Am J Cancer Res 2016;6(7):1537-1546
www.ajcr.us /ISSN:2156-6976/ajcr0030129

Original Article

MicroRNAs-208b-3p, 204-5p, 129-2-3p and 3065-5p as predictive markers of oral leukoplakia that progress to cancer

Elizabeth Philipone[1,2], Angela J Yoon[1,2], Shuang Wang[3], Jing Shen[4], Yen Chen Kevin Ko[1], Jill M Sink[1], Andrew Rockafellow[5], Nathanel A Shammay[5], Regina M Santella[4]

[1]Division of Oral and Maxillofacial Pathology, Columbia University College of Dental Medicine, New York, NY, USA; [2]Departments of Pathology and Cell Biology, Columbia University College of Physicians and Surgeons, New York, NY, USA; [3]Department of Biostatistics, Columbia University Mailman School of Public Health, New York, NY, USA; [4]Department of Environmental Health Sciences, Columbia University Mailman School of Public Health, [5]Columbia University College of Dental Medicine, New York, NY, USA

Received April 8, 2016; Accepted May 1, 2016; Epub July 1, 2016; Published July 15, 2016

Abstract: Leukoplakia is the most common precursor lesion of oral squamous cell carcinoma (OSCC). Currently, the risk of progression to OSCC is assessed based on histopathologic examination alone. However, this method fails to identify the subset of microscopically innocuous leukoplakia that ultimately transforms to OSCC. The aim of this study was to determine if microRNAs (miRNAs) can be utilized to identify non- and low-grade dysplastic oral lesions at risk for cancer progression. A retrospective study of genome-wide miRNA expression level analyses was performed in the training cohort (n=20) using deep sequencing formalin-fixed paraffin embedded incisional biopsy tissues from patients with oral leukoplakic lesions diagnosed with non- or low-grade dysplasia and known clinical outcome. The promising miRNA candidates were then evaluated in the validation cohort (n=80) using quantitative real-time PCR (qRT-PCR). Four promising miRNAs-208b-3p, 204-5p, 129-2-3p and 3065-5p were identified. Combining these four miRNAs as a panel with age and histologic diagnosis ($p<0.004$), our final model had a predictive value for the area under the receiver operating characteristic (ROC) curve (AUC) of 0.792, sensitivity of 76.9% and specificity of 73.7% to accurately identify non- and low-grade dysplastic lesions at risk of cancer progression, which is a significant improvement over histopathologic examination alone (AUC of 0.645). While further investigation is needed, discovery of predictive markers that can accurately identify histologically innocuous oral lesions at high risk for progression to OSSC will significantly improve clinical outcome by means of early intervention.

Keywords: microRNA, deep sequencing, qRT-PCR, predictive markers, oral leukoplakia, oral epithelial dysplasia, oral squamous cell carcinoma

Introduction

Oral squamous cell carcinoma (OSCC) is the 6th leading cause of cancer related deaths worldwide [1, 2]. Over 30,000 people in the United States are diagnosed with OSCC each year [1-4]. Half of those newly diagnosed with oral cancer will die of the disease [1-4]. Since the majority of OSCC develop from a precursor lesion [4, 5], accurate identification, followed by complete surgical removal of the precursor lesion will effectively halt its malignant progression, offering the best hope at reducing OSCC associated morbidity and mortality.

Most precursor lesions of OSCC clinically present as a white patch known as leukoplakia [4-16]. Upon microscopic examination, most leukoplakia (80%) appear to be non-dysplastic or low-grade dysplastic lesions, whereupon the patients are diagnosed with epithelial hyperplasia and/or hyperkeratosis and placed under 'close observation' without further treatment [14]. Moderate to high-grade epithelial dysplasia is seen in ~17% of leukoplakia [15], in which case the lesion is completely removed by surgery [16]. The remaining 3% of cases are OSCC [14]. While moderate to high-grade epithelial dysplasia will most likely progress to OSCC with-

in 4-5 years, some of the non-dysplastic and low-grade dysplastic lesions also undergo malignant transformation [8-16]. Hence histopathologic examination alone is insufficient for identification of leukoplakic lesions that will progress of OSCC, especially if the leukoplakia lacks atypical cellular features at the time of initial histopathologic assessment.

It is these histologically innocuous lesions that are at greatest need for intervention since without the presence of significant histologic dysplasia, these lesions could be erroneously interpreted as being harmless and not managed properly. Currently there are no established guidelines for managing non- and low-grade dysplastic lesions. Treatment options vary from no treatment to complete surgical excision to laser ablation [17]. Moreover, there is no established guideline for patient follow-up duration or frequency [17]. Accurate identification of non-dysplastic and low-grade dysplastic lesions at risk for malignant transformation will ensure proper clinical management of these premalignant lesions.

We assessed microRNA (miRNA)-based markers that can increase the predictive value when combined with histopathologic examination to accurately identify non- or low-grade dysplastic lesion that will ultimately progress to OSCC. miRNAs are small non-coding RNA molecules that function as post-transcriptional regulators capable of causing both gene silencing and activation [18]. They are involved in multiple critical biological processes, including cellular proliferation, apoptosis, differentiation, and carcinogenesis [19-22]. A limited number of studies have examined differential miRNA expression levels in precancerous lesions in the esophagus, cervix, lung and stomach [23-28]. While there are reports of miRNA expression profiles in oral leukoplakia that progressed to OSCC [29-36], only a couple of these studies focused on low-grade dysplastic lesions that progressed to cancer [34-36]. Moreover, while the majority of other studies of oral leukoplakia utilized TaqMan low density arrays, we employed genome wide deep sequencing to determine a miRNA expression profile unique to histologically innocuous (non- and low-grade dysplastic lesions) leukoplakias that ultimately progress to cancer. The advantage of this next-generating sequencing technology is that it is capable of detecting all miRNAs present in the samples, thereby increasing testing coverage [37]. Development of a clinically applicable method to identify histologically innocuous precancerous lesions will be a valuable modality for clinicians in guiding patient management.

Materials and methods

Subjects

Following institutional review board approval at the Columbia University Medical Center, we conducted a retrospective search of our pathology database to identify 100 adult patients ≥21 years old, with a clinical leukoplakia diagnosed as 'epithelial hyperplasia', 'epithelial hyperplasia with hyperkeratosis', 'epithelial atypia limited to the basal cell region' or 'mild epithelial dysplasia' prior to 2008 and which also had a minimum of 5-year follow-up information available. Only the patients who had an incisional biopsy of the leukoplakia were selected. As an excisional biopsy could account for the reason why the leukoplakia did not progress to cancer, the patients who had excisional biopsy were excluded from the study. Identified patients were stratified into the following two groups; Group 1 'Progressive Group' (patients with leukoplakia that progressed to OSCC within 5 years) and Group 2 'Non-Progressive Group' (patients with leukoplakia that did not progress to OSCC within 5 years). The age and gender of the patient, histologic diagnosis, and the location of lesion were recorded. Archived formalin-fixed paraffin-embedded (FFPE) tissue blocks were retrieved for all subjects.

Total RNA extraction

Ten 10-μm sections were obtained from archived FFPE precursor tissue samples for all subjects. For each sample, a representative section was stained with H&E and reviewed by a pathologist to confirm the histologic diagnosis. Total RNA was isolated from tissues using RNeasy FFPE kit (Qiagen Inc., Valencia, CA) following the manufacturer's protocol and yield was quantitated by Nanodrop, and samples were stored at -80°C.

microRNA quality control, library preparation and sequencing

Deep sequencing analysis was performed as previously described [38] in the training set of 20 patients; 10 from Group 1 (progressive group) and 10 from Group 2 (non-progressive group). For quantification and quality control,

total RNAs were tested using the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) and the Qubit 2.0 Fluorometer (Life Technologies, Grand Island, NY). Total RNA (minimum of 2 μg) ranging from 18 to 30 nt were gel-purified and ligated to 3' and 5'adaptors. Ligation products were reverse transcribed, then amplified for 16 cycles using the adaptor primers, and the fragments around 150 bp were isolated from PAGE-gels using a TruSeq Small RNA Sample Prep Kit (Illumina, San Diego, CA). Libraries were sequenced on an Illumina HiSeq 2500 platform–50 SR that allows for 1x50 base-pair single-end reads in the New York Genome Center. The data was deposited in the publicly available Gene Expression Omnibus (GEO) database (GSE62809)

microRNA mapping and differential expression analysis

Adaptors were removed and low quality tags were filtered with FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). Reads were processed with the pipeline miraligner, which mapped them to the miRBase v.20 sequences. The number of reads per miRNA was first assessed and analyzed using the bioconductor package DESeq to compare miRNA expression between the two groups. In addition to normalizing between the samples, DESeq performed a statistical test of differential expression under the hypothesis of a negative binomial distribution of the reads. To select the top miRNAs with prognostic value, those with significant (unadjusted $p<0.05$) and >1.25 log2-fold change of the different mean normalized expression levels between the two groups were identified. The association between selected miRNAs and their role in carcinogenesis was also investigated through a literature search.

qRT-PCR for microRNAs

Selected top-ranked miRNAs were quantified using TaqMan MicroRNA Assays Kits (Applied Biosystems, Foster City, CA) as described previously [38], in the validation set with 80 patients, consisting of 40 patients in Group 1 (progressive cases) and 40 patients in Group 2 (non-progressive cases). Input RNA was reverse transcribed using the TaqMan miR Reverse Transcription Kit and miR-specific stem-loop primers for the selected four miRs in a small-scale RT reaction. For quantification, diluted RT product was combined with PCR assay reagents and real-time PCR carried out on an ABI7900HT thermocycler. The endogenous control (RNU48) showed even expression levels between the two groups. Test samples were assayed in duplicate with the laboratory blinded to survival status and with 5% triplication after relabeling. The coefficient of variation was calculated and values <5% was considered acceptable. Data was analyzed with SDS Relative Quantification Software version 2.2.2 (Applied BioSystems) to determine the threshold cycle (Ct). The fold change was determined by the $2^{-\Delta\Delta Ct}$ method [39]. A two-sample *t* test was used to compare the normalized expression levels between the two groups.

microRNA target prediction

MicroRNA target analysis was performed by the TargetScanHuman release 7.0 (www.targetscan.org) to compare potential target genes affected by the top four miRNAs with the NCBI reference (human genome build 36). The biologic processes, molecular functions, cellular components, and protein classes of these genes were examined utilizing the Panther web site (Supplemental Table 1).

