DECODING THE GENOMIC CONTROL OF IMMUNE REACTIONS
DECODING THE GENOMIC CONTROL OF IMMUNE REACTIONS
The Novartis Foundation is an international scientific and educational charity (UK Registered Charity No. 313574). Known until September 1997 as the Ciba Foundation, it was established in 1947 by the CIBA company of Basle, which merged with Sandoz in 1996, to form Novartis. The Foundation operates independently in London under English trust law. It was formally opened on 22 June 1949.

The Foundation promotes the study and general knowledge of science and in particular encourages international co-operation in scientific research. To this end, it organizes internationally acclaimed meetings (typically eight symposia and allied open meetings and 15–20 discussion meetings each year) and publishes eight books per year featuring the presented papers and discussions from the symposia. Although primarily an operational rather than a grant-making foundation, it awards bursaries to young scientists to attend the symposia and afterwards work with one of the other participants.

The Foundation’s headquarters at 41 Portland Place, London W1B 1BN, provide library facilities, open to graduates in science and allied disciplines. Media relations are fostered by regular press conferences and by articles prepared by the Foundation’s Science Writer in Residence. The Foundation offers accommodation and meeting facilities to visiting scientists and their societies.

Information on all Foundation activities can be found at http://www.novartisfound.org.uk

The Australian Phenomics Facility (APF) is a leading organization that provides a unique approach to the study of physiology and health via the elucidation of the genome–phenome code. Many of the health problems we face today—obesity, cancer, autoimmune diseases, allergy, cardiovascular disease—stem from a discordance between genetics and environment. The APF has pioneered technologies to produce libraries of mice with informative gene alterations, and ways of using those libraries for new genome–phenome discoveries.

The APF is located in Canberra, Australia. Information on all APF activities can be found at http://www.apf.edu.au.
Contents

Novartis Foundation Symposium on Decoding the genomic control of immune reactions, held at University House, The Australian National University, Canberra, 7–9 March 2006, in collaboration with the Australian Phenomics Facility and The Australian National University

Editors: Gregory Bock (Organizer) and Jamie Goode

This symposium was based on a proposal made by Chris Goodnow

Chris Goodnow  Chair’s Introduction   1

David A. Hume, Christine A. Wells and Timothy Ravasi  Transcriptional regulatory networks in macrophages   2

Discussion   18

Christian Schönbach  The RIKEN mouse transcriptome: lessons learned and implications for the regulation of immune reactions   25

Discussion   34

Steven A. Stacker, Rae H. Farnsworth, Tara Karnezis, Ramin Shayan, Darrin P. Smith, Karri Paavonen, Natalia Davydo, Carol Caesar, Rachael Inder, Megan E. Baldwin, Bradley K. McColl, Sally Roufail, Richard A. Williams, Richard A. Hughes, Kari Alitalo and Marc G. Achen  Molecular pathways for lymphangiogenesis and their role in human disease   38

Discussion   44

General discussion I   50

Jason G. Cyster  Specifying the patterns of immune cell migration   54

Discussion   61

Capucine Picard, Laurent Abel and Jean-Laurent Casanova  Human monogenic disorders that confer predisposition to specific infections   65

Discussion   73
W. J. Britton, S. L. Fernando, B. M. Saunders, R. Sluyter and J. S. Wiley
The genetic control of susceptibility to *Mycobacterium tuberculosis* 79
*Discussion* 89

Bernard Malissen, Ying Wang, Michael Mingueneau and Marie Malissen
Th2 lymphoproliferative disorders resulting from defective LAT signalosomes 93
*Discussion* 100

Carola G. Vinuesa and Matthew C. Cook  Genetic analysis of systemic autoimmunity 103
*Discussion* 120

Gunasegaran Karupiah, Vijay Panchanathan, Isaac G. Sakala and Geeta Chaudhri  Genetic resistance to smallpox: lessons from mousepox 129
*Discussion* 136

