Genomics Data 8 (2016) 93–96

Contents lists available at ScienceDirect

Genomics Data

journal homepage: www.elsevier.com/locate/gdata

Data in Brief

Expression of microRNAs in HPV negative tonsil cancers and their regulation of PDCD4

Samantha Khoury [a], Alireza Ahadi [a], Xiaoying Zhang [b,c], Nham Tran [a,b,*]

[a] Centre for Health Technologies, Faculty of Engineering and Information Technology, University of Technology, Sydney, Australia
[b] The Sydney Head and Neck Cancer Institute, Sydney Cancer Centre, Royal Prince Alfred Hospital, Australia
[c] Department of Infectious Diseases and Immunology, University of Sydney, NSW, Australia

ARTICLE INFO

Article history:
Received 11 April 2016
Accepted 12 April 2016
Available online 18 April 2016

ABSTRACT

Global rates of tonsil cancer have been increasing since the turn of the millennia, however we still have a limited understanding of the genes and pathways which control this disease. This array dataset which is linked to our publication (Zhang et al., 2015) describes the profiling of human miRNAs in tonsil and normal adjacent tissues. With this dataset, we identified a list of microRNA (miRNA) which were highly over represented in tonsil cancers and showed that several miRNAs were able to regulate the tumour suppressor PDCD4 in a temporal manner. The dataset has been deposited into Gene Expression Omnibus (GSE75630).

Specifications	
Organism/cell line/tissue	Human tonsil SCC and normal adjacent tissues
Sex	Male and female
Sequencer or array type	LNA Exiqon array
Data format	Raw: GenePix Results GPR File
Experimental factors	Tumour vs. normal
Experimental features	Extraction of total RNA from tumour and normal tissues followed by miRNA profiling using custom LNA oligonucleotide arrays
Consent	Level of consent allowed for reuse if applicable
Sample source location	Sydney, Australia

1. Direct link to deposited data

http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE75630.

2. Experimental design, materials and methods

2.1. Patient cohort

The cohort consisted of 43 patients (39 males, 4 females) treated for tonsillar cancer at Royal Prince Alfred Hospital Sydney, Australia between 2002 and 2006. The mean age was 57 years (range 39–80). Seventeen Tonsil Squamous Cell Carcinomas (SCCs) and matched microscopically normal adjacent (2 cm outside the surgical margin) tissues proved suitable for the profiling analyses (see Fig. 1 for experimental overview).

2.2. Isolation of RNA from fresh tissue or cultured cells

Approximately 100 mg of fresh frozen tissue was diced and homogenized with a scalpel blade and then mixed with 1 mL of Trizol reagent (Invitrogen, USA). Total RNA was then extracted using isopropanol precipitation and quantified using a NanoDrop ND 1000 (Thermo Fisher Scientific, USA). Samples with ratios of 260/280 in the range of 1.71 to 2.1 were used for the downstream arrays.

2.3. Total RNA labeling

MicroRNAs were labeled at the 3′-end with a P-CU-C3-Cy3 RNA linker by RNA ligation as described [2,3]. 60 μL ligation reaction was prepared with 6 μg of total RNA, 0.1 mM ATP, 20 mM $MgCl_2$, 3.5 mM DTT, 10 mg/mL BSA, 10% DMSO, 50 mM HEPES, pH 7.8, 250 ng of P-CU-C3-Cy3 (GeneLink, USA) and 20 units of T4 RNA ligase (NEB, USA). The reaction was incubated on ice for 2 h followed by precipitation at −70 °C for 20 min with 0.3 M sodium acetate, 0.5 mg/mL glycogen (Life Technologies, USA) and 2 volumes of 100% ethanol to remove any unbound RNA-linkers. Each labeled sample was dissolved in 30 μL of 400-fold diluted ULS labeled reference set, then mixed with 300 μL Church and Gilbert hybridization buffer. This mixture was denatured at 95 °C for 2 min before hybridization. A mixture of 371 synthetic DNA reference oligonucleotides (Sigma-Genosys, Australia) containing complementary sequences to all LNA probes, was randomly labeled using the ULYSIS labeling kit (Invitrogen, USA) and then filtered using a MicroSpin™ G-25 column (Amersham, USA). Aliquots of a 400-fold dilution of labeled reference set were stored at −20 °C until needed.