Statistical analysis

Univariate logistic regression models were first used to test if the selected miRNAs were independent factors of 5-year progression status. A multiple logistic regression model was then used to build a prediction model. Histologic findings were incorporated into the model by assigning a numeric value of '1' to diagnoses of hyperkeratosis, '2' to epithelial hyperplasia with hyperkeratosis, '3' to epithelial atypia and '4' to mild epithelial dysplasia. Using the multiple logistic regression model, we constructed receiver-operating characteristic (ROC) curves and calculated the area under the curve (AUC) and the sensitivity and specificity with 0.5 predicted probability of progression to OSCC as the cutoff point. $P<0.05$ was considered statistically significant. Statistical analyses were conducted using SAS 9.3 (SAS Institute).

Results

Deep sequencing, mapping of miRNAs and selection of predictive miRNAs

Age, gender, histopathologic diagnosis and location of lesion are listed in **Table 1** for the 20

Table 1. Demographic and clinicopathologic characteristics of training and validation sets

	Training Set (n=20)		Validation Set (n=77)		
	Non-progressive Group (n=10)	Progressive Group (n=10)	Non-progressive Group (n=39)	Progressive Group (n=38)	*P*-value*
Age mean (SD)	58.9 (11.1)	63.2 (23.2)	59.62 (12.6)	66.5 (17.5)	0.050
Gender % male	40%	50%	33%	35%	0.813
Histologic Diagnosis					0.058
Hyperkeratosis	1 (10%)	2 (20%)	17 (43%)	11 (28%)	
Epithelial hyperplasia and hyperkeratosis	3 (30%)	3 (30%)	9 (23%)	6 (15%)	
Epithelial atypia of the basal cell layer	4 (40%)	2 (20%)	10 (25%)	9 (23%)	
Mild epithelial dysplasia	2 (20%)	3 (30%)	4 (10%)	14 (35%)	
Location					0.256
High risk site (tongue, floor of mouth)	7 (70%)	7 (70%)	21 (53%)	26 (65%)	
Low risk site (buccal mucosa, vestibule, gingiva, palate, lip mucosa)	3 (30%)	3 (30%)	19 (47.5%)	14 (35%)	

*Two-sample t-test was used to compare means between non-progressive and progressive groups and Pearson's Chi-square test was used to compare proportions between non-progressive and progressive groups within the validation set (n=80).

Table 2. Mean expression levels of the 4 selected miRNAs and the fold change between the two groups by deep sequencing in the testing set with 20 patients and qRT-PCR in the validation set with 77 patients

miRNAs	Methods	Non-progressive Group mean (SD)	Progressive Group mean (SD)	Fold-change	*P*-value**
miR-129-2-3p	Deep Seq (n=20)	13.28 (9.67)	2.75 (2.78)	0.21	0.0001
	qRT-PCR (n=77)	6.81 (1.98)	6.21 (1.42)	1.36	0.361
miR-204-5p	Deep Seq (n=20)	848.03 (1117)	239.16 (537.9)	0.28	0.007
	qRT-PCR (n=77)	3.55 (1.83)	2.88 (1.70)	1.60	0.250
miR-208b-3p	Deep Seq (n=20)	51.86 (58.46)	132.73 (169.1)	2.56	0.003
	qRT-PCR (n=77)	6.24 (1.89)	5.43 (1.99)	1.73	0.049
miR-3065-5p	Deep Seq (n=20)	14.03 (9.36)	5.01 (3.47)	0.36	0.005
	qRT-PCR (n=77)	6.23 (1.78)	6.61 (1.79)	0.86	0.564

*Mean count for deep sequencing and mean normalized Ct value for qRT-PCR. **p-value is unadjusted for multiple comparisons and is based on two-sample t-test comparing the means of the normalized miRNA expression levels.

subjects in the training set and 80 subjects in the validation. For deep sequencing, the total number of reads obtained ranged from 3,229,855 to 39,216,964 for the 20 tissue samples. After removing adaptors and filtering out low quality tags, 973,424 - 31,446,445 clean reads were obtained (~30.14-80.19%). From these reads, the mapping rate to miRNA ranged from 0.89 to 12.71, and the rate to miRalligner, which distinguish -3p and -5p sequences, ranged from 1.48 to 14.75. The length distribution analysis revealed a peak at 22 nt, which is the size of most known microRNAs. For each sample, ~30 thousand to 11 million sequence reads mapped to the human genome were obtained, which included miRNA, rRNA, Mt_rRNA, snoRNA, snRNA, tRNA. A total of 1755 miRNAs were detected in each sample using the miRaligner and the number of reads of each miRNA ranged from 0 to 138,335.

A total of 4 miRNAs (miRs-129-2-3p, 204-5p, 208b-3p and 3065-5p) were identified to be differentially expressed in the progressive group (Group 1) compared to non-progressive group (Group 2) with at least 1.25 log2 fold change with unadjusted $p<0.05$ and to be expressed in close to (at least 80%) if not all samples. The mean expression levels and the fold changes of the 4 selected miRNAs are listed in **Table 2**. There were no miRNAs with false discovery rate adjusted $p<0.05$. One of the miRNAs (miR-208b-3p) was overexpressed in the progressive group and 3 were underexpressed (miRs-129-2-3p, 204-5p and 3065-5p) in the progressive group as shown in **Figure 1**.

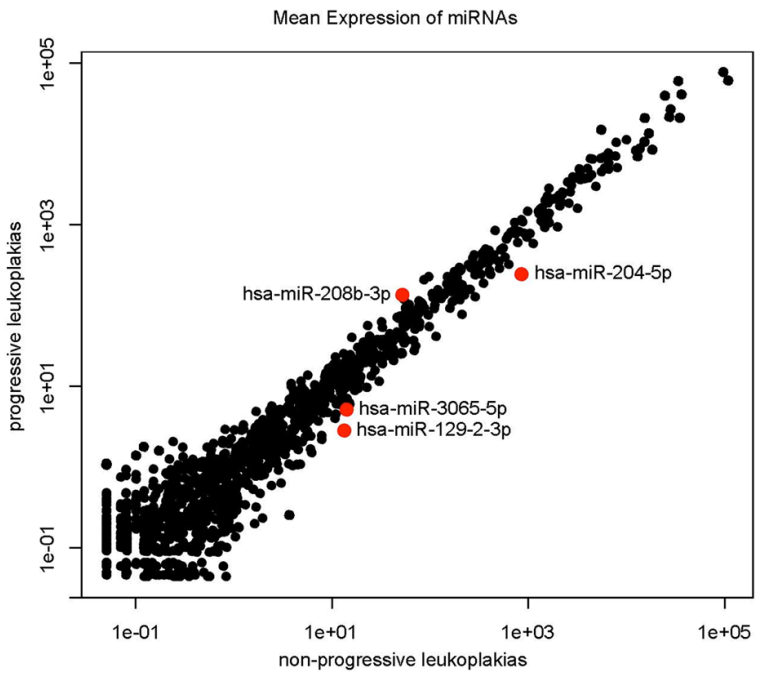

Figure 1. Scatterplot demonstrating overexpression of miRNA-208-3p and underexpression of miRNAs-204-5p, 3065-5p and 129-2-3p in the Progressive Group (Group 1) compared to the Non-Progressive Group (Group 2) in the training set (n=20) as measured by deep sequencing.

Putative target genes for the selected microRNAs

The biologic characteristics of 8, 74, 59, and 96 conserved genes potentially targeted by miRNAs-129-2-3p, 204-5p, 208b-3p and 3065-5p, respectively, were evaluated using TargetScanHuman release 7.0 (www.targetscan.org) and NCBI database (human genome build 36). Enriched genes were associated with multiple biologic pathways including pathways involving regulation of transcription, VEGF activity, Activin, IL-6, Notch signaling pathways, p53 regulation, Wnt/beta-catenin signaling cascade, cell proliferation, cell differentiation, cellular adhesion, apoptosis, cell migration, tumor suppression, and oncogenic signaling (Supplemental Table 1).

Evaluation of the top four miRNA expression profiles by RT-qPCR in the validation set

To confirm overexpression of miR-208b-3p and underexpression of miRs-204-5p, 129-2-3p and 3065-5p, these four miRNAs were quantified using qRT-PCR in the validation set with an additional 40 patients that progressed to OSCC (Group 1) and 40 patients that did not progress to OSCC (Group 2).