Anny Fortin, Eduardo Diez, Janet E. Henderson, Jeffrey S. Mogil, Philippe Gros and Emil Skamene  The AcB/BcA recombinant congenic strains of mice: strategies for phenotype dissection, mapping and cloning of quantitative trait genes 141
*Discussion* 153

Gundula Min-Oo, Mary M. Stevenson, Anny Fortin and Philippe Gros  Genetic control of host–pathogen interactions in mice 156
*Discussion* 165

Douglas Young and Anne O'Garra  *Mycobacterium tuberculosis* and its ability to resist immunity 169
*Discussion* 177

Grant Morahan and Robert W. Williams  Systems genetics: the next generation in genetics research? 181
*Discussion* 188

Rick Maizels  Regulation of the immune system in metazoan parasite infections 192
*Discussion* 204

Chris Goodnow  Closing remarks 208

Contributor Index 210

Subject Index 212
Participants

Warren S. Alexander The Walter and Eliza Hall Institute of Medical Research, and Cancer & Haematology Division, 1G Royal Parade, Parkville, 3050 Victoria, Australia

Gabrielle Belz The Walter and Eliza Hall Institute, 1G Royal Parade, Parkville, 3050 Victoria, Australia

Warwick J. Britton Centenary Institute of Cancer Medicine and Cell Biology, and Department of Medicine, Central Clinical School, University of Sydney, Bld DO6, Sydney, NSW 2006, Australia

Jean-Laurent Casanova Laboratoire de Génétique Humaine des Maladies Infectieuses, Université de Paris René Descartes-INSERM U550, Faculté de Médecine Necker, 156 rue de Vaugirard, F-75015 Paris, France

Matthew Cook Australian National University Medical School, Department of Immunology, The Canberra Hospital, Canberra, Australia

Jason Cyster Howard Hughes Medical Institute and Department of Microbiology and Immunology, University of California, San Francisco, San Francisco, CA 94143-0414, USA

Simon Foote The Menzies Research Institute, 17 Liverpool Street, Hobart 7000, Tasmania, Australia

Anny Fortin Emerillon Therapeutics Inc., 416 de Maisonneuve West, Suite 1000, Montreal, H3A 1L2, Canada

Christopher C. Goodnow (Chair) John Curtin School of Medical Research, The Australian National University, Canberra, ACT 2061, Australia

Philippe Gros Department of Biochemistry, Center for the Study of Host Resistance, and Cancer Center, McGill University, Montreal, PQ, Canada H3G 1Y6
David Hume Institute for Molecular Biosciences, The University of Queensland, Brisbane, QLD 4072, Australia

Gunasegaran Karupiah Immunology and Genetics Division, John Curtin School of Medical Research, Building 54, Australian National University, Canberra, ACT 2000, Australia

Jim Kaufman Institute for Animal Health, Compton, Newbury, Berkshire RG20 7NN, UK

Kong-Peng Lam Laboratory for Immunology, Centre for Molecular Medicine, 20 Biopolis Way, #08-01 Centros, Singapore 138668

Charles Mackay Garvan Institute of Medical Research, 384 Victoria St, Darlinghurst, NSW 2010, Australia

Rick M. Maizels Institute of Immunology and Infection Research, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK

Bernard Malissen Centre d’Immunologie de Marseille-Luminy, INSERM-CNRS, Parc Scientifique et Technologique de Luminy—Case 906, F-13009 Marseille, France

Grant Morahan Western Australia Institute for Medical Research, Perth Campus, Level 6, MRF Building, Rear, 50 Murray Street, Perth, WA 6000, Australia

Anthony Scalzo Centre for Experimental Immunology, Lions Eye Institute, 2 Verdun St, Nedlands, WA 6009, Australia

Christian Schönbach Nanyang Technological University, School of Biological Sciences, Division of Genomics and Genetics, 60 Nanyang Drive, Singapore 637551

Steven Stacker Angiogenesis Laboratory, Ludwig Institute for Cancer Research, Post Office Royal Melbourne Hospital, Parkville, Victoria 3050, Australia

