlation within sample type*: for one sample type, the correlation between technical replicates (2 by 2) are computed, then the mean of all correlations obtained is computed.

[00275] *Approximation of the number of detected sequences by sample*: When only one array is run for a sample, a very large or very small number of detected sequences may suggest a technical problem.

[00276] *Effect of the block and array on the normalized signal*: based on the positive controls used for normalization, an ANOVA analysis is done to show the effect of the block and the array on the normalized signal. We check that the normalized signal obtained for one control on a particular block and a particular array is not due to the particular block or array.

[00277] *Standard deviation for positive control within one particular array and among all arrays:* The standard deviation for each positive control among all arrays is computed, as well as the mean of the standard deviation obtained for each control within one particular array. The normalization process may allow computation of similar “intra” and “inter” array variation (mean of standard deviation obtained within each array and standard deviation among all arrays).

[00278] *Ability of the miRNA profiles to distinguish sample type:* A hierarchical classification among all arrays is applied, taking into account all predictions (the distance used is based on Pearson correlation and the agglomerate method is “Ward’s minimum variant method”). A good clustering by sample type (technical replicates) indicates data of good quality.

Results of microarray validation

[00279] *Control HeatMap:* The positive, block, and labeling/purification controls gave approximately the same normalized signals. We found that the specificity of all of the arrays was excellent.

[00280] *Correlation, standard deviation, approximate number of detected sequences.* As shown in Table 13, the correlations obtained for tissues with two technical replicates were excellent. A higher or lower number of detected sequences were observed in some tissues compared to the typical number of detected sequences (see Table 13, shaded rows). That result may be due to the degradation state of the RNA obtained.

Table 13				
	<i>Correlation (mean)</i>	<i>Number of detected sequences</i>	<i>Number of Arrays kept for analysis</i>	<i>Number of Arrays not computed</i>
ASCC-1	-	322	1	0
SCC-1	0,983	~400	2	0
SCC-2	0,969	~300	2	0
ADC-1a	-	97	1	0
normal	-	-	0	1

SCC-3a	0,962	~300	2	0
SCC-4	0,99	~100	2	0
SCC-1b	0,986	~250	2	0
SCC-3b	-	291	1	1
ADC-1b	-	74	1	0
SCC-5	0,978	~450	2	0
ASCC-1b	-	624	1	0
CIN1	-	-	0	1
SCC-6	-	123	1	0
CIN3	-	-	0	1
SCC-7	-	610	1	0
SCC-8	-	249	1	0
cx-normal-1	-	109	1	0
cx-normal-3	-	66	1	0
cx-normal-4	-	365	1	0
Mean / Total	0,978	961	23	14,81%

[00281] *Effect of the array and block on the normalized signal obtained.*
Based on the ANOVA analysis, we observed no effect of the array and block on the positive control normalized signal.

[00282] *Standard deviation of positive controls.* The intra-array standard deviation was 0.29, and the inter-array standard deviation was 0.32. Both numbers are low and suggest little array effect on the normalized signals.

[00283] *Ability of the miRNA profiles to distinguish between sample types.*
All tissues that were analyzed in duplicate clustered together in this analysis, suggesting that the miRNA profiles did distinguish between sample types.

[00284] Based on the quality criteria considered, this microarray experiment was technically successful. For some tissues, however, a low number of

sequences were detected, suggesting that tissue or RNA degradation may have affected some samples.

5.4 Example 4: mRNA Expression in Cervical Cell Lines and Clinical Samples

Cell lines and tumor samples

[00285] The cell lines shown in Table 5 were used in this experiment. Total RNA from the cell lines was prepared as described in Example 1.

[00286] In addition, total RNA from two normal cervix and one squamous cervical tumor (SCC) were purchased from Applied Biosystems (Foster City, CA). See Table 14.

Table 14. Total RNA from Ambion

sample	Applied Biosystems
normal cervix – Ambion #1	AM6992 lot no.07060421
normal cervix – Ambion #2	AM7276 lot no. 03030243
cervix tumor – Ambion #1 (squamous)	AM7277 lot 03030253

[00287] Total RNA was prepared from fresh frozen samples using TRIzol® Reagent (Invitrogen; Carlsbad, CA) according to manufacturer's protocols. All RNA samples showed a good A260/280 ratio.

Table 15. Clinical samples

<i>Sample name</i>	<i>Sample type</i>
SCC-1c	Frozen
SCC-3b	Frozen
SCC-4b	Frozen
SCC-5	Frozen
SCC-8	Frozen

SCC-9	Frozen
SCC-10	Frozen
ADC-1b	Frozen
ASCC-1b	Frozen
CIN1#1	Frozen
CIN1#2	Frozen
CIN2#1	Frozen
cx-normal-6	Frozen
cx-normal-8	Frozen
cx-normal-9	Frozen
cx-normal-12	Frozen

cDNA Synthesis – Cell Lines and Ambion Samples

[00288] Reverse transcription was performed using 1 µg total RNA. In each case, the cDNA preparation was performed in a 50 µL reaction volume using random hexamers and TaqMan® reverse transcription reagents (Applied Biosystems; Foster City, CA) according to manufacturer's instructions. In parallel, the same reactions were performed without the reverse transcriptase as no-RT controls.

cDNA Synthesis – Clinical Samples

[00289] Reverse transcription was performed using 0.5 µg total RNA in a 20 µL reaction volume using random hexamers and the High Capacity cDNA RT kit from Applied Biosystems, Inc. (Foster City, CA) according to manufacturer's protocol. In parallel, the same reactions were performed without the reverse transcriptase as no-RT controls. In order to evaluate the variation of cDNA synthesis within a sample, three parallel cDNA syntheses were performed on a selection of RNA samples.

Selection of Reference Genes

[00290] For the cell lines, ACTB and TBP were used as reference genes. For the clinical samples, ACTB, TBP, and GAPDH were used as reference genes for the clinical samples.

Selection of mRNA Targets

[00291] The mRNA targets are shown in Table 16.

Table 16: mRNA targets

gene	alias	amplicon size	exon boundary	comment
CDKN2A	p16 ^{ink4}	70	1-2	all major isoforms
BIRC5	survivin	86	1-2	all major isoforms
TOP2A		72	23-24	
MCM5		70	16-17	
KRT19	CK19	64	2-3	
EPCAM	TACSTD1	82	2-3	
MMP2		86	10-11	
MMP9		67	1-2	
MCM2		82	2-3	
VEGFC		93	4-5	
TERT		94	3-4	
PCNA		117	5-6	
RPSA	67LR	121	7	
MAPK3	Erk-1	64	2-3	
IGF2BP3	L523S, IMP-3	97	12-13	
PIK3CA		104	6-7	
POU4F1	brn-3a	104	1-2	
MKI67	Ki-67	66	8-9	

PCR

[00292] All PCR reactions were performed in triplicate in a 25 μ L reaction volume on a Stratagene MX3000p instrument using a template concentration corresponding to 2 ng total RNA/reaction. TaqMan® Universal PCR Master Mix (Applied Biosystems Inc, Foster City, CA) was used for all reactions. The primers and

probes used in the PCR reactions are shown in Tables 17 and 18. All probes were FAM-labeled and all reactions were run in singleplex.