Only two of the four selected microRNAs showed changes in expression level consistent with that of the deep sequencing; miR-208b-3p was overexpressed and miR-3065-5p was underexpressed in the progressive group (**Table 2**). This finding is consistent with the putative role of these miRNAs; miR-208b-3p with oncogenic function, and miR-3065-5p with tumor suppressor role (**Table 3**). Although statistically insignificant, the other two miRNAs, miR-129-2-3p and miR-204-5p, showed overexpression in the progressive group (significant underexpression of these two microRNAs were demonstrated in the deep sequencing analysis), despite their presumptive role as tumor suppressors.

Multiple logistic regression analysis was performed to construct a predictive model of cancer progression using the top 4 miRNAs together with age, gender, histologic diagnosis, and location of the lesion. After model selection procedures, the final predictive model included the four miRNA marker panel (miRs-208b-3p, 3065-5p, 129-2-3p and 204-5p), age and histologic diagnosis. In the final model the endogenous control was undetectable in three samples. Since normalization could not be performed for miRNA expression level in those samples, they were eliminated (n=77). The AUC of the ROC curve of the final model ($p<0.004$ for the 6 predictors combined) was 0.792, with a sensitivity of 76.9% and specificity of 73.7% (**Figure 2**). Interestingly, this predictive value was stronger than that histology alone (AUC of the ROC curve = 0.646 with sensitivity of 57.5% and specificity of 65.0%), which is utilized in current clinical practice as a sole source of oral cancer progression predictor. As an exploratory study, we performed Maximum Likelihood

Table 3. Univariate analysis of microRNAs associated with cancer progression in the validation set (n=77)

miRNAs	OR (95% CI)	*P*-value[a]	Chromosomal Location	Putative function	Expression in Progressive Group
miR-129-2-3p	1.785 (1.686-1.694)	0.361	11p11.2	Tumor suppressor	increased
miR-204-5p	2.184 (1.78-2.243)	0.250	9q21.12	Tumor suppressor	increased
miR-208b-3p	1.472 (1.236-2.061)	0.049	14q11.2	Oncogenic	increased
miR-3065-5p	1.605 (1.569-1.469)	0.564	17q25.3	Tumor suppressor	decreased

[a]$p<0.05$ is considered to be significant.

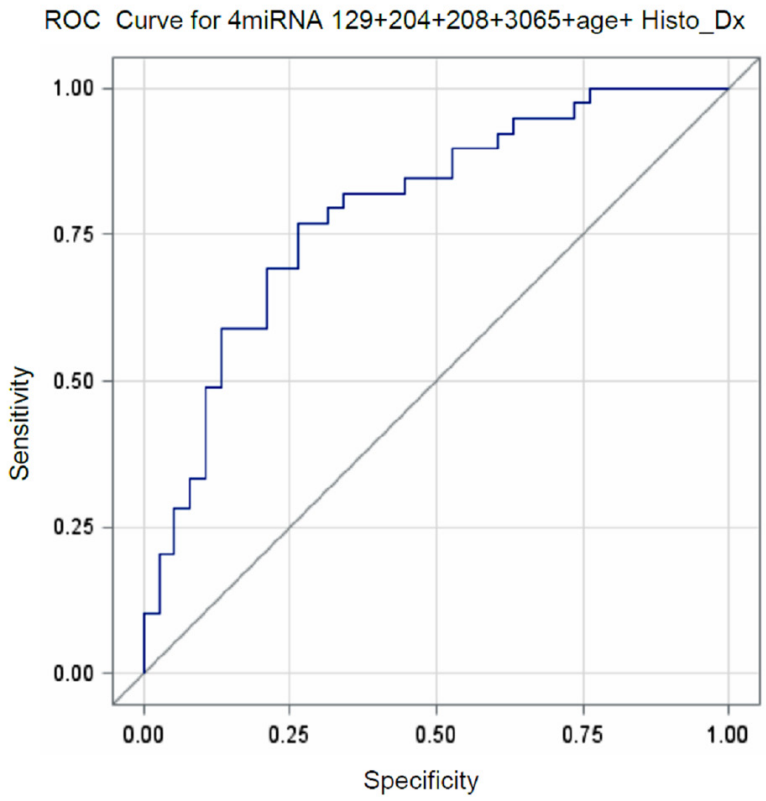

Figure 2. ROC curve for miRNA levels combined with age and histologic diagnosis that can identify progressive oral precursor lesions in the validation set (n=77).

Estimates analysis for the 6 predictors included in the final model in an attempt to generate a meaningful risk score formula that can be utilized in a clinical setting. The following formula was constructed from the analysis:

Score of cancer progression = -1.837 – (0.147 x miR-129b-3p) – (0.320 x miR-204-5p) – (0.312 x miR-208b-3p) + (0.348 x miR-3065-5p) + (0.048 x age) + (0.160 x histologic diagnosis level2) + (0.259 x histologic diagnosis level3) + (1.803 x histologic diagnosis level4)

Risk of cancer progression = exp(Score)/{1+exp(Score)}

From this analysis, -1.837 is the intercept, and the higher risk of cancer progression was associated with over expression of miR-3065-5p and underexpression of miRs-129-2-3p, 204-5p and 208b-3p, as well as older age. For the histologic diagnosis, increase in assigned number [1=epithelial hyperplasia, 2= epithelial hyperplasia with hyperkeratosis, 3=cellular atypia, 4=epithelial dysplasia] was associated with increased risk of cancer progression. When the risk score of cancer progression was calculated for all subjects in the validation set (n=77), the mean risk score in the Progressive group was 0.773 (63% likelihood) versus those in the Non-Progressive group -0.680 (38% likelihood). Using the mean (-0.282) as a cut-off point, 31 of the 39 progressive cases (80%) were correctly identified as high risk for cancer progression. In the non-progressive group, 24 of the 38 cases (63%) were properly identified as minimal risk of cancer progression.

Discussion

The deep sequencing analyses in the training set of 20 subjects revealed the overexpression of miR-208b-3p and underexpression of miR-204-5p, 129-2-3p and 3065-5p in the non- and low-grade dysplastic lesions that ultimately pro-

gressed to cancer. Indeed, the biologic characteristics of conserved genes potentially targeted by these miRNAs are associated with multiple critical biological processes related to carcinogenesis and tumor suppression. miR-208-3p has a presumptive oncogenic role and its overexpression has been linked to increased cellular proliferation, cell cycle progression, and tumorigenicity in hepatic tissue [40] and esophageal squamous cells [41]. *SOX6* tumor suppressor gene is directly targeted by miR-208, and hence, overexpression of miR-208 leads to downregulation of SOX6 protein, which results in downregulation of p21, upregulation of cyclin D1, and deregulation Rb via phosphorylation [41].

miR-3065-5p is a relatively novel microRNA with a limited number of reports. Expression of miR-3065-5p was much lower in metastatic prostate cancer cells compared to those that did not metastasize. Hence, miR-3065-5p is thought to have a tumor-suppressive effect, contributing to reduction of cell migration and tissue invasion [42]. miR-129-2-3p is a negative regulator of *SOX4*, an oncogene [4]. Hypermethylation of its promoter region and subsequent gene silencing of miR-129-2-3p leads to overexpression of SOX4 and Cdk6, an event reported in tumorigenesis of endometrium [43], stomach [44], and liver [45]. Epigenetic alteration, such as hypermethylation of the promoter region, has been reported to be an early event in oral carcinogenesis [46]. Similarly, miR-204-5p is thought to have a tumor suppressor function. Its downregulation via hypermethylation is associated with increased metastasis and decreased overall survival in breast cancer [47], endometrial carcinoma [48], and colorectal cancer [49]. Researchers have shown that miR-204 is located at the genomic imbalanced 9q21.1-22.3 locus associated with genetic predisposition for head and neck cancer [50].

In the validation cohort (n=77), miR-208-5p expression levels showed a significant increase in the progressive group compared to that of the non-progressive group, concordant with the deep sequencing data. There was underexpression of miR-3065-5p in the progressive group, as observed in the deep sequencing analysis, although it was not statically significant. In contrast to the deep sequencing data, overexpression of miR-129-2-3p and miR-204-5p were detected in the progressive group in the validation cohort. Similar discrepancies between the deep sequencing and qRT-PCR data have been reported by Schee et al [50]. In theory, lack of specificity of the qRT-PCR assay for miR-129-2-3p and miR-204-5p might explain such discrepancies [50]. Alternatively, the small sample size in our training set may have led to false positives.

Despite the discrepancies in expression profile of microRNAs between the training and validation cohorts, when the four miRNAs were combined as a panel together with age and histologic diagnosis, the AUC of the ROC curve was 0.792, with a sensitivity of 76.9% and specificity of 73.7% ($p<0.004$) to identify non- and low-grade dysplastic lesions that progress to cancer. As previously mentioned, histopathologic examination alone has limitations in identifying leukoplakic lesion that will progress of OSCC, especially if the leukoplakia lacks significant atypical cellular features at the time of initial histopathologic assessment. Limitations of histopathology were also observed in our study; the predictive value of the histopathology alone was 0.645 (AUC of the ROC curve). If the four miRNA marker panel is added to the histopathologic examination, which can be obtained via qRT-PCR with relative ease, together with patient's age, the predictive power to identify non- and low-grade dysplastic lesions at higher risk for malignant transformation will increase from 0.645 to 0.792.