Stephen Turner Department of Microbiology and Immunology, The University of Melbourne, Parkville, Victoria 3010, Australia

Carola Vinuesa John Curtin School of Medical Research, The Australian National University, Canberra, ACT 2061, Australia
Edward K. Wakeland Center for Immunology, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235, USA

Douglas Young Centre for Molecular Microbiology and Infection, Division of Investigative Sciences, Flowers Building, Room 3.42, South Kensington Campus, Imperial College London, London SW7 2AZ, UK
Chair’s introduction

Chris Goodnow

*John Curtin School of Medical Research, The Australian National University, Canberra, ACT 2061, Australia*

There are lots of meetings in immunology—too many, probably—where we discuss all the latest hot topics. There are also lots of meetings on genomics. But what attracted me about this meeting is the unique format, with an equal balance between people presenting work and then the discussions that follow. This allows us to address some of the bigger issues. This is particularly important at this time in immunology. Rather than focusing on all the details, we have an opportunity to step back and try to identify the strategic gaps in the field.

We know the genome sequence is a tremendous toolkit. It is a parts list of the whole immune system. But we are increasingly becoming aware that it isn’t entirely clear which way is best for using this genome sequence to conquer some of the big issues in immunology that have been there for a long time. If we are going to achieve something in this meeting that will be different from any other kind of meeting, it will be to tease out some of the issues about what does work. What sorts of approaches are working for different people here? What kinds of conceptual issues are on the boil? Where are the conceptual gaps and technical gaps? These are the big strategic issues that the unique combination of people here will be well placed to addressed. I hope our discussions will tease out some of these issues.

Some of the issues that might be illuminated by the papers are as follows. What sequences in the genome encode different kinds of response, such as resistance to infection, memory, tolerance and immunopathology? There are issues about reductionist approaches, and the need to focus on individual genes to get at mechanism, versus the problem of trying to understand the system as a whole. What is ‘systems biology’, and how can we use this concept? There are issues of bottom–up approaches, starting with gene sequences and trying to work up to whole immune systems or organs, versus top–down approaches, of starting with populations of humans or animals and working down to mechanisms underlying variation in immune responses. I look forward to a good meeting.
Transcriptional regulatory networks in macrophages

David A. Hume*, Christine A. Wells† and Timothy Ravasi‡

*ARC Special Research Centre for Functional and Applied Genomics and CRC for Chronic Inflammatory Diseases, Institute for Molecular Bioscience, University of Queensland, Q4072, Australia, †School of Biomolecular and Biomedical Science, (The Eskitis Institute for Cell and Molecular Therapies) Griffith University, Nathan Campus Kessels Rd, Q4111, Australia and ‡Department of Bioengineering, Jacobs School of Engineering, University of California, San Diego. 9500 Gilman Drive, La Jolla, CA 92093-0412, USA

Abstract. The functions of macrophages in the innate immune system require the constitutive expression of a wide range of myeloid-specific genes, including various pattern recognition receptors, as well as the inducible expression of a suite of genes required to initiate inflammation and eliminate pathogens. Our overall aim is to understand the transcriptional networks that underlie both macrophage-specific transcription and the response to pathogen components such as lipopolysaccharide (LPS). The approaches used include detailed functional analysis of specific promoters, such as that of the CSF1 receptor, global cDNA microarray expression profiling, high throughput real-time PCR analysis of all the transcription factors encoded by the mammalian genome, full length cDNA library construction and sequencing, CAGE analysis to identify specific promoters used in macrophages and motif analysis to detect candidate cis-acting elements in co-regulated genes in macrophages. This review discusses some of the progress in moving towards a transcriptional network model for mouse macrophage activation by LPS, as well as insight into the role of alternative promoter usage and polyadenylation in generating functional protein variants that impact on signalling in macrophages.