* Corresponding author at: Centre for Health Technologies, Faculty of Engineering and Information Technology, University of Technology, Sydney, Australia.
E-mail address: nham.tran@uts.edu.au (N. Tran).

http://dx.doi.org/10.1016/j.gdata.2016.04.007

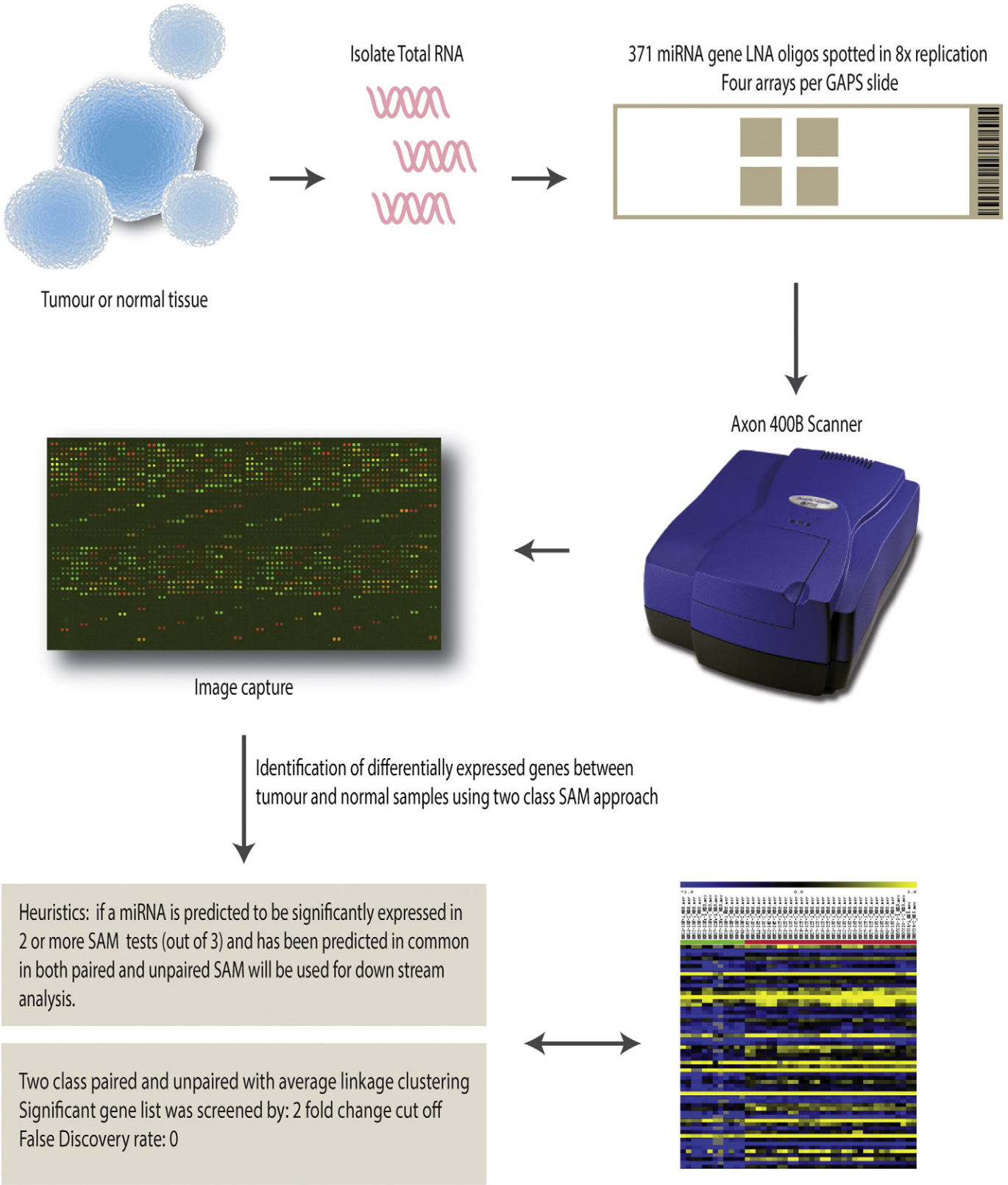

Fig. 1. Experimental overview of the project. Total RNA was extracted from cancer and normal tissues and the miRNA population was assessed using a custom printed LNA oligonucleotide array.