For clinical practicality, we conducted an exploratory study to assess the feasibility of developing a parsimonious risk score formula from our final model. The main purpose of risk score formula is to translate miRNA expression levels assessed by qRT-PCR at the time of initial biopsy into a score that reflects the patient's risk of cancer progression. Using the risk score formula, 31 of the 39 progressive cases (80%) were accurately identified as high risk for cancer progression. Conversely, 24 of the 38 non-progressive cases (63%) were properly identified as minimal risk of cancer progression. Those cases identified as high risk would have been treated aggressively initially which would potentially have prevented carcinoma development. The 63% of the low cancer progression risk group could have potentially been spared from

unnecessary surgical excision. Further investigation is being planned to determine if our model can be utilized as a predictive modality for progression in early oral leukoplakia. This study is limited by small sample size. Also there are limitations in terms of generalizability as this is a single institutional study. However, our study is unique in that it assesses genome-wide miRNA expression profiles to identify a panel of miRNAs that may be utilized as a cancer progression predictive modality in non- and low-grade dysplastic lesions. We have plans to perform an internal validation and test the repeatability of the final risk score model in a larger set of samples. We will then test the model in a multicenter setting prior to conducting a large-scale prospective study. If validated, we will have obtained a useful predictive modality that can be applied in the clinic, which will guide appropriate management of patients at risk of developing oral cancer.

Acknowledgements

We thank Irina Gurvich for the technical assistance. This work was supported by the NIH/NIDCR R03-DE-023820 (to E. Philipone) and P30ES009089.

Disclosure of conflict of interest

None.

Address correspondence to: Elizabeth Philipone, Columbia University Medical Center, 630 West 168th Street, PH15W-1562, New York, NY 10032, USA. Tel: 212-305-3295; Fax: 212-305-5958; E-mail: ep2464@cumc.columbia.edu

References

[1] Parkin DM, Pisani P and Ferlay J. Global cancer statistics. CA Cancer J Clin 1999; 49: 33-64, 31.

[2] Kademani D. Oral cancer. Mayo Clin Proc 2007; 82: 878-887.

[3] Rodu B and Cole P. Oral cavity and pharynx-throat cancer in the United States, 1973-2003. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2007; 104: 653-658.

[4] Sciubba JJ. Oral cancer. The importance of early diagnosis and treatment. Am J Clin Dermatol 2001; 2: 239-251.

[5] Lee JJ, Hong WK, Hittelman WN, Mao L, Lotan R, Shin DM, Benner SE, Xu XC, Lee JS, Papadimitrakopoulou VM, Geyer C, Perez C, Martin JW, El-Naggar AK and Lippman SM. Predicting cancer development in oral leukoplakia: ten years of translational research. Clin Cancer Res 2000; 6: 1702-1710.

[6] Bouquot JE and Gorlin RJ. Leukoplakia, lichen planus, and other oral keratoses in 23,616 white Americans over the age of 35 years. Oral Surg Oral Med Oral Pathol 1986; 61: 373-381.

[7] Bouquot JE. Oral leukoplakia and erythroplakia: a review and update. Pract Periodontics Aesthet Dent 1994; 6: 9-17; quiz 19.

[8] Schepman KP, van der Meij EH, Smeele LE and van der Waal I. Malignant transformation of oral leukoplakia: a follow-up study of a hospital-based population of 166 patients with oral leukoplakia from The Netherlands. Oral Oncol 1998; 34: 270-275.

[9] Silverman S Jr, Gorsky M and Lozada F. Oral leukoplakia and malignant transformation. A follow-up study of 257 patients. Cancer 1984; 53: 563-568.

[10] Cowan CG, Gregg TA, Napier SS, McKenna SM and Kee F. Potentially malignant oral lesions in northern Ireland: a 20-year population-based perspective of malignant transformation. Oral Dis 2001; 7: 18-24.

[11] Mehanna HM, Rattay T, Smith J and McConkey CC. Treatment and follow-up of oral dysplasia - a systematic review and meta-analysis. Head Neck 2009; 31: 1600-1609.

[12] Speight PM. Update on oral epithelial dysplasia and progression to cancer. Head Neck Pathol 2007; 1: 61-66.

[13] Bouquot JE SP, Farthing PM. Epithelial dysplasia of the oral mucosa-diagnostic problems and prognostic features. Curr Diagn Pathol 2006; 12: 11-21.

[14] Waldron CA and Shafer WG. Leukoplakia revisited. A clinicopathologic study 3256 oral leukoplakias. Cancer 1975; 36: 1386-1392.

[15] Gupta PC, Mehta FS, Daftary DK, Pindborg JJ, Bhonsle RB, Jalnawalla PN, Sinor PN, Pitkar VK, Murti PR, Irani RR, Shah HT, Kadam PM, Iyer KS, Iyer HM, Hegde AK, Chandrashekar GK, Shiroff BC, Sahiar BE and Mehta MN. Incidence rates of oral cancer and natural history of oral precancerous lesions in a 10-year follow-up study of Indian villagers. Community Dent Oral Epidemiol 1980; 8: 283-333.

[16] Hsue SS, Wang WC, Chen CH, Lin CC, Chen YK and Lin LM. Malignant transformation in 1458 patients with potentially malignant oral mucosal disorders: a follow-up study based in a Taiwanese hospital. J Oral Pathol Med 2007; 36: 25-29.

[17] Nankivell P and Mehanna H. Oral dysplasia: biomarkers, treatment, and follow-up. Curr Oncol Rep 2011; 13: 145-152.

[18] Wu BH, Xiong XP, Jia J and Zhang WF. MicroRNAs: new actors in the oral cancer scene. Oral Oncol 2011; 47: 314-319.
[19] Bandres E, Bitarte N, Arias F, Agorreta J, Fortes P, Agirre X, Zarate R, Diaz-Gonzalez JA, Ramirez N, Sola JJ, Jimenez P, Rodriguez J and Garcia-Foncillas J. microRNA-451 regulates macrophage migration inhibitory factor production and proliferation of gastrointestinal cancer cells. Clin Cancer Res 2009; 15: 2281-2290.
[20] Kozaki K, Imoto I, Mogi S, Omura K and Inazawa J. Exploration of tumor-suppressive microRNAs silenced by DNA hypermethylation in oral cancer. Cancer Res 2008; 68: 2094-2105.
[21] Liu X, Jiang L, Wang A, Yu J, Shi F and Zhou X. MicroRNA-138 suppresses invasion and promotes apoptosis in head and neck squamous cell carcinoma cell lines. Cancer Lett 2009; 286: 217-222.
[22] Luo H, Zou J, Dong Z, Zeng Q, Wu D and Liu L. Up-regulated miR-17 promotes cell proliferation, tumour growth and cell cycle progression by targeting the RND3 tumour suppressor gene in colorectal carcinoma. Biochem J 2012; 442: 311-321.
[23] Feber A, Xi L, Luketich JD, Pennathur A, Landreneau RJ, Wu M, Swanson SJ, Godfrey TE and Litle VR. MicroRNA expression profiles of esophageal cancer. J Thorac Cardiovasc Surg 2008; 135: 255-260; discussion 260.
[24] Maru DM, Singh RR, Hannah C, Albarracin CT, Li YX, Abraham R, Romans AM, Yao H, Luthra MG, Anandasabapathy S, Swisher SG, Hofstetter WL, Rashid A and Luthra R. MicroRNA-196a is a potential marker of progression during Barrett's metaplasia-dysplasia-invasive adenocarcinoma sequence in esophagus. Am J Pathol 2009; 174: 1940-1948.
[25] Bansal A, Lee IH, Hong X, Anand V, Mathur SC, Gaddam S, Rastogi A, Wani SB, Gupta N, Visvanathan M, Sharma P and Christenson LK. Feasibility of mcroRNAs as biomarkers for Barrett's Esophagus progression: a pilot cross-sectional, phase 2 biomarker study. Am J Gastroenterol 2011; 106: 1055-1063.
[26] Pereira PM, Marques JP, Soares AR, Carreto L and Santos MA. MicroRNA expression variability in human cervical tissues. PLoS One 2010; 5: e11780.
[27] Mascaux C, Laes JF, Anthoine G, Haller A, Ninane V, Burny A and Sculier JP. Evolution of microRNA expression during human bronchial squamous carcinogenesis. Eur Respir J 2009; 33: 352-359.
[28] Zhang Z, Li Z, Gao C, Chen P, Chen J, Liu W, Xiao S and Lu H. miR-21 plays a pivotal role in gastric cancer pathogenesis and progression. Lab Invest 2008; 88: 1358-1366.
[29] Cervigne NK, Reis PP, Machado J, Sadikovic B, Bradley G, Galloni NN, Pintilie M, Jurisica I, Perez-Ordonez B, Gilbert R, Gullane P, Irish J and Kamel-Reid S. Identification of a microRNA signature associated with progression of leukoplakia to oral carcinoma. Hum Mol Genet 2009; 18: 4818-4829.
[30] Zhu G, He Y, Yang S, Chen B, Zhou M and Xu XJ. Identification of Gene and MicroRNA Signatures for Oral Cancer Developed from Oral Leukoplakia. Biomed Res Int 2015; 2015: 841956.
[31] Maimaiti A, Abudoukeremu K, Tie L, Pan Y and Li X. MicroRNA expression profiling and functional annotation analysis of their targets associated with the malignant transformation of oral leukoplakia. Gene 2015; 558: 271-277.
[32] De Sarkar N, Roy R, Mitra JK, Ghose S, Chakraborty A, Paul RR, Mukhopadhyay I and Roy B. A quest for miRNA bio-marker: a track back approach from gingivo buccal cancer to two different types of precancers. PLoS One 2014; 9: e104839.
[33] Roy R, De Sarkar N, Ghose S, Paul RR, Ray A, Mukhopadhyay I and Roy B. Association between risk of oral precancer and genetic variations in microRNA and related processing genes. J Biomed Sci 2014; 21: 48.
[34] Brito JA, Gomes CC, Guimaraes AL, Campos K and Gomez RS. Relationship between microRNA expression levels and histopathological features of dysplasia in oral leukoplakia. J Oral Pathol Med 2014; 43: 211-216.
[35] Yang Y, Li YX, Yang X, Jiang L, Zhou ZJ and Zhu YQ. Progress risk assessment of oral premalignant lesions with saliva miRNA analysis. BMC Cancer 2013; 13: 129.
[36] Xiao W, Bao ZX, Zhang CY, Zhang XY, Shi LJ, Zhou ZT and Jiang WW. Upregulation of miR-31* is negatively associated with recurrent/newly formed oral leukoplakia. PLoS One 2012; 7: e38648.
[37] Weng L, Wu X, Gao H, Mu B, Li X, Wang JH, Guo C, Jin JM, Chen Z, Covarrubias M, Yuan YC, Weiss LM and Wu H. MicroRNA profiling of clear cell renal cell carcinoma by whole-genome small RNA deep sequencing of paired frozen and formalin-fixed, paraffin-embedded tissue specimens. J Pathol 2010; 222: 41-51.
[38] Yoon AJ, Wang S, Shen J, Robine N, Philipone E, Oster MW, Nam A and Santella RM. Prognostic value of miR-375 and miR-214-3p in early stage oral squamous cell carcinoma. Am J Transl Res 2014; 6: 580-592.
[39] Catuogno S, Esposito CL, Quintavalle C, Cerchia L, Condorelli G and De Franciscis V. Recent Advance in Biosensors for microRNAs