Table 17: Primers and probes for mRNA target PCR reactions

CDKN2A		SEQ ID
forward	5'-CATAGATGCCGCGGAAGGT-3'	318
reverse	5'-CCCGAGGTTTCTCAGAGCCT-3'	319
probe	FAM-CCTCAGACATCCCCGATTGAAAGAACC-TAMRA	320
BIRC5		
forward	5'-CTTTCTCAAGGACCACCGCA-3'	321
reverse	5'-GCCTCGGCCATCCGCT-3'	322
probe	FAM-CATTCAAGAACTGGCCCTTCTTGGAGG-TAMRA	323
KRT19		
forward	5'-AGATCGACAATGCCCCGT-3'	324
reverse	5'-AGAGCCTGTTCCGTCTCAAA-3'	325
probe	FAM-TGGCTGCAGATGACTTCCGAACCA-TAMRA	326
EPCAM		
forward	5'-TCATTGCTCAAAGCTGGCTG-3'	327
reverse	5'-AAACTTGGGAGAAGAGCAAAACC-3'	328
probe	FAM-AAATGTTTGGTGATGAAGGCAGAAATGAATGG-TAMRA	329
VEGFC		
forward	5'-TTCATTCCATTATTAGACGTTCCCT-3'	330
reverse	5'-GATTATTCCACATGTAATTGGTGGG-3'	331
probe	FAM-CCAGCAACACTACCACAGTGTCAGGCA-TAMRA	332
PCNA		
forward	5'-TTAAATTGTCACAGACAAGTAATGTCG-3'	333
reverse	5'-TGGCTTTTGTAAAGAAGTTCAGGTAC-3'	334
probe	FAM-TGGTTCATTTCATCTCTATGGTAACAGCTTCCTCCT-TAMRA	335
MMP9		

forward	5' -CCCTGGAGACCTGAGAACCA-3'	336
reverse	5' -AACCATAGCGGTACAGGTATTCCT-3'	337
probe	FAM-TCTCACCAGACAGGCAGCTGGCA-TAMRA	338
MMP2		
forward	5' -CCTGAGATCTGCAAACAGGACAT-3'	339
reverse	5' -CCAAATGAACCGGTCCTTGA-3'	340
probe	FAM-TTGATGGCATCGCTCAGATCCGTG-TAMRA	341
IGF2BP3		
forward	5' -GCTAAAGTGAGGATGGTGATTATCACT-3'	342
reverse	5' -ACTAACAAAGTTTTCTTCTTTAATTTTCCAT-3'	343
probe	FAM-ACCAGAGGCTCAGTTCAAGGCTCAGGGAA-TAMRA	344

Table 18: Primer/probe kits used for mRNA target PCR reactions

	Applied Biosystems Item Number
ACTB	Hs99999903_m1
GAPDH	Hs00266705_g1
TBP	Hs00427621_m1
TOP2A	Hs03063307_m1
MCM5	Hs01052142_m1
MKI67	Hs010332443_m1
POU4F1	Hs00366711_m1
MCM2	Hs00170472_m1
PIK3CA	Hs00180679_m1
MAPK3	Hs00385075_m1
RPSA	Hs03046712_g1
TERT	Hs99999022_m1

[00293] Primer and probe concentrations, as well as threshold settings used on the Stratagene MX3000 are shown in Table 19.

Table 19: mRNA target PCR conditions

target	Master Mix 25 μ L reaction			Threshold setting
	[forward]	[reverse]	[probe]	
CDKN2A	300 nM	300 nM	100 nM	500
BIRC5	100 nM	900 nM	150 nM	500
EPCAM	200 nM	200 nM	200 nM	500
KRT19	200 nM	200 nM	200 nM	500
VEGFC	500 nM	500 nM	200 nM	500
PCNA	500 nM	500 nM	200 nM	500
MMP2	500 nM	500 nM	200 nM	500
MMP9	500 nM	500 nM	200 nM	500
IGF2BP3	500nM	500nM	200nM	250
RPSA 20X mix	1,25 μ L			500
MAPK3 20X mix	1,25 μ L			500
MCM5 20X mix	1,25 μ L			250
TOP2A 20X mix	1,25 μ L			500
TERT 20X mix	1,25 μ L			250
MCM2 20X mix	1,25 μ L			500
MKI67 20X mix	1,25 μ L			500
PIK3CA 20X mix	1,25 μ L			150
POU4F1 20X mix	1,25 μ L			250
ACTB mix	1,25 μ L			500
TBP mix	1,25 μ L			500
GAPDH mix	1,25 μ L			500

[00294] All reactions, except for EPCAM, were cycled as follows: 10 minutes at 95°, followed by 40 cycles of 20 seconds at 95°C and 1 minute at 60°C. For EPCAM, the reaction was cycled as follows: 10 minutes at 95°, followed by 40 cycles of 20 seconds at 95°C and 1 minute at 64°C.

Expression and Statistical Analysis

[00295] For the analysis of relative expression, GenEx 4.4.2 software was used (multiD analysis, Gothenburg, Sweden). The GenEx software uses the $\Delta\Delta C_t$ formula, compensating for differences in PCR efficiency. All fold-change values are

calculated relative to one normal sample from Ambion (AM6992). The GenEx statistical module was used for descriptive statistics and t-test analysis.

Results

[00296] *PCR Specificity.* Specificity was determined using no-RT controls and analysis of amplicon size on an agarose gel. All amplicons tested were of the expected size, and we observed no amplification in no-RT controls, except for the no-RT controls for PCNA and RPSA, which was at very low levels and may have been due to contaminating DNA.

[00297] *PCR Efficiency.* In order to estimate the PCR efficiency for target and control mRNAs, a pool of cDNA from all of the cell lines was diluted in three-fold steps to generate a standard curve. PCR efficiency was calculated by the MX3000p software (Stratagene). Most assays had a similar efficiency, as shown in Table 20, except for MMP2, IGF2BP3, PIK3CA and POU4F1, which had very low expression levels in the cell lines, so it was not possible to generate a standard curve. We found that MMP9 was not expressed in cell lines at all.

Table 20: PCR efficiency in pooled cDNA from cell lines

gene	PCR efficiency (%)
ACTB	95
TBP	95
GAPDH	95
CDKN2A	95
BIRC5	80
TOP2A	90
MCM5	95
KRT19	80
EPCAM	90
MMP2	too low expression
MMP9	no expression in cell lines
MCM2	75
VEGFC	80
TERT	too low expression
PCNA	75
RPSA	95
MAPK3	80
IGF2BP3	too low expression
PIK3CA	too low expression
POU4F1	too low expression
MKI67	90

[00298] *Expression in cell lines.* In this experiment, CDKN2A and MKI67 were highly elevated compared to the normal cervix sample from Ambion. BIRC5 and TOP2A also show elevated levels. See Tables 21 and 22. MMP2, MMP9, TERT, and POU4F1 generated very high or no Ct values. IGF2BP3 was expressed at fairly high levels in all but one cell line.

Table 21. Fold-changes in expression levels: Cell lines vs. normal#1 (Ambion)

cell line	CDKN2A	BIRC5	TOP2A	MCM5	MKI67
CaSki	589,3	11,4	16,5	7,2	128,8
SiHa	241,6	5,3	3,3	1,3	47,0
C4-1	294,6	5,1	2,7	0,5	85,6
C4-2	899,5	17,7	49,5	22,9	507,9
sw756	684,8	6,3	9,5	0,9	217,1
ME-180	455,7	4,2	4,4	0,7	98,0
C33-A	523,7	4,3	12,0	3,9	67,8
HeLa S3	487,1	5,1	20,8	3,9	137,4

Table 22. Fold-changes in expression levels: Cell lines vs. normal#1 (Ambion)

cell line	KRT19	EPCAM	MCM2	RPSA	PCNA	MAPK3	VEGFC
CaSki	2,86	1,27	0,27	0,82	1,51	0,40	0,23
SiHa	0,02	0,02	0,13	0,20	0,36	0,21	0,06
C4-1	2,11	2,19	0,06	0,80	1,02	0,61	0,26
C4-2	4,79	1,47	0,97	3,49	3,79	2,72	0,12
sw756	No Ct	0,20	0,14	0,58	1,46	0,36	0,17
ME-180	2,05	0,36	0,16	0,28	0,81	0,37	No Ct
C33-A	No Ct	0,03	0,31	1,90	0,55	0,78	1,72
HeLa S3	0,06	0,03	0,61	1,27	1,70	0,36	0,60

[00299] *Expression in clinical samples.* All of the clinical samples, except for CIN3, SCC6, and SCC7, generated reliable Ct values for all of the reference genes. CIN3, SCC6, and SCC7 were therefore excluded from the analysis. *See* Tables 23 and 24.

