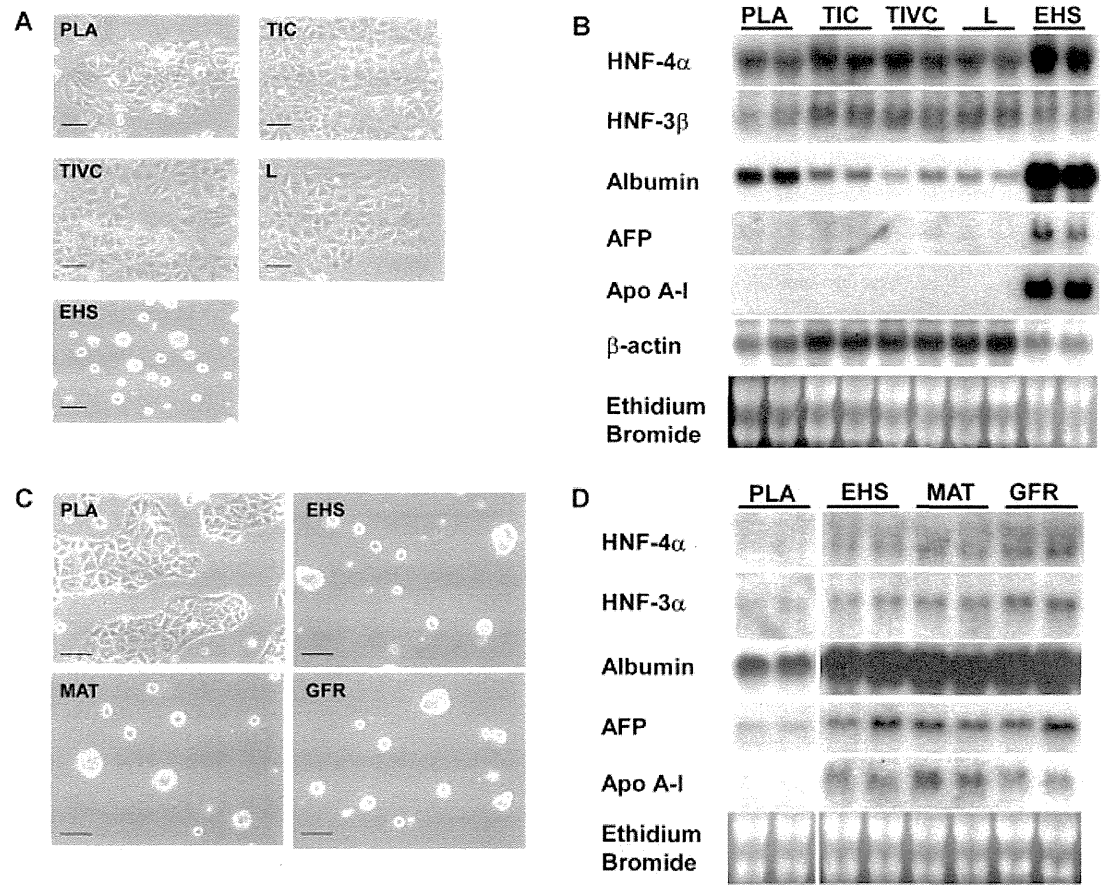


Fig. 6. Neither individual major ECM constituents of EHS-gel nor reduced growth factors mediated liver phenotypic expression of FLC-4 cells cultured on EHS-gel. FLC-4 cells were plated at 40% density and cultured for 48 h on uncoated (PLA), TIC-coated (TIC), type IV collagen-coated (TIVC), laminin-coated (L), and EHS-gel-coated (EHS) plastic dishes. Cell appearance and morphology are shown after 48 h. Scale bar = 100 μm (A). Total RNA was extracted, and mRNA expression levels of HNF-4α, HNF-3β, albumin, AFP, apo A-I, and β-actin were analyzed by northern blotting. Ethidium bromide-stained rRNA was used as a standard (B). FLC-4 cells were plated at 40% density and cultured for 48 h on uncoated (PLA), EHS-gel-coated (EHS), commercial EHS-gel (Matrigel)-coated (MAT), and GFR-EHS-gel-coated (GFR) plastic dishes. Cell appearance and morphology are shown. Scale bar = 100 μm (C). Total RNA was extracted, and mRNA expression levels of HNF-4α, HNF-3α, albumin, AFP, and apo A-I were analyzed by northern blotting. Ethidium bromide-stained rRNA was used as a standard (D).

A previous study demonstrated a central role for HNF-4α in liver differentiation, since introduction of HNF-4α to a hepatoma cell line induced redifferentiation (Hayashi et al., 1999). Furthermore, HNF-4α participates both in constitutive and drug-induced expression of CYP genes, such as CYP3A4, a major CYP expressed in human liver (Tirona et al., 2003). Therefore, it has been postulated that HNF-4α acts as a core transcription factor in the maintenance of liver function (Watt et al., 2003). Thus, we considered HNF-4α as a suitable marker for assessment of hepatocyte differentiation in this study. In FLC-4 and FLC-7 cells, the increased HNF-4α gene expression in cells grown on EHS-gel was thought to trigger expression of target genes that are involved in many hepatic functions (e.g., apo A-I) (Fraser et al., 1997; Nagaki and Moriwaki, 2008).

Our results showed that HNF-3α expression was also induced in FLC-4 and FLC-7 cells cultured on EHS-gel. The mRNA level of CYP3A7 was significantly elevated by growth on EHS-gel in FLC-4 cells. CYP3A7 is expressed abundantly in fetal liver, and its constitutive gene expression is regulated by HNF-3, as also reported for CYP3A4 (Saito et al., 2001). CYP3A4, which is abundantly expressed in adult liver, tended to increase in FLC-4 cells cultured on EHS-gel. Albumin expression and HNF-3α binding to albumin promoter were reported to be positively regulated by an ECM gel substratum (Caron, 1990; DiPersio et al., 1991). In particular, HNF-3α can transactivate the expression of albumin and AFP (DiPersio et al., 1991; Nagaki and Moriwaki, 2008). Increased expression of albumin, AFP, and CYP3A7 would therefore be attributable to the increased HNF-3α gene expression induced by culture on EHS-gel. In contrast, the gene expression of HNF-3β was downregulated in FLC-4 cells grown on EHS-gel (Fig. 3). Although HNF-3β is important in liver development, a study using HNF-3β conditional knockout mice revealed that expression of this gene is not necessary for maintaining liver phenotype (Sund et al., 2000). Finally, Sund et al. (2000) reported that HNF-3β knockout did not affect mRNA levels of apolipoproteins and albumin. We concluded that HNF-3β was not significant as a transcription factor for the maintenance of liver functions.

One of the most important findings of this study is that spherical FLC-4 cells showed highly differentiated functions similar to that of human liver (Fig. 4). Although monolayer FLC-4 cells were well-differentiated, conversion to a three-dimensional configuration similar to that of hepatocytes in the liver increased FLC-4 cell functions to human hepatocyte-like levels. This suggests that spherical cultured FLC-4 cells are suitable candidates for an artificial liver support system. Even in

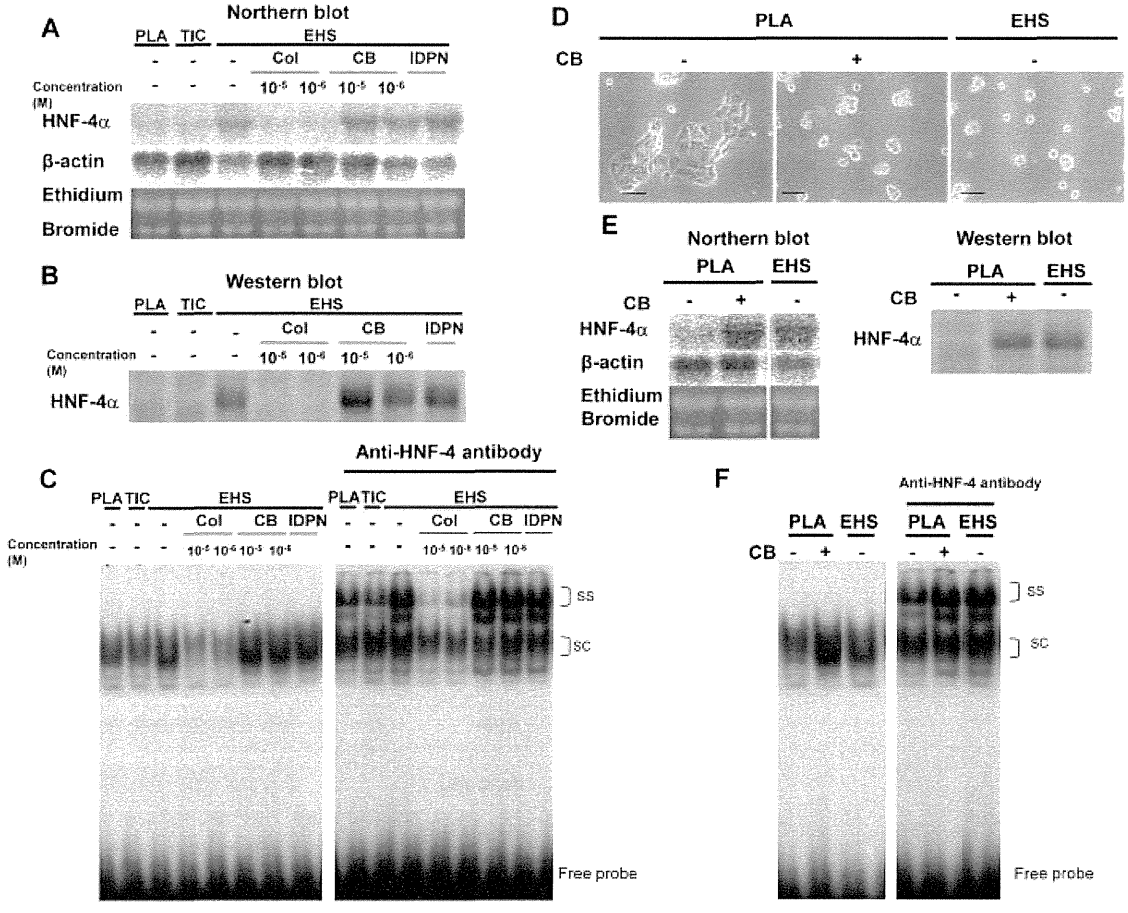


Fig. 7. Signals of cell shape were transmitted to the nucleus through microtubule organization in spherical FLC-4 cells. FLC-4 cells were plated at 40% density and cultured for 24 h on uncoated (PLA), TIC-coated (TIC), or EHS-gel-coated (EHS) plastic dishes. The cells cultured on EHS-gel were then treated with 10^{-5} or 10^{-6} M colchicine (col), 10^{-5} or 10^{-6} M cytochalasin B (CB), or 1% IDPN (w/v) for an additional 24 h. Cells cultured on plastic or TIC-coated dishes were not treated. Total RNA was extracted, and the HNF-4α and β-actin mRNA levels were investigated by northern blotting (A). The nuclear HNF-4α protein level was evaluated by Western blotting (B). EMSA was performed using a radiolabeled DR-1 double-stranded oligonucleotide (C). In EMSA experiments, a supershift assay using anti-HNF-4 antibody was also performed. "SS" and "SC" indicate supershifted HNF-4 complex and shifted complex, respectively. FLC-4 cells were plated at 40% density, and cultured for 24 h on uncoated (PLA) or EHS-gel-coated (EHS) plastic dishes, and then treated with 1×10^{-6} M cytochalasin B (CB) for an additional 24 h. The cell appearance and morphology after treatment with cytochalasin B are shown. Scale bar = 100 μm (D). Total RNA was extracted, and the HNF-4α and β-actin mRNA levels were evaluated by northern blotting. After nuclear protein extraction, the HNF-4α level was determined by Western blotting (E). EMSA was carried out using a radiolabeled DR-1 double-stranded oligonucleotide (F). A supershift assay was carried out with anti-HNF-4 antibody.

spherical FLC-4 cells, the C/EBPα mRNA level was relatively low (Fig. 3A). Transcription of the albumin gene is known to be regulated by C/EBPα (Friedman et al., 1989), which has also been reported to inhibit proliferation and to be one of the determinants of terminal differentiation of the liver (Friedman et al., 1989; Hayashi et al., 1999; Schrem et al., 2004; Nagaki and Moriwaki, 2008). We are now attempting to introduce the *C/EBPα* gene into FLC-4 cells to further enhance their hepatic phenotype.

An important question in the induction of liver phenotypic expression by growth on EHS-gel is which culture conditions significantly influence differentiation of FLC-4 cells. EHS-gel is well known to contain several ECM components as well as growth factors. EHS-gel contains laminin, TIVC, entactin, and heparin sulfateproteoglycan. Laminin and TIVC account for 87% of EHS-gel (56% and 31%, respectively) (Mannuzza, 1994). However, neither laminin nor TIVC alone affected liver-specific gene expression in FLC-4 cells (Fig. 6). These results suggested that spherical cell shape and/or combinations of ECM components, rather than any single ECM, were important for liver phenotypic expression. It is also well known that certain growth factors are able to induce cellular growth, proliferation, and differentiation (Jung et al., 1999; Suzuki et al., 2003). Fibroblast growth factor-2 (FGF2) has been shown to be an inducer during organogenesis; in particular, it promotes cells located in the medial mesoderm to differentiate into liver progenitors (Sturm et al., 2004). However, we noticed no difference in liver-specific gene expression between FLC-4 cells cultured on EHS-gel and GFR-EHS-gel (Fig. 6). Therefore, growth factors in EHS-gel may not be involved in EHS-gel-dependent induction of liver phenotypic expression. FLC-4 cells cultured on EHS-gel in a serum-free Dulbecco's Modified Eagle's medium, the cells were not able to grow and survive (unpublished data), indicating that the growth factors present in EHS-gel were not sufficient for promoting cell survival and liver differentiation. However, as GFR-EHS-gel still contains some growth factors, we could not exclude the possibility that these remaining growth factors were involved in EHS-gel-dependent induction of liver-specific gene expression. On the other hand, we successfully showed that spherical FLC-4 cells without actin

filament showed strongly enhanced gene expression of HNF-4α (Fig. 7D–F). This clearly indicates that EHS-gel is not required for the induction of HNF-4α and that spherical cell shape per se is a strong determinant of liver differentiation in cultured FLC-4 cells.

Our results also raise the question of how cell morphology signals reach the nucleus. ECM has been widely reported to cause reorganization of the cytoskeleton. The cytoskeleton consists of microtubules, actin filaments, and intermediate filaments: disruption of microtubules by colchicine completely suppressed the induction of gene expression and activity of HNF-4α in FLC-4 cells cultured on EHS-gel, while disruption of actin filaments or intermediate filaments had no effect on HNF-4α gene expression (Fig. 7). Microtubules are known to be critical cytoskeletal components that sustain cell polarity within the cells and are consequentially tightly linked to the maintenance of differentiated functions (Eng et al., 2006). This suggests that signals of spherical cell shape are transmitted through microtubules. The mechanism by which these signals are transmitted along microtubules to the nucleus and the nature of the molecules responsible for the signal transduction are of particular interest.