Detection in Cancer. Cancers (Basel) 2011; 3: 1877-1898.
[40] Yu P, Wu D, You Y, Sun J, Lu L, Tan J and Bie P. miR-208-3p promotes hepatocellular carcinoma cell proliferation and invasion through regulating ARID2 expression. Exp Cell Res 2015; 336: 232-241.
[41] Li H, Zheng D, Zhang B, Liu L, Ou J, Chen W, Xiong S, Gu Y and Yang J. Mir-208 promotes cell proliferation by repressing SOX6 expression in human esophageal squamous cell carcinoma. J Transl Med 2014; 12: 196.
[42] Watahiki A, Wang Y, Morris J, Dennis K, O'Dwyer HM, Gleave M, Gout PW and Wang Y. MicroRNAs associated with metastatic prostate cancer. PLoS One 2011; 6: e24950.
[43] Huang YW, Liu JC, Deatherage DE, Luo J, Mutch DG, Goodfellow PJ, Miller DS and Huang TH. Epigenetic repression of microRNA-129-2 leads to overexpression of SOX4 oncogene in endometrial cancer. Cancer Res 2009; 69: 9038-9046.
[44] Yu X, Luo L, Wu Y, Yu X, Liu Y, Yu X, Zhao X, Zhang X, Cui L, Ye G, Le Y and Guo J. Gastric juice miR-129 as a potential biomarker for screening gastric cancer. Med Oncol 2013; 30: 365.
[45] Lu CY, Lin KY, Tien MT, Wu CT, Uen YH and Tseng TL. Frequent DNA methylation of MiR-129-2 and its potential clinical implication in hepatocellular carcinoma. Genes Chromosomes Cancer 2013; 52: 636-643.
[46] Ha PK and Califano JA. Promoter methylation and inactivation of tumour-suppressor genes in oral squamous-cell carcinoma. Lancet Oncol 2006; 7: 77-82.
[47] Li W, Jin X, Zhang Q, Zhang G, Deng X and Ma L. Decreased expression of miR-204 is associated with poor prognosis in patients with breast cancer. Int J Clin Exp Pathol 2014; 7: 3287-3292.
[48] Bao W, Wang HH, Tian FJ, He XY, Qiu MT, Wang JY, Zhang HJ, Wang LH and Wan XP. A TrkB-STAT3-miR-204-5p regulatory circuitry controls proliferation and invasion of endometrial carcinoma cells. Mol Cancer 2013; 12: 155.
[49] Yin Y, Zhang B, Wang W, Fei B, Quan C, Zhang J, Song M, Bian Z, Wang Q, Ni S, Hu Y, Mao Y, Zhou L, Wang Y, Yu J, Du X, Hua D and Huang Z. miR-204-5p inhibits proliferation and invasion and enhances chemotherapeutic sensitivity of colorectal cancer cells by downregulating RAB22A. Clin Cancer Res 2014; 20: 6187-6199.
[50] Schee K, Lorenz S, Worren MM, Gunther CC, Holden M, Hovig E, Fodstad O, Meza-Zepeda LA and Flatmark K. Deep Sequencing the MicroRNA Transcriptome in Colorectal Cancer. PLoS One 2013; 8: e66165.

Supplemental Table 1. Conserved Genes Potentially Targeted by miRNA's -129-2-3p, 204-5p, 208b-3p and 3065-5p

Gene ID[a]	Gene Name	PANTHER Family/Subfamily	PANTHER Protein Class
COLQ	Acetylcholinesterase collagenic tail peptide	ACETYLCHOLINESTERASE COLLAGENIC TAIL PEPTIDE (PTHR24023:SF6)	Transporter, surfactant, receptor, extracellular matrix structural protein, antibacterial response protein
REEP1	Receptor expression-enhancing protein 1	RECEPTOR EXPRESSION-ENHANCING PROTEIN 1 (PTHR12300:SF33)	Transporter, receptor
DDX5	Probable ATP-dependent RNA helicase DDX5	ATP-DEPENDENT RNA HELICASE DDX5-RELATED (PTHR24031:SF219)	RNA helicase
FZD8	Frizzled-8	FRIZZLED-8 (PTHR11309:SF89)	signaling molecule, G-protein coupled receptor
PURB	Transcriptional activator protein Pur-beta	TRANSCRIPTIONAL ACTIVATOR PROTEIN PUR-BETA (PTHR12611:SF4)	transcription factor, DNA binding protein
LCOR	Ligand-dependent corepressor	LIGAND-DEPENDENT COREPRESSOR (PTHR21545:SF14)	DNA-directed RNA polymerase
TMBIM6	Bax inhibitor 1	BAX INHIBITOR 1 (PTHR23291:SF34)	-
SCN4A	Sodium channel protein type 4 subunit alpha	SODIUM CHANNEL PROTEIN TYPE 4 SUBUNIT ALPHA (PTHR10037:SF223)	voltage-gated calcium channel
PACS2	Phosphofurin acidic cluster sorting protein 2	PHOSPHOFURIN ACIDIC CLUSTER SORTING PROTEIN 2 (PTHR13280:SF15)	-
ASB7	Ankyrin repeat and SOCS box protein 7	ANKYRIN REPEAT AND SOCS BOX PROTEIN 7 (PTHR24123:SF25)	-
PDCD4	Programmed cell death protein 4	PROGRAMMED CELL DEATH PROTEIN 4 (PTHR12626:SF3)	translation elongation factor
LRP6	Low-density lipoprotein receptor-related protein 6	LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN 6 (PTHR10529:SF109)	Receptor, extracellular matrix glycoprotein
PHF21B	PHD finger protein 21B	PHD FINGER PROTEIN 21B (PTHR24102:SF18)	transcription cofactor, ubiquitin-protein ligase
SCN2B	Sodium channel subunit beta-2	SODIUM CHANNEL SUBUNIT BETA-2 (PTHR13869:SF3)	voltage-gated sodium channel, voltage-gated ion channel, cell adhesion molecule
NAV2	Neuron navigator 2	NEURON NAVIGATOR 2 (PTHR12784:SF6)	-
VAV3	Guanine nucleotide exchange factor VAV3	GUANINE NUCLEOTIDE EXCHANGE FACTOR VAV3 (PTHR22826:SF97)	signaling molecule, guanyl-nucleotide exchange factor
SNRK	SNF-related serine/threonine-protein kinase	SNF-RELATED SERINE/THREONINE-PROTEIN KINASE (PTHR24343:SF201)	non-receptor serine/threonine protein kinase
GATA4	Transcription factor GATA-4	TRANSCRIPTION FACTOR GATA-4 (PTHR10071:SF154)	zinc finger transcription factor, nuclease
AAK1	AP2-associated protein kinase 1	AP2-ASSOCIATED PROTEIN KINASE 1 (PTHR22967:SF57)	non-receptor serine/threonine protein kinase
TBC1D22B	TBC1 domain family member 22B	TBC1 DOMAIN FAMILY MEMBER 22B (PTHR22957:SF321)	Hydrolase, G-protein modulator
FAM110B	Protein FAM110B	PROTEIN FAM110B (PTHR14758:SF2)	-
SOX6	Transcription factor SOX-6	TRANSCRIPTION FACTOR SOX-6 (PTHR10270:SF89)	HMG box transcription factor, nucleic acid binding
PAK2	Serine/threonine-protein kinase PAK 2	SERINE/THREONINE-PROTEIN KINASE PAK 2 (PTHR24361:SF281)	-
ZNF704	Zinc finger protein 704	ZINC FINGER PROTEIN 704 (PTHR13006:SF7)	transcription cofactor
TRAF4	TNF receptor-associated factor 4	TNF RECEPTOR-ASSOCIATED FACTOR 4 (PTHR10131:SF89)	signaling molecule
USP13	Ubiquitin carboxyl-terminal hydrolase 13	-	-
DYNLL2	Dynein light chain 2, cytoplasmic	DYNEIN LIGHT CHAIN 2, CYTOPLASMIC (PTHR11886:SF54)	enzyme modulator, microtubule family cytoskeletal protein
ANKRD52	Serine/threonine-protein phosphatase 6 regulatory ankyrin repeat subunit C	SERINE/THREONINE-PROTEIN PHOSPHATASE 6 REGULATORY ANKYRIN REPEAT SUBUNIT C (PTHR24158:SF17)	-
MMP16	Matrix metalloproteinase-16	MATRIX METALLOPROTEINASE-16 (PTHR10201:SF26)	-