Table 23. Fold-changes in expression levels: Clinical samples vs. normal#1 (Ambion)

Tissue samples	CDKN2A	BIRC5	TOP2A	MCM5	MMP9	MKI67
CIN1#1	21,6	9,5	26,1	23,0	70,9	0,2
CIN1#2	24,0	1,0	1,3	1,9	1,1	0,8
CIN2#1	30,8	3,5	3,7	0,4	2,0	1,2
SCC1c	304,3	7,0	65,7	31,6	14,0	137,9
SCC3b	202,2	5,5	79,5	39,8	6,4	107,0
SCC4b	282,1	9,9	39,3	13,4	3,4	46,4
SCC5	248,5	13,9	77,7	90,5	57,7	352,1
SCC8	292,0	4,0	26,1	11,9	8,8	48,7
SCC9	1,4	3,6	34,4	14,1	4,2	15,8
SCC10	261,4	4,0	15,7	1,5	11,9	9,6
ADC1b	76,6	13,1	82,3	21,3	12,2	
ASCC1b	131,6	6,7	43,4	66,6	1,9	144,1
Tumor AM7277 Ambion	382,8	2,5	57,8	47,9	1,4	84,4
Normal AM7276 Ambion	3,6	0,6	1,9	1,6	0,1	1,8
normal#6	4,8	3,1	6,9	1,0	0,8	4,8
normal#8	6,3	3,7	3,5	2,1	0,8	6,2
normal#9	4,5	2,6	2,7	2,4	0,4	5,2
normal#12	8,7	1,4	1,0	1,1	4,2	2,6

Table 24. Fold-changes in expression levels: Clinical samples vs. normal#1 (Ambion)

Tissue samples	KRT19	EPCAM	MMP2	MCM2	RPSA	PCNA	MAPK3
CIN1#1	1,1	0,8	0,02	1,5	6,2	1,7	0,9
CIN1#2	2,8	3,2	0,26	0,3	4,9	1,3	1,8
CIN2#1	1,4	1,0	0,23	0,3	3,1	2,4	3,3
SCC1c	2,8	1,8	0,09	4,7	4,6	1,7	0,8
SCC3b	2,2	2,7	0,30	4,5	2,8	4,4	1,4
SCC4b	4,7	17,6	0,41	5,2	3,6	1,0	2,2
SCC5	4,6	2,5	0,09	10,7	4,6	2,1	1,9
SCC8	7,5	1,6	0,32	2,1	3,3	1,2	0,5
SCC9	11,9	0,9	0,08	1,7	5,6	0,8	0,7
SCC10	5,0	1,6	0,39	1,0	6,0	8,1	2,4
ADC1b	2,7	31,8	0,38	2,6	1,5	1,6	3,8
ASCC1b	7,5	9,5	0,16	4,9	2,8	4,4	0,6
Tumor AM7277 Ambion	0,3	0,1	0,07	3,3	0,7	0,6	1,6
Normal AM7276 Ambion	0,5	1,2	0,11	1,2	1,4	0,4	3,0
normal#6	1,5	0,3	0,39	0,5	4,4	2,8	3,8
normal#8	2,0	0,9	0,40	0,7	2,2	2,8	7,6
normal#9	4,2	32,5	0,60	1,0	10,5	3,6	6,2
normal#12	2,3	7,2	1,31	0,6	8,0	2,3	4,3

[00300] The expression patterns for the cell lines in this experiment was similar to the expression patterns in the tumors. CDKN2A showed very high fold-changes in all SCC samples except for SCC9. SCC9 was further evaluated by immunohistological staining, and interestingly, the SCC9 sample was CDKN2A negative on the protein level. It is possible that SCC9 is HPV negative. MKI67, TOP2A, and MCM5 also showed high fold-changes in many of the tumor specimens.

[00301] There was a divergence in the results between the two CIN1 samples in this experiment. Because the CIN samples originate from biopsies without further micro-dissection, it is possible that the samples contain differing amounts of normal cells.

[00302] In this experiment, VEGFC, TERT, PIK3CA, POU4F1, and IGF2BP3 showed low to absent expression in all tumor and normal samples. Interestingly, IGF2BP3 was only detected in tumor samples, although with very high Ct values, and not in any of the normal samples.

[00303] Log2 fold-change in expression values relative to the Ambion normal sample were plotted for all tumor samples, except for CIN samples, and for all of the normal samples. Those results are shown in Figure 3A and 3B. Error bars represent standard deviation (SD).

[00304] *Statistical analysis.* Tumor samples (n=10) and normal samples (n=6) were further analysed for significant differences in relative expression levels, using the t-test. Because SCC9 alone among tumor samples did not appear to express CDKN2A, SCC9 was excluded in the t-test. All genes fulfilled the normal distribution criteria for the t-test to be valid. The results are shown in Table 25.

Table 25. Statistical (t-test) analysis of expression data – tumor vs. normal

gene	expression pattern	p-value
CDKN2A	tumor > normal	0,000000012
BIRC5	tumor > normal	0,0015591
TOP2A	tumor > normal	0,0000002
MCM5	tumor > normal	0,0000901
MMP9	tumor > normal	0,0015077
MKI67	tumor > normal	0,0000787
KRT19		0,1037645 NS
EPCAM		0,7343875 NS
MMP2	tumor < normal	0,0315609

MCM2	tumor > normal	0,0002907
RPSA		0,8822517 NS
PCNA		0,8408861 NS
MAPK3	tumor < normal	0,0116822

NS = not significant

[00305] In this experiment, expression levels of CDKN2A, MKI67, TOP2A, and MCM5 were clearly elevated, and that elevation was statistically significant. Increased expression of at least those four mRNAs, and possibly others, correlate with tumors.

5.5 Example 5: mRNA Expression in Liquid PAP Specimens

Materials and Methods

[00306] *Samples.* Three to six week old clinical liquid PAP specimens in PreservCyt transport media (Cytec) were used in this study. Four mL of each cell suspension was centrifuged at 2200xg for 15 minutes. The cell pellet was mixed with 700 μ L QIAzol lysis reagent (Qiagen; Hilden, Germany). Total RNA was extracted using miRNeasy RNA extraction kit (Qiagen; Hilden, Germany), according to manufacturer's protocol. The RNA concentration of each sample was measured with a NanoDrop instrument (Thermo Scientific; Wilmington, DE).

[00307] *cDNA Synthesis.* Reverse transcription was performed using 10 μ L total RNA (0.1 – 0.5 μ g) in a 20 μ L reaction volume using random hexamers and the High Capacity cDNA RT kit (Applied Biosystems, Inc.; Foster City, CA) according to manufacturer's protocol. In parallel, the same reactions were performed without the reverse transcriptase for no RT controls.

[00308] *Selection of Reference mRNAs and Target mRNAs.* ACTB, GAPDH and TBP were selected as references for this experiment. CDKN2A and MKI67 were selected as target mRNAs.

[00309] *PCR Reactions.* All PCR reactions were performed in triplicate in a 25 μ L reaction volume on a Stratagene MX3000p instrument using a cDNA concentration corresponding to about 2 ng total RNA/reaction. TaqMan® Universal PCR

Master Mix (Applied Biosystems, Inc; Foster City, CA) was used for all reactions. Primer and probe sequences, concentrations and thermal cycling conditions were the same as discussed in Example 4. All probes were FAM-labeled and all reactions were run in singleplex.

[00310] *Expression Analysis.* GenEx 4.4.2 software (multiD analysis, Gothenburg, Sweden) to analyse relative expression. The GenEx software uses the $\Delta\Delta C_t$ formula, compensating for differences in PCR efficiency.