In conclusion, we have characterized a hepatoma cell line, FLC-4, that was capable of responding to a particular three-dimensional cell morphology by inducing HNF-4α, which is considered a central regulator of liver differentiation. We found that the enhanced gene expression in spherical FLC-4 cells was similar to that in human liver. Three-dimensional cell shape per se determined the liver phenotypic expression of FLC-4 cells. Our results indicated that microtubule organization was essential for coupling three-dimensional cell shape and the regulation of differentiation in the FLC-4 cell line. We do not yet understand how the physical information on cell morphology is transformed into a chemical signal, HNF-4α. We are currently investigating the changes in global gene expression profiles of three-dimensional FLC-4 cells.

References

Bhalla S, Ozalp C, Fang SS, Xiang LJ, Kemper K. 2004. Ligand-activated pregnane X receptor interferes with HNF-4 signaling by targeting a common coactivator PGC-1 alpha—Functional implications in hepatic cholesterol and glucose metabolism. J Biol Chem 279:45139–45147.

Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. Anal Biochem 72:248–254.

Caron JM. 1990. Induction of albumin gene transcription in hepatocytes by extracellular matrix proteins. Mol Cell Biol 10:1239–1243.

Chamuleau RA, Deurholt T, Hoekstra R. 2005. Which are the right cells to be used in a bioartificial liver? Metab Brain Dis 20:327–335.

Chomczynski P, Sacchi N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. Anal Biochem 162:156–159.

DiPersio CM, Jackson DA, Zaret KS. 1991. The extracellular matrix coordinately modulates liver transcription factors and hepatocyte morphology. Mol Cell Biol 11:4405–4414.

Eng CH, Huckaba TM, Gundersen GG. 2006. The formin mDia regulates GSK3beta through novel PKCs to promote microtubule stabilization but not MTOC reorientation in migrating fibroblasts. Mol Biol Cell 17:5004–5016.

Fraser JD, Keller D, Martinez V, SantisoMere D, Straney R, Briggs MR. 1997. Utilization of recombinant adenovirus and dominant negative mutants to characterize hepatocyte nuclear factor 4-regulated apolipoprotein AI and CIII expression. J Biol Chem 272:13892–13898.

Friedman AD, Landschulz WH, McKnight SL. 1989. CCAAT/enhancer binding protein activates the promoter of the serum albumin gene in cultured hepatoma cells. Genes Dev 3:1314–1322.

Hamilton GA, Jolley SL, Gilbert D, Coon DJ, Barros S, LeCluyse EL. 2001. Regulation of cell morphology and cytochrome P450 expression in human hepatocytes by extracellular matrix and cell–cell interactions. Cell Tissue Res 306:85–99.

Hascilowicz T, Kosuge M, Matsuura T, Matsufuji S, Murai N. 2005. Two-dimensional protein analysis of functional liver cells for bioartificial liver. Jikeikai Med J 52:109–114.

Hasumura S, Sujino H, Nagamori S, Kameda H. 1988. [Establishment and characterization of a human hepatocellular carcinoma cell line JHH-4]. Hum Cell 1:98–100.

Hayashi Y, Wang W, Ninomiya T, Nagano H, Ohta K, Itoh H. 1999. Liver enriched transcription factors and differentiation of hepatocellular carcinoma. Mol Pathol 52:19–24.

Homma S, Nagamori S, Fujise K, Hasumura S, Sujino H, Matsuura T, Shimizu K, Niiya M, Kameda H. 1990. [Establishment and characterization of a human hepatocellular carcinoma cell line JHH-7 producing alpha-fetoprotein and carcinoembryonic antigen—Changes in secretion of AFP and CEA from JHH-7 cells after heat treatment]. Hum Cell 3:152–157.

Iwahori T, Matsuura T, Maehashi H, Sugo K, Saito M, Hosokawa M, Chiba K, Masaki T, Aizaki H, Ohkawa M, Suzuki T. 2003. CYP3A4 inducible model for in vitro analysis of human drug metabolism using a bioartificial liver. Hepatology 37:665–673.

Jung JN, Zheng MH, Goldfarb M, Zaret KS. 1999. Initiation of mammalian liver development from endoderm by fibroblast growth factors. Science 284:1998–2003.

Mannuzza J. 1994. Growth factor reduced Matrigel basement membrane matrix: A matrix preparation with greatly reduced concentrations of growth factors and proteinases. Cell Line 4:1–6.

Marek CJ, Cameron GA, Elrick LJ, Hawksworth GM, Wright MC. 2003. Generation of hepatocytes expressing functional cytochromes P450 from a pancreatic progenitor cell line in vitro. Biochem J 370:763–769.

Nagaki M, Moriwaki H. 2008. Transcription factor HNF and hepatocyte differentiation. Hepatol Res 38:961–969.

Nagamori S, Hasumura S, Matsuura T, Aizaki H, Kawada M. 2000. Developments in bioartificial liver research: Concepts, performance, and applications. J Gastroenterol 35:493–503.

Oda H, Nozawa K, Hitomi Y, Kakinuma A. 1995. Laminin-rich extracellular matrix maintains high level of hepatocyte nuclear factor 4 in rat hepatocyte culture. Biochem Biophys Res Commun 212:800–805.

Oda H, Yoshida Y, Kawamura A, Kakinuma A. 2008. Cell shape, cell–cell contact, cell–extracellular matrix contact and cell polarity are all required for the maximum induction of CYP2B1 and CYP2B2 gene expression by phenobarbital in adult rat cultured hepatocytes. Biochem Pharmacol 75:1209–1217.

Rodriguez-Antona C, Bort R, Jover R, Tindberg N, Ingelman-Sundberg M, Gomez-Lechon MJ, Castell JV. 2003. Transcriptional regulation of human CYP3A4 basal expression by CCAAT enhancer-binding protein alpha and hepatocyte nuclear factor-3 gamma. Mol Pharmacol 63:1180–1189.

Saito T, Takahashi Y, Hashimoto H, Kamataki T. 2001. Novel transcriptional regulation of the human CYP3A7 gene by Sp1 and Sp3 through nuclear factor kappa B-like element. J Biol Chem 276:38010–38022.

Schrem H, Klempnauer J, Borlak J. 2004. Liver-enriched transcription factors in liver function and development. Part II: The C/EBPs and D site-binding protein in cell cycle control, carcinogenesis, circadian gene regulation, liver regeneration, apoptosis, and liver-specific gene regulation. Pharmacol Rev 56:291–330.

Sidhu JS, Liu F, Omiecinski CJ. 2004. Phenobarbital responsiveness as a uniquely sensitive indicator of hepatocyte differentiation status: Requirement of dexamethasone and extracellular matrix in establishing the functional integrity of cultured primary rat hepatocytes. Exp Cell Res 292:252–264.

Stoffel M, Duncan SA. 1997. The maturity-onset diabetes of the young (MODY1) transcription factor HNF4 alpha regulates expression of genes required for glucose transport and metabolism. Proc Natl Acad Sci USA 94:13209–13214.

Sturm J, Keese M, Zhang H, Bonninghoff R, Magdeburg R, Vajkoczy P, Dono R, Zeller R, Gretz N. 2004. Liver regeneration in FGF-2-deficient mice: VEGF acts as potential functional substitute for FGF-2. Liver Int 24:161–168.

Sund NJ, Ang SL, Sackett SD, Shen W, Daigle N, Magnuson MA, Kaestner KH. 2000. Hepatocyte nuclear factor 3beta (Foxa2) is dispensable for maintaining the differentiated state of the adult hepatocyte. Mol Cell Biol 20:5175–5183.

Suzuki A, Iwama A, Miyashita H, Nakauchi H, Taniguchi H. 2003. Role for growth factors and extracellular matrix in controlling differentiation of prospectively isolated hepatic stem cells. Development 130:2513–2524.

Tirona RG, Lee W, Leake BF, Lan LB, Cline CB, Lamba V, Parviz F, Duncan SA, Inoue Y, Gonzalez FJ, Schuetz EG, Kim RB. 2003. The orphan nuclear receptor HNF4alpha determines PXR- and CAR-mediated xenobiotic induction of CYP3A4. Nat Med 9:220–224.

Trauner M, Boyer JL. 2003. Bile salt transporters: Molecular characterization, function, and regulation. Physiol Rev 83:633–671.

Watt AJ, Garrison WD, Duncan SA. 2003. HNF4: A central regulator of hepatocyte differentiation and function. Hepatology 37:1249–1253.

Zamule SM, Strom SC, Omiecinski CJ. 2008. Preservation of hepatic phenotype in lentiviral-transduced primary human hepatocytes. Chem Biol Interact 173:179–186.

Conservation and divergence in Toll-like receptor 4-regulated gene expression in primary human versus mouse macrophages

Kate Schroder[a,1,2], Katharine M. Irvine[a,1,2], Martin S. Taylor[b,c,1], Nilesh J. Bokil[a], Kim-Anh Le Cao[a], Kelly-Anne Masterman[a], Larisa I. Labzin[a], Colin A. Semple[c], Ronan Kapetanovic[d], Lynsey Fairbairn[d], Altuna Akalin[e], Geoffrey J. Faulkner[d], John Kenneth Baillie[d], Milena Gongora[a], Carsten O. Daub[f], Hideya Kawaji[f], Geoffrey J. McLachlan[a,g], Nick Goldman[b], Sean M. Grimmond[a], Piero Carninci[f], Harukazu Suzuki[f], Yoshihide Hayashizaki[f], Boris Lenhard[h], David A. Hume[d,2], and Matthew J. Sweet[a,2]

[a]Institute for Molecular Bioscience, University of Queensland, Brisbane 4072, Australia; [b]European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Trust Genome Campus, Cambridge CB10 1SD, United Kingdom; [c]Medical Research Council Human Genetics Unit, Western General Hospital, Edinburgh EH4 2XU, United Kingdom; [d]Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Midlothian EH25 9RG, United Kingdom; [e]Department of Physiology and Biophysics and the Institute for Computational Biomedicine, Weill Cornell Medical College of Cornell University, New York, NY 10065; [f]RIKEN Omics Science Center, RIKEN Yokohama Institute, Kanagawa 230-0045, Japan; [g]Department of Mathematics, University of Queensland, Brisbane 4072, Australia; and [h]Medical Research Council Clinical Sciences Centre, Institute of Clinical Sciences, Imperial College London, London W12 0NN, United Kingdom

Edited by Shizuo Akira, Osaka University, Osaka, Japan, and approved February 24, 2012 (received for review June 23, 2011)

Evolutionary change in gene expression is generally considered to be a major driver of phenotypic differences between species. We investigated innate immune diversification by analyzing interspecies differences in the transcriptional responses of primary human and mouse macrophages to the Toll-like receptor (TLR)-4 agonist lipopolysaccharide (LPS). By using a custom platform permitting cross-species interrogation coupled with deep sequencing of mRNA 5′ ends, we identified extensive divergence in LPS-regulated orthologous gene expression between humans and mice (24% of orthologues were identified as "divergently regulated"). We further demonstrate concordant regulation of human-specific LPS target genes in primary pig macrophages. Divergently regulated orthologues were enriched for genes encoding cellular "inputs" such as cell surface receptors (e.g., TLR6, IL-7Rα) and functional "outputs" such as inflammatory cytokines/chemokines (e.g., CCL20, CXCL13). Conversely, intracellular signaling components linking inputs to outputs were typically concordantly regulated. Functional consequences of divergent gene regulation were confirmed by showing LPS pretreatment boosts subsequent TLR6 responses in mouse but not human macrophages, in keeping with mouse-specific TLR6 induction. Divergently regulated genes were associated with a large dynamic range of gene expression, and specific promoter architectural features (TATA box enrichment, CpG island depletion). Surprisingly, regulatory divergence was also associated with enhanced interspecies promoter conservation. Thus, the genes controlled by complex, highly conserved promoters that facilitate dynamic regulation are also the most susceptible to evolutionary change.

evolution | inflammation | innate immunity | pattern-recognition receptor | transcriptional regulation

Rodents are widely used as models for humans in experimental pathology. The assumption that mice provide a reliable model for understanding and treating human disease is a particular concern for immunological studies, as humans and mice diverged 65 to 75 Mya (1), and pathogens evolve quickly and exert strong pressure on the immune system to coevolve. Accordingly, many aspects of innate and acquired immunity are different between mouse and human (reviewed in ref. 2). High-profile commentaries have highlighted limitations of mice for immunological studies and emphasized the urgent need to identify critical points of divergence between the species (3, 4).

Macrophages and other cells of the innate immune system recognize distinct microbial products through families of pattern recognition receptors such as the Toll-like receptors (TLRs) (5). Although the TLR family exhibits divergence in expression, sequence, and gene number among species, the basic biology and downstream signaling is generally thought to be conserved. The most widely studied pattern recognition receptor is TLR4, which recognizes bacterial lipopolysaccharide (LPS) from Gram-negative bacteria, as well as several host-derived danger signals. TLR4 is essential for host resistance to a range of bacterial infections, but dysregulated responses through this receptor lead to systemic inflammation (e.g., septic shock). Some species differences in the transcriptional response to inflammatory stimuli such as LPS are already appreciated. For example, the inducible nitric oxide synthase (*iNOS*) gene, the product of which generates the antimicrobial agent nitric oxide, is robustly induced by IFN-γ and LPS in macrophages from mice, but not humans (6) or several other mammalian species (7). Conversely, the gene encoding the antimicrobial peptide cathelicidin is TLR-inducible via a vitamin D-dependent pathway in human but not mouse macrophages (8). Moreover, mice are relatively resistant to LPS-mediated toxicity (9), and many therapeutic agents that reduce inflammation and mortality in mouse septic shock models show no benefit clinically (10) or even increase mortality (e.g., nitric oxide antagonists) (11).

Multiple mechanisms can contribute to phenotypic differences between species. These include the evolution of nonorthologous genes, functional evolution of orthologous genes (e.g., changes in amino acid sequence affecting function), and evolution in transcriptional regulation of orthologous genes. Extensive gene ex-

Author contributions: K.S., K.M.I., M.S.T., D.A.H., and M.J.S. designed research; K.S., K.M.I., N.J.B., K.-A.M., L.I.L., R.K., and L.F. performed research; K.S., K.M.I., M.S.T., G.J.F., M.G., C.O.D., H.K., G.J.M., N.G., S.M.G., P.C., H.S., and Y.H. contributed new reagents/analytic tools; K.S., K.M.I., M.S.T., K.-A.L.C., C.A.S., A.A., J.K.B., B.L., D.A.H., and M.J.S. analyzed data; and K.S., K.M.I., M.S.T., D.A.H., and M.J.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE19492). The mRNA deep sequencing data reported in this paper is available for download from http://fantom.gsc.riken.jp/4/download/Supplemental_Materials/Schroder_et_al_2012/. The analyzed interspecies comparative data reported in this paper is available for perusal and interrogation at http://www.macgate.qfab.org.

[1]K.S., K.M.I., and M.S.T. contributed equally to this work.

[2]To whom correspondence may be addressed. E-mail: k.schroder@imb.uq.edu.au, katharine.irvine@uq.edu.au, david.hume@roslin.ed.ac.uk, or m.sweet@imb.uq.edu.au.

See Author Summary on page 5925 (volume 109, number 16).

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1110156109/-/DCSupplemental.

pression divergence between even closely related species clearly exists (12–18), but our understanding of how gene regulation evolves, and the impact of this on phenotypic diversity is limited. A number of mechanisms could contribute to expression divergence between species, including divergence in *cis* (e.g., promoter/enhancer sequence change) or in *trans* (e.g., altered transcription factor expression, activation or motif specificity, chromatin remodeling, or cellular environment). As with protein coding sequence, gene regulatory sequences evolve as a consequence of mutation and the subsequent action of genetic drift and natural selection. Adaptive evolution is particularly prevalent in immune response genes as a result of sustained selective pressure exerted by rapid pathogen evolution (19–21).

