ALEXA – A microarray design platform for alternative expression analysis

Malachi Griffith, Michelle Tang, Obi Griffith, Ryan Morin, Susanna Chan, Jennifer Asano, Thomas Zeng, Stephane Flibotte, Adrian Ally, Agnes Baross, Martin Hirst, Steven Jones, Gregg Morin, Isabella Tai and Marco Marra

British Columbia Cancer Agency
Vancouver, British Columbia, Canada

Canada's Michael Smith Genome Sciences Centre
www.bcgsc.ca

1. Abstract

Objective: (1) To create an open source platform called 'ALEXA' for the design of microarrays capable of profiling alternative mRNA isoforms generated by 'alternative expression' (transcript initiation, splicing and polyadenylation) events. (2) To illustrate the utility of this approach by applying it to the identification of differentially expressed isoforms associated with 5-FU resistance in human colorectal cancer cell lines.
Design: We generated a prototype human microarray design and used it to profile alternative isoforms expressed in RNA samples extracted from 5-FU sensitive and resistant colorectal cancer cell lines. To assess the performance of our approach, we profiled the same RNAs with Affymetrix Human Exon arrays and compared the results between platforms.
Methods: Creation of ALEXA arrays involved extracting, scoring, filtering and annotating oligonucleotide probes that corresponded to the exons, introns, exon boundaries and exon-exon junctions of EnsEMBL genes. Prototype arrays created with the ALEXA platform were synthesized by NimbleGen. RNA samples were isolated in triplicate from colorectal cancer cell lines and profiled on the NimbleGen and Affymetrix array platforms according to the manufacturer's instructions. The resulting data were analyzed to identify isoforms differentially expressed between 5-FU sensitive and resistant samples.
Results: The ALEXA platform was used to pre-compute genome-wide microarray designs for ten species consisting of a total of ~100 million scored and annotated oligonucleotide probes. Differential exon and gene expression values from ALEXA and Affymetrix arrays were highly concordant (Pearson correlations of 0.87 and 0.67 for gene and exon values, respectively), although each platform provided some unique information. Using the combination of Affymetrix and ALEXA microarrays we identified ~50 candidate differentially expressed genes and ~25 differentially expressed mRNA isoforms that were associated with acquired 5-FU resistance. For example, we observed a reciprocal differential expression (DE) pattern for two isoforms of the gene *UMPS*, a gene that is directly involved in 5-FU metabolism.
Conclusion: The use of microarray technology to profile mRNA transcripts generated by alternative expression events is an area of rapid development. We created a computational platform to facilitate these experiments and verified its effectiveness experimentally. Our custom microarray data were comparable or superior to Affymetrix exon arrays in terms of reproducibility, sensitivity and specificity and provided additional information on the connectivity and boundaries of exons. We identified several differentially expressed isoforms with potential relevance to 5-FU resistance in colorectal cancer. The ALEXA platform is available from the website: www.AlexaPlatform.org.

ALEXA platform description

The ALEXA microarray design platform consists of Perl and R modules interacting with a mySQL relational database and is capable of creating custom array designs for any EnsEMBL annotated species. The design process consists of: (1) extracting probe sequences corresponding to exons, introns, exon junctions, exon boundaries and random sequences with minimal homology to any known sequence; (2) filtering probes according to their specificity and thermodynamic parameters; (3) selecting gene targets and generating an array design file for submission to a custom array manufacturer.

Pre-computed array designs

Using the ALEXA approach, we pre-computed alternative expression microarray designs representing ~100 million probe sequences for ten species residing in EnsEMBL. The designs currently available, the number of genes targeted and the number of probes available after filtering are as follows :

C. familiaris (dog)
- 14.1k genes, 2.1m probes

C. elegans (worm)
- 20.0k genes, 1.6m probes

D. melanogaster (fly)
- 14.6k probes, 700k probes

D. rerio (zebrafish)
- 16.3k genes, 1.8m probes

G. gallus (chicken)
- 17.2k genes, 2.3m probes

H. sapiens (human)
- 22.7k genes, 3.1m probes

M. musculus (mouse)
- 22.3k genes, 3.0m probes

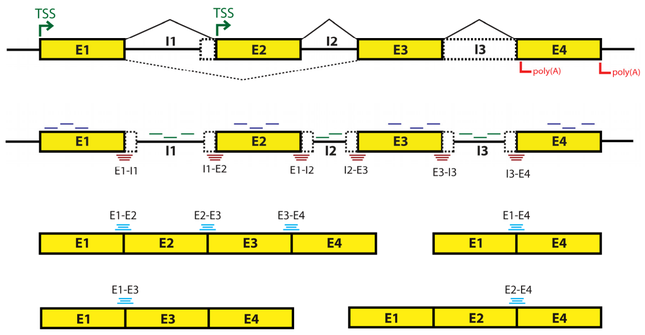
P. troglodytes (chimp)
- 24.5k genes, 2.9m probes

