

**Michael Olivier**

*Physiol Genomics* 16:182-183, 2004. doi:10.1152/physiolgenomics.00194.2003

---

**You might find this additional information useful...**

---

This article has been cited by 1 other HighWire hosted article:

**Rigid spin-labeled nucleoside C: a nonperturbing EPR probe of nucleic acid conformation**

P. Cekan, A. L. Smith, N. Barhate, B. H. Robinson and S. Th. Sigurdsson

*Nucleic Acids Res.*, October 1, 2008; 36 (18): 5946-5954.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)

Medline items on this article's topics can be found at <http://highwire.stanford.edu/lists/artbytopic.dtl> on the following topics:

Medicine .. Cardiovascular Genetics

Medicine .. Gene Expression

Medicine .. Genetics

Medicine .. Diabetes Genetics

Medicine .. Genes

Physiology .. Humans

Updated information and services including high-resolution figures, can be found at:

<http://physiolgenomics.physiology.org/cgi/content/full/16/2/182>

Additional material and information about *Physiological Genomics* can be found at:

<http://www.the-aps.org/publications/pg>

---

This information is current as of April 23, 2009 .

## From SNPs to function: the effect of sequence variation on gene expression.

### Focus on “A survey of genetic and epigenetic variation affecting human gene expression”

**Michael Olivier**

Human and Molecular Genetics Center, Department of Physiology,  
Medical College of Wisconsin, Milwaukee, Wisconsin 53226

THE RECENT COMPLETION of the human genome sequence has shifted research efforts in genomics toward understanding the function of the human genome, its regulation, and how sequence variation contributes to disease. Large numbers of sequence variants throughout the human genome have been identified, and efforts are currently underway to understand the overall relationship between sequence variation on a genomic level, and the goal of identifying a subset of single nucleotide polymorphisms (SNPs) that will capture the vast majority of genetic diversity found in the human population. The hope is that this subset could then be used to identify genomic regions and SNPs, in genome-wide analyses, that predispose human beings to common disorders such as obesity, diabetes, or cardiovascular disorders.

SNPs are, of course, already routinely used in human studies to test individual genes or genomic regions for their association with disease phenotypes. A number of SNPs have been identified in several genes that contribute to the complex etiology of diseases such as diabetes and hypertension. However, these studies often fail to verify causality of individual SNPs for the disease phenotype. Testing for functionality of a SNP is not a simple task. Testing amino acid-altering coding SNPs for their effect on protein function or testing promoter and splice site variants for their effect on gene transcription often requires elaborate expression constructs and analysis using *in vitro* systems. This analysis cannot be done for large numbers of genes or variants, and it fails to test the functionality of SNPs in intronic regions, unknown regulatory elements, or intergenic regions. Our current understanding of how the genome regulates gene expression and function is limited, and our knowledge mainly stems from our investigations of DNA sequences adjacent to individual genes. Since these regions only account for about 5% of the human genome, we have little knowledge of (and no high-throughput methodology to investigate) the vast majority of the human genome sequence.

In this release of *Physiological Genomics*, Pastinen et al. (1) propose an alternative path to investigate the effects of SNPs on gene expression *in vivo*. Rather than expressing SNPs in a reporter construct, they investigate the effect of SNPs on the amount of mRNA produced directly in the tissue of interest. RNA from cell lines heterozygous for the selected SNP is tested for allelic imbalance in the expression of the two alleles. Normally, both alleles of autosomal genes are expected to be expressed at equal levels. Any deviation from this equal ratio

would suggest that the SNP allele in question (or another SNP allele in linkage disequilibrium with the one ascertained) somehow affects expression levels. The authors also show that the same correlation can be found when intronic SNPs are used (by looking at hnRNA), and they even show a haplotype for one gene (*BTN3A2*) that strongly affects gene expression levels of the two alleles.

This manuscript is the first description of a novel high-throughput pipeline for analyzing large numbers of SNPs and genes for such allelic imbalances, and this novel approach may help researchers in correlating SNPs with a direct cellular effect on a genome-wide scale. Although this does not answer the question of how these SNPs exert their effect, it will help in understanding why anonymous SNPs and haplotypes show association to specific disease phenotypes, and it may provide a first glimpse at potential intervention strategies to correct the expression imbalance. A comprehensive analysis of these allelic imbalances in gene expression may allow us to design specific therapeutics to correct the alteration in gene expression and therefore treat the resulting disease without explicitly knowing the causative SNP itself and its direct mode of action. In addition, showing a correlation between alterations in gene expression and genetic variation in individual genes will help to confirm results from association studies. It will provide more confidence for a notion that a particular gene plays an important role in the pathogenesis of a complex disorder such as diabetes and thus will allow more confident use of genetic information in managing the disease in affected individuals.

Other researchers have previously shown allelic imbalances in gene repression. In these studies as well as in the current reported analyses, 20–50% of all genes tested show signs of allelic imbalance in gene expression, and, in almost all cases, no obvious reason for this effect on gene expression is evident from the SNP itself. Given this high percentage of genes showing alterations in gene expression, the approach presented in the paper by Pastinen et al. (1) may provide significant insights into the biology of complex disorders in humans for a large percentage of genes, not just for a few rare cases.

Nonetheless, not all diseases are caused by altered gene expression levels. Thus analyzing the effect of SNPs on gene expression levels will not always explain the mechanism by which disease-associated SNP alleles cause the phenotypic changes. Alterations in DNA sequence may affect RNA stability, protein function, or other cellular mechanisms. Furthermore, SNPs associated with disease phenotypes may reside in (as of yet unknown) *trans*-regulatory elements, affecting gene function on other chromosomes. Studying the effect on expression of neighboring genes (and not of the gene actually regulated elsewhere in the genome) would show no correlation.

Article published online before print. See web site for date of publication (<http://physiolgenomics.physiology.org>).

Address for reprint requests and other correspondence: M. Olivier, Human and Molecular Genetics Center, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226 (E-mail: [molivier@mcw.edu](mailto:molivier@mcw.edu)).

In summary, the proposed high-throughput approach may not allow us to understand the mechanism by which all SNPs cause their effect, but its application to genome-wide gene expression analyses will allow researchers to confirm association results with a directly measurable effect on gene function, a significant step toward using genetic information to understand, treat, and possibly prevent common human disorders. Developing similar high-throughput approaches for the analysis of alternative mechanisms by which SNPs can

cause disease will be one of the remaining challenges for genomic research.

#### REFERENCES

1. **Pastinen T, Sladek R, Gurd S, Sammak A, Ge B, Lepage P, Lavergne K, Villeneuve A, Gaudin T, Brändström H, Beck A, Verner A, Kingsley J, Harmsen E, Labuda D, Morgan K, Vohl MC, Naumova AK, Sinnett D, and Hudson TJ.** A survey of genetic and epigenetic variation affecting human gene expression. *Physiol Genomics* 16: 184–193, 2003. First published October 28, 2003; 10.1152/physiolgenomics.00163.2003.