2.4. LNA Microarrays

A commercial LNA-modified oligonucleotides library (Exiqon, Denmark) based on miRBase release 7.1, covering 371 human and mouse miRNA was utilized for expression profiling. Features were deposited onto GAPS II slides (Amersham, USA) at a concentration of 10 μM (Australian Genome Research Facility, Australia). Individual miRNA LNA probes were printed four times on each array. In addition,

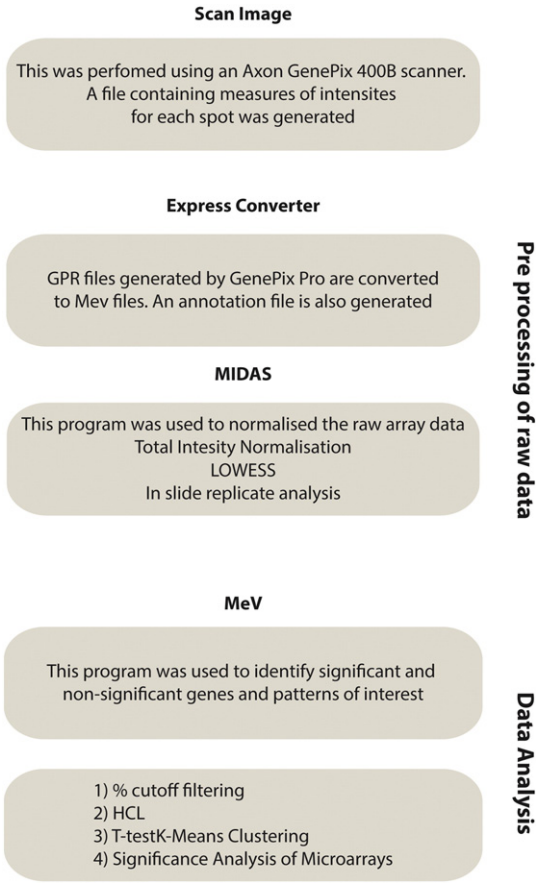

Fig. 2. Data processing pipeline using the TM4 suite of tools.

all samples were arrayed in technical duplicate. Pre-hybridization of array slide was performed in 3× SSC, 0.1% SDS, 0.2% BSA at 60 °C for 1 h. Each slide was then rinsed in full DEPC treated water, followed by 100% ethanol. Hybridization was then performed in disposable reaction chambers (ABGene, USA). The combined hybridization mixture of 330 μL was injected into the chamber, and then incubated in a hybridization oven with constant rotation at 5 rpm for 3 h at 52 °C. Slides were then washed briefly in 4× SSC, twice in 2× SSC plus 0.1% SDS, twice in 0.2% SSC and twice in 0.1% SSC. These were then scanned with a Genepix 4000B Scanner (Axon Instruments, USA). Scanning wavelengths were set between 600 and 700 for the red channel (wavelength 635) and 500–700 for the green channel (wavelength 532). Measures of intensity were assigned to each spot on the array using the software GenePix® Pro Version 5.0 (Axon Instruments, USA).

2.5. Microarray data analysis

Raw data manipulation and downstream statistical analyses were performed using the TM4 suite (http://www.tm4.org) [4]. The first step involved the GenePix Results GPR File conversion using Express Converter Version 1.7 (Institute for Genomic Research, USA). The Microarray Data Analysis System (MIDAS) version 2.19 (Institute for Genomic Research, USA) program was used to normalise raw experimental array data [4]. In brief, three types of normalisation were performed. Firstly, total intensity normalisation was performed, followed by Lowess normalisation and finally, in-slide replicate normalisation. The Multi-experimental Viewer (MeV) Version 4.0 (Institute for Genomic Research, USA) program was used to identify genes and expression patterns of interest. Normalised expression data was loaded onto the program and percentage cut-offs were used to remove genes with invalid expression values, such as buffers. The following operations were performed on the data to extract gene expression information;