2007 Decoding the genomic control of immune reactions. Wiley, Chichester (Novartis Foundation Symposium 281) p 2–24

The mononuclear phagocyte system

The mononuclear phagocyte system (MPS) is defined as a family of cells comprising bone marrow progenitors, blood monocytes and tissue macrophages. Macrophages are a major cell population in most of the tissues in the body, and their numbers increase further in inflammation, wounding and malignancy (Hume 2006, Hume et al 2002). Through their endocytic and cytotoxic activities, they provide a first line of defence against pathogens in the innate immune system. They contribute to antigen processing and presentation, and in turn their effector
functions are activated in response to products of the acquired immune system, notably antibodies (via Fc receptors), and T cell products (e.g. interferon γ and interleukin 4). Apart from their roles in immunity, macrophages also contribute to many aspects of homeostasis, vascularization, normal development, tumour progression and wound healing in part through their adaptation to recognize and remove cells undergoing apoptosis (Henson & Hume 2006, Lichanska & Hume 2000) and through their secretion of growth factors and proteolytic enzymes (Hume 2006, Hume et al 2002).

Mononuclear phagocytes as they appear in tissues share a number of features (Hume 2006, Hume et al 2002):

- Stellate morphology and ultrastructural evidence of endocytic activity observed by light and electron microscopy.
- Expression of histochemical markers such as non-specific esterase, lysosomal hydrolases and ecto-enzymes.
- Evidence of endocytic activity.
- The presence of cell surface proteins (such as F4/80, CD14 and CSF1R) defined by monoclonal antibodies.

A database of images of tissue macrophages detected using the F4/80 monoclonal antibody in all mouse organs can be found at www.macrophages.com. The precise anatomical location of these cells supports a physiological role for tissue macrophages in the development, structure and homeostasis of organs. Not all macrophages express the F4/80 marker. To find an alternative we have produced the MacGreen transgenic mouse lines, in which macrophage-restricted expression of an EGFP transgene is directed by the promoter of the CSF1R (c-fms) gene. The proliferation, differentiation and survival of macrophages is controlled by macrophage colony-stimulating factor (CSF-1), and the receptor is expressed in a myeloid-restricted manner. These mice provide a resource for many aspects of macrophage biology, including the isolation of macrophage populations from specific organs (MacDonald et al 2005, Sasmono et al 2003). A library of images showing the location of macrophages detected using the Fms-EGFP transgene is also available at www.macrophages.com.

Our ability to monitor this population of cells in real time and in vivo raises the interesting question of exactly how their numbers are monitored, controlled and replenished; the cellular ‘systems biology’ of the mononuclear phagocyte system. At least in the normal steady state, the availability of CSF-1 appears to be regulated via macrophage-dependent endocytosis and destruction, which provides a simple link between macrophage numbers and the production of cells by the marrow. Accordingly, the level of CSF-1 rises substantially in mice that lack the CSF-1 receptor (Dai et al 2002). The control is undoubtedly more complex than that; since CSF-1 expression is regulated in many different tissues (Sweet & Hume 2003).
and at least some macrophage populations may be controlled by local proliferation (Hume 2006). Furthermore, several other growth factors can regulate macrophage production from bone marrow (Hume 2006).

Mammalian cells are complex systems. Their functions require the appropriate interactions of millions of individual components in appropriate order in time and space. Although biological scientists commonly study cells as if they were static entities, individual cells never exist in a steady state. Like an entire multicellular organism, each individual cell is born by cell division, adapts to its environment to carry out a particular function that is determined by its genetically programmed response to that environment, and then dies. If all the components of a biological system interacted with each other, the complexity would be impenetrable. However, as in most complex systems, individual components within a cell interact and organize in defined modules that we recognize as structures, pathways and regulatory frameworks. The traditional approach to understanding cell biology has been to identify those individual modules, dissect their components, and then attempt to determine how they can be reassembled. Systems Biology can be considered an alternative approach, in which one actually measures all the components and their interactions within a biological system, and then assembles them into modules (Aderem 2005). The advantage of this approach is that it makes no assumptions; the disadvantage is that it is constrained by the scale of information that can be obtained.