In the present study, we developed a custom platform that permits interspecies gene expression comparison and, in parallel, performed genome-wide transcription start site (TSS) identification by deeply sequencing mRNA 5′ ends on matched samples from human and mouse. This study represents the most comprehensive and detailed study to date of a gene expression time course, measured in parallel and anchored to experimentally defined promoters, in two mammalian species. In so doing, we systematically identified conserved and divergent TLR4-regulated gene expression, with demonstrated biological consequences, and defined promoter properties that are associated with evolutionary divergence in innate immune genes.

Results

LPS-Regulated mRNA Regulation in Human and Mouse Macrophages. Interspecies comparisons that use whole-genome arrays exaggerate transcriptional divergence because of low transcript coverage and poor internal replication. We therefore undertook a rigorous, two-step approach. We first interrogated whole-genome microarrays (Illumina; Affymetrix) to identify transcripts significantly regulated over a time course of LPS stimulation in mouse or human macrophages. LPS-regulated genes identified in this screen were then interrogated using custom-designed mouse and human arrays. These afforded a high level of internal replication, comprehensive coverage of transcript variants, and superior probe design to commercially available whole-genome arrays (Table S1).

We chose 2 h LPS stimulation as representative of direct response genes (validated for the early response gene, *TNF*), 6 h as representative of the peak of inflammatory gene responses, and 24 h to sample the subsequent "resolution" phase during which feedback control is exerted (22, 23). To minimize differences caused by cellular environment, we compared primary macrophage populations derived under very similar conditions ex vivo: human monocyte-derived macrophages (HMDMs) and mouse bone marrow-derived macrophages (BMMs). Both are differentiated ex vivo with colony-stimulating factor (CSF)-1 for 7 d. We also included a second mouse macrophage population, thioglycollate-elicited peritoneal macrophages (TEPMs), that are derived from mouse monocytes in vivo and are commonly used as a cellular model of inflammatory macrophages. Our choice of human and mouse macrophages reflects the most widely used primary macrophage populations in studies of innate immunity.

For one-to-one orthologues, we adopted a stringent approach to capturing interspecies divergence by selecting a "representative profile" for each gene (summarized in Fig. S1; see *SI Glossary of Technical Terms* for definitions of all technical terms). This is defined as the most similar expression profile between HMDM and either mouse macrophage population, for any transcript (or set of transcripts, wherein individual probes detect multiple transcripts) emanating from that gene. Probe, transcript, and representative profile data and statistical analyses are available from the Macrophage Comparative Expression Gateway (MacGate, http://www.macgate.qfab.org), and summary data are available in Dataset S1*A*. Although our study focused on conservation and divergence in LPS regulation between human and mouse genes displaying one-to-one orthology relationships, we also analyzed coregulation between genes with expanded orthology relationships (e.g., one-to-many, many-to-many) and found regulatory divergence between and within species was common, reflecting the trend for paralogous genes to evolve specialized functions (Fig. S2).

Conservation in LPS-Regulated Gene Expression in Human and Mouse Macrophages. The relative importance of selection vs. random genetic drift in shaping gene expression patterns is controversial (24). Under a simple model of neutral drift, we would expect extensive gene expression divergence between species, approaching saturation (24). In fact, gene regulation over the time course was strikingly similar between the human (i.e., HMDM) and both mouse (i.e., BMM, TEPM) macrophage populations (Fig. 1*A*). Permutation testing of summed normalized expression difference (SDiff, the sum of absolute differences between orthologous expression profile values over all time points; adapted from ref. 25)

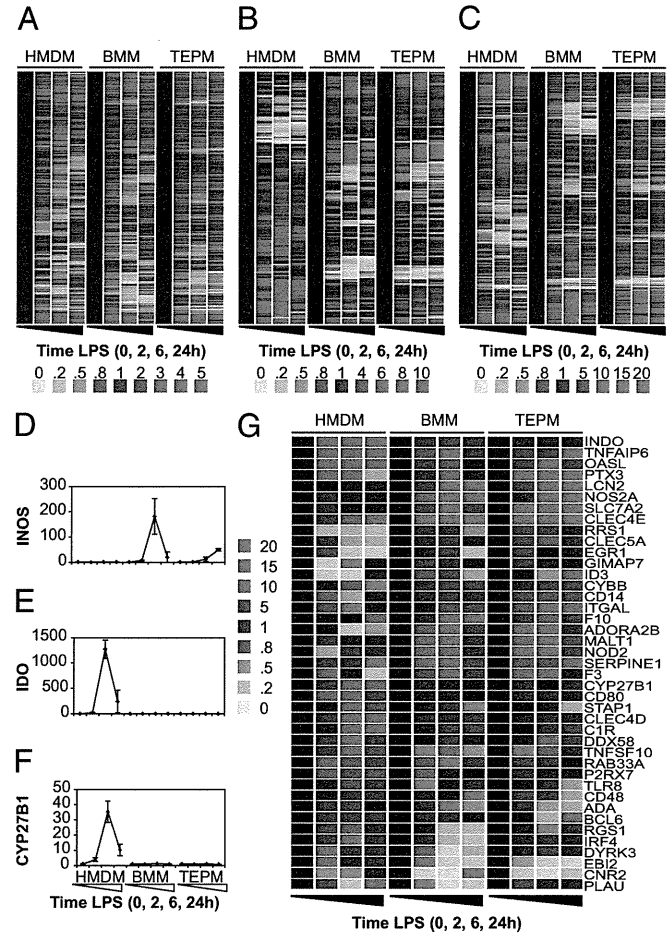


Fig. 1. Global overview of LPS regulation between human and mouse macrophage populations. Clustered representative profiles for (*A*) all one-to-one orthologues significantly regulated by LPS in any macrophage population, (*B*) all profile-divergent orthologues, and (*C*) all time point-divergent orthologues. Known examples of expression divergence between the species were confirmed, including (*D*) *INOS* ($P = 1.41 \times 10^{-7}$), (*E*) *IDO* ($P = 1.46 \times 10^{-24}$), and (*F*) *CYP27B1* ($P = 2.31 \times 10^{-7}$; microarray data average fold induction ± SEM, $n = 3$–4, significance assessed by ANOVA). (*G*) Hierarchical clustering of profile-divergent orthologues involved in immune defense or antimicrobial pathways. Genes encoding cytokine, chemokine, and growth factor ligands and receptors, as well as matrix metalloproteinases, were subtracted from this list as they are considered separately in Fig. 3*A*. Gene fold induction is indicated by the color-code key.

revealed that this similarity was significantly greater than would be expected by chance ($P < 0.05$) for 30.30% (759 of 2,505) regulated gene pairs. These results are consistent with the prevailing action of purifying selection constraining regulatory divergence. The overall similarities in the LPS-regulated mRNA expression profiles of the human and mouse macrophages support the view that the experimental systems are broadly comparable and that the transcription milieu triggered by LPS is also broadly conserved.

Identification of Interspecies Divergence Between Primary Human and Mouse Macrophages. Given the overall conservation in LPS responses, regulatory divergence should be viewed against a background of overall similarity. By using the stringent approach described earlier, representative profiles for 186 (7.4%) LPS-regulated genes, whose profile in HMDM was significantly different ($P < 0.05$, two-way ANOVA with variables of sample and time) to both mouse populations (i.e., BMM and TEPM), were classified as differentially regulated across the LPS time course ("profile divergent"; Fig. 1*B* and Fig. S1). To capture more subtle differences in gene regulation between the species, we also performed *t* tests to compare regulation at each time point ("time-point divergent"; Fig. 1*C*). A total of 23.9% (598 of 2,505) of regulated genes exhibited differential profile or time-point regulation between human and both mouse macrophage populations ["divergently regulated" (DR); Fig. S1 and Dataset S1*B*]. As expected, our analysis captured TLR4 target genes already known to be differentially regulated between human and mouse macrophages including *INOS/NOS2* (Fig. 1*D*), *IDO* (Fig. 1*E*), and *CYP27B1* (Fig. 1*F*). A number of genes known to be involved in immune defense and antimicrobial pathways showed profile divergence between species (Fig. 1*G* and Dataset S1*C*).

Macrophages exhibit a degree of plasticity with respect to phenotype and function. It was therefore conceivable that some of the differences attributed to species divergence actually reflected heterogeneity between different macrophage populations. Our choice of specific mouse macrophage populations was designed to minimize such effects. CSF-1–replete BMMs are a proliferating cell population, and LPS treatment causes growth inhibition. Consequently, LPS regulates gene expression programs associated with the cell cycle, as well as inflammatory and antimicrobial responses in BMMs. TEPMs are post-proliferative, so LPS does not cause growth inhibition in these cells. We therefore predicted overlapping but distinct LPS-regulated gene expression programs in these cells. Indeed, BMMs and TEPMs did exhibit striking differences in LPS-regulated gene expression (Fig. S3), as we previously documented for individual LPS target genes (26). Importantly, our analysis required that the HMDM profile was distinct from those of both mouse macrophage populations. This selection criterion enabled us to identify a wide spectrum of LPS target genes in the mouse, and to ensure that these genes were differentially regulated in the primary human macrophage population examined.

To confirm that the aforementioned strategy had successfully identified true species differences, we next tested expression of individual human-specific TLR4 target genes by quantitative PCR across an extensive panel of human and mouse primary cells and cell lines that were not used in our initial expression profiling analysis. The human-specific LPS target genes *CCL20*, *CXCL13*, *IL-7R*, *P2RX7*, and *STAT4* were all robustly up-regulated in the human monocyte/macrophage populations examined compared with the mouse macrophage populations (Fig. 2 *A–E*). In each of these cases, the lack of induction by LPS in mouse is supported by Cap-Analysis Gene Expression (CAGE) data from the Functional Annotation of the Mouse (FANTOM) 3 project (http://fantom3.gsc.riken.jp/) (27), as well as previous data with the use of the Affymetrix array platform. Importantly, not all DR genes showed an all-or-nothing response. Indeed, differences in the timing or the magnitude of the response were frequently observed (Fig. 1*C* and Fig. S1). In the case of the *P2RX7* gene,

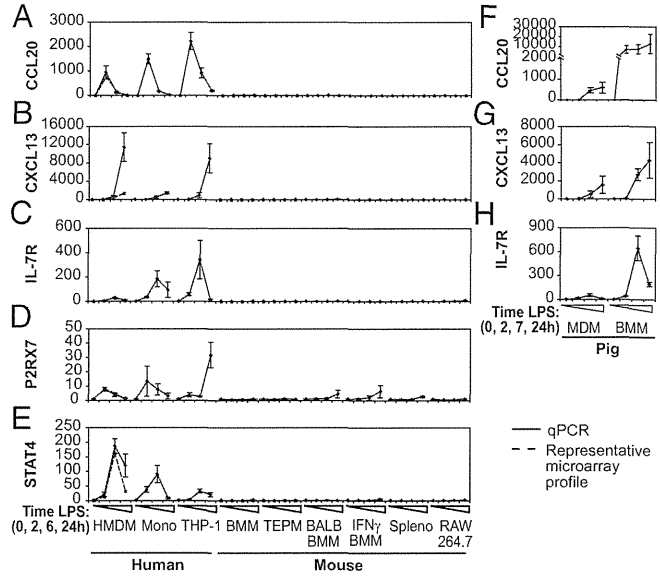


Fig. 2. Validation of human-specific TLR4 target genes and confirmed coregulation in porcine macrophages. LPS-regulated mRNA expression of selected human-specific TLR4 target genes was quantified by quantitative PCR in cells from human (HMDMs, primary $CD14^+$ monocytes, PMA-differentiated THP-1 macrophage-like cells) or mouse (BMMs, TEPMs, BMMs derived from BALB/c mice, C57BL/6 BMMs primed overnight with IFN-γ, C57BL/6 splenocytes, macrophage-like cell line RAW264.7). Human-specific LPS-regulated gene expression is displayed for *CCL20* (A), *CXCL13* (*B*), *IL-7R* (*C*), *P2RX7* (*D*), and *STAT4* (*E*). Quantitative PCR data represent average fold induction ± range (BALB/c BMMs, $n = 2$) or ± SEM ($n \geq 3$, all other profiles) of measurements with independent samples. Microarray data are shown only for HMDMs, BMMs, and TEPMs (dotted line, average fold induction ± SEM, $n = 3$–4). Expression of *CCL20* (*F*), *CXCL13* (*G*), and *IL-7R* (*H*) were further investigated by profiling LPS-dependent gene expression in MDMs and BMMs derived from pigs; data represent average fold induction ± SEM of measurements with three or four independently prepared samples.

for example, rapid up-regulation was apparent in all human myeloid populations (2 h after stimulation). In mouse cells, this early response was not apparent, but up-regulation did occur for some cell populations at a later time point (24 h after stimulation; Fig. 2*D*). To further ensure that apparent species differences were not merely a result of differences in the cellular systems used, we also sampled pig macrophages. In this species, it is possible to generate both BMMs and monocyte-derived macrophages (MDMs) in the presence of human CSF-1. LPS robustly induced *CCL20*, *CXCL13*, and *IL-7R* in MDMs and BMMs from pigs (Fig. 2 *F–H*). *IDO* and *STAT4* were also LPS-inducible in pig macrophages, whereas *iNOS* was not (28). Thus, regardless of their origin (MDM or BMM), pig macrophages resembled HMDMs rather than mouse BMMs or TEPMs in their TLR4-regulated responses, thus confirming, at least for these genes, that cell-type differences are not sufficient to explain the divergent regulation we observe.