Results

[00311] Total RNA yield varied between 0.3 μ g – 4 μ g. Four of the samples yielded insufficient concentrations of RNA to be used for cDNA synthesis. Ten out of twenty-one cDNAs were further analyzed for mRNA expression; the remaining samples resulted in C_t values that were too high to be reliable. C_t values for the analyzed samples are shown in Table 26. The relative log2 fold-changes are plotted in Figure 4.

Table 26: C_t values

Sample ID	ACTB	GAPDH	TBP	CDKN2A	MKI67
HPV0056, HSIL	24.19	26.21	33.73	31,37	34,53
	24.16	26.28	34.31	31,3	34,16
	24.46	26.46	33.58	31,14	34,72
HPV0059, HSIL	24.79	27.20	35.30	28,9	35,14
	24.54	27.42	35.18	29,18	35,2
	25.03	27.24	35.27	28,99	35,74
HPV0060, HSIL	25.38	24.92	31.75	28,91	32,41
	25.63	25.10	31.67	28,95	32,31
	25.65	25.13	31.59	29,09	32,76
HPV0061, HSIL	25.88	25.98	31.69	26,98	29,68
	25.79	25.93	32.09	27,01	29,52
	26.05	25.91	31.93	26,89	29,86
HPV0062, HSIL	25.64	28.14	33.87	29,61	36,34

	25.48	28.05	35.43	29,56	35,64
	25.86	28.10	34.76	29,66	36,16
HPV0065, LSIL	29.72	28.07	35.75	32,44	34,86
	29.62	27.87	35.92	32,22	34,98
	30.01	27.82	34.90	32,84	34,82
	27.17	27.72	35.23	32,41	35,27
HPV0066, LSIL	27.29	27.56	34.97	32,18	35,21
	27.31	27.67	35.02	32,41	34,34
	26.56	29.05	34.97	31,21	No Ct
HPV0069, negative	26.60	28.72	36.00	31,04	39,36
	26.97	28.88	35.15	31,11	No Ct
	28.32	29.21	34.89	31,9	36,69
HPV0071, negative	28.48	28.90	34.50	32,14	35,78
	28.70	29.17	34.69	32,07	36,8
	28.37	27.09	32.97	31,06	35,18
HPV0074, negative	28.60	27.02	32.74	31,05	36,25
	28.46	27.00	33.01	31,3	36,26
	28.90	29.19	34.03	32,78	38,61
HPV0078, negative	28.91	29.16	34.63	32,94	36,39
	28.93	29.29	34.61	33,03	39,07

LSIL = low grade squamous intraepithelial lesion

HSIL = high grade squamous intraepithelial lesion

[00312] This experiment demonstrates that mRNA expression markers can be detected in liquid PAP specimens.

5.6 Example 6: microRNA detection in clinical samples by RT-PCR

Materials and Methods

[00313] *Samples.* Total RNA from normal cervix was purchased from Applied Biosystems, Inc. (Foster City, CA; ABI AM 6992; “Ambion sample”). Total RNA from cervical tumor and normal specimens were prepared as in Example 3.

[00314] *MicroRNAs Selected for Analysis.* miR-21 was analyzed for each of the samples shown in Table 27. miR-1290 was analyzed for a subset of the samples, as shown in Table 27. RNU44, U47, and RNU48 were used for normalization.

Table 27. RNA samples and microRNAs selected for expression analysis

Sample name	Sample description	miR-21	miR-1290
SCC-1c	Frozen	x	x
SCC-3b	Frozen	x	x
SCC-4b	Frozen	x	x
SCC-5	Frozen	x	x
SCC-8	Frozen	x	x
SCC-9	Frozen	x	x
SCC-10	Frozen	x	x
SCC-11	Frozen		x
SCC-12	Frozen		x
SCC-13	FFPE		x
SCC-14	FFPE		x
SCC-15	FFPE		x
SCC-16	FFPE		x
SCC-17	FFPE		x

Sample name	Sample description	miR-21	miR-1290
ASCC-1b	Frozen	x	
AIS-1	FFPE		x
AIS-2	FFPE		x
ADC-1b	Frozen	x	
ADC-2	FFPE		x
ADC-3	FFPE		x
ADC-4	FFPE		x
CIN3-2	FFPE	x	
cx-normal-6	Frozen	x	
cx-normal-8	Frozen	x	
cx-normal-9	Frozen	x	
cx-normal-12	Frozen	x	x
cx-normal-13	Frozen		x
cx-normal-14	Frozen		x
cx-normal-15	Frozen		x
cx-normal-16	Frozen		x
cx-normal-17	Frozen		x
cx-normal-19	Frozen		x
cx-normal-20	Frozen		x
normal Ambion	ABI frozen	x	x
hyperplasia-1 (benign)	Frozen		x

SCC = Squamous Cervical Carcinoma, ADC = Cervical Adenocarcinoma,

ASCC = Adeno-Squamous Cervical Carcinoma, AIS = Adenocarcinoma in situ

[00315] *RT-PCR Reactions.* All microRNA RT-PCR assays were purchased from Applied Biosystems Inc (Foster City, CA), including those for normalization genes. The cDNA synthesis and PCR reactions were performed according to manufacturer's protocols. All PCR reactions were run on an MX3000 instrument (Stratagene).

[00316] *Expression analysis.* For the analysis of relative expression, the GenEx 4.4.2 software (multiD analysis, Gothenburg, Sweden) was used. The GenEx software uses the $\Delta\Delta C_t$ formula, compensating for differences in PCR efficiency. The GenEx statistical module was used for t-test analysis.

Results

[00317] MiR-205 and miR-1290 show a statistically significant difference in expression levels between tumor and normal samples. *See* Figures 5 and 6. Interestingly, miR-205 is markedly down-regulated in adeno-carcinoma specimens, which originate from glandular endothelial cells, while it is upregulated in squamous carcinoma specimens, which are derived from epithelial cells. This result is consistent with reports discussing other forms of cancer, such as lung cancer.

[00318] These results demonstrate that RT-PCR can be used to detect increases in expression of microRNAs in cervical specimens.

5.8 Example 7: Bioinformatic Analysis to Identify microRNAs

[00319] In order to identify the microRNAs detected with the probes shown, e.g., in Tables 1 and 11, small RNA sequencing (smRNASeq) datasets were analysed using the probe sequences to identify expressed microRNAs detected by those sequences. The analysis identified 44 sequences with precise ends, corresponding to 37 arms (i.e., some of the sequences appear to be isomirs, or multiple candidate microRNAs from a single arm). Those 44 candidate microRNA sequences are show in Table 28.

Table 28: microRNA candidate sequences corresponding to probes

<u>Arm name</u>	<u>microRNA candidate sequence 5' -> 3'</u>	<u>SEQ ID</u>
12726-L	TCCCCCAACCCACAGCACACAC	345
12730-R	CCCGGAGAGCGGAGCACAACACA	346
12730-R	CCGGAGAGCGGAGCACAAC	347
13108-L	CCAAGGAAGGCAGCAGGC	348
13122-L	GATGGAATTTCTTAAAGG	349
13124-L	GGAGGGGAGGAGACATG	350
13181-L	GCAGTGACTGTTTCAGACGTCCA	351
13207-R	TGTCTTTCCTTGTTGGAGCAGG	352
13209-L	CAGCAGGCGAGTTACCTCAA	353
13227-L	GAGGAGGACTGGGCCCTA	354
13229-R	AGCCGCTCTTCTCCCTGCCCACA	355
13229-R	AGCCGCTCTTCTCCCTGCCCACAG	356
13231-L	TGGGGAGCGGCCCCCGGG	357
13247-L	GAGGTCGGGAGGGGAAGGCGGCT	358
13252-L	TCAAGGAGCTCACAGTC	359
13254-R	GCATGAGTGGTTCAGTGGT	360
13267-L	GTGGGCTGGGCTGGGCTGGGC	361
13274-L	GGAGGACCCTGAGGGAGGGTGGG	362
13274-L	TGAGGGAGGGTGGGAGC	363
13283-L	TGGCAGCAAGGAAGGCAGGGGTC	364
13291-L	GAGGGAAGGAGGGAGGAA	365
13296-L	CAGGGCAGAGGGCACAGGAATCTGA	366
13325-R	GGGAAGAGCCCAGCGCC	367
13339-L	ACCCTCAGTCCGTATTGGTCTCT	368
13504-R	GTCTCCCAGAGCAGGGACGCTTT	369
25-R	TTAGAAAAAGAGGGGGTGAGG	370

