Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Immunology

Genetics of gene expression in immunity to infection

Benjamin P Fairfax and Julian C Knight

Mapping gene expression as a quantitative trait (eQTL mapping) can reveal local and distant associations with functionally important genetic variation informative for disease. Recent studies are reviewed which have demonstrated that this approach is particularly informative when applied to diverse immune cell populations and situations relevant to infection and immunity. Context-specific eQTL have now been characterised following endotoxin activation, induction with interferons, mycobacteria, and influenza, together with genetic determinants of response to vaccination. The application of genetical genomic approaches offers new opportunities to advance our understanding of gene–environment interactions and fundamental processes in innate and adaptive immunity.

Addresses
Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK

Corresponding authors: Fairfax, Benjamin P (bfairfax@well.ox.ac.uk, benjamin.p.fairfax@googlemail.com) and Knight, Julian C (julian@well.ox.ac.uk)

Current Opinion in Immunology 2014, **30**:63–71

This review comes from a themed issue on **Immunogenetics and transplantation**

Edited by **Luis B Barreiro** and **Lluis Quintana-Murci**

For a complete overview see the Issue and the Editorial

Available online 31th July 2014

http://dx.doi.org/10.1016/j.coi.2014.07.001

Introduction

Individuals vary in susceptibility to infectious disease and only a minority develop the most severe complications [1]. Responses to pathogen invasion and immunity to infection require coordinated regulation of gene expression. At any given locus gene expression may vary markedly between individuals. This variation has a strong genetic component, with polymorphisms associating with gene expression known as expression quantitative trait loci (eQTL) [2,3]. Genome-wide association studies (GWAS) of disease susceptibility demonstrate the majority of identified risk loci involve non-coding single nucleotide variants (SNVs) and these are thought to exert primarily regulatory effects on gene expression [4•]. The identification of regulatory genetic variants and eQTL is therefore of significant biomedical interest. Whilst variance in expression is highly gene specific, eQTL contribute to phenotypic diversity with regulatory variation proposed as the predominant driver of local population adaptive changes [5]. Quantitative changes in gene expression influence the outcomes of immune responses and genes implicated in immune-related disease demonstrate higher heritability for expression, as do those under balancing selection, reflecting the selective pressure of infectious disease [6•]. In this review we highlight recent insights from eQTL analysis in settings relevant to immunity and infection. This work has illustrated the extent of context-specific associations, informing our understanding of gene–environment interactions and population-wide variance in expression of key genes implicated in immunity to infection that may help refine our understanding of immunogenetics.

Informativeness of eQTL mapping

The eQTL approach builds upon the principles of genetical genomics [7••], whereby analysis of intermediate phenotypes (such as transcript abundance) between those at the organismal level (i.e. disease trait) and underlying DNA variation provides insights into the functional correlates of associated variation. This is particularly useful for non-Mendelian immune and infectious disease phenotypes where interaction between polygenic variants and environmental factors may be required for disease phenotype manifestation. In brief, eQTL can be subdivided into those that show association with gene expression locally (typically within 1Mb region, likely *cis*-acting) and those that demonstrate association to distant genes — known as *trans* eQTL. Such associations identify putative functional regulatory genetic variants and the specific genes or gene networks that may be modulated. *trans* eQTL have the potential to provide unbiased discovery of novel pathways and processes, especially those that influence multiple genes (discussed further below) although effect sizes are typically smaller and, given genome-wide testing, *trans* associations are harder to resolve when accounting for Type-1 error rates. Recent meta-analyses involving large cohorts for whom peripheral blood eQTLs have been mapped [8••] has made notable progress into the identification of reproducible *trans* eQTL however. There is also evidence that induction of gene expression with immune stimulants may enrich for *trans*-effects by harmonizing expression patterns and reducing variance [9••]. The intersection of eQTL and GWAS has been conducive to hypothesis generation regarding the identity of causative genes underlying disease associations [10]. Although local eQTL effect sizes increase as distance to the gene transcription start site falls, it is increasingly appreciated many local eQTL act over longer ranges through distal enhancer effects, making resolution of GWAS difficult

[8••,11]. This is illustrated by recent work demonstrating that obesity risk polymorphisms located in the *FTO* gene have functional effects across over 400 kb on the homeobox gene *IRX3*, proposed to be the causative mechanism for obesity susceptibility [12•].