R. norvegicus (rat)
- 18.8k genes, 2.8m probes

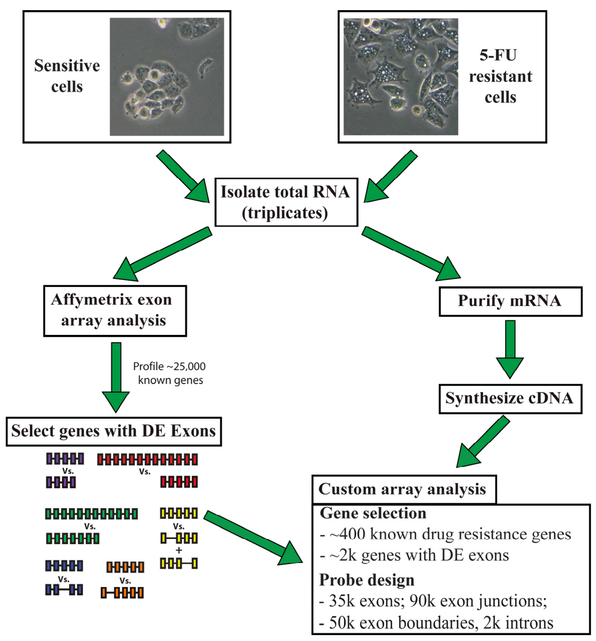
S. cerevisiae (yeast)
- 6.7k genes, 30k probes

2. Methods

Array design strategy

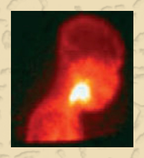


Experimental overview



Alternative expression and cancer

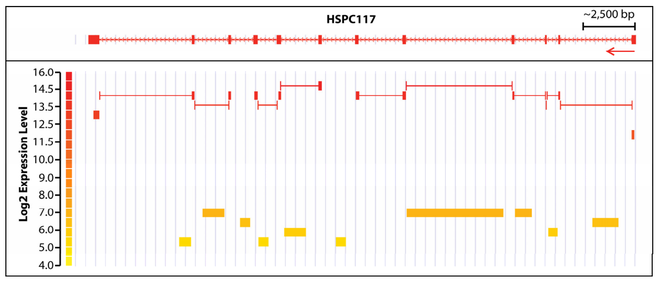
Alternative expression (alternative transcript initiation, splicing and polyadenylation) generates multiple isoforms for most human genes. Specific isoforms may represent useful therapeutic targets or diagnostic markers. The image below shows a throat and neck cancer which has been targeted by a therapeutic antibody specific to a single alternate CD44 exon that is aberrantly expressed in the cancer (Venables, 2006).



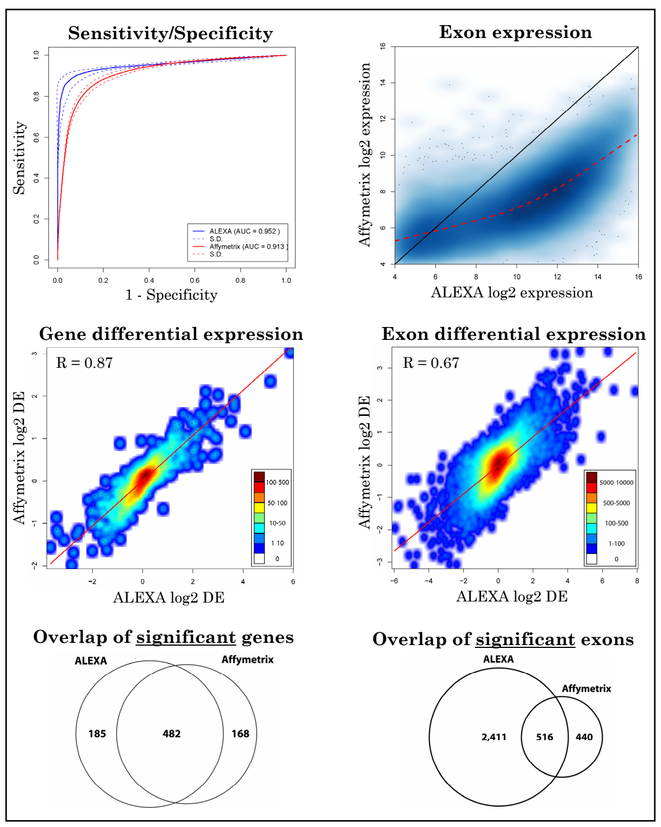
3. Results – Platform validation

The utility of the ALEXA approach was assessed by using a prototype human array to profile the expression of alternative mRNA isoforms in 5-fluourouracil (5-FU) sensitive and resistant colorectal cancer cell lines and comparing the results to those from Affymetrix's 'GeneChip® Human Exon 1.0 ST' array. ALEXA array data were comparable or superior to Affymetrix exon arrays in terms of reproducibility, sensitivity and specificity and signal-to-noise ratio (~3 times higher in ALEXA array data).