DNAJC14	DnaJ homolog subfamily C member 14	DNAJ HOMOLOG SUBFAMILY C MEMBER 14 (PTHR24078:SF191)	-
NMU	Neuromedin-U	NEUROMEDIN-U (PTHR15390:SF0)	-
GLS	Glutaminase kidney isoform, mitochondrial	GLUTAMINASE KIDNEY ISOFORM, MITOCHONDRIAL (PTHR12544:SF30)	hydrolase
AMER2	APC membrane recruitment protein 2	APC MEMBRANE RECRUITMENT PROTEIN 2 (PTHR22237:SF1)	-
ZNF431	Zinc finger protein 431	ZINC FINGER PROTEIN 431 (PTHR24384:SF41)	KRAB box transcription factor
HMGB1	High mobility group protein B1	HIGH MOBILITY GROUP PROTEIN B1-RELATED (PTHR13711:SF164)	HMG box transcription factor, signaling molecule chromatin/ chromatin-binding protein
KLF12	Krueppel-like factor 12	-	-
SOX7	Transcription factor SOX-7	TRANSCRIPTION FACTOR SOX-7 (PTHR10270:SF210)	HMG box transcription factor, nucleic acid binding
LARP7	La-related protein 7	LA-RELATED PROTEIN 7 (PTHR22792:SF66)	ribonucleoprotein
CDC42	Cell division control protein 42 homolog	CELL DIVISION CONTROL PROTEIN 42 HOMOLOG (PTHR24072:SF136)	small GTPase
TMSB4Y	Thymosin beta-4, Y-chromosomal	THYMOSIN BETA-4, Y-CHROMOSOMAL (PTHR12021:SF8)	-
KIAA1024L	UPF0258 protein KIAA1024-like	UPF0258 PROTEIN KIAA1024-LIKE (PTHR31530:SF4)	-
UHMK1	Serine/threonine-protein kinase Kist	SERINE/THREONINE-PROTEIN KINASE KIST (PTHR23139:SF14)	mRNA splicing factor
NHLRC2	NHL repeat-containing protein 2	NHL REPEAT-CONTAINING PROTEIN 2 (PTHR13833:SF49)	-
TCF4	Transcription factor 4	TRANSCRIPTION FACTOR 4 (PTHR11793:SF10)	basic helix-loop-helix transcription factor, nucleic acid binding
TM2D1	TM2 domain-containing protein 1	TM2 DOMAIN-CONTAINING PROTEIN 1 (PTHR21016:SF1)	-
ZNF66	Putative zinc finger protein 66	ZINC FINGER PROTEIN 66-RELATED (PTHR24384:SF105)	KRAB box transcription factor
TPD52L3	Tumor protein D55	TUMOR PROTEIN D55 (PTHR19307:SF5)	-
XKR4	XK-related protein 4	XK-RELATED PROTEIN 4 (PTHR32129:SF14)	-
KIAA0101	PCNA-associated factor	PCNA-ASSOCIATED FACTOR (PTHR15679:SF8)	-
TRUB1	Probable tRNA pseudouridine synthase 1	TRNA PSEUDOURIDINE SYNTHASE 1-RELATED (PTHR13767:SF2)	nucleic acid binding, lyase isomerase
PCDH11X	Protocadherin-11 X-linked	PROTOCADHERIN-11 X-LINKED-RELATED (PTHR24027:SF15)	cadherin
LYSMD3	LysM and putative peptidoglycan-binding domain-containing protein 3	AND PUTATIVE PEPTIDOGLYCAN-BINDING DOMAIN-CONTAINING PROTEIN 3-RELATED (PTHR20932:SF5)	-
HERC5	E3 ISG15--protein ligase HERC5	E3 ISG15--PROTEIN LIGASE HERC5 (PTHR11254:SF344)	ubiquitin-protein ligase
ATF7	Cyclic AMP-dependent transcription factor ATF-7	CYCLIC AMP-DEPENDENT TRANSCRIPTION FACTOR ATF-7 (PTHR19304:SF25)	transcription factor, nucleic acid binding
CDK6	Cyclin-dependent kinase 6	CYCLIN-DEPENDENT KINASE 6 (PTHR24056:SF130)	non-receptor serine/threonine protein kinase
FAM46A	Protein FAM46A	PROTEIN FAM46A (PTHR12974:SF25)	-
XPO4	Exportin-4	EXPORTIN-4 (PTHR12596:SF1)	transfer/carrier protein
ZBTB34	Zinc finger and BTB domain-containing protein 34	ZINC FINGER AND BTB DOMAIN-CONTAINING PROTEIN 34 (PTHR24375:SF44)	KRAB box transcription factor
CLDN10	Claudin-10	CLAUDIN-10 (PTHR12002:SF13)	tight junction
WRB	Tail-anchored protein insertion receptor WRB	TAIL-ANCHORED PROTEIN INSERTION RECEPTOR WRB (PTHR11760:SF21)	ribosomal protein
NUDT15	Probable 8-oxo-dGTP diphosphatase NUDT15	8-OXO-DGTP DIPHOSPHATASE NUDT15-RELATED (PTHR16099:SF5)	-
HOPX	Homeodomain-only protein	HOMEODOMAIN-ONLY PROTEIN (PTHR21408:SF1)	-
PMP2	Myelin P2 protein	MYELIN P2 PROTEIN (PTHR11955:SF64)	-
SH2D6	SH2 domain-containing protein 6	SH2 DOMAIN-CONTAINING PROTEIN 6 (PTHR14098:SF16)	signaling molecule
DFFB	DNA fragmentation factor subunit beta	DNAATION FACTOR-RELATED PROTEIN 4 (PTHR13067:SF2)	nuclease

C1orf192	UPF0740 protein C1orf192	UPF0740 PROTEIN C1ORF192 (PTHR34639:SF1)	-
MRPL16	39S ribosomal protein L16, mitochondrial	39S RIBOSOMAL PROTEIN L16, MITOCHONDRIAL (PTHR12220:SF13)	ribosomal protein
CSRNP3	Cysteine/serine-rich nuclear protein 3	CYSTEINE/SERINE-RICH NUCLEAR PROTEIN 3 (PTHR13580:SF13)	-
EEF1E1	Eukaryotic translation elongation factor 1 epsilon-1	EUKARYOTIC TRANSLATION ELONGATION FACTOR 1 EPSILON-1-RELATED (PTHR11260:SF249)	Transferase, signaling molecule, reductase translation elongation factor, epimerase/racemase, cytoskeletal protein
HIGD1A	HIG1 domain family member 1A, mitochondrial	HIG1 DOMAIN FAMILY MEMBER 1A, MITOCHONDRIAL (PTHR12297:SF5)	-
FBXO30	F-box only protein 30	F-BOX ONLY PROTEIN 30 (PTHR15933:SF13)	-
EIF1AX	Eukaryotic translation initiation factor 1A, X-chromosomal	EUKARYOTIC TRANSLATION INITIATION FACTOR 1A, X-CHROMOSOMAL (PTHR21668:SF4)	translation initiation factor
TMEM170B	Transmembrane protein 170B	TRANSMEMBRANE PROTEIN 170B (PTHR22779:SF4)	-
CCDC50	Coiled-coil domain-containing protein 50	COILED-COIL DOMAIN-CONTAINING PROTEIN 50 (PTHR22115:SF1)	-
CBWD1	COBW domain-containing protein 1	COBW DOMAIN-CONTAINING PROTEIN 1-RELATED (PTHR13748:SF38)	-
DUSP23	Dual specificity protein phosphatase 23	DUAL SPECIFICITY PROTEIN PHOSPHATASE 23 (PTHR23339:SF26)	protein phosphatase
CNOT6	CCR4-NOT transcription complex subunit 6	CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 6 (PTHR12121:SF33)	Exoribonuclease, nuclease
CCL2	C-C motif chemokine 2	C-C MOTIF CHEMOKINE 2 (PTHR12015:SF98)	chemokine
DUSP18	Dual specificity protein phosphatase 18	DUAL SPECIFICITY PROTEIN PHOSPHATASE 18 (PTHR10159:SF313)	-
NPY5R	Neuropeptide Y receptor type 5	NEUROPEPTIDE Y RECEPTOR TYPE 5 (PTHR24242:SF217)	G-protein coupled receptor
PGAP1	GPI inositol-deacylase	GPI INOSITOL-DEACYLASE (PTHR15495:SF15)	transporter
PPM1A	Protein phosphatase 1A	PROTEIN PHOSPHATASE 1A (PTHR13832:SF262)	protein phosphatase, kinase inhibitor
SPIN4	Spindlin-4	SPINDLIN-4 (PTHR10405:SF9)	-
NFIA	Nuclear factor 1 A-type	NUCLEAR FACTOR 1 A-TYPE (PTHR11492:SF6)	transcription factor, nucleic acid binding
ABCA6	ATP-binding cassette sub-family A member 6	ATP-BINDING CASSETTE SUB-FAMILY A MEMBER 6 (PTHR19229:SF13)	ATP-binding cassette (ABC) transporter
CD9	CD9 antigen	CD9 ANTIGEN (PTHR19282:SF163)	membrane-bound signaling molecule, receptor, cell adhesion molecule
UBL3	Ubiquitin-like protein 3	UBIQUITIN-LIKE PROTEIN 3 (PTHR13169:SF0)	ubiquitin-protein ligase
QKI	Protein quaking	PROTEIN QUAKING (PTHR11208:SF54)	transcription cofactor, mRNA splicing factor
CCDC73	Coiled-coil domain-containing protein 73	COILED-COIL DOMAIN-CONTAINING PROTEIN 73 (PTHR28660:SF1)	-
CSTF2	Cleavage stimulation factor subunit 2	CLEAVAGE STIMULATION FACTOR SUBUNIT 2 (PTHR23139:SF57)	mRNA splicing factor
ZNF460	Zinc finger protein 460	ZINC FINGER PROTEIN 460 (PTHR24384:SF15)	KRAB box transcription factor
SRSF7	Serine/arginine-rich splicing factor 7	-	-
SPAG11A	Sperm-associated antigen 11A	SPERM-ASSOCIATED ANTIGEN 11A (PTHR14081:SF1)	-
GABRA4	Gamma-aminobutyric acid receptor subunit alpha-4	GAMMA-AMINOBUTYRIC ACID RECEPTOR SUBUNIT ALPHA-4 (PTHR18945:SF393)	GABA receptor, acetylcholine receptor, GABA receptor
C1QTNF1-AS1	Protein C1QTNF1-AS1	-	-
POU2F1	POU domain, class 2, transcription factor 1	POU DOMAIN, CLASS 2, TRANSCRIPTION FACTOR 1 (PTHR11636:SF47)	homeobox transcription factor
SNX24	Sorting nexin-24	SORTING NEXIN-24 (PTHR15813:SF10)	-
SETD9	SET domain-containing protein 9	SET DOMAIN-CONTAINING PROTEIN 9 (PTHR33524:SF1)	-
RFX7	DNA-binding protein RFX7	DNA-BINDING PROTEIN RFX7 (PTHR12619:SF2)	transcription cofactor, nuclease