3371-L	TGGGGTGTGGAGGGGAGG	371
3744-R	AGGGGAGCAGGGAGGAA	372
3995-L	TGGCCTGACGTGAGGAGGAGG	373
4417-R	ACGGACAGGGAACTTTTGAT	374
4440-L	GCCCAGTGCTCTGAATGTCAAA	375
4440-L	TCTGCCCAGTGCTCTGAATGTCA	376
4440-R	GGCGGGAGTAACTATGAC	377
4440-R	CGGGTAAACGGCGGGAGTAACT	378
4498-L	AGCAGGCGCACGGCCGTCTGGATC	379
4498-L	GCACGGCCGTCTGGATCTCC	380
5192-L	GAGGAAGGAAGGGGAAA	381
6216-L	CAGTGCTCTGAATGTCAAAGTGAAGA	382
6216-R	GGGTAAACGGCGGGAGTA	383
6235-R	AAATGGATTTTTGGAGCAG	384
6930-R	TGCAAGATCAGAGGGGAGA	385
7578-L	GGGGCTGTAGCTCAGGG	386
8339-R	GCCGGTACAGTGAAAAT	387
8339-R	GCCGGTACAGTGAAAATG	388

[00320] All publications, patents, patent applications and other documents cited in this application are hereby incorporated by reference in their entireties for all purposes to the same extent as if each individual publication, patent, patent application or other document were individually indicated to be incorporated by reference for all purposes.

[00321] While various specific embodiments have been illustrated and described, it will be appreciated that changes can be made without departing from the spirit and scope of the invention(s).

WHAT IS CLAIMED IS:

1. A method for detecting the presence of cervical dysplasia in a human cervical sample, the method comprising:

hybridizing nucleic acids of the sample with at least a first polynucleotide that is complementary to a first target RNA in the sample or to a complement thereof and detecting a first complex comprising the first polynucleotide hybridized to at least one nucleic acid selected from the first target RNA, a DNA amplicon of the first target RNA, and a complement of the first target RNA, wherein the first target RNA is capable of specifically hybridizing to a nucleic acid having a sequence selected from SEQ ID NO.:1 to 41 and 133 to 211, wherein detection of a level of the first complex that is greater than a normal level of the first complex indicates the presence of cervical dysplasia in the sample.

2. The method of claim 1, wherein the first target RNA comprises a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NO.:1 to 41 and 133 to 211.

3. The method of any one of the preceding claims, wherein the first target RNA comprises at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388.

4. The method of any one of the preceding claims, wherein the sample comprises RNA that has been separated from DNA.

5. The method of any of the preceding claims, wherein the first target RNA in its mature form comprises fewer than 30 nucleotides.

6. The method of any of the preceding claims, wherein the first target RNA is a microRNA.

7. The method of any one of the preceding claims, wherein the method further comprises hybridizing nucleic acids of the sample with a second polynucleotide that is complementary to a second target RNA in the sample or to a complement thereof, wherein the second target RNA is different from the first target RNA, and detecting a second complex comprising the second polynucleotide hybridized to at least one nucleic acid selected from the second target RNA, a DNA amplicon of the second target RNA, and a complement of the second target RNA.

8. The method of claim 7, wherein the second target RNA is a microRNA.
9. The method of claim 7, wherein the second target RNA is an mRNA.
10. The method of claim 9, wherein the mRNA is selected from CDKN2A, MKI67, TOP2A, MCM5, BIRC5, MMP9, and MCM2.
11. The method of any one of claims 7 to 9, wherein detection of a level of at least one of the first and second complexes that is greater than the normal level of the respective complex indicates the presence of cervical dysplasia in the sample.
12. The method of claim 11, wherein detection of levels of the first and second complexes that are both greater than the normal levels of the first and second complexes indicates the presence of cervical dysplasia in the sample.
13. The method of claim 7, wherein the second target RNA is capable of specifically hybridizing to a nucleic acid having a sequence selected from SEQ ID NO.:1 to 41 and 133 to 211, and wherein the first target RNA and the second target RNA are different.
14. The method of claim 7, wherein the second target RNA comprises a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NO.:1 to 41 and 133 to 211, and wherein the first target RNA and the second target RNA are different.
15. The method of claim 7, wherein the second target RNA comprises at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388, and wherein the first target RNA and the second target RNA are different.
16. The method of any one of claims 7 to 15, wherein the method further comprises hybridizing nucleic acids of the sample with a third polynucleotide that is complementary to a third target RNA in the sample or to a complement thereof, wherein the third target RNA is different from the first target RNA and the second target RNA, and detecting a third complex comprising the third polynucleotide hybridized to at least one nucleic acid selected from the third target RNA, a DNA amplicon of the third target RNA, and a complement of the third target RNA.
17. The method of claim 16, wherein detection of a level of at least one of the first, second, and third complexes that is greater than the normal level of the respective complex indicates the presence of cervical dysplasia in the sample.

18. The method of claim 17, wherein detection of levels of the first, second, and third complexes that are greater than their normal levels of the first, second, and third complexes indicates the presence of cervical dysplasia in the sample.

19. The method of claim 16, wherein the third target RNA is capable of specifically hybridizing to a nucleic acid having a sequence selected from SEQ ID NO.:1 to 41 and 133 to 211, and wherein the first, second, and third target RNAs are different.

20. The method of claim 16, wherein the third target RNA comprises a sequence that is complementary to at least 15 contiguous nucleotides of a sequence selected from SEQ ID NO.:1 to 41 and 133 to 211, and wherein the first, second, and third target RNAs are different.

21. The method of claim 16, wherein the third target RNA comprises at least 15 contiguous nucleotides of a sequence selected from SEQ ID NOs: 345 to 388, and wherein the first, second, and third target RNAs are different.

22. The method of any one of claims 16 to 21, wherein the method comprises detecting at least five complexes, each comprising a polynucleotide hybridized to a nucleic acid selected from a target RNA, a DNA amplicon of a target RNA, and a complement of a target RNA.

23. The method of claim 1, wherein at least one target RNA is capable of specifically hybridizing to a sequence selected from SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7 and 8.

24. The method of claim 1, wherein at least one target RNA is capable of specifically hybridizing to a sequence selected from SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12.

25. The method of claim 1, wherein at least one target RNA is capable of specifically hybridizing to a sequence selected from SEQ ID NOs: 1, 5, 7, and 32.

26. The method of claim 1, wherein at least one target RNA is capable of specifically hybridizing to a sequence selected from SEQ ID NOs: 142, 151, 153, 193, 194, 205, 172, 208, 210, and 211.

27. The method of any one of the preceding claims, wherein at least one polynucleotide is complementary to at least 10 contiguous nucleotides of a target RNA or its complement.

28. The method of any one of the preceding claims, wherein at least one polynucleotide is complementary to at least 15 contiguous nucleotides of a target RNA or its complement.

29. The method of any one of the preceding claims, wherein at least one polynucleotide comprises a region that is not complementary to a target RNA or its complement.

30. A synthetic polynucleotide comprising a first region, wherein the first region comprises a sequence of at least 8 contiguous nucleotides that is identical or complementary to a sequence of at least 8 contiguous nucleotides of one of SEQ ID NOs: 1 to 7, 9 to 37, 133 to 201, and 345 to 388.