Functional Impact of Regulatory Divergence in TLR4 Responses. We next assessed the functional consequences of species differences in TLR4 responses. A number of Gene Ontology terms were significantly overrepresented in DR orthologues (Table S2), many of which correspond to secreted or plasma membrane proteins (156 genes; Dataset S1*D*), including a number of matrix metalloproteinases and cytokine, chemokine, and growth factor ligands and receptors. Fig. 3*A* shows a heat map for profile divergent orthologues within this subset of secreted or plasma membrane proteins. Such DR orthologues include *CCL20*, *CXCL13*, *IL-7R*, and *P2RX7* (Fig. 2 *A–D*). We confirmed hu-

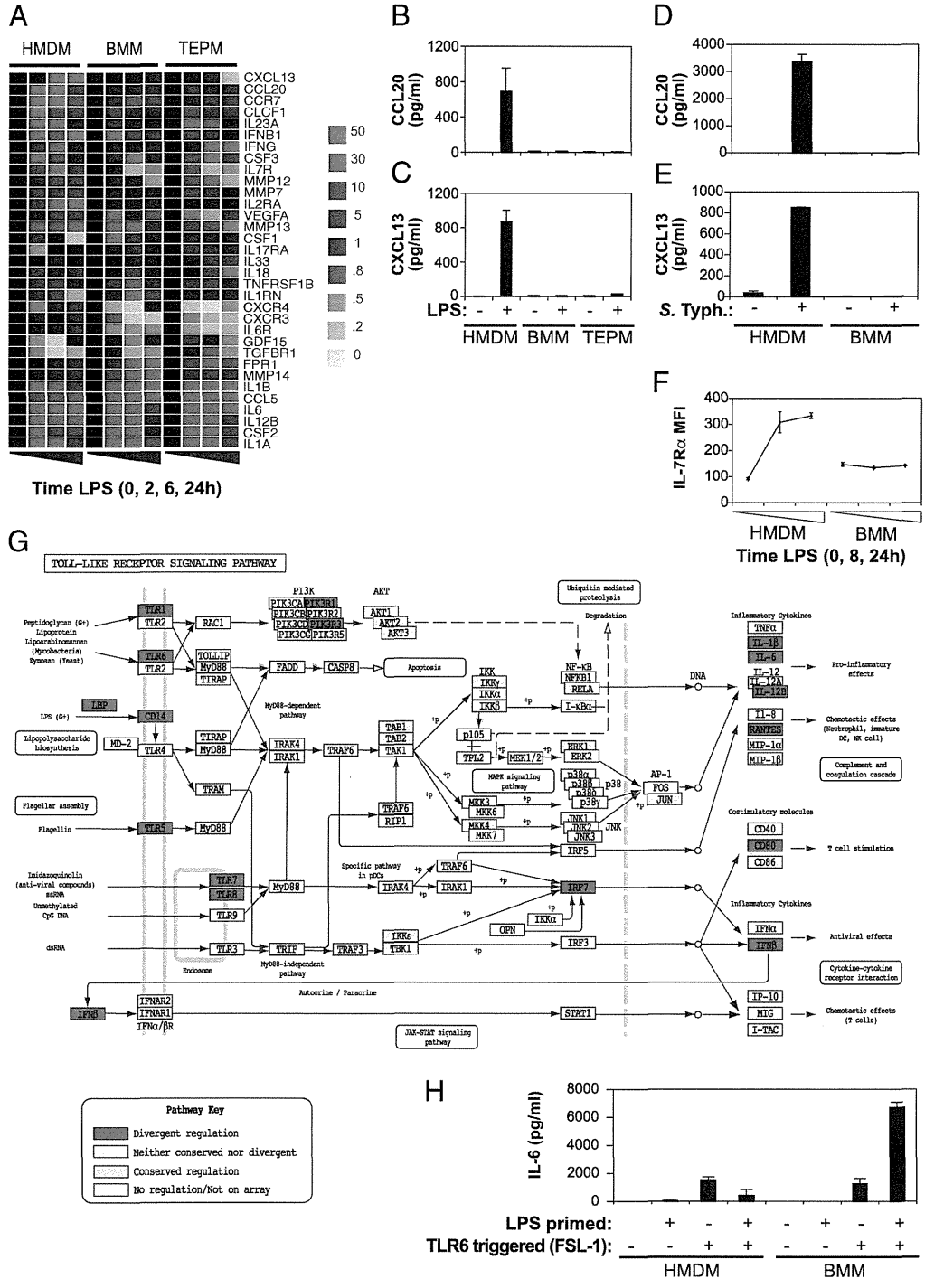


Fig. 3. Functional consequences of regulatory divergence in macrophage inputs and outputs. (*A*) Hierarchical clustering of DR orthologues that encode matrix metalloproteinases or cytokines, chemokines, and growth factors and their receptors. Gene fold induction is indicated by the color-code key. Human-specific production of the DR secreted proteins, CCL20 (*B* and *D*) and CXCL13 (*C* and *E*) in response to 24 h LPS treatment (*B* and *C*) or 24 h *S.* Typhimurium infection (*D* and *E*) was validated by ELISA. Data represent mean ± SD of technical replicates ($n = 4$, LPS-stimulated HMDMs or BMMs) or mean ± range of technical replicates ($n = 2$, LPS-stimulated TEPMs and *Salmonella*-infected HMDMs and BMMs). Profiles are representative of two to four experiments with independent cellular sources. (*F*) Cell-surface IL-7Rα/CD127 expression on HMDMs and BMMs over a time course of LPS stimulation was measured by flow cytometry [average mean fluorescence intensity (MFI) ± SD of $n = 4$ technical replicates, prepared, stimulated, and harvested in parallel]. Data are representative of two to five experiments with independent cellular sources. (*G*) The reference Kyoto Encyclopedia of Genes and Genomes pathway for TLR signaling was color-coded according to regulatory conservation (yellow) or divergence (profile or time point divergence; red) between species. LPS-regulated orthologues displaying neither striking conservation nor divergence are shaded in gray, and genes that were not significantly LPS-regulated or not represented on the focused microarray are white. This striking separation of DR genes according to pathway input/output vs. pathway core is highly unlikely to occur by chance ($P = 4 \times 10^{-7}$). (*H*) Macrophages were primed with LPS or medium for 11 h, washed, and restimulated with the TLR2/6 ligand FSL-1 for 8 h. IL-6 production was measured by ELISA. Data are average ± SD of three technical replicates, prepared, stimulated, and harvested in parallel, and are representative of three or four independent experiments.

man-specific secretion of CCL20 and CXCL13 in response to LPS (Fig. 3 *B* and *C*). Similarly, infection with the Gram-negative pathogen *Salmonella enterica* serovar Typhimurium (*S.* Typhimurium) triggered CCL20 and CXCL13 production in human but not mouse macrophages (Fig. 3 *D* and *E*). We also confirmed induction of cell surface expression of the cytokine receptor IL-7Rα (CD127) in HMDMs but not BMMs (Fig. 3*F*). In confirmation of the ontology enrichment analysis, annotation of the core TLR signaling pathway itself according to regulatory divergence highlights that signaling components were generally conserved in their regulation, whereas regulatory divergence was concentrated in cellular inputs and outputs (Fig. 3*G*). To examine potential regulatory divergence in TLR signaling and TLRs themselves, we interrogated a broad, hand-curated set of TLR-associated genes on the focused arrays. Twenty of these were DR, including TLRs (e.g., *TLR6*, *TLR8*) and several negative feedback regulators (e.g., *IRAK3*, *JDP2*, *SOCS1*, and *ATF3*; Fig. S4 *A–D* and Dataset S1*E*). The prevalence of regulatory divergence in immune receptors suggests that LPS may differentially affect sensitivity to a second immune challenge. Interspecies differences in ligand sensitization are likely to be important in the context of infection or inflammation in which multiple receptors are engaged. To verify this model, we selected the pathogen recognition receptor TLR6 for functional analysis. *TLR6* mRNA was up-regulated by LPS in mouse but not human macrophages (Fig. 3*G* and Fig. S4*A*; also see MacGate, http://www.macgate.qfab.org). As predicted, LPS priming boosted IL-6 production in response to the TLR2/TLR6 agonist FSL-1 in mouse macrophages, whereas this effect was not apparent in human macrophages (Fig. 3*H*).

Species Divergence in Transcription Factor Gene Regulation. LPS differentially regulated 49 transcriptional regulators between the species (Fig. S4*E* and Dataset S1*F*). For example, *HEY1* and *HESX1* were strikingly induced only in HMDMs (Fig. S4 *F* and *G*). As with mRNA regulation (Fig. 2*E*), STAT4 protein was markedly LPS-inducible in HMDMs and the human monocytic cell line THP-1, but not in BMMs (Fig. S4*H*). Such data imply that inducibly expressed, DR transcription factors may affect downstream gene expression, thus contributing to regulatory divergence. However, the extent of regulatory divergence between species declined with time when LPS-regulated genes were partitioned according to the time point of their peak regulation (DR genes were 10.5% of total regulated genes peaking at 2 h, 10.3% at 6 h, and 4% at 24 h after LPS stimulation; Fisher exact test, $P = 0.9101$, $P = 1.73 \times 10^{-5}$, and $P = 7.858 \times 10^{-4}$, respectively, for early/mid, mid/late, and early/late comparisons; χ^2 test, $P = 1.27 \times 10^{-4}$). This argues against inducible transcription factors acting on downstream genes as a major driver of regulatory divergence, although it is still possible that there is some contribution from this mechanism. Contraction in interspecies regulatory divergence across time could reflect differing mechanisms for feedback control of LPS-regulated gene expression, which ultimately achieves the same end (Fig. S4 *A–C*). It also suggests that evolution selects for the final LPS-induced cellular state (e.g., 24 h and beyond) rather than the specific mechanisms to achieve this state (2–6 h). Interestingly, the kinetics of peak LPS regulation appeared to be more rapid in HMDMs compared with mouse macrophages [502 human and 355 mouse profiles showed peak LPS regulation at the earliest (2 h) time point; Fisher exact test, odds ratio of 1.52, $P = 4.052 \times 10^{-8}$ for 2 h vs. not-2 h data partitions].

We also examined the time course of regulatory divergence between the mouse macrophage populations. As these cells are genetically identical, regulatory divergence must be mediated by differences in the basal or inducible *trans* environment. In stark contrast to the interspecies comparison, regulatory divergence between the two mouse macrophage populations modestly increased with time (1.27 and 1.65 fold for 2 h and 6 h vs. 24 h after LPS stimulation; Fisher exact test, $P = 0.08576$, $P = 1.42 \times 10^{-5}$, and $P = 0.1171$, respectively, for early/mid, mid/late, and early/late comparisons; χ^2 test, $P = 5.14 \times 10^{-5}$). The distinct temporal signature in inter- vs. intraspecies gene regulatory divergence further argues against a major contribution from macrophage heterogeneity in the species divergence signature. It also suggests that differences in *cis* (i.e., DNA sequence changes near the DR gene) and/or the *trans* environment (e.g., availability of specific transcriptional regulators) are likely to make a substantial contribution to interspecies regulatory divergence.

Paradox of Promoter Conservation and Expression Divergence. To examine the contribution of promoter sequence evolution to regulatory divergence between species, we precisely defined TSSs using genome-wide deep CAGE performed in parallel with expression profiling. This yielded 5.3 million human macrophage and 2.6 million mouse macrophage tags that define TSSs to single nucleotide resolution (Fig. S5 *A* and *B*; data available at http://fantom.gsc.riken.jp/4/download/Supplemental_Materials/Schroder_et_al_2012/). CAGE-defined TSSs were projected onto multiple genome sequence alignments of as many as nine placental mammals with deeply sequenced genomes (Fig. S5*C*). Each gene was assigned a primary TSS in each species, being the TSS with the most CAGE tags over the LPS time course, after correcting for depth of library coverage. We initially investigated the role of TSS switching in expression divergence (i.e., the primary TSS for a gene in one species is not orthologous to the primary TSS in the other species), but found that the DR gene set was not significantly enriched for genes showing evidence of TSS switching between species (odds ratios of 1.12 and 1.33 for profile and time-point divergent genes, respectively; Fisher test, $P > 0.2$ in both cases). Subsequent analyses focused on the stringently filtered set of 1,210 TSSs that occur in orthologous positions and represent the primary TSSs in human and mouse.

As we hypothesized that *cis*-regulatory sequence evolution would be a major driver of gene regulatory divergence, we analyzed sequence conservation in promoters of DR vs. non-DR genes. Several approaches confirmed a very surprising finding: promoters driving DR genes were demonstrably more conserved than non-DR promoters. Relative to non-DR genes, DR gene "TSS regions" (−5 kb to +5 kb) exhibited a greater fraction of human:mouse-aligning nucleotides than those of non-DR genes (Fig. 4 *A* and *B*). We next analyzed the extent of core promoter sequence divergence between species. DR gene core promoter substitution rates (R_{core}, −200 to −1 nt relative to TSS; further information is provided in *SI Glossary of Technical Terms*) were lower than non-DR substitution rates (Kolmogorov–Smirnov test, $D = 0.18$, $P = 3.4 \times 10^{-4}$; Table S3), and R_{core} was negatively correlated with expression divergence (Spearman correlation, $\rho = -0.15$, $P = 3.6 \times 10^{-6}$). One possible explanation for this finding was that DR genes are biased toward regions in the genome that are highly conserved. To address this possibility, we quantified evolutionary constraint on a per-promoter basis, correcting for the local neutral rate of evolution. We thus calculated R_{core} to the substitution rate in nearby interspersed repetitive elements (R_{ire}) to give ω_{core} (i.e., R_{core}/R_{ire}), a measure analogous to protein coding sequence ω (i.e., $R_{synonymous}/R_{nonsynonymous}$). The ω_{core} significantly negatively correlated with regulatory divergence (SDiff; Spearman $\rho = -0.13$, $P = 5 \times 10^{-6}$; Table S3), and the promoters of DR genes were significantly biased to lower values of ω_{core} (Table S3 and Fig. 4*C*). Thus, higher sequence conservation is associated with increasing regulatory divergence, even when the local substitution rate was accounted for. Even at a wider (1 Mb) interval, DR genes were significantly enriched for association with conserved noncoding regions (as defined in ref. 29) relative to non-DR genes (one-tailed Wilcoxon test, $P = 0.0038$). Together, these data indicate that regulatory sequences

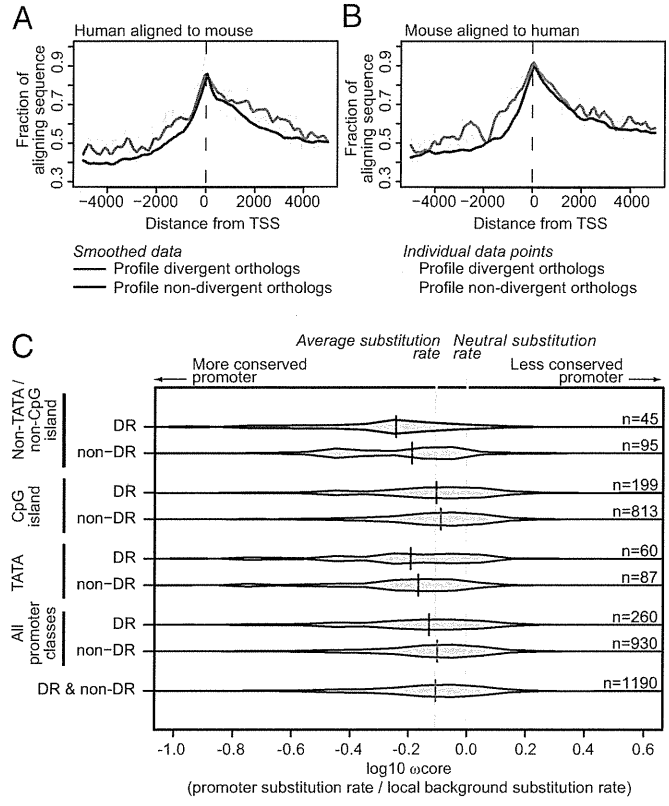


Fig. 4. The promoters of DR genes exhibit a higher fraction of aligning sequence and lower substitution rates than non-DR promoters. (*A* and *B*) Spatial profile of human/mouse aligning sequence. The fraction of human TSS region nucleotides aligned with mouse (*A*) and mouse TSS aligned with human (*B*) for profile-divergent genes (red) vs. non–profile-divergent subset (black), for which an orthologous macrophage TSS was defined in the other species (mouse or human) by CAGE. (*C*) Distributions of core promoter corrected substitution rate ratios [$\log10(\omega_{core})$] for primary TSSs that are orthologous between human and mouse (n = 1,190). $\log10(\omega_{core}) < 0$ indicates predominantly purifying selection; $\log10(\omega_{core}) > 0$ indicates predominantly diversifying selection. Also shown is the distribution median (vertical black/red bar), the fraction of all TSSs represented by the distribution (red to black fraction of vertical bar), and the median of all TSSs (pale red). Distributions are shown for specific promoter types (TATA, CpG island, neither TATA nor CpG island).

of DR genes are under greater evolutionary constraint than those associated with non-DR genes.