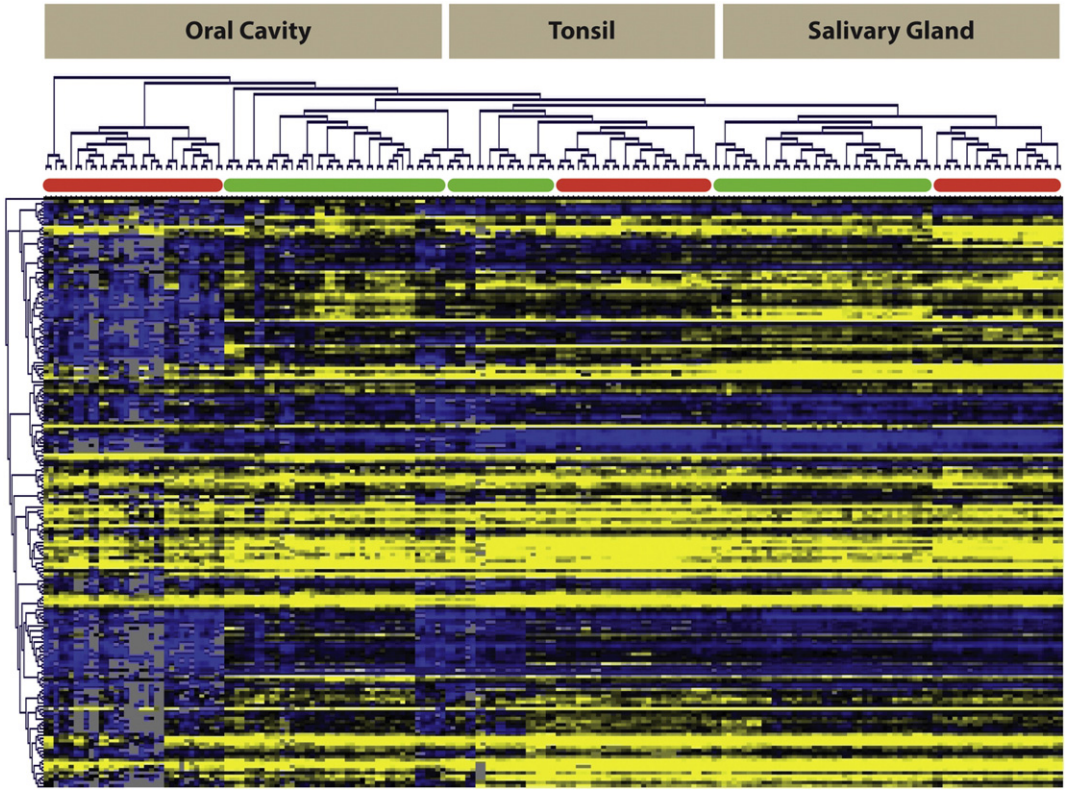

Fig. 3. Unsupervised HCL incorporating miRNA signatures from our previous study in salivary gland tumours [3] and other oral cancers [1]. HCL was able to separate the various head and neck tumours into distinct clusters. Red bars denote tumour tissue, and green bars represent adjacent normal tissue.

hierarchical clustering, K-means clustering (KMC), *t*-test and Significance Analysis for Microarrays (SAM). Refer to Fig. 2.

3. Discussion

Using this dataset we identified eleven differentially expressed miRNAs (>2.0 fold) using the statistical analysis of microarrays (SAM) algorithm [5]. Of these, nine miRNAs were up-regulated and two down-regulated in cancers relative to normal tissue. Of particular note, miR-499, miR-372, miR-18a, miR-21, and miR-30d were up-regulated while let-7c and miR-198 down-regulated. Expression data for selected miRNAs were then validated using two independent patient cohorts which strongly supported the array data. This analysis was then extended to determine if these tonsil miRNAs signatures could be used as potential diagnostic markers. HCL incorporating miRNA signatures from the previous study in salivary gland tumours [3] and other oral cancers was performed. This analysis was able to define the different types of head and neck tumours and cancers (Fig. 3). Clustering was extremely robust, clearly separating tumour from normal groups and separating the different head and neck tumour types.

References

[1] X. Zhang, H. Gee, B. Rose, C.S. Lee, J. Clark, M. Elliott, J.R. Gamble, M.J. Cairns, A. Harris, S. Khoury, et al., Regulation of the tumour suppressor PDCD4 by miR-499 and miR-21 in oropharyngeal cancers. BMC Cancer 16 (1) (2015) 86.
[2] J.M. Thomson, J. Parker, C.M. Perou, S.M. Hammond, A custom microarray platform for analysis of microRNA gene expression. Nat. Methods 1 (1) (2004) 47–53.
[3] X. Zhang, M. Cairns, B. Rose, C. O'Brien, K. Shannon, J. Clark, J. Gamble, N. Tran, Alterations in miRNA processing and expression in pleomorphic adenomas of the salivary gland. Int. J. Cancer 124 (12) (2009) 2855–2863.
[4] A.I. Saeed, V. Sharov, J. White, J. Li, W. Liang, N. Bhagabati, J. Braisted, M. Klapa, T. Currier, M. Thiagarajan, et al., TM4: a free, open-source system for microarray data management and analysis. Biotechniques 34 (2) (2003) 374–378.
[5] V.G. Tusher, R. Tibshirani, G. Chu, Significance analysis of microarrays applied to the ionizing radiation response. Proc. Natl. Acad. Sci. U. S. A. 98 (9) (2001) 5116–5121.