31. The synthetic polynucleotide of claim 30, wherein the first region comprises a sequence of at least 9 contiguous nucleotides that is identical or complementary to a sequence of at least 9 contiguous nucleotides of one of SEQ ID NOs: 1 to 7, 9 to 37, 133 to 201, and 345 to 388.

32. The synthetic polynucleotide of claim 30, wherein the first region comprises a sequence of at least 10 contiguous nucleotides that is identical or complementary to a sequence of at least 10 contiguous nucleotides of one of SEQ ID NOs: 1 to 7, 9 to 37, 133 to 201, and 345 to 388.

33. The synthetic polynucleotide of claim 30, wherein the first region comprises a sequence of at least 12 contiguous nucleotides that is identical or complementary to a sequence of at least 12 contiguous nucleotides of one of SEQ ID NOs: 1 to 7, 9 to 37, 133 to 201, and 345 to 388.

34. The synthetic polynucleotide of any one of claims 30 to 33, wherein the polynucleotide comprises a detectable label.

35. The synthetic polynucleotide of claim 34, wherein the detectable label is a FRET label.

36. The synthetic polynucleotide of any one of claims 30 to 35, wherein the first region is identical or complementary to a region of a target RNA.

37. The synthetic polynucleotide of claim 36, wherein the polynucleotide further comprises a second region that is not identical or complementary to a region of the target RNA.

38. A composition comprising a plurality of synthetic polynucleotides, wherein at least one polynucleotide comprises a first region comprising a sequence of at least 8 contiguous nucleotides that is identical or complementary to a sequence of at least 8 contiguous nucleotides of one or more of SEQ ID NOs: 1 to 7, 9 to 37, 133 to 201, and 345 to 388.

39. The composition of claim 38, wherein at least two polynucleotides of the plurality of synthetic polynucleotides comprise a first region comprising a sequence of at least 8 contiguous nucleotides that is identical or complementary to a sequence of at least 8 contiguous nucleotides of one or more of SEQ ID NOs: 1 to 7, 9 to 37, 133 to 201, and 345 to 388, and wherein the first regions of the at least two polynucleotides are different.

40. The composition of claim 38, wherein at least three polynucleotides of the plurality of synthetic polynucleotides comprise a first region comprising a sequence of at least 8 contiguous nucleotides that is identical or complementary to a sequence of at least 8 contiguous nucleotides of one or more of SEQ ID NOs: 1 to 7, 9 to 37, 133 to 201, and 345 to 388, and wherein the first regions of the at least three polynucleotides are different.

41. The composition of claim 38, wherein at least five polynucleotides of the plurality of synthetic polynucleotides comprise a first region comprising a sequence of at least 8 contiguous nucleotides that is identical or complementary to a sequence of at least 8 contiguous nucleotides of one or more of SEQ ID NOs: 1 to 7, 9 to 37, 133 to 201, and 345 to 388, and wherein the first regions of the at least five polynucleotides are different.