Despite the predominant trend for evolutionary constraint within DR gene promoters, the small population of core promoters with substitution rates significantly in excess of the neutral substitution rate ($\omega_{core} > 1$, <5% FDR), indicative of positive selection, were modestly enriched for DR genes (odds ratio = 1.4; Dataset S1*G*). Thus, for a small number of DR genes, positive selection on *cis*-regulatory sequences within the proximal promoter is likely to drive gene regulatory divergence. Conversely, a small population of non-DR genes also exhibited strong evidence of positive selection, demonstrating that rapid promoter evolution may not always lead to a measurable impact on gene expression, at least in our restricted sampling of cells and conditions.

Regulatory Divergence Is Associated with Specific Promoter Architecture. DR genes exhibited lower substitution rates 20 to 35 bases upstream of the TSS, and a series of pronounced peaks and troughs in sequence conservation downstream of the TSS (Fig. 4), which are hallmarks of TATA box-containing promoters (30). Indeed, promoters of DR genes were enriched for the presence of a TATA box. 17.2% of human and 14.8% of mouse DR gene promoters contained a TATA box, compared with 9.1% of human and 7.7% of mouse nondivergent gene promoters (Fisher test, $P = 1.2 \times 10^{-7}$ and $P = 4.8 \times 10^{-7}$ for human and mouse comparisons, respectively). To identify other features of DR genes, we also surveyed transcription factor binding sites (TFBSs) within proximal promoters. No position constraint was applied in this analysis, as recent evidence suggests that functional TFBSs are often not conserved across species (31). Although our approach identified TFBSs in DR and concordantly regulated genes that are commonly associated with LPS-regulated transcription (e.g., Rel, Stat, Ets), we did not find any strong association between specific TFBSs and regulatory divergence (Table S4). As described earlier, the one obvious exception was the TATA motif, which was more likely to be present in DR genes (odds ratio, 1.42; $P = 3.32 \times 10^{-3}$; Table S4). However, even when using this approach that applied no position constraint, the TATA box was present in fewer than 50% of DR genes. Thus, not all DR genes contain a TATA box in their promoter regions.

Profile divergent genes (Fisher test, $P = 2.67 \times 10^{-7}$ and $P = 9.69 \times 10^{-10}$ for human and mouse comparisons, respectively) and time point-divergent genes (Fisher test, $P = 4.58 \times 10^{-7}$ and $P = 1.39 \times 10^{-7}$ for human and mouse comparisons, respectively) were also significantly depleted of 5′ CpG islands, compared with nondivergent genes. TATA enrichment and CpG island depletion were not only associated with the degree of regulatory divergence (ranking by SDiff or ANOVA P value), but also the extent to which gene expression changes over the LPS time course (dynamic range, calculated as the difference between maximum fold induction and maximum fold suppression over the time course; Fig. 5 *A* and *B*). Conversely, genes exhibiting a relatively small dynamic range and low interspecies divergence were enriched for CpG island promoters and depleted of TATA promoters. TATA enrichment and CpG island depletion are clearly not redundant signals, as conditioning on one does not remove the effect of the other (Fig. 5 *C* and *D*). A gene's dynamic range, as well as its maximum LPS-dependent regulation (i.e., induction or suppression from unstimulated), was clearly associated with regulatory divergence, such that highly regulated genes were enriched for regulatory divergence (Fig. 5 *E* and *F*). In total, these results indicate that TATA-containing and non-CpG island promoters confer a high degree of inherent transcriptional plasticity (i.e., the capacity of a gene to change its transcription level under different environmental or evolutionary conditions), as TATA-containing and non-CpG island promoters were associated with both a large magnitude of response to LPS and expression divergence between species. TATA boxes tend to be associated with slowly evolving promoter sequences, and CpG islands with rapidly evolving promoter sequences (30), so it is important to note that the greater sequence constraint observed in DR gene promoters (Table S3) held true even within a promoter class (Fig. 4*C*). Thus, the effects observed are not just a manifestation of differences in promoter architecture between DR and non-DR genes, but rather are a general property of DR macrophage gene promoters. These distinct properties ultimately validate our approach for identifying DR genes.

Discussion

Some differences between mouse and human LPS responses at the whole-organism level are already documented (9). In this study, we demonstrate differential gene regulation in response to LPS in human and mouse macrophages that is strongly biased toward genes encoding cell surface receptors and secreted molecules. Individual examples of such DR genes were further examined in porcine BMMs and MDMs, in which LPS profiles for both macrophage populations reflected human-specific up-regulation, tempting speculation that the short generation time of

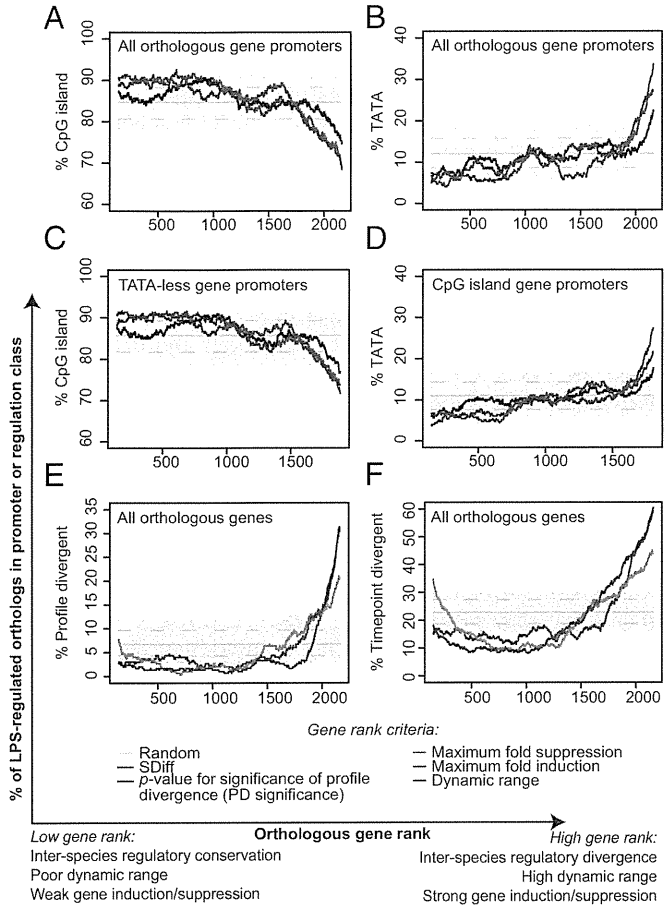


Fig. 5. DR genes are enriched for TATA box promoters, are depleted of CpG islands, and tend to exhibit a large dynamic range of expression in response to LPS. (*A* and *B*) Orthologous genes were ranked based on variation between species (SDiff or *P* value for significance of profile divergence; high rank is high significance and low *P* value) or dynamic range, such that genes with a high variation within or between species achieve a high rank, and analyzed according to percentage of genes with (*A*) CpG island-type orthologous gene promoters or (*B*) TATA box-containing orthologous gene promoters. (*C* and *D*) analyses were performed as in *A* and *B*, but considering only orthologues without a TATA box in the promoter (*C*) or only those that had CpG island-containing promoters (*D*). (*E* and *F*) Genes were ranked based on dynamic range, maximum fold suppression, or maximum fold induction and analyzed for percentage of DR genes according to (*E*) profile divergence or (*F*) time-point divergence. For *A*–*F*, 100 randomly permuted rankings of the data are shown (gray curves; nonrandom rankings protruding from the gray curves represent $P < 0.01$). Also indicated are confidence intervals (95%, dashed gray lines; 99%, dotted gray lines) and average for the whole dataset (solid gray lines). Sliding 300-gene windows are shown.

mice may accelerate regulatory evolution in this species. The prevalence of regulatory divergence in cell surface and secreted proteins predicts that differences in LPS regulation underpin divergent cellular and physiological functional responses to LPS, or may function to sensitize the cell to further stimuli that would be present during in vivo inflammation. We verified this at the cellular level by examining downstream responses to one DR "input," TLR6 (Fig. 3*H*). Notably, more than 150 other genes encoding inputs and outputs were also differentially regulated, which suggest further differences in macrophage functional responses to inflammation/infection in humans and mice, and/or the mechanisms used to achieve them.

Although it is known that humans are physiologically more sensitive than mice to LPS (9), the contribution of particular DR genes to this phenomenon remains to be established. Our study used a maximal stimulatory dose of LPS to elicit all LPS-regulated transcriptional responses; thus, our data do not yield a direct comparison of LPS sensitivity between human and mouse macrophages. However, our data are suggestive of mechanisms contributing to differential LPS sensitivity between species. We observed more rapid peak regulation kinetics in human vs. mouse macrophages, suggesting that HMDMs may be somewhat "primed" for immediate/early LPS responses, which in turn suggests increased sensitivity. We also observed more rapid and greater induction of several negative feedback regulators of the TLR pathway in mouse macrophages (IRAK3/IRAK-M and JDP2; Fig. S4 *B* and *C*). Enhanced feedback regulation may dampen the primary LPS response in mouse macrophages, thereby contributing to the relatively reduced endotoxin sensitivity observed in mice compared with humans. Elevated expression of some of the "human-specific" TLR4 target genes, for example IL-7R (32) and CXCL13 (33), has been linked to pathological human macrophage inflammatory responses. Similarly, the chemokine CCL20, which was specifically induced in human macrophages by LPS or whole pathogen infection in our study, is clearly linked to a range of human autoimmune diseases by virtue of its ability to attract pathogenic Th17 cells to a site of in vivo inflammation (34). Furthermore, polymorphisms in several DR genes, including coding sequence variants in *P2RX7* (35) and a promoter polymorphism in *IL7R* (36), have been associated with infectious and inflammatory diseases. Such data support the contention that DR genes are critical components of the human innate immune system, and are likely to contribute to pathological inflammatory processes.

Although some differences between species in antimicrobial gene regulation have previously been documented, it is clear from our study that multiple genes within such pathways were coregulated in a species-specific fashion. For example, mouse-specific LPS induction of *Inos* (Fig. 1*D*) paralleled that of mRNAs encoding proteins regulating iNOS function (the arginine transporter *Slc7a2* and the degradative enzyme arginase; also see MacGate, http://www.macgate.qfab.org). Conversely, *IDO* and *KYNU*, which encodes an enzyme downstream of IDO in the tryptophan metabolic pathway, were coinduced in a human-specific manner (Fig. 1*E*; also see MacGate, http://www.macgate.qfab.org). Importantly, our approach also identified previously unrecognized regulatory divergence in other antimicrobial pathways, such as *Lcn2* (mouse-specific) that encodes a secreted protein interfering with iron acquisition by bacterial siderophores (37), and two genes specifically induced in human macrophages (P2X7 receptor, *P2RX7*, and CXCL13, Fig. 2 *B* and *E*) that are required for effective clearance of *Mycobacterium tuberculosis* (38, 39). Given the evolutionary arms race between host and pathogen, we postulate that many other DR genes are likely to have key roles in direct antimicrobial responses of macrophages.

Despite significant divergence in gene expression profiles, LPS-regulated gene expression in human and mouse macrophages was far more similar than would be expected if expression diverges randomly by neutral drift alone. This is consistent with prior expectation and previous studies on other systems suggesting that gene regulation is often conserved by purifying selection (13, 18, 40, 41). Furthermore, expression divergence declined over time, suggesting that, even in the face of extensive species differences at early time points, the two biological systems ultimately reached a level of commonality, at least at the cellular level. This in turn suggests that evolution selects for the final LPS-reprogrammed cellular state (at 24 h and beyond) rather than the means to achieve this state. Thus, robustness in innate immune pathways may ultimately enable a level of functional conservation in at least some biological responses.

Recent expression quantitative trait locus studies have highlighted the importance of *cis*-regulatory changes, and particularly

promoter region polymorphisms, in determining gene expression level within the human population (42, 43). The strongest expression quantitative trait locus effects are preferentially clustered around gene TSSs. We thus hypothesized that gene expression divergence would be mirrored to some extent by sequence divergence in *cis*-regulatory sequences, particularly in core promoters. Indeed, numerous promoter polymorphisms have been identified that alter the expression of inflammation-associated TLR4 target genes (e.g., *IL1*, *IL6*, *TNF*, *IL10*) (44). This was indeed the case for a small number of DR genes, which exhibited significantly faster than neutral substitution rates in their promoters. These genes appear to possess very rapidly evolving promoters, suggestive of sustained or repeated rounds of diversifying selection. Notable among them is *DDX58*, encoding the protein RIG-I, an RNA helicase that recognizes viral RNAs and is a key regulator of antiviral innate immunity (45). However, we consistently found the opposite of the expected trend in the data set as a whole; interspecies regulatory divergence was associated with constraint in promoter sequences. This was seen in the per-promoter substitution rates (including the ω estimates that correct for local background effects), the fraction of nucleotides aligning between species, and the association of highly conserved proximal noncoding regions with DR genes.

DR genes are significantly enriched for TATA-containing promoters. Such promoters are known to possess a higher density of regulatory motifs relative to other promoters (46). This likely explains the observation that such promoters evolve rapidly in their transcriptional regulation (14), despite their propensity for enhanced conservation at the nucleotide level. The higher levels of promoter constraint we observed for DR relative to non-DR genes also applied across promoter "categories"; DR genes with TATA promoters have more constrained promoter regions than non-DR genes with TATA promoters, and DR genes have more constrained *cis*-regulatory sequences for the CpG island promoter and non-CpG island promoter data partitions (Fig. 4*C*). Although enhanced promoter sequence conservation seems to be a generalized feature of DR genes, it is important to note that not all DR genes are characterized by the presence of a TATA box and/or lack of CpG islands. Indeed, when applying a position constraint on the motif, only 17.2% of human and 14.8% of mouse DR genes contained a TATA box. *STAT4* and *CCL20*, for example, were both induced in a human-specific fashion (Fig. 2 *A* and *E*), yet only the latter is a TATA-containing promoter in human, pig, and mouse. In some cases, differences in the actual TATA box may directly contribute to expression divergence. The *IL-7Rα* promoter, for example, is a TATA-containing promoter in human and pig, but not mouse, which is consistent with its human/pig-specific induction by LPS.

Dynamic range was also strongly correlated to the extent of regulatory divergence. The most regulated genes in each species were among the least likely to be similarly regulated in the other. A large dynamic range in mRNA expression in response to a TLR agonist indicates a gene is highly sensitive to changes in the *trans*-regulatory environment. It follows, then, that such genes might also be sensitive to the many alterations to the *trans* environment concomitant with species divergence, consistent with previous studies in yeast (14, 17, 46). Our data support such a model for higher eukaryotes, and we extend this model by showing the connection between greater transcriptional plasticity (i.e., highly dynamic, evolvable transcriptional responses) and higher levels of promoter constraint. Consider a promoter that responds to multiple *trans*-regulatory inputs. The integration of more inputs suggests more *cis*-regulatory sites, both in distant enhancers and in the core promoter, where these signals are ultimately assimilated for gene expression. A greater number of *cis*-regulatory sites leads to a higher fraction of constrained sites and promoter/*cis*-regulatory sequence conservation. Such a gene is also likely to be hypersensitive to evolutionary perturbation of the *trans*-acting milieu or specific *cis*-regulatory sites.