To understand the function of a cell in perfect detail, we would need to know the identity of every macromolecule and metabolite, its location, abundance and chemical status across the lifetime of each cell. Since cells in a population influence each other’s behaviour in many ways, we would also need to know the same information about its neighbours; every other cell in the body! This is clearly not possible. However, the problem can be simplified if one takes the view that the entire control framework of a mammalian cell is encoded in its genome, and exerted through the selective production of RNA, the process of transcription. If we know the identity of every transcript produced by the genome and the way that transcription changes with time, we can infer the fundamental control pathways. Certainly, different tissues express distinct transcriptional profiles—transcriptional lineage markers—that can be used to identify sub-populations of cells in complex tissues, for example the stem cell populations profiled in the Stem Cell Gene Anatomy Project (see www.scgap.org; Challen et al 2005, 2004). The process of lineage differentiation is arguably the process of restricting the transcriptional network of a cell to those genes that define its function. Ultimately, the life of that cell can be viewed by a transcriptional network.

Transcription is itself controlled by RNA; the precise way this feedback control operates may be very complex. RNA may act directly on the induction, processing
or stability of another transcript. Non-coding RNAs are a major, regulated, output of the mammalian genome (Katayama et al 2005, Ravasi et al 2006). Alternatively, the RNA may encode a protein. That protein may participate directly in transcription control in the nucleus, or it may be a secreted protein that directly or indirectly initiates a signalling cascade that ultimately regulates transcription. Hence, the control system of biological systems is ultimately a network of linked transcriptional switches. The completion of multiple eukaryote genome sequences and the rapid progress towards definition of the complete transcriptional output (the transcriptome) has given us access to all of the components of the network. The capacity to analyse the complete transcriptional profile of cells, and the advent of technologies that allow us to capture protein/RNA and DNA interactions within this transcriptional context offers us the tools to identify the connectivities within the network, and ultimately, the control system.

There are many reasons why macrophages are an ideal cell type in which to apply a systems approach to transcription control in a mammalian context:

1. They can be obtained as reasonably homogeneous primary cell cultures; from peripheral blood monocytes in humans, or by cultivation of bone marrow cells or peritoneal lavage in laboratory animals.
2. There are cell lines available in both mouse and human that replicate many of the characteristics of primary macrophages and which can be transiently or stably transfected.
3. The range of mammalian biologies/pathways that can be studied using macrophages is substantial; they can be regarded in significant measure as the archetypal cell.
4. They alter their function in distinct ways in response to many different extracellular signals acting through multiple distinct signalling pathways, and most importantly these changes in the macrophage regulatory networks can be measured over time (Sester et al 2005, Rehli et al 2005, Schroder et al 2004, Wells et al 2003a).
5. There are substantial natural, or introduced, genetic influences on the function of macrophages (Wells et al 2003a, Beutler 2005, Beutler et al 2005, Fortier et al 2005). These provide perturbations of the network which can be analysed to establish connections between nodes.
6. Since macrophages are both our primary line of defence against pathogens and mediators of much of the pathology of infectious, inflammatory and malignant disease, a fundamental understanding of the way their function is controlled is likely to translate into rational development of human therapies.
7. Their differentiation from progenitor cells, which involves stable epigenetic changes, can also be studied in detail in vitro (Tagoh et al 2002).
**Macrophage activation and deactivation**