Trans eQTL provide unbiased insights into immunological pathways

trans eQTL, whilst typically more difficult to detect, often leave a signature upon expression across multiple different loci — so called 'master regulatory regions'. By resolving local effects of these polymorphisms on gene expression, the likely upstream causative gene can be identified permitting unbiased delineation of physiologically relevant downstream pathways and avoiding the use of genomic interference technologies such as gene knockouts. An illustrative example being recently identified in blood where a lupus associated variant associated in *trans* with several interferon response genes [8••]. By combining sequencing and ChIP-seq data the same SNV was found to have local effects upon expression of the transcription factor IKZF1 with the genes in *trans* being enriched for IKZF1 binding. In addition to mechanistically resolving a role for IKZF1 in lupus predisposition, it also suggested that many genes implicated in lupus pathogenesis are subtly dysregulated in healthy carriers of a predisposing susceptibility variant. Similarly, many *trans* associations involve SNVs within the Major Histocompatibility Complex (MHC) [8••,9••,13•,14•], which is 10-fold enriched for such associations [14•], with evidence these are context specific and correlate to MHC class II gene expression [9••]. Another intriguing master-regulatory *trans* eQTL resolved only in B-cells has local effects upon expression of the transcription factor KLF4, a gene implicated in pluripotency and breast cancer susceptibility [15] and robustly expressed in monocytes. Carriage of the minor allele of rs61414050 was associated with upregulation of several innate pro-inflammatory genes and, whilst it should be noted that this allele was relatively rare (minor allele frequency in population of European ancestry 4.2%) and therefore replication of this effect is required, such widespread changes in expression are likely to have functional outcomes.

eQTL vary in effect size across tissue and cell type

The majority of early human eQTL studies focused on Epstein-Barr Virus transformed lymphoblastoid cell lines (LCLs) which, although informative [16•], have significant limitations due to immortalization and physiological relevance. Notably, immune cells are characterised by a high degree of inter-cellular cross-talk and antigen presentation. Use of LCL monocultures fails to capture the effects of this cross-talk on expression and also diminishes the detection of other physiologically relevant context-specific effects. The latter may depend on the specific cell and tissue type analysed [13•,17••,18•], population [19•], sex [20], age [20,21], geography [22], disease [23] and environmental context [9••,24,25••], all of which may impact immunity.

Peripheral blood mononuclear cells (PBMCs) form an easily accessibly primary substrate for eQTL analysis. However, they represent a heterologous assortment of multiple cell subsets, each with markedly different expression profiles and separate roles in immunity; the proportional composition varying inter-individually and intra-individually over time. eQTL for one cell subtype, which may only form a small proportion of total cell count, are at risk of not being detected upon analysis of whole blood or PBMCs. This is a primary motivator behind methods to identify cell specific eQTL. These involve attempts to bioinformatically deconvolute the signal of specific subtypes from whole blood [26], or use physical methods of separation. To date only a modest number of studies have been published interrogating different primary immune cell types from peripheral blood and tissues (Table 1), although this will be advanced by ongoing efforts such as the ImmVar project (www.immvar.org). From these initial studies it is apparent that specific analysis of lymphoid and myeloid cell types demonstrates widespread variation in eQTL activity according to cell type. Intriguingly, eQTL frequently demonstrate marked difference in effect sizes between different peripheral blood cell types, even when the gene expression is similar. In a few circumstances this effect is opposing in nature between cells, with polymorphisms simultaneously associated with increased expression in one cell type and reduced expression in an alternative cell type. Just as eQTL can elicit changes in the balance of gene expression between immune effector and regulatory genes within one cell type, this differential activity of immune eQTL may result in the altered balance of gene expression between different arms of the immune system. Where the gene encodes a surface receptor, differential expression might be anticipated to result in variation in proportion of cell subset types recruited by the respective marker ligand. Whilst this is currently untested, recent studies have demonstrated such effects for polymorphisms associated with contrasting effects between monocytes and T-cells on expression of *CD52*, encoding the target of the immunosuppressant Alemtuzumab, and the gene *SELL* encoding L-selectin, a protein which plays a crucial role in leukocyte trafficking to areas of inflammation [27] between B cells and monocytes (Figure 1). The *CD52* observation is of significance in the field of immunosuppression in organ transplantation and in the management of multiple sclerosis where Alemtuzumab is used to deplete T-cells. A polymorphism regulating expression of surface CD52 between T-cells and monocytes might be expected to alter the proportional depletion of T-cells and monocytes. Carriage of the minor allele of rs2223286, a SNP intronic to *SELL* that disrupts a consensus motif for FOS and is located in a