LPS induces transcriptional changes in hundreds of genes and their products, indicative of complex gene regulatory networks (22, 23, 47). Widespread divergence in TLR-regulated transcription factor expression between mouse and human (Fig. S4 *E–H*) predicts species-specific transcriptional networks leading to regulatory divergence. There are some correlative suggestions of this from our analysis. For example, *HEY1*, which was specifically induced by LPS in human macrophages (Fig. S4*F*), was recently identified as a negative feedback regulator of *IL6* in LPS-activated human macrophages (48). In our study, peak *IL6* induction in HMDMs was significantly lower and faster than in either BMMs or TEPMs (MacGate, http://www.macgate.qfab.org), implicating this transcription factor in feedback regulation in humans but not mice. Similarly, *STAT4* was LPS-inducible in human but not mouse macrophages (Fig. 2*E* and Fig. S4*H*), as was the known STAT4 target gene, *IRF4* (49) (MacGate, http://www.macgate.qfab.org). Nonetheless, regulatory divergence significantly contracted as the LPS time course proceeded. This finding argues against a major input from inducible *trans* factors to regulatory divergence, although we cannot discount the possibility that they make some contribution. Instead, we suggest that different sets of LPS-inducible transcription factors in human vs. mouse enable conserved regulation of late response genes and/or that DR inducible feedback regulators ultimately limit functional divergence over time. It is also possible that many of the TLR4-inducible transcription factors are not involved in secondary LPS responses, but rather are induced to permit responses to other stimuli in the in vivo setting. Importantly, the finding that divergence is most pronounced in primary LPS responses suggests changes in *cis* and/or the basal *trans* environment are the major drivers of regulatory divergence. We have provided evidence for the former, but it is likely that the latter also contributes.

In summary, we have performed the most stringent human: mouse comparison of innate immune gene regulation reported to date to the best of our knowledge. Many LPS-responsive genes in mouse and human macrophages are conserved in their regulation, but we also identified and validated considerable divergence in TLR4 responses between these species. We suggest this regulatory divergence is likely to be one contributing factor to phenotypic differences between species in innate immunity. Importantly, our findings show that transcriptional plasticity in response to a stimulus within species and variability between species are intrinsically linked. The capacity for expression variation in both cases is associated with specific promoter properties (high promoter conservation, TATA box enrichment, and CpG island depletion). Paradoxically, differences in transcriptional regulation between species may be taken as evidence that the absolute level of the transcript, and its protein product, is crucial for innate immunity. By extension, functional analyses of TLR4 target genes uniquely regulated in human macrophages are likely to provide important insights into human infectious and inflammatory diseases. Furthermore, by revealing control points not already used in humans, the alternative evolutionary path trodden by the mouse may also provide opportunities for therapeutic intervention in such diseases.

Materials and Methods

A summary of the methods used is presented in the subsequent sections; further details are provided in *SI Materials and Methods*.

Ethics Statement. Before undertaking the studies described, approvals for all experiments using primary human and mouse cells were obtained from the relevant University of Queensland Ethics Committee.

Preparation of Primary Macrophages. HMDMs were differentiated in vitro from $CD14^+$ monocytes isolated from peripheral blood mononuclear cells using MACS technology (Miltenyi Biotech), and BMMs were derived from bone marrow progenitors. HMDMs and BMMs were differentiated from progenitors in complete media containing recombinant CSF-1 for 7 d. Mouse TEPMs were harvested by PBS solution lavage 5 d after i.p. injection of thioglycollate broth. All mice were C57BL/6 unless otherwise stated.

Design, Hybridization, and Signal Processing of Custom Microarrays. By using in-house and public access microarrays, we identified LPS-regulated genes in HMDMs, BMMs, or TEPMs. We then designed human and mouse custom 15,000 probe microarrays (Agilent Technologies) targeting genes regulated by LPS in any of these macrophage populations, for which orthology could be predicted with high confidence. RNA from three independent cell preparations (BMMs, TEPMs) or four independent biological replicates (HMDMs, in which each replicate RNA was a pool of RNA from two independent blood donors; $n = 8$ donors in total) was analyzed by microarray.

Custom Microarray Data Analysis. Processed signal intensities were quantile normalized in R using the Affymetrix package in Bioconductor, and median-centered. Because of the high level of internal replication for the quantification of transcript expression as a result of multiple probes targeting the same transcript(s), probe-level data were collapsed into transcript-level data if profiles from multiple probes were highly correlated. For one-to-one orthologues, all transcript-level profiles significantly regulated by LPS were compared with all transcript-level profiles for the orthologous gene. A two-way ANOVA with two fixed effects (time and species) was fitted to normalized, log10 data averaged over the replicates, and representative profiles were chosen by taking the least divergent profile pair (highest Benjamini–Hochberg corrected P value in either the HMDM/BMM or HMDM/TEPM comparison). If this P value was below a threshold of 0.05, the orthologous genes were classified as exhibiting significant profile divergence between the species. A second statistical analysis was performed on the representative profiles for human and mouse to identify more subtle differences in expression at single time points (more than threefold difference in expression and Student two-tailed t test P value <0.05). All P values for time point or profile divergence calculated were corrected for multiple testing using the Benjamini–Hochberg method. Expression profile similarity was further assessed by using SDiff (25), calculated as the sum of absolute differences between orthologous expression profile values over all time points. Concordantly regulated genes were defined as those with SDiff scores lower than 0.8 and lower than 0.6 (corresponding to $P < 0.05$ for randomly permuted data) for HMDM/BMM and HMDM/TEPM profile comparisons, respectively. All microarray data (probe-level, transcript profile-level, and representative human/mouse profiles) are freely available through MacGate (http://www.macgate.qfab.org), and are also accessible via the National Center for Biotechnology Information Gene Expression Omnibus repository (GSE19492 SuperSeries containing human and mouse data for Agilent custom and Illumina whole-genome microarrays). Custom microarrays used in this study are freely available for order from Agilent Technologies (probe files available via GSE19492).

TSS Identification and Evolutionary Analyses. CAGE libraries from HMDM (0, 2, 6, 24 h LPS) and BMM (0, 6 h LPS) total RNA were prepared as described and subjected to deep sequencing to identify TSSs (50). BMM deep CAGE data were merged with comparable data from FANTOM3 for LPS-stimulated BMMs (Fig. S5) (27). CAGE tags were mapped to genome assemblies (51) (Hg18, Mm9) and mapped to probe sets via Ensembl and Refseq transcripts. CAGE data are freely accessible through FANTOM (http://fantom.gsc.riken.jp/4/download/Supplemental_Materials/Schroder_et_al_2012/). By this method, at least one TSS was defined for 89% (2,238 of 2,505) of human and 92% (2,297 of 2,505) of mouse genes on the targeted microarray (Fig. S5). The EPO-9 whole genome alignment dataset from Ensembl was used for coordinate transformations between species to extract orthologous sequence alignments from the reference genomes for available eutherian mammals and for the calculation of substitution rates.

TATA Box and CpG Island Mapping. CAGE-defined TSSs were defined as "TATA-TSS" if they had a match to the Jaspar TATA box matrix (TFMscan, options: –l, –p, 6), with the 5′ end of the match located 25 to 35 nt upstream of the TSS reference position (the interval that shows TATA box enrichment over the whole dataset). Genes were defined as "TATA-genes" if they had one or more TATA-TSS. Mouse (mm9) and human (hg18) CpG islands were downloaded from the University of California, Santa Cruz, genome table browser, and CpG island genes were defined as those with a CpG island overlapping their core promoter (–300 to +100 with respect to the TSS reference position).

ACKNOWLEDGMENTS. We thank the Institute for Molecular Bioscience Special Research Centre Microarray Facility for performing the microarray hybridizations, as well as David Wood and Dr. Cas Simons (Queensland Facility for Advanced Bioinformatics, University of Queensland) for building the MacGate Web site. This work was supported by C. J. Martin Fellowship 490993 from the Australian National Health and Medical Research Council (to K.S.); Australian Research Council under the ARC Centres of Excellence program (ARC Centre of Excellence in Bioinformatics, K.-A.L.C.); Australian Research Council Future Fellowship FT100100657 (to M.J.S.); honorary Australian National Health and Medical Research Council Senior Research Fellowship APP1003470 (to M.J.S.); Australian National Health and Medical Research Council Grant 631531, a research grant to the RIKEN Omics Science Center from the Ministry of Education, Culture, Sports, Science and Technology, Japan (MEXT); and an Innovative Cell Biology by Innovative Technology Project grant from MEXT (to Y.H.), which supported Cap Analysis of Gene Expression data collection and mapping.

1. Waterston RH, et al.; Mouse Genome Sequencing Consortium (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature* 420:520–562.
2. Mestas J, Hughes CC (2004) Of mice and not men: Differences between mouse and human immunology. *J Immunol* 172:2731–2738.
3. Davis MM (2008) A prescription for human immunology. *Immunity* 29:835–838.
4. Hayday AC, Peakman M (2008) The habitual, diverse and surmountable obstacles to human immunology research. *Nat Immunol* 9:575–580.
5. Medzhitov R (2007) Recognition of microorganisms and activation of the immune response. *Nature* 449:819–826.
6. Schneemann M, et al. (1993) Nitric oxide synthase is not a constituent of the antimicrobial armature of human mononuclear phagocytes. *J Infect Dis* 167:1358–1363.
7. Jesch NK, et al. (1997) Expression of inducible nitric oxide synthase and formation of nitric oxide by alveolar macrophages: An interspecies comparison. *Environ Health Perspect* 105(suppl 5):1297–1300.
8. Liu PT, et al. (2006) Toll-like receptor triggering of a vitamin D-mediated human antimicrobial response. *Science* 311:1770–1773.
9. Copeland S, Warren HS, Lowry SF, Calvano SE, Remick D; Inflammation and the Host Response to Injury Investigators (2005) Acute inflammatory response to endotoxin in mice and humans. *Clin Diagn Lab Immunol* 12:60–67.
10. Poli-de-Figueiredo LF, Garrido AG, Nakagawa N, Sannomiya P (2008) Experimental models of sepsis and their clinical relevance. *Shock* 30(suppl 1):53–59.
11. López A, et al. (2004) Multiple-center, randomized, placebo-controlled, double-blind study of the nitric oxide synthase inhibitor 546C88: Effect on survival in patients with septic shock. *Crit Care Med* 32:21–30.
12. King MC, Wilson AC (1975) Evolution at two levels in humans and chimpanzees. *Science* 188:107–116.
13. McCarroll SA, et al. (2004) Comparing genomic expression patterns across species identifies shared transcriptional profile in aging. *Nat Genet* 36:197–204.
14. Tirosh I, Weinberger A, Carmi M, Barkai N (2006) A genetic signature of interspecies variations in gene expression. *Nat Genet* 38:830–834.
15. Khaitovich P, et al. (2004) A neutral model of transcriptome evolution. *PLoS Biol* 2: E132.
16. Liao BY, Zhang J (2006) Evolutionary conservation of expression profiles between human and mouse orthologous genes. *Mol Biol Evol* 23:530–540.
17. Landry CR, Lemos B, Rifkin SA, Dickinson WJ, Hartl DL (2007) Genetic properties influencing the evolvability of gene expression. *Science* 317:118–121.
18. Barreiro LB, Marioni JC, Blekhman R, Stephens M, Gilad Y (2010) Functional comparison of innate immune signaling pathways in primates. *PLoS Genet* 6:e1001249.
19. Christophides GK, et al. (2002) Immunity-related genes and gene families in Anopheles gambiae. *Science* 298:159–165.
20. Kosiol C, et al. (2008) Patterns of positive selection in six Mammalian genomes. *PLoS Genet* 4:e1000144.
21. Obbard DJ, Welch JJ, Kim KW, Jiggins FM (2009) Quantifying adaptive evolution in the Drosophila immune system. *PLoS Genet* 5:e1000698.
22. Gilchrist M, et al. (2006) Systems biology approaches identify ATF3 as a negative regulator of Toll-like receptor 4. *Nature* 441:173–178.
23. Nilsson R, et al. (2006) Transcriptional network dynamics in macrophage activation. *Genomics* 88:133–142.
24. Bedford T, Hartl DL (2009) Optimization of gene expression by natural selection. *Proc Natl Acad Sci USA* 106:1133–1138.
25. Lemos B, Bettencourt BR, Meiklejohn CD, Hartl DL (2005) Evolution of proteins and gene expression levels are coupled in Drosophila and are independently associated with mRNA abundance, protein length, and number of protein-protein interactions. *Mol Biol Evol* 22:1345–1354.
26. Lattin JE, et al. (2008) Expression analysis of G protein-coupled receptors in mouse macrophages. *Immunome Res* 4:5.

27. Carninci P, et al.; FANTOM Consortium; RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group) (2005) The transcriptional landscape of the mammalian genome. *Science* 309:1559–1563.
28. Kapetanovic R, et al. (2012) Pig bone marrow-derived macrophages resemble human macrophages in their response to bacterial lipopolysaccharide. *J Immunol*, 10.4049/jimmunol.1102649.
29. Prabhakar S, Noonan JP, Pääbo S, Rubin EM (2006) Accelerated evolution of conserved noncoding sequences in humans. *Science* 314:786.
30. Taylor MS, et al. (2006) Heterotachy in mammalian promoter evolution. *PLoS Genet* 2: e30.
31. Schmidt D, et al. (2010) Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. *Science* 328:1036–1040.
32. Pickens SR, et al. (2011) Characterization of interleukin-7 and interleukin-7 receptor in the pathogenesis of rheumatoid arthritis. *Arthritis Rheum* 63:2884–2893.
33. Carlsen HS, Baekkevold ES, Morton HC, Haraldsen G, Brandtzaeg P (2004) Monocyte-like and mature macrophages produce CXCL13 (B cell-attracting chemokine 1) in inflammatory lesions with lymphoid neogenesis. *Blood* 104:3021–3027.
34. Comerford I, et al. (2010) An immune paradox: How can the same chemokine axis regulate both immune tolerance and activation?: CCR6/CCL20: A chemokine axis balancing immunological tolerance and inflammation in autoimmune disease. *Bioessays* 32:1067–1076.
35. Fernando SL, et al. (2007) A polymorphism in the P2X7 gene increases susceptibility to extrapulmonary tuberculosis. *Am J Respir Crit Care Med* 175:360–366.
36. Gregory SG, et al.; Multiple Sclerosis Genetics Group (2007) Interleukin 7 receptor alpha chain (IL7R) shows allelic and functional association with multiple sclerosis. *Nat Genet* 39:1083–1091.
37. Flo TH, et al. (2004) Lipocalin 2 mediates an innate immune response to bacterial infection by sequestrating iron. *Nature* 432:917–921.
38. Britton WJ, Fernando SL, Saunders BM, Sluyter R, Wiley JS (2007) The genetic control of susceptibility to Mycobacterium tuberculosis. *Novartis Found Symp* 281:79–89.
39. Khader SA, et al. (2009) In a murine tuberculosis model, the absence of homeostatic chemokines delays granuloma formation and protective immunity. *J Immunol* 183: 8004–8014.
40. Sun Y, et al. (2007) Cross-species transcriptional profiles establish a functional portrait of embryonic stem cells. *Genomics* 89:22–35.
41. Chaix R, Somel M, Kreil DP, Khaitovich P, Lunter GA (2008) Evolution of primate gene expression: drift and corrective sweeps? *Genetics* 180:1379–1389.
42. Dimas AS, et al. (2009) Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science* 325:1246–1250.
43. Pickrell JK, et al. (2010) Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature* 464:768–772.
44. Smith AJ, Humphries SE (2009) Cytokine and cytokine receptor gene polymorphisms and their functionality. *Cytokine Growth Factor Rev* 20:43–59.
45. Yoneyama M, et al. (2004) The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat Immunol* 5:730–737.
46. Tirosh I, Barkai N (2008) Two strategies for gene regulation by promoter nucleosomes. *Genome Res* 18:1084–1091.
47. Amit I, et al. (2009) Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. *Science* 326:257–263.
48. Hu X, et al. (2008) Integrated regulation of Toll-like receptor responses by Notch and interferon-gamma pathways. *Immunity* 29:691–703.
49. Lehtonen A, et al. (2005) Differential expression of IFN regulatory factor 4 gene in human monocyte-derived dendritic cells and macrophages. *J Immunol* 175: 6570–6579.
50. Maeda N, et al. (2008) Development of a DNA barcode tagging method for monitoring dynamic changes in gene expression by using an ultra high-throughput sequencer. *Biotechniques* 45:95–97.
51. Faulkner GJ, et al. (2008) A rescue strategy for multimapping short sequence tags refines surveys of transcriptional activity by CAGE. *Genomics* 91:281–288.