The life of a macrophage ‘family’ starts with the division of a pluripotent progenitor, and subsequent division of successive progeny associated with progressive transcription of genes that are required for the mature function of macrophages, and which distinguish the macrophage from other cell types. Individual cells receive further signals that direct them to adhere to endothelial cells, transmigrate, and adapt to whatever environment they encounter. If that environment is a site of inflammation, rather than a place of normal macrophage residence, the cell will be ‘activated’ to eliminate the challenge, be it dead cells or pathogens. Subsequently, the cell must be deactivated and either die or migrate out of the site to permit resolution and restoration of normal tissue architecture (or in some cases, a scar). This life story requires the macrophage to respond to hundreds, perhaps thousands of distinct signals produced by other cells, including other macrophages, as well as pathogen products. Because macrophages are able to cause damage to normal tissue, their arsenals must be unleashed only when required to deal with a pathogen or other emergency. This requires a very rigorous mechanism for self/non-self discrimination. This distinction is mediated in large measure by so-called pattern recognition receptors. The toll-like receptors (TLRs), which recognize cell wall components and nucleic acid molecules that are conserved amongst pathogen classes, are the archetypal pattern recognition receptors, but many other classes of receptors contribute to selective pattern recognition.

Of all the pathogen components, lipopolysaccharide (LPS, also known as endotoxin), a structural constituent of Gram-negative bacterial cell walls, has been studied in most depth. LPS action on macrophages requires the TLR4 receptor, and involves multiple signalling pathways leading to regulated transcription of thousands of genes. The system has been so widely studied that a survey of the literature alone can generate a list of hundreds of target genes discovered one at a time (Hume et al 2002). In considering the response of macrophages to LPS, it is important to identify a number of features that we feel are generic to all transcriptional networks and which provides challenges to design and interpretation:

1. The initiation response to LPS at a single cell level is, to some extent at least, all-or-nothing (Sester et al 1999, Sweet & Hume 1996). It is not clear whether individual signalling pathways linked to TLR4 have distinct thresholds, nor how a presumably analogue signal (extent of receptor occupancy) is translated into a digital outcome.

2. At some level, each gene has its own threshold for transcriptional activation. This can be seen in the observation that individual genes have LPS dose response curves that differ by orders of magnitude (Costelloe et al 1999).
Individual genes also have a degree of all-or-nothing transcriptional activity at the single cell level, determined in part by stable epigenetic modifications and in part by chance (Hume 2000).

No signal is received in isolation. Other signals alter the dose response curve, magnitude and nature of the LPS response. Three of particular importance are the growth factor, CSF1 (Sester et al 2005), the major T cell product interferon γ (Schroder et al 2004), and glucocorticoids, which are commonly used therapeutically to ablate macrophage activation and inflammation (Ogawa et al 2005).

The response of cells to any signal is influenced by individual genetic variation in the species (Wells et al 2003a). In the case of macrophage response to LPS and other pathogen-associated molecules, this difference is particularly prevalent because the nature of the macrophage response is under strong selection pressure from pathogens.

The response is a sequential cascade of transcriptional regulatory events. Individual genes (notably the proinflammatory cytokines) are induced transiently and then repressed (Wells et al 2003a). The transient nature of the transcriptional regulatory response to LPS (and any other pathogen challenge) is absolutely crucial to resolution.

The last point highlights an interesting feature of the control architecture of the macrophage transcriptional network. We have reviewed elsewhere the observation that many of the most LPS-inducible transcripts are actually not effectors of immunity, but feedback controllers that act at every level of the signalling pathway from the receptor to transcription, as well as repressors of the effectors (such as protease inhibitors and antiinflammatory cytokines) (Wells et al 2005). One might consider that this multilevel control exists to ensure robustness, but in fact mutation in any individual control leads to disordered or excessive macrophage activation. A prediction of the observation that the feedback control is itself inducible, and the system is sensitive to that feedback control, is that it will be relatively insensitive to inhibitors that act on the afferent pathways (Wells et al 2005).