INO80D	INO80 complex subunit D	INO80 COMPLEX SUBUNIT D (PTHR16198:SF2)	-
PAFAH1B2	Platelet-activating factor acetylhydrolase IB subunit beta	PLATELET-ACTIVATING FACTOR ACETYLHYDROLASE IB SUBUNIT BETA (PTHR11852:SF1)	-
STIM2	Stromal interaction molecule 2	STROMAL INTERACTION MOLECULE 2 (PTHR15136:SF2)	-
UBE2E2	Ubiquitin-conjugating enzyme E2 E2	UBIQUITIN-CONJUGATING ENZYME E2 E2 (PTHR24068:SF62)	ligase
ZNF264	Zinc finger protein 264	ZINC FINGER PROTEIN 264 (PTHR24384:SF96)	KRAB box transcription factor
PCDH11Y	Protocadherin-11 Y-linked	PROTOCADHERIN-11 X-LINKED-RELATED (PTHR24027:SF15)	cadherin
ST8SIA3	Sia-alpha-2,3-Gal-beta-1,4-GlcNAc-R:alpha 2,8-sialyltransferase	SIA-ALPHA-2,3-GAL-BETA-1,4-GLCNAC-R:ALPHA 2,8-SIALYLTRANSFERASE (PTHR11987:SF36)	-
RBPJ	Recombining binding protein suppressor of hairless	RECOMBINING BINDING PROTEIN SUPPRESSOR OF HAIRLESS (PTHR10665:SF3)	transcription factor, nucleic acid binding
RBMS3	RNA-binding motif, single-stranded-interacting protein 3	RNA-BINDING MOTIF, SINGLE-STRANDED-INTERACTING PROTEIN 3 (PTHR24012:SF473)	-
ZNF573	Zinc finger protein 573	ZINC FINGER PROTEIN 573 (PTHR24377:SF159)	-
FSD1L	FSD1-like protein	FSD1-LIKE PROTEIN (PTHR24099:SF8)	-
RBM7	RNA-binding protein 7	RNA-BINDING PROTEIN 7 (PTHR13798:SF4)	nucleic acid binding
ARRDC3	Arrestin domain-containing protein 3	ARRESTIN DOMAIN-CONTAINING PROTEIN 3 (PTHR11188:SF49)	-
RNASE6	Ribonuclease K6	RIBONUCLEASE K6 (PTHR11437:SF4)	Endoribonuclease, nuclease, hydrolase, enzyme modulator
DERL2	Derlin-2	DERLIN-2 (PTHR11009:SF5)	receptor
ZNF736	Zinc finger protein 736	ZINC FINGER PROTEIN 736-RELATED (PTHR24384:SF98)	KRAB box transcription factor
UBE2W	Ubiquitin-conjugating enzyme E2 W	UBIQUITIN-CONJUGATING ENZYME E2 W (PTHR24067:SF108)	-
SLC25A31	ADP/ATP translocase 4	ADP/ATP TRANSLOCASE 4 (PTHR24089:SF381)	amino acid transporter, mitochondrial carrier protein, transfer/carrier protein, ribosomal protein, calmodulin
UFM1	Ubiquitin-fold modifier 1	UBIQUITIN-FOLD MODIFIER 1 (PTHR15825:SF0)	-
GPRASP2	G-protein coupled receptor-associated sorting protein 2	G-PROTEIN COUPLED RECEPTOR-ASSOCIATED SORTING PROTEIN 2 (PTHR15712:SF10)	-
C16orf72	UPF0472 protein C16orf72	UPF0472 PROTEIN C16ORF72 (PTHR31624:SF3)	-
DNM2	Dynamin-2	DYNAMIN-2 (PTHR11566:SF23)	Hydrolase, small GTPase, microtubule family cytoskeletal protein
PIK3CB	Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit beta isoform	PHOSPHATIDYLINOSITOL 4,5-BISPHOSPHATE 3-KINASE CATALYTIC SUBUNIT BETA ISOFORM (PTHR10048:SF33)	kinase
ANKRD13A	Ankyrin repeat domain-containing protein 13A	ANKYRIN REPEAT DOMAIN-CONTAINING PROTEIN 13A (PTHR12447:SF4)	-
NR3C2	Mineralocorticoid receptor	MINERALOCORTICOID RECEPTOR (PTHR24084:SF22)	nuclear hormone receptor, receptor, nucleic acid binding
AP2A2	AP-2 complex subunit alpha-2	AP-2 COMPLEX SUBUNIT ALPHA-2 (PTHR22780:SF24)	transmembrane receptor regulatory/adaptor protein
MMGT1	Membrane magnesium transporter 1	MEMBRANE MAGNESIUM TRANSPORTER 1 (PTHR21181:SF7)	-
SHC1	SHC-transforming protein 1	SHC-TRANSFORMING PROTEIN 1 (PTHR10337:SF2)	signaling molecule
FAM175B	BRISC complex subunit Abro1	BRISC COMPLEX SUBUNIT ABRO1 (PTHR31728:SF3)	-
TMEM64	Transmembrane protein 64	TRANSMEMBRANE PROTEIN 64 (PTHR12677:SF24)	-
ARX	Homeobox protein ARX	HOMEOBOX PROTEIN ARX (PTHR24329:SF356)	homeobox transcription factor, DNA binding protein
CHD5	Chromodomain-helicase-DNA-binding protein 5	CHROMODOMAIN-HELICASE-DNA-BINDING PROTEIN 5 (PTHR10799:SF583)	DNA helicase, helicase
GPT2	Alanine aminotransferase 2	ALANINE AMINOTRANSFERASE 2 (PTHR11751:SF311)	transaminase