Reconstruction of Monocyte Transcriptional Regulatory Network Accompanies Monocytic Functions in Human Fibroblasts

Takahiro Suzuki[1], Mika Nakano-Ikegaya[1☯], Haruka Yabukami-Okuda[1☯], Michiel de Hoon[1], Jessica Severin[1], Satomi Saga-Hatano[1], Jay W. Shin[1], Atsutaka Kubosaki[1], Christophe Simon[1], Yuki Hasegawa[1], Yoshihide Hayashizaki[1,2], Harukazu Suzuki[1]*

1 Omics Science Center (OSC), RIKEN Yokohama Institute, Yokohama, Kanagawa, Japan, **2** Division of Genomic Information Resources, Supramolecular Biology, International Graduate School of Arts and Sciences, Yokohama City University, Yokohama, Kanagawa, Japan

Abstract

Transcriptional regulatory networks (TRN) control the underlying mechanisms behind cellular functions and they are defined by a set of core transcription factors regulating cascades of peripheral genes. Here we report SPI1, CEBPA, MNDA and IRF8 as core transcription factors of monocyte TRN and demonstrate functional inductions of phagocytosis, inflammatory response and chemotaxis activities in human dermal fibroblasts. The Gene Ontology and KEGG pathway analyses also revealed notable representation of genes involved in immune response and endocytosis in fibroblasts. Moreover, monocyte TRN-inducers triggered multiple monocyte-specific genes based on the transcription factor motif response analysis and suggest that complex cellular TRNs are uniquely amenable to elicit cell-specific functions in unrelated cell types.

Citation: Suzuki T, Nakano-Ikegaya M, Yabukami-Okuda H, de Hoon M, Severin J, et al. (2012) Reconstruction of Monocyte Transcriptional Regulatory Network Accompanies Monocytic Functions in Human Fibroblasts. PLoS ONE 7(3): e33474. doi:10.1371/journal.pone.0033474

Editor: Raffaella Bonecchi, Università degli Studi di Milano, Italy

Received June 19, 2011; **Accepted** February 15, 2012; **Published** March 13, 2012

Funding: This work was supported in by a research grant for RIKEN Omics Science Center from MEXT to Yoshihide Hayashizaki, a grant of the Innovative Cell Biology by Innovative Technology (Cell Innovation Program) from the MEXT, Japan to Yoshihide Hayashizaki, Strategic Programs for R&D (President's discretionary fund) (super immune cell) for Yoshihide Hayashizaki, and the ENCODE project from National Institutes of Health (NIH). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: rgscerg@gsc.riken.jp

☯ These authors contributed equally to this work.

Introduction

A transcriptional regulatory network (TRN), in part, defines the functional properties of a cell-type [1]. Therefore, the ability to reengineer cellular TRNs would provide vast opportunities to treat patients suffering from impaired function of physiologically important cell types such as pancreatic β-cells [2]. However, methods to identify and induce cell-specific networks are still immature and further exploration into gene regulation is much required.

Numerous attempts to induce cell reprogramming have been reported in recent years [3,4,5,6,7,8]. Cell reprogramming, however, does not project a natural biological process, and approaches to identify a set of defined factors often deemed *hit-or-miss* or lack supporting evidence for selecting the candidate genes to induce reprogramming. Additionally, it is unknown whether every cell type can be reprogrammed. In contrast, cells reengineered to implement cell-specific function(s) would be of great value since rectifying a specific cellular function is often preferred compared to a complete replacement of impaired cells [9].

We previously described a TRN of differentiating THP-1 (human acute myeloid leukemia) cell line [1]. Utilizing two transcriptome analyses, sequencing-based CAGE [10] and microarray technologies, we characterized the dynamic regulatory activities of transcription factor binding sites and inferred the motif occupancy during a cellular differentiation. Our analysis suggested that functional characteristics of a cell-type are maintained by its specific TRN. This study also led us to believe that TRNs are driven by multiple transcription factors (TFs), and therefore various combinations of TRN elements are required to induce cellular functions.

In the present study, in order to artificially reconstruct a cellular TRN in human fibroblasts, we adopted a monocyte system mainly due to several supporting evidence of monocyte TRN [1,11] and therapeutically applicable functions such as phagocytosis, inflammatory response and chemotaxis, which are critical for host defense and targeting cancer cells [12]. Despite the intricate connectivity of regulatory networks, TRNs are known to form a chain-of-command hierarchy [13] where a core set of TFs drive the expressions of down-stream peripheral genes and act as key modulators to induce cell-specific functions. To identify these core TFs for the monocyte network, we first integrated gene expression profiling and a text-mining strategy. Furthermore, utilizing the lentivirus system and performing a perturbation matrix analysis we then further narrowed down to *SPI1*, *CEBPA*, *IRF8* and *MNDA* as core inducers of monocyte TRN (the most up-stream elements). Ectopic expression of these four factors activated monocytic functions and revealed key monocyte-specific pathways in dermal

fibroblasts. This work demonstrates that reconstruction of functional TRN can be achieved by introducing core TRN elements into unrelated cell types, such as human fibroblasts.

Results

Isolation of Core Transcriptional Regulatory Network Elements of Monocytes

Based on a hierarchical-gene model [13], TRNs are thought to be driven by a defined set of core elements which represent the most upstream TFs. Core TFs control and maintain the expression of downstream peripheral genes which make up the rest of the TRN. To reconstruct the monocyte TRN, we set out to identify the core elements of monocyte TRN in two steps. First, we carried out a gene expression profiling together with literature-based text mining in order to isolate a broader set of monocyte core TRN elements. We profiled mRNA expression profiles of human $CD14^+$ monocytes and human dermal fibroblasts using the Illumina microarray system. A pair-wise analysis revealed 49 highly expressed TFs in monocytes as compared to fibroblasts (2-fold or greater, $p\text{-value}<0.05$; Figure 1, left flow). Then, we implemented a literature-based text mining system to weigh the contextual relevancy of every transcription factor in monocytes (Figure 1, right flow; Table S1). We initially text mined for key biological-processes such as "differentiation", "reprogramming" and "transformation" in all available abstracts in MEDLINE. Among the extracted abstracts, we searched for co-existence of the word "monocytes" against all functionally annotated TFs (Table S2) [14] and ranked the genes based on the highest number of co-occurrences (Table 1). Finally, we integrated both the microarray data and the text-mined gene list by adding the inverse of ranks and selected top 20 TFs as core TRN-elements of monocytes.

Identification of Monocyte TRN Inducers

In order to investigate the hierarchical structure of the broader 20-monocyte TRN-elements, we performed a lentivirus-based perturbation-matrix analysis [11]. Briefly, full-length cDNAs of the selected 16 out of 20 TRN-elements were recombined into a lentivirus vector encoding Venus (enhanced yellow fluorescent protein) (Figure S1A). We excluded *EGR2*, *BTG2*, and *STAT5A*

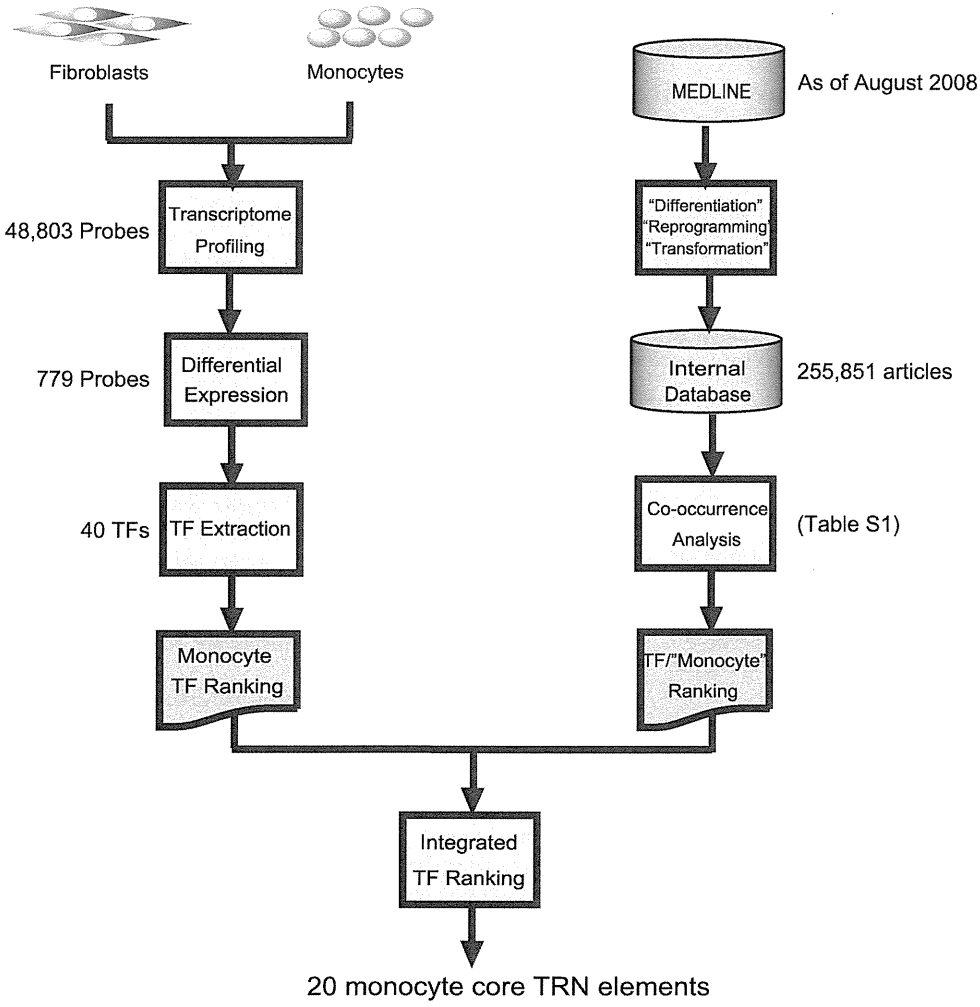


Figure 1. Core TRN elements isolation workflow. Key TFs were selected based on differential gene expression analysis and text mining. Left flow represents steps of the differential gene expression analysis. Gene expression profiling of primary human monocytes and that of human dermal fibroblasts were archived by using Human WG-6 v3.0 Expression beadschips (n = 3). Right flow represents the text mining. The text mining ranked the co-occurrence of "Monocyte" and "TF Name". Finally, those two methods were integrated and isolated top 20 TFs as core monocyte TRN elements.
doi:10.1371/journal.pone.0033474.g001

Table 1. The monocyte core TRN elements.

			Rank		Motif Activity	
PROBE_ID	SYMBOL	IS	Rfc	Rco	Fibroblast	Monocyte
ILMN_1669523	FOS	1.500	1	2	0.000986	−0.03624
ILMN_1764709	MAFB	0.511	2	90	−0.00494	−0.03165
ILMN_2086077	JUNB	0.396	16	3	0.000986	−0.03624
ILMN_1751607	FOSB	0.343	7	5	0.000986	−0.03624
ILMN_1743199	EGR2	0.338	3	218	−0.00575	−0.00267
ILMN_1666594	IRF8	0.272	4	45		
ILMN_1696463	SPI1	0.268	8	7	−0.00535	0.038682
ILMN_1727402	HCLS1	0.203	5	291		
ILMN_2392043	SPI1	0.193	20	7		
ILMN_1770085	BTG2	0.170	6	291		
ILMN_1715715	CEBPA	0.118	9	147	0.010685	0.025339
ILMN_1738992	MNDA	0.111	10	90		
ILMN_1719543	MAF	0.094	12	97		
ILMN_1753547	STAT5A	0.083	25	23	−0.01595	−0.02842
ILMN_1680624	CREG1	0.080	13	291		
ILMN_2216582	LYL1	0.075	14	291		
ILMN_1800078	LMO2	0.070	15	291	−0.00162	0.007074
ILMN_2214678	MXD1	0.065	18	105		
ILMN_1720829	ZFP36	0.058	19	174		
ILMN_1782305	NR4A2	0.047	23	291		

Rfc = differential expression fold change rank.
Rco = co-occurrence rank.
IS = Importance Score (1/Rfc+1/Rco).
doi:10.1371/journal.pone.0033474.t001

due to a technical difficulty of lentivirus vector construction and one duplicated *SPI1* transcript variant. Each TRN-element or TF-lentivirus was then transduced onto human dermal fibroblasts for eight days and cells expressing high levels of Venus were isolated using FACS sorter. Total RNA from cells expressing unique TRN-element was used to profile endogenous transcripts of the all 19-monocyte TRN-elements (Figure 2A and Figure S2). This perturbation-matrix analysis revealed that within the investigated 16 TRN-elements, *SPI1*, *CEBPA* and *NR4A2* induced the highest numbers of other TRN-elements including *EGR2*, *BTG2*, and *STS5A*, suggesting that these three factors represent the top-line elements of the monocyte TRN structure. We also took notice that *NR4A2*, which induced the expression of eight TRN-elements, was upregulated by *CEBPA* suggesting that *NR4A2* is a down-stream gene of *CEBPA*. Moreover, the endogenous expression of *SPI1*, *IRF8* and *MNDA* could not be sufficiently expressed by any of the exogenously transduced TRN-elements, indicating that these three factors are not down-stream of the examined TRN-elements. Taken together, the perturbation-matrix analysis revealed that *SPI1*, *CEBPA*, *MNDA* and *IRF8* represent as the top-line elements of the monocyte TRN (Figure 2B) referred hereinafter as "TRN-inducers".

Transduction of Four TRN-Inducers Enhances Monocytic Functions

To functionally characterize the role of TRN-inducers, we constructed *SPI1*, *CEBPA*, *IRF8* and *MNDA* into DsRed, Venus, BFP (Blue fluorescence protein), and blasticidine-resistant lentivirus vectors, respectively (Figure S1A), and transduced the pooled virus into human dermal fibroblasts. Since *SPI1* is a well-known master gene of monocytes [15] and our perturbation-matrix analysis also indicated a large-scale regulation by *SPI1*, we also transduced *SPI1* alone into human fibroblasts (FIB-SPI1). In combination with fluorescent proteins and blasticidine selection, we purified the four-factor transduced fibroblasts (FIB-4Fs) using FACS (Figure S1B). Four factors or SPI1 alone remarkably induced cell morphology changes as compared with mock-lentivirus transduced fibroblasts (FIB-mock) (Figure 3A). FIB-mock maintained their spindle shape and visible F-actin fibers representative of fibroblasts-like structures, whereas, FIB-SPI1 and FIB-4Fs showed circular shapes and lacked F-actin fibers, suggesting that FIB-SPI1 and FIB-4Fs retained different cellular features than FIB-mock control.