**Transcriptome data sets and the macrophage transcriptional network**

Macrophages as a cell population provide one of the most complex sources of transcripts in any cell type. In fact, we have performed comparative transcriptional profiling of LPS-stimulated mouse macrophages compared to 17 day embryos and embryonic stem cells (Fig. 1). The macrophage and embryonic stem (ES) cell populations in combination cover the large majority of the protein-coding transcripts on a complex microarray platform. The complexity of macrophages as a source of transcripts was also exploited by the FANTOM consortium.
in its efforts to identify the full diversity of transcripts encoded by the mouse genome (Wells et al 2003b). Arising from that effort, we have recently analysed the impact of alternative initiation, splicing and polyadenylation on a selected part on the macrophage proteome, focusing specifically on signalling molecules that contribute to the LPS response (Wells et al 2006). This analysis built upon the original finding that LPS acts upon an internal promoter in the IRAK-2 signalling molecule to generate inducible expression of a natural feedback repressor, a component of control network mentioned above. Using splicing arrays, we confirmed the expression of many additional examples of variant forms of signalling molecules that could form part of the control architecture.

Table 1 contains a summary of the microarray studies on macrophages that are currently available in the public domain. Our own analysis has focused on a particular mouse cellular system. Bone marrow-derived macrophages cultured in CSF-1 are exposed to LPS across a time course from 1 h to 18–24 h. Our interest in this system is partly based upon an interest in interactions between the CSF-1 and LPS signalling pathways (Sester et al 1999, 2005, Sweet & Hume 2003). The system is interesting because LPS, in common with agonists of other TLRs, blocks CSF-1-induced proliferation, but substitutes for CSF-1 in maintaining survival (Sester et al 1999, 2005). Aside from extensive array profiling on a number of different platforms (Nilsson et al 2006, Wells et al 2006, 2003a, symatlas.gnf.org), and in a number of different mouse strains which produce radically different transcription profiles...

FIG. 1. A Venn diagram showing the overlap in transcription profiles between macrophages (BMM), embryonic tissues (E17.5dpc) and embryonic stem (ES) cells. RNA was prepared from bone marrow-derived macrophages grown in CSF-1 (BMM), a pool of 17 day mouse embryos, and embryonic stem cells, each from 129/Sv mice, and compared on the Mouse Compugen microarray platform. The diagram shows the numbers of genes expressed above the significant detection threshold. 1522 were not detected in any sample. Note that macrophages and ES cells each express genes that are unique to that cell population, and in combination, express the large majority of genes on the chip.
<table>
<thead>
<tr>
<th>Description</th>
<th>Database</th>
<th>Species</th>
<th>Platform</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time course of LPS-stimulated bone marrow derived macrophages</td>
<td><a href="http://www.macrophages.com">www.macrophages.com</a></td>
<td>Mouse</td>
<td>RIKEN 20K cDNA microarray</td>
<td>Wells et al 2003 (PMID: 12826024)</td>
</tr>
<tr>
<td>Bone marrow derived macrophages stimulated for 16h with interleukin (IL)4</td>
<td><a href="http://www.macrophages.com">www.macrophages.com</a></td>
<td>Mouse</td>
<td>NIA 15K mouse cDNA microarrays</td>
<td>Rehli et al 2005 (PMID: 15908341)</td>
</tr>
<tr>
<td>Time course of LPS-stimulated bone marrow derived macrophages</td>
<td><a href="http://www.macrophages.com">www.macrophages.com</a></td>
<td>Mouse</td>
<td>RIKEN 60K cDNA microarray</td>
<td>Nillson et al 2006 (PMID: 16698233)</td>
</tr>
<tr>
<td>Time course of LPS-stimulated bone marrow derived macrophages</td>
<td><a href="http://fantom3.gsc.riken.jp">http://fantom3.gsc.riken.jp</a></td>
<td>Mouse</td>
<td>CAGE libraries</td>
<td>Kawaji et al 2006 (PMID: 16381948)</td>
</tr>
<tr>
<td>GM-CSF matured human blood monocytes with or without LPS or glucocorticoid</td>
<td><a href="http://linkage.garvan.unsw.edu.au/public/microarrays/Arthritis_Inflammation/index.html">http://linkage.garvan.unsw.edu.au/public/microarrays/Arthritis_Inflammation/index.html</a></td>
<td>Human</td>
<td>Affymetrix GeneChip technology</td>
<td>N/A</td>
</tr>
<tr>
<td>Description</td>
<td>Database</td>
<td>Species</td>
<td>Platform</td>
<td>Reference</td>
</tr>
<tr>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------------------------</td>
<td>---------</td>
<td>------------------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>Systems Biology and Medicine Database (range of macrophage cell types)</td>
<td><a href="http://www.lsbm.org">www.lsbm.org</a></td>
<td>Human</td>
<td>Affymetrix GeneChip technology</td>
<td>N/A</td>
</tr>
<tr>
<td>Time course of <em>Salmonella typhimurium</em> induced gene expression on RAW264.7 macrophage cell line</td>
<td><a href="http://cmdr.ubc.ca/salmonellaarray">http://cmdr.ubc.ca/salmonellaarray</a></td>
<td>Mouse</td>
<td>Atlas mouse cDNA arrays</td>
<td>Rosenberger et al 2000 (PMID: 10820271)</td>
</tr>
<tr>
<td>Human monocytes exposed to a range of pathogens</td>
<td><a href="http://web.wi.mit.edu/young/pathogens">http://web.wi.mit.edu/young/pathogens</a></td>
<td>Human</td>
<td>Affymetrix GeneChip technology x</td>
<td>Nau et al 2002 (PMID: 11805289)</td>
</tr>
</tbody>
</table>
(Wells et al 2003a), we have also used this system in the context of analysis of the promoters of the mouse genome.