VWA8	von Willebrand factor A domain-containing protein 8	VON WILLEBRAND FACTOR A DOMAIN-CONTAINING PROTEIN 8 (PTHR21610:SF9)	-
TPPP	Tubulin polymerization-promoting protein	TUBULIN POLYMERIZATION-PROMOTING PROTEIN (PTHR12932:SF18)	non-motor microtubule binding protein
SLC25A35	Solute carrier family 25 member 35	SOLUTE CARRIER FAMILY 25 MEMBER 35 (PTHR24089:SF274)	amino acid transporter, mitochondrial carrier protein, transfer/carrier protein, ribosomal protein, calmodulin
DNAJC13	DnaJ homolog subfamily C member 13	DNAJ HOMOLOG SUBFAMILY C MEMBER 13 (PTHR24078:SF25)	-
AKAP1	A-kinase anchor protein 1, mitochondrial	A-KINASE ANCHOR PROTEIN 1, MITOCHONDRIAL (PTHR22948:SF18)	signaling molecule, nuclease
HSPH1	Heat shock protein 105 kDa	HEAT SHOCK PROTEIN 105 KDA (PTHR19375:SF18)	Hsp70 family chaperone
FOXC1	Forkhead box protein C1	FORKHEAD BOX PROTEIN C1 (PTHR11829:SF166)	transcription factor, DNA binding protein
EPHA7	Ephrin type-A receptor 7	EPHRIN TYPE-A RECEPTOR 7 (PTHR24416:SF377)	-
RNF217	Probable E3 ubiquitin-protein ligase RNF217	E3 UBIQUITIN-PROTEIN LIGASE RNF217-RELATED (PTHR11685:SF113)	ubiquitin-protein ligase
EZR	Ezrin	EZRIN (PTHR23281:SF13)	actin family cytoskeletal protein
NOVA1	RNA-binding protein Nova-1	RNA-BINDING PROTEIN NOVA-1 (PTHR10288:SF150)	mRNA splicing factor, ribonucleoprotein, enzyme modulator
SGIP1	SH3-containing GRB2-like protein 3-interacting protein 1	-	-
RAB10	Ras-related protein Rab-10	RAS-RELATED PROTEIN RAB-10 (PTHR24073:SF483)	-
ZCCHC24	Zinc finger CCHC domain-containing protein 24	ZINC FINGER CCHC DOMAIN-CONTAINING PROTEIN 24 (PTHR15439:SF4)	nuclease
COX5A	Cytochrome c oxidase subunit 5A, mitochondrial	CYTOCHROME C OXIDASE SUBUNIT 5A, MITOCHONDRIAL (PTHR14200:SF11)	oxidase
NR3C1	Glucocorticoid receptor	GLUCOCORTICOID RECEPTOR (PTHR24084:SF4)	nuclear hormone receptor, receptor, nucleic acid binding
NCOA7	Nuclear receptor coactivator 7	NUCLEAR RECEPTOR COACTIVATOR 7 (PTHR23354:SF68)	-
RAB22A	Ras-related protein Rab-22A	RAS-RELATED PROTEIN RAB-22A (PTHR24073:SF401)	-
CAMK2D	Calcium/calmodulin-dependent protein kinase type II subunit delta	CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE TYPE II SUBUNIT DELTA (PTHR24347:SF108)	non-receptor serine/threonine protein kinase
FAM168B	Myelin-associated neurite-outgrowth inhibitor	MYELIN-ASSOCIATED NEURITE-OUTGROWTH INHIBITOR (PTHR31844:SF2)	-
TMEM178B	Transmembrane protein 178B	TRANSMEMBRANE PROTEIN 178B (PTHR32005:SF1)	-
FRS2	Fibroblast growth factor receptor substrate 2	FIBROBLAST GROWTH FACTOR RECEPTOR SUBSTRATE 2 (PTHR21258:SF40)	-
ZBTB20	Zinc finger and BTB domain-containing protein 20	-	-
ZNF423	Zinc finger protein 423	ZINC FINGER PROTEIN 423 (PTHR24402:SF210)	KRAB box transcription factor
ELAVL3	ELAV-like protein 3	-	-
ESRRG	Estrogen-related receptor gamma	ESTROGEN-RELATED RECEPTOR GAMMA (PTHR24084:SF21)	nuclear hormone receptor, receptor, nucleic acid binding
ATG7	Ubiquitin-like modifier-activating enzyme ATG7	UBIQUITIN-LIKE MODIFIER-ACTIVATING ENZYME ATG7 (PTHR10953:SF3)	transfer/carrier protein, ligase
SMIM13	Small integral membrane protein 13	SMALL INTEGRAL MEMBRANE PROTEIN 13 (PTHR36877:SF1)	-
ZNF521	Zinc finger protein 521	ZINC FINGER PROTEIN 521 (PTHR24402:SF222)	KRAB box transcription factor
AP1S1	AP-1 complex subunit sigma-1A	AP-1 COMPLEX SUBUNIT SIGMA-1A (PTHR11753:SF14)	vesicle coat protein
KHDRBS3	KH domain-containing, RNA-binding, signal transduction-associated protein 3	KH DOMAIN-CONTAINING, RNA-BINDING, SIGNAL TRANSDUCTION-ASSOCIATED PROTEIN 3 (PTHR11208:SF29)	transcription cofactor, mRNA splicing factor

JAK2	Tyrosine-protein kinase JAK2	TYROSINE-PROTEIN KINASE JAK2 (PTHR24418:SF179)	non-receptor tyrosine protein kinase
PPP3R1	Calcineurin subunit B type 1	CALCINEURIN SUBUNIT B TYPE 1 (PTHR23056:SF43)	protein phosphatase, calmodulin
FARP1	FERM, RhoGEF and pleckstrin domain-containing protein 1	FERM, RHOGEF AND PLECKSTRIN DOMAIN-CONTAINING PROTEIN 1 (PTHR12673:SF105)	guanyl-nucleotide exchange factor
NRBF2	Nuclear receptor-binding factor 2	NUCLEAR RECEPTOR-BINDING FACTOR 2 (PTHR14964:SF2)	-
SMOC1	SPARC-related modular calcium-binding protein 1	SPARC-RELATED MODULAR CALCIUM-BINDING PROTEIN 1 (PTHR12352:SF13)	calcium-binding protein
EPHA5	Ephrin type-A receptor 5	EPHRIN TYPE-A RECEPTOR 5 (PTHR24416:SF17)	-
MYO10	Unconventional myosin-X	UNCONVENTIONAL MYOSIN-X (PTHR13140:SF276)	G-protein modulator, actin binding motor protein, cell junction protein
ATF2	Cyclic AMP-dependent transcription factor ATF-2	CYCLIC AMP-DEPENDENT TRANSCRIPTION FACTOR ATF-2 (PTHR19304:SF9)	transcription factor, nucleic acid binding
PPARGC1A	Peroxisome proliferator-activated receptor gamma coactivator 1-alpha	PEROXISOME PROLIFERATOR-ACTIVATED RECEPTOR GAMMA COACTIVATOR 1-ALPHA (PTHR15528:SF10)	transcription cofactor
FBN2	Fibrillin-2	FIBRILLIN-2 (PTHR24039:SF26)	signaling molecule, extracellular matrix glycoprotein, extracellular matrix structural protein, cell adhesion molecule, annexin, calmodulin
CHP1	Calcineurin B homologous protein 1	CALCINEURIN B HOMOLOGOUS PROTEIN 1 (PTHR23056:SF17)	protein phosphatase, calmodulin
RHOT1	Mitochondrial Rho GTPase 1	MITOCHONDRIAL RHO GTPASE 1 (PTHR24072:SF124)	small GTPase
ANKRD13C	Ankyrin repeat domain-containing protein 13C	ANKYRIN REPEAT DOMAIN-CONTAINING PROTEIN 13C (PTHR12447:SF0)	-
JARID2	Protein Jumonji	PROTEIN JUMONJI (PTHR10694:SF44)	zinc finger transcription factor
RHOBTB3	Rho-related BTB domain-containing protein 3	RHO-RELATED BTB DOMAIN-CONTAINING PROTEIN 3 (PTHR24072:SF139)	small GTPase
AGO2	Protein argonaute-2	PROTEIN ARGONAUTE-2 (PTHR22891:SF0)	-
SOX11	Transcription factor SOX-11	TRANSCRIPTION FACTOR SOX-11 (PTHR10270:SF113)	HMG box transcription factor, nucleic acid binding
PDE4A	cAMP-specific 3',5'-cyclic phosphodiesterase 4A	CAMP-SPECIFIC 3',5'-CYCLIC PHOSPHODIESTERASE 4A (PTHR11347:SF74)	phosphodiesterase
SLC43A1	Large neutral amino acids transporter small subunit 3	LARGE NEUTRAL AMINO ACIDS TRANSPORTER SMALL SUBUNIT 3 (PTHR20766:SF0)	amino acid transporter
SIRT1	NAD-dependent protein deacetylase sirtuin-1	NAD-DEPENDENT PROTEIN DEACETYLASE SIRTUIN-1 (PTHR11085:SF15)	chromatin/chromatin-binding protein, deacetylase
LRRC40	Leucine-rich repeat-containing protein 40	LEUCINE-RICH REPEAT-CONTAINING PROTEIN 40 (PTHR23155:SF424)	-
SLC16A6	Monocarboxylate transporter 7	MONOCARBOXYLATE TRANSPORTER 7 (PTHR11360:SF20)	transporter
INO80D	INO80 complex subunit D	INO80 COMPLEX SUBUNIT D (PTHR16198:SF2)	-
HOXC8	Homeobox protein Hox-C8	HOMEOBOX PROTEIN HOX-C8 (PTHR24326:SF255)	homeobox transcription factor, DNA binding protein
SPRYD7	SPRY domain-containing protein 7	SPRY DOMAIN-CONTAINING PROTEIN 7 (PTHR20951:SF2)	-
PPM1K	Protein phosphatase 1K, mitochondrial	PROTEIN PHOSPHATASE 1K, MITOCHONDRIAL (PTHR13832:SF267)	protein phosphatase, kinase inhibitor
PHOX2B	Paired mesoderm homeobox protein 2B	PAIRED MESODERM HOMEOBOX PROTEIN 2B (PTHR24329:SF301)	homeobox transcription factor, DNA binding protein
DVL3	Segment polarity protein dishevelled homolog DVL-3	SEGMENT POLARITY PROTEIN DISHEVELLED HOMOLOG DVL-3 (PTHR10878:SF6)	signaling molecule, enzyme modulator
TNRC6B	Trinucleotide repeat-containing gene 6B protein	TRINUCLEOTIDE REPEAT-CONTAINING GENE 6B PROTEIN (PTHR13020:SF32)	-
MLXIP	MLX-interacting protein	MLX-INTERACTING PROTEIN (PTHR15741:SF23)	basic helix-loop-helix transcription factor, nucleic acid binding

miRNAs-208b-3p, 204-5p, 129-2-3p and 3065-5p predicted gene targets according to TargetScanHuman (www.target scan.org) Release 7.0: August 2015.