Expression data for an example housekeeping gene



Comparison of Affymetrix and ALEXA data



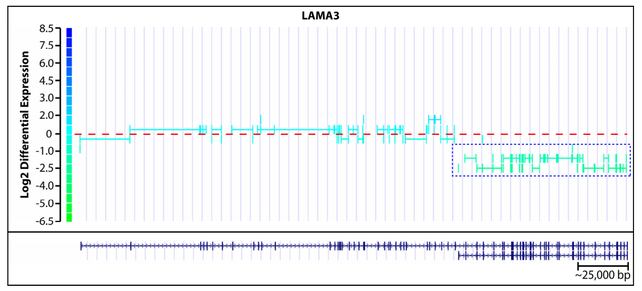
Summary of DE events associated with 5-FU resistance

	DE event type	# Events profiled	# Significant DE events (> 4-fold)	# Within ORF	# Affecting known feature (domain, signal peptide, etc.)
Affymetrix	Gene-level	2,507	78	N/A	N/A
	Exon	49,681	1117	978	589
	Intron	65,327	25	20	0
	Total	117,515	**1,220**	998	589
ALEXA	Gene-level	2,507	233	N/A	N/A
	Exon	32,164	2,703	2,537	1,544
	Canonical junction	27,046	2,310	2,260	1,277
	Exon skip	69,761	191	180	103
	Exon boundary	52,402	253	219	100
	Intron	472	0	0	0
	Total	184,354	**5690**	5196	3024

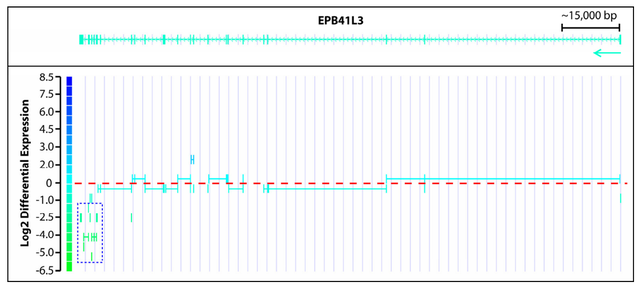
4. Results – DE events associated with 5-FU resistance

Genes and exons differentially expressed between 5-FU sensitive and resistant cells (log2 sensitive minus log2 resistant) were identified by both platforms, but ALEXA arrays provided additional information on the connectivity and boundaries of exons. Alternative expression (AE) events (alternative exon junctions or boundaries) identified by ALEXA were significantly enriched for known AE events represented in publicly available mRNA and EST databases. The following figures depict known and predicted isoforms identified by the ALEXA approach as differentially expressed between 5-FU sensitive and resistant cells.

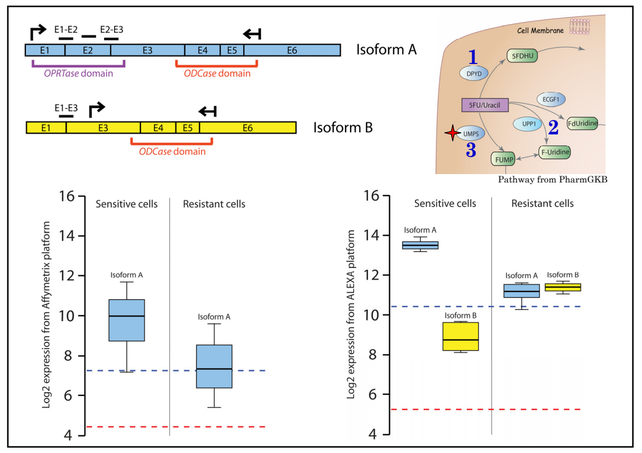
Differential expression of a known *LAMA3* isoform



Differential expression of a novel *EPB41L3* (*Dal-1*) isoform



Reciprocal DE of two known *UMPS* isoforms



5. Availability

Citation: Malachi Griffith, Michelle Tang, Obi Griffith, Ryan Morin, Susanna Chan, Jennifer Asano, Thomas Zeng, Stephane Flibotte, Adrian Ally, Agnes Baross, Martin Hirst, Steven Jones, Gregg Morin, Isabella Tai and Marco Marra. *ALEXA – A microarray design platform for alternative expression analysis*. Submitted to Nature Methods 21st Aug 2007.
Download: www.AlexaPlatform.org

6. Acknowledgments

Thanks | In addition to the authors listed above I would like to thank the GSC MGC Closure Group, Joseph Connors, Noushin Farnoud, Anca Petrescu and Trevor Pugh

Funding | I am grateful for support provided by he Natural Sciences and Engineering Research Council, the Michael Smith Foundation for Health Research and Genome British Columbia

NSERC CRSNG · Michael Smith Foundation for Health Research · GenomeBritishColumbia