The transcriptional output of any cell is influenced by the interaction between nuclear proteins and DNA sequences in the vicinity of the gene (promoters, enhancers, repressors). The identification of promoters is dependent upon accurate determination of the transcription start site. Recently, this has become possible on a genome-wide scale using cap analysis of gene expression (CAGE), essentially high throughput 5′ RACE/SAGE, technology (Carninci et al 2006). This combined with sequences of full length cDNAs and other polling technologies (diTags) has enabled the first genome-wide view of transcription initiation. To expedite access to this information, we have created a portal at www.macrophages.com/bioinfoweb. For the analysis of macrophages, CAGE sequencing has been performed in multiple independent libraries across the time course of activation by CSF-1 and LPS, providing an index of relative promoter use. Such data highlight the fact that a large proportion of the protein-coding genes in the genome have multiple promoters expressed in different cell types and tissues, or, as in the case of IRAK2, regulated differently in the same cell (Carninci et al 2005, 2006). Figures 2 and 3 show examples of two key macrophage-expressed transcription factors, STAT1 and NF-κB (p105). In each case, the presence of alternative promoters allows the locus to generate alternative protein isoforms that may have quite distinct biological activities.

The CAGE data provide us with a survey of the set of transcription start sites (TSS) utilized specifically in macrophages. Importantly, CAGE tags also provide a measure of frequency of promoter usage—information lacking from normalized EST or FL-cDNA datasets. The sequences flanking these start sites can be extracted and analysed as a class for the presence of over-represented sequence motifs for broad classes of transcription factors. Not surprisingly, given the known architecture of TATA-less myeloid promoters (Rehli et al 1999, Ross et al 1998) such an analysis reveals over-representation of purine-rich motifs recognized by the Ets transcription factor family amongst constitutive macrophage-specific genes, as well sites for the LPS-inducible transcription factor NF-κB among inducible genes (Carninci et al 2006). A more penetrating analysis can be done using either experimentally-validated binding sites for known transcription factors (e.g. Transfac, Jaspar) or a priori motif detection (Nilsson et al 2006). The final piece in the puzzle is the identity of the transcription factors that are actually present and likely to bind the candidate motifs in the promoters.

The FANTOM project has also provided a comprehensive set of candidate transcriptional regulators encoded by the mouse genome. Transcription factors (TF) were amongst the most highly alternatively spliced class in the F3 dataset—many new variants of known TF have been discovered (Ravasi et al 2003, Kummerfeld & Teichmann 2006). Aside from known transcription factors, additional candidates