

# Chapter XV

## Alternative Isoform Detection Using Exon Arrays

**Jacek Majewski**

*McGill University and Genome Québec Innovation Centre, Canada*

**David Benovoy**

*McGill University and Genome Québec Innovation Centre, Canada*

**Tony Kwan**

*McGill University and Genome Québec Innovation Centre, Canada*

### ABSTRACT

*Eukaryotic genes have the ability to produce several distinct products from a single genomic locus. Recent developments in microarray technology allow monitoring of such isoform variation at a genome-wide scale. In our research, we have used Affymetrix Exon Arrays to detect variation in alternative splicing, initiation of transcription, and polyadenylation among humans. We demonstrated that such variation is common in human populations and has an underlying genetic component. Here, we use our study to illustrate the use of Exon Arrays to detect alternative isoforms, to outline the analysis involved, and to point out potential problems that may be encountered by researchers using this technology.*

### INTRODUCTION

Alternative pre-mRNA splicing is a process allowing the production of several distinct gene isoforms from a single genomic locus. The most common type of alternative splicing events in mammals results in cassette exons, where each such exon can be either included or excluded from the mature mRNA. Other events include alternative use of donor or acceptor splice sites, and intron retention. In addition, processes such as alternative promoter usage and alternative polyadenylation, resulting in differences in initiation and termination of the transcript, respectively, further diversify eukaryotic transcriptomes and proteomes. Such processes have been suggested to be at least partly responsible for mammalian

complexity, which is otherwise difficult to explain in view of our relatively low number of genomic loci - less than 25,000 genes in humans, versus approximately 20,000 in the nematode worm *C. elegans* (Claverie, 2001). It is estimated that a high percentage of mammalian genes is alternatively spliced, and this frequency is highest in specialized and complex tissues, such as the brain and the liver. Differences in splicing patterns have been shown to exist across species, and within populations of the same species. In humans, splicing defects are known to result in numerous genetic disorders (Faustino & Cooper, 2003) and may confer susceptibility to complex genetic diseases. Thus, the process of alternative splicing attracts the interest of researchers across the entire biomedical sciences spectrum, ranging from evolutionary biology, through development, to medicine.

In recent years, alternative transcript investigation in a genome-wide context has been carried out using expressed sequence tag libraries (ESTs). Generally, ESTs (short cDNA sequence reads) are mapped to the genomic sequence, and different isoforms can be inferred from incongruence of splicing patterns (Modrek, Resch, Grasso, & Lee, 2001). However, EST library analyses are prone to sequencing errors, biased towards highly expressed genes, and influenced by cancer-derived ESTs, which may not generally be present in healthy tissues.

More recently, microarray platforms have been proposed as a tool for studying gene expression at the isoform level (Black & Graveley, 2006; Lee & Roy, 2004; Zhang et al., 2006). Splicing sensitive microarrays employ a number of exon body oligonucleotide probes, or exon junction probes, or a combination of the two designs, to determine mRNA levels at the resolution of a single exon or splice site. The Affymetrix GeneChip® Human Exon 1.0 ST Array is the first commercially available microarray product designed for genome-wide, exon level expression analysis. The array relies on targeting multiple probes to individual exons and allows simultaneous, exon-level detection of expression intensity for 1.4 million probesets covering over 1 million known and predicted human exons. The Exon Array is a flexible tool, which can be used to perform the function of classical expression arrays and concurrently provide supplementary information on isoform changes. However, because of the complexity of the design, statistical analysis of the data becomes much more intensive, both at the theoretical and computational level. The simplest illustration is the multiple testing problem; whereas in classical expression arrays the number of tests is of the order of the number of genes, in an exon array, the number of tests is over ten-fold higher – and can vary between a few hundred thousand to over 1 million (if computationally predicted exons are included). The statistical approaches need to be able to distinguish between whole gene expression differences and isoform differences, which introduces a new level of complexity. The robustness of measurement is also an issue, since the exon array has on average four probes per probeset, whereas Affymetrix expression arrays relied on more than 10 probes per probeset to estimate expression.

In our lab, we have used the Exon Arrays, to investigate differences in splicing patterns among humans. We were able to demonstrate the existence of common variation in splicing, polyadenylation sites, transcription initiation. We also demonstrated genetic linkage and allelic association of the isoform variation to common genetic single nucleotide polymorphisms (SNPs). Our results show that the effects of genetic variants on gene expression are much more complex than previously believed, and constitute an important step towards understanding the functional consequences of such variation.

In this chapter, we will use our experiments as a case study, in order to outline the flow of the analysis required to process Exon Arrays. We will also outline problems which may be encountered by potential users of the chips, and describe solutions which we have developed to overcome such problems. We will discuss current advances in statistical analysis and propose future improvements to optimize both the array design and statistical solutions.

14 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:  
[www.igi-global.com/chapter/alternative-isoform-detection-using-exon/21537](http://www.igi-global.com/chapter/alternative-isoform-detection-using-exon/21537)

## Related Content

---

### Comparative Study of Fuzzy Entropy with Relative Spike Amplitude Features for Recognizing Wake-Sleep Stage 1 EEGs

Natarajan Sriraam, B. R. Purnimaand Uma Maheswari Krishnaswamy (2015). *International Journal of Biomedical and Clinical Engineering* (pp. 12-25).

[www.irma-international.org/article/comparative-study-of-fuzzy-entropy-with-relative-spike-amplitude-features-for-recognizing-wake-sleep-stage-1-eegs/138224](http://www.irma-international.org/article/comparative-study-of-fuzzy-entropy-with-relative-spike-amplitude-features-for-recognizing-wake-sleep-stage-1-eegs/138224)

### Use of Telemedicine Systems and Devices for Patient Monitoring

Dionisia Damigou (2009). *Medical Informatics: Concepts, Methodologies, Tools, and Applications* (pp. 733-739).

[www.irma-international.org/chapter/use-telemedicine-systems-devices-patient/26253](http://www.irma-international.org/chapter/use-telemedicine-systems-devices-patient/26253)

### Bioinformatics-Inspired Algorithms for 2D-Image Analysis-Application to Synthetic and Medical Images Part I: Images in Rectangular Format

Perambur S. Neelakantaand Deepti Pappusetty (2012). *International Journal of Biomedical and Clinical Engineering* (pp. 14-38).

[www.irma-international.org/article/bioinformatics-inspired-algorithms-image-analysis/73691](http://www.irma-international.org/article/bioinformatics-inspired-algorithms-image-analysis/73691)

### Bioinformatics-Inspired Algorithms for 2D-Image Analysis-Application to Synthetic and Medical Images Part I: Images in Rectangular Format

Perambur S. Neelakantaand Deepti Pappusetty (2012). *International Journal of Biomedical and Clinical Engineering* (pp. 14-38).

[www.irma-international.org/article/bioinformatics-inspired-algorithms-image-analysis/73691](http://www.irma-international.org/article/bioinformatics-inspired-algorithms-image-analysis/73691)

### Automatic Detection of Irritable Bowel Syndrome for 3D Images Using Supervoxel and Graph Cut Algorithm

Geetha Vaithianathanand Rajkumar E. (2021). *International Journal of Biomedical and Clinical Engineering* (pp. 1-13).

[www.irma-international.org/article/automatic-detection-of-irritable-bowel-syndrome-for-3d-images-using-supervoxel-and-graph-cut-algorithm/282491](http://www.irma-international.org/article/automatic-detection-of-irritable-bowel-syndrome-for-3d-images-using-supervoxel-and-graph-cut-algorithm/282491)