42. A kit comprising a synthetic polynucleotide of any one of claims 30 to 37.

43. A kit comprising a composition of any one of claims 38 to 41.

44. The kit of claim 42 or claim 43, wherein the kit further comprises at least one polymerase.

45. The kit of any one of claims 42 to 44, wherein the kit further comprises dNTPs.

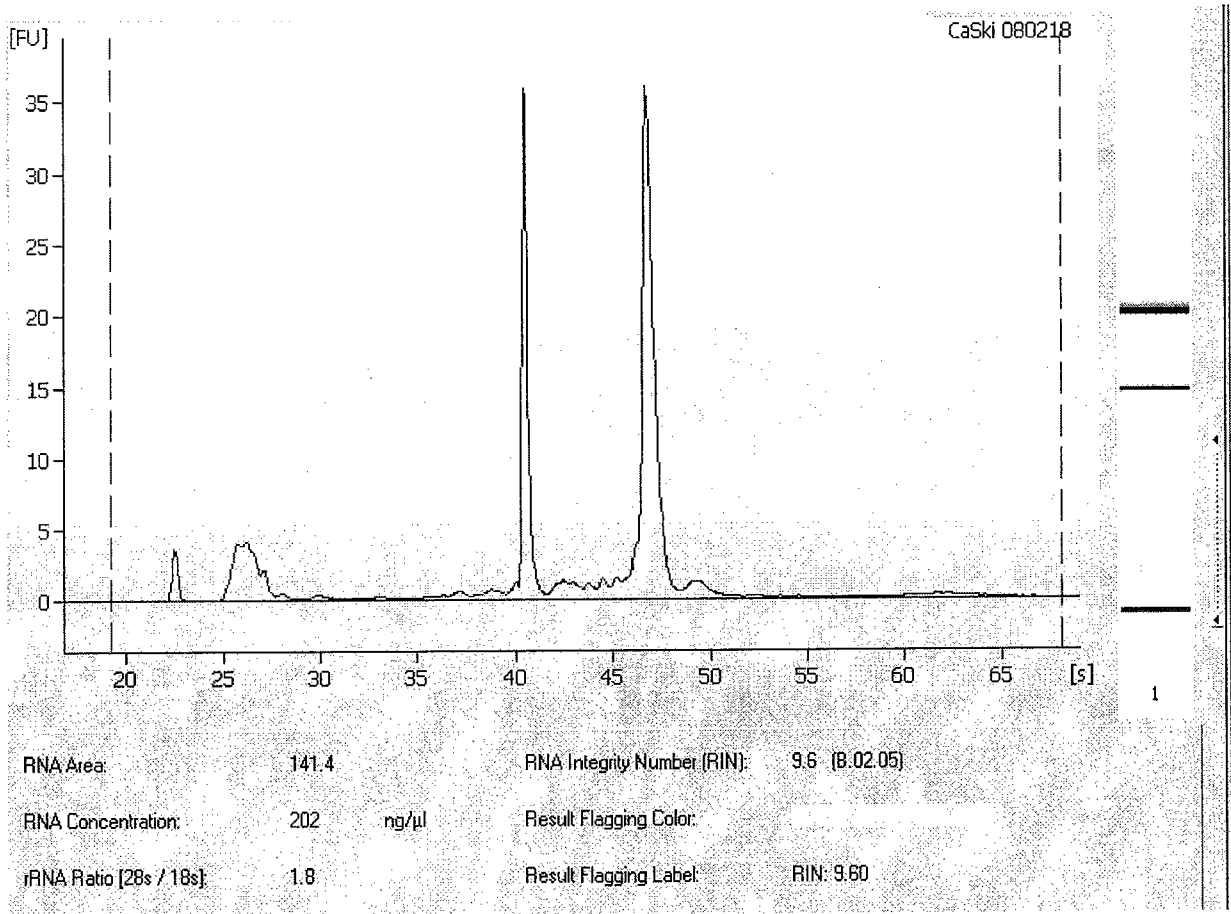


Fig. 1

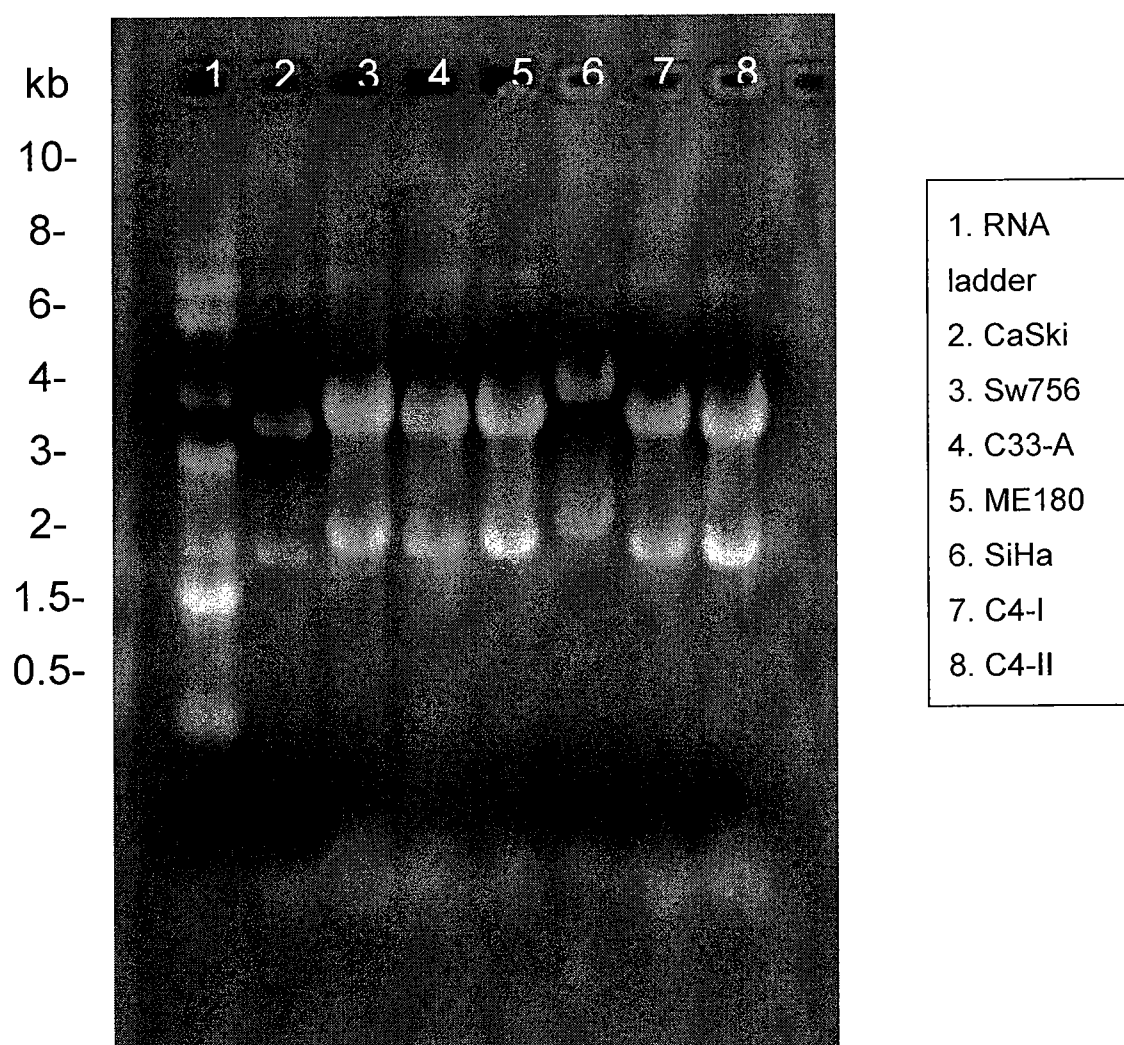
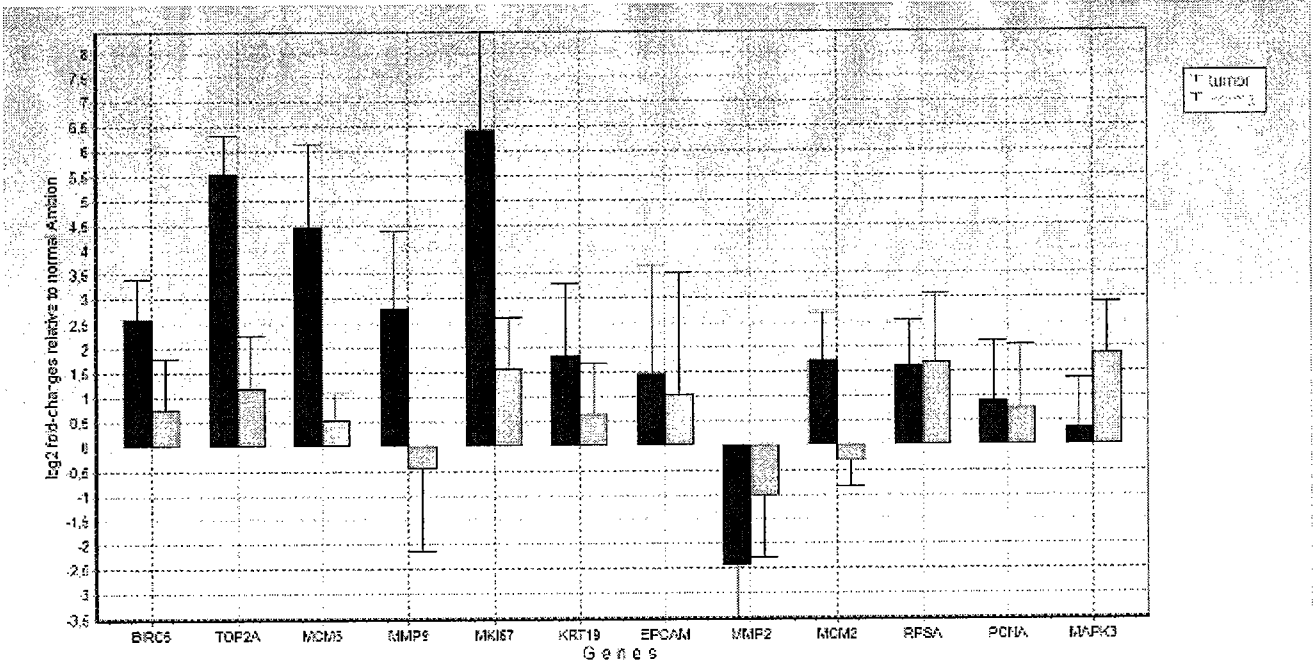


Fig. 2

A



B

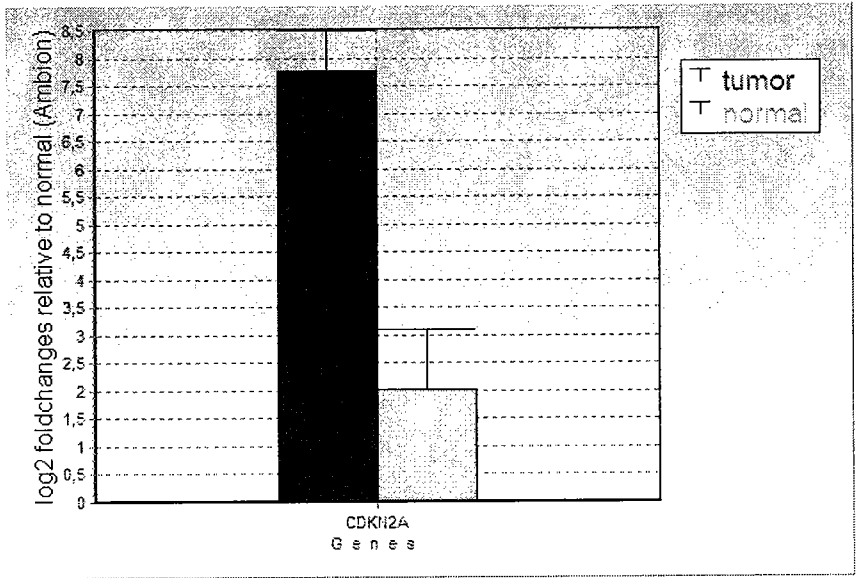


Fig. 3

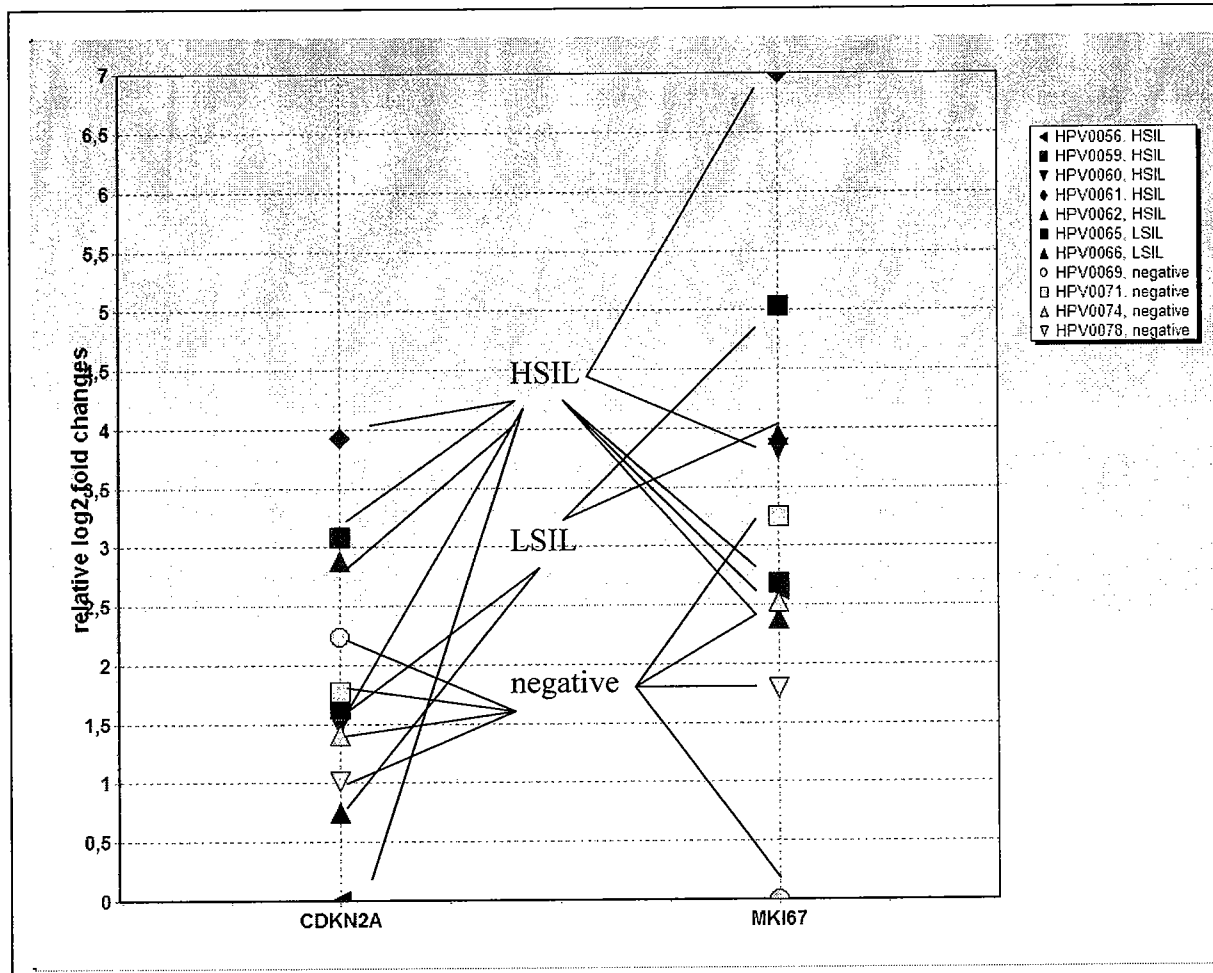


Fig. 4

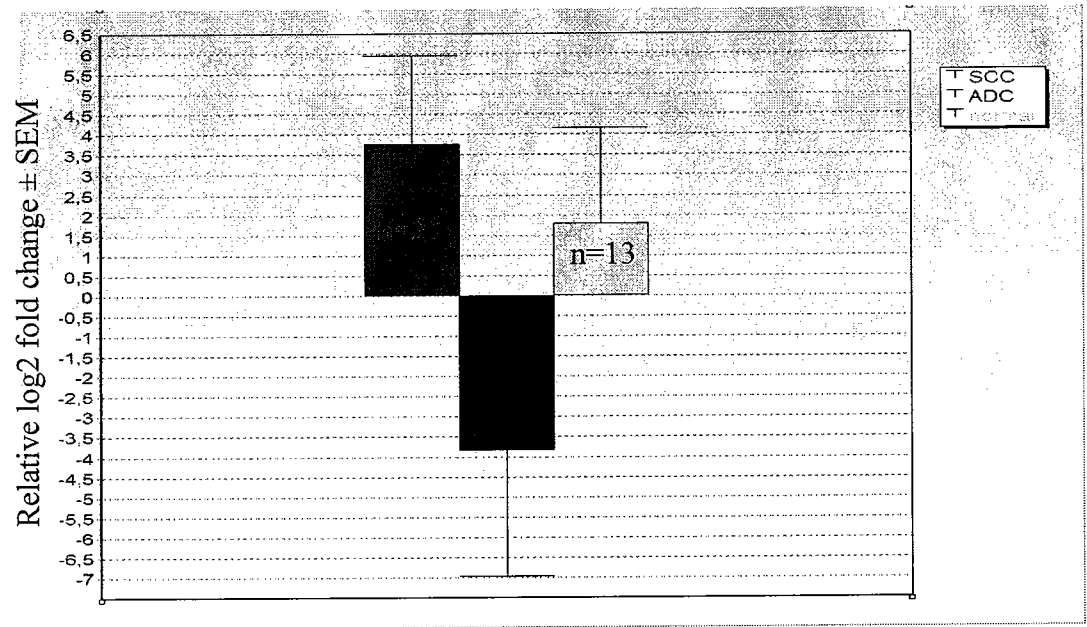


Fig. 5

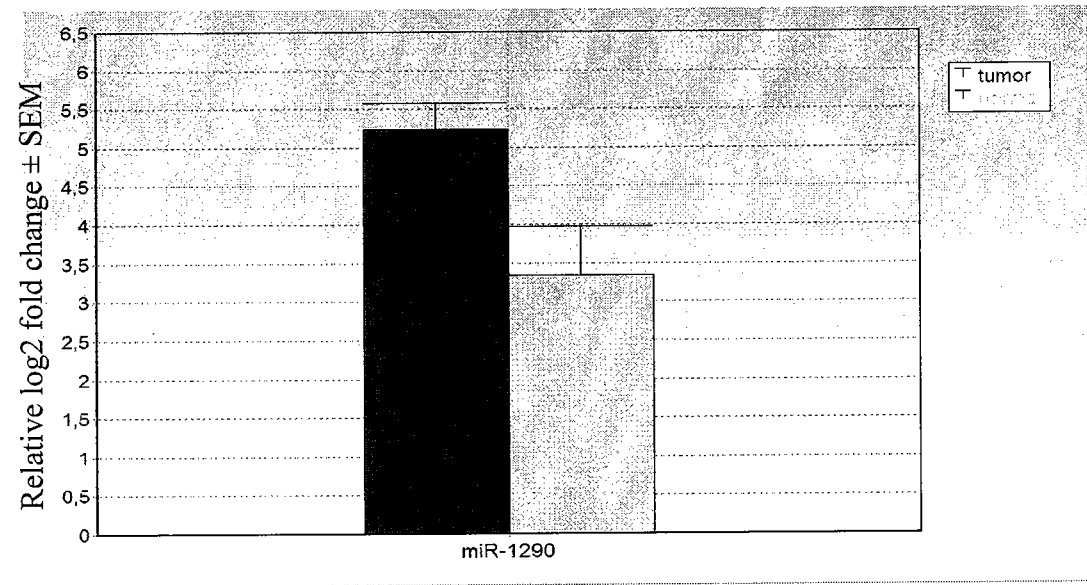


Fig. 6