Since phagocytosis has been implicated as one of the key functions of monocytes/macrophages [16], we cultured FIB-SPI1, FIB-4Fs and FIB-mock in the presence of 2.0 μm red fluorescent latex beads for 2, 4 and 6 hours at the final concentration of 0.002v/v%. A confocal microscopy imaging revealed strong phagocytosis activity in FIB-4Fs cells than FIB-mock, whereas FIB-SPI1 showed only a weak phagocytosis activity (Figure 3B). Further flow cytometric analysis validated the result of the confocal microscopy imaging by showing a higher signal-to-background ratio of ingested beads in FIB-4Fs cells than FIB-mock over the time course. In contrary, FIB-SPI1 showed slightly higher intensity as compared to FIB-mock (Figure 3C and D). A similar result was obtained by changing the beads concentration to 0.001v/v% (data not shown).

Next we investigated the inflammatory response to bacterial lipopolysaccharide (LPS) as another functional characteristic of monocytes/macrophages [17]. We treated FIB-mock, FIB-SPI1 and FIB-4Fs cells with or without LPS at the final concentration of 10 μg/ml for 24 hours followed by qRT-PCR to compare the expression changes of eight inflammatory responsive genes (*TNF*, *IL6*, *IL1A*, *IL1B*, *IL8*, *CCL2*, *CXCL10*, *IFNB1*; Figure 3E). The qRT-PCR analysis revealed that IL6 gene was significantly upregulated in both FIB-SPI1 and FIB-4Fs in response to LPS, however, *TNF*, *IL8* and *IFNB1* did not show any significant difference among the three samples (Figure 3E). Interestingly, in response to LPS, four genes (*IL1A*, *IL1B*, *CCL2* and *CXCL10*) were significantly upregulated only by FIB-4Fs cells and not by FIB-SPI1 and FIB-mock cells, suggesting that the inflammatory response was activated by the additional three factors and not by *SPI1* alone. Fibroblasts are found in a variety of connective tissues and have the role to synthesize the extracellular matrix and collagen [18]. However, depending on the tissue origin, fibroblasts can respond to LPS and/or possess phagocytosis activity [19,20]. Here, we could not detect any phagocytosis or LPS response and hence we can conclude that monocytic functions were newly adapted, rather than enhanced, in these fibroblasts.

To further confirm monocyte functions in reconstructed cells, we also carried out cytokine secretion analysis by using the proteome profiler (Figure 4). Although GROα (CXCL1) and IL-8 were slightly increased in FIB-mock (2.12- and 2.23-fold increase, respectively), C5a, IP-10 (CXCL10) and RANTES (CCL5) were strongly upregulated in FIB-4Fs in response to LPS (2,66, 7.45, and 8.85 fold increase, respectively). Although, the protein expression of GROα, sICAM-1 (CD54), IL-1ra, IL-6, IL-8, MCP-1 (CCL2) and Serpin E1 (PAI-1) were not induced in response to LPS, these cytokines were already secreted in FIB-4Fs as compared to FIB-mock, suggesting that the four factors can increase the basal level of inflammatory cytokines. These results suggest that the FIB-4Fs adopted the enhanced inflammatory response in comparisons to FIB-mock.

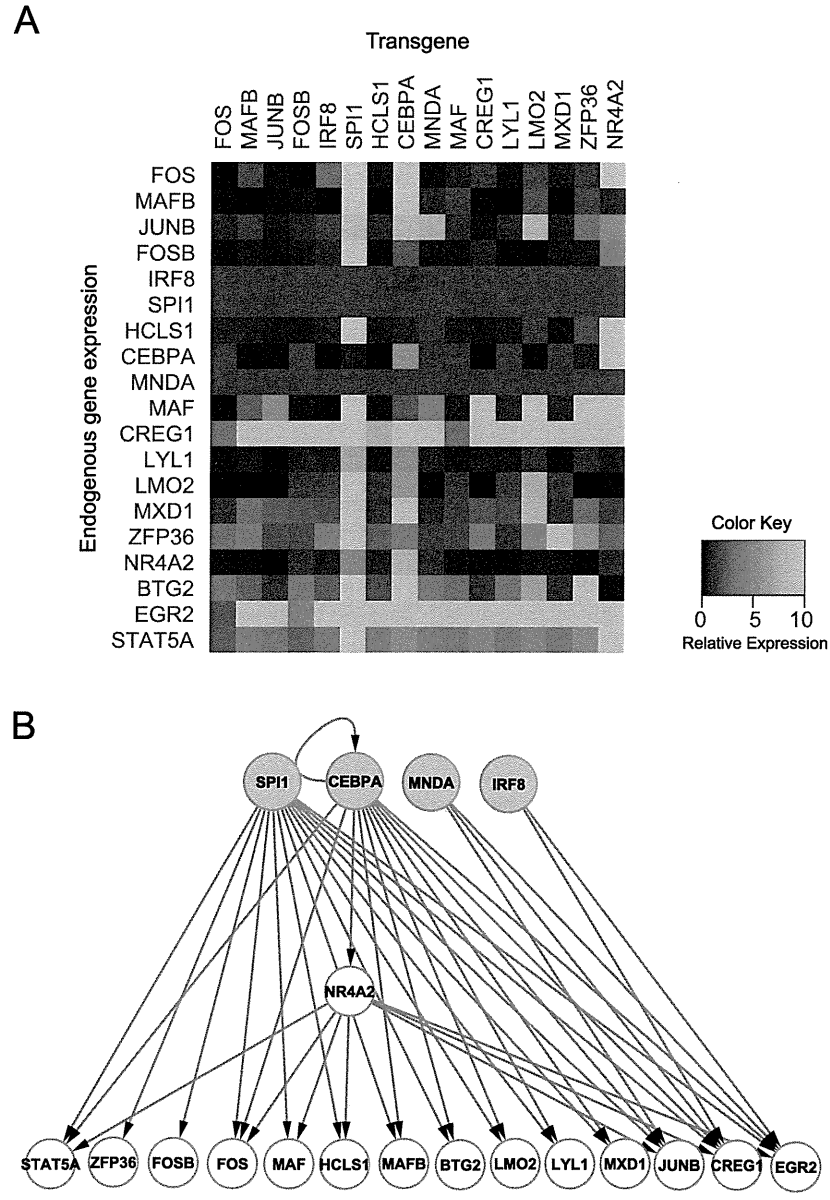


Figure 2. Regulatory relationship of monocyte's core TRN elements. (A) Relative expressions to monocyte are represented in colors. Higher relative expression is depicted as intensifying green color. The value of each relative expression is the average of biological replicates (n = 3). (B) Illustration of hierarchical network of monocyte core TRN elements. Each node (circle) indicates monocyte core TRN elements and green nodes represent the identified TRN inducers. When the TRN inducers or NR4A2 upregulate the gene expression to more than 5% of that of monocyte, an edge was drawn. An edge from an upper node to lower node indicates positive regulation.
doi:10.1371/journal.pone.0033474.g002

In order to investigate differential chemotactic response to CCL2 (also known as MCP-1) in FIB-4Fs, we seeded the cells in FluoroBlok transwell inserts and added the CCL2 protein into lower chambers at the final concentration of 5 μM and cultured for 16 hours. We stained the cells with Calcein-AM and measured fluorescent signal intensity at the bottom of the transwell inserts. The relative signal intensity was calculated by setting the samples without CCL2 as the basal chemotaxis activity (Figure 5). While we observed no chemotactic response of cells transduced with mock control, the FIB-4Fs cells displayed a significant chemotactic response (p-*value*<0.01, Figure 5). Taken together, the chemotactic response to CCL2, differential gene expression characteristics of functional monocytes, together with their morphology and marker profile, support their adaptation of monocyte TRN in dermal fibroblasts.

Validation of the Monocyte Network Reconstruction

We assessed the dynamics of monocyte marker expression by flow cytometry after transduction with FIB-4F, FIB-SPI1 and FIB-mock (Figure 6A). The analysis revealed that both FIB-4Fs and FIB-SPI1 significantly increased the percentage of CD14 (conventional monocyte marker), and the hematopoietic marker, CD45 positive cells (65.8±3.0% or 2.8±0.31% and 43.3±1.1% or 5.6±0.3%, respectively). However, the difference between FIB-

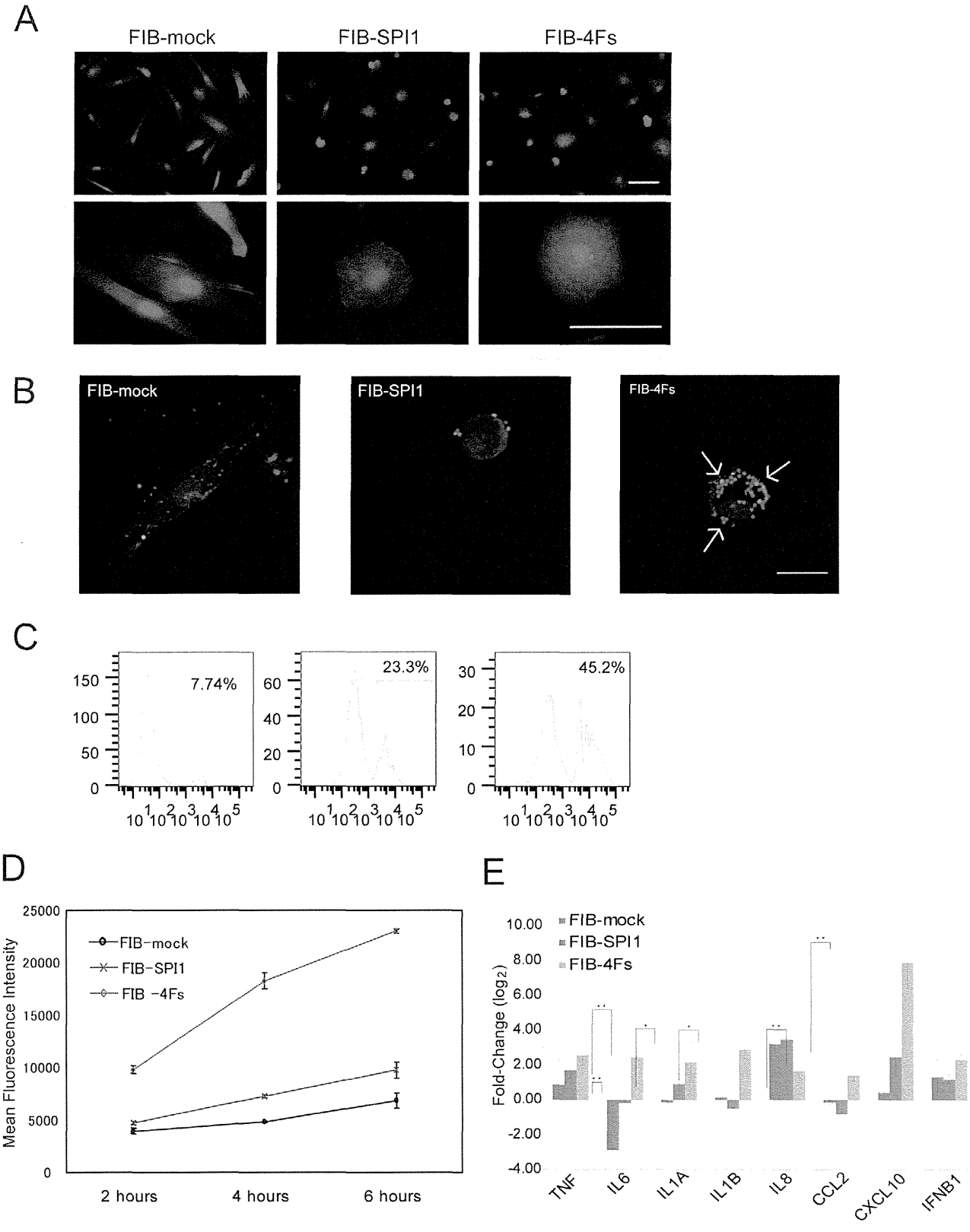
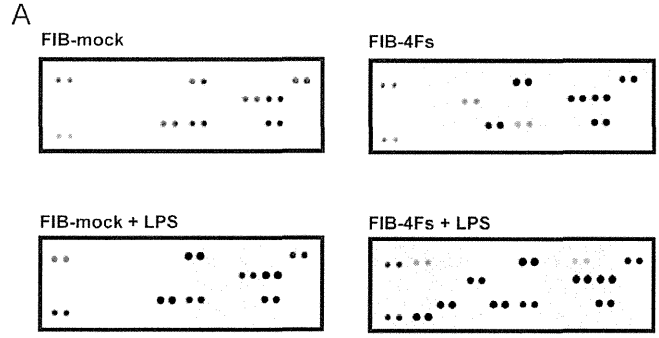


Figure 3. Cell feature assessments in reconstructed fibroblasts. (A) Morphological changes were visualized by microscopy. The cells were stained with Pallodin-Rhodamin (Yellow), Hoechst33342 (Blue), and Whole Cell Stains (Red). (B–D) Phagocytic latex beads were visualized (B) and the mean intensity of the ingested beads were confirmed by flow cytometry (C–D). (B) Beads-ingested cells are indicated by white arrows. Red color represents DID cell membrane staining, blue color represents nuclei, and green color represents the latex beads. (C) Beads ingested cells was

quantified based on the beads fluorescence by flow cytometry in FIB-mock, FIB-SPI1 and DIB-4Fs 2 houes after beads addition.. The cells were cultured with 0.002 v/v%. Vertical and horizontal axes represent cell count and fluorescent intensity, respectively. Beads ingested cells were gated. The analysis was performed in triplicates, showing the similar results.. (D) The flow cytometric analysis of the phagocytosis was summarized. Black, blue, and red lines represent FIB-mock, FIB-SPI1, and FIB-4Fs, respectively. The mean fluorescent intensity of ingested beads was measured by flow cytometry. The vertical axis and horizontal axis represent mean beads fluorescent intensity and incubation time, respectively. Error bars represent standard deviation (s.d.) (E) The expression change of *TNF*, *IL6*, *IL1A*, *IL1B*, *IL8*, *CCL2*, *CXCL10* and *IFNB1* induced by LPS treatment were measured by qRT-PCR in FIB-mock, FIB-SPI1, and FIB-4Fs. The cells were treated with LPS for 24 hours at the final concentration of 10 μg/μl. The bar represents the relative expression of LPS-treated cells as compared to untreated cells and error bar represents s.d. These experiments were repeated three times. * represents *P*-value≤0.05, ** represents *P*-value≤0.01 (t-test). The scale bar is 50 μm.
doi:10.1371/journal.pone.0033474.g003



Array Layout

PC	PC	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8	8	PC	PC
		9	9	10	10	11	11	12	12	13	13	14	14	15	15	16	16		
		17	17	18	18	19	19	20	20	21	21	22	22	23	23	24	24		
		25	25	26	26	27	27	28	28	29	29	30	30	31	31	32	32		
PC	PC	33	33	34	34	35	35	36	36									NC	NC

1:C5a 2:CD40 Ligand 3:G-CSF 4:GM-CSF 6:I-309 7:sICAM-1 8:IFNg
9:IL-1a 10:IL-1b 11:IL-1ra 12:IL-2 13:IL-4 14:IL-5 15:IL-6 16:IL-8
17:IL-10 18:IL-12 p70 19:IL-13 20:IL-16 21:IL-17 22:IL17E 23:IL-23 24:IL-27
25:IL-32a 26:IP-10 27:I-TAC 28:MCP1 29:MIF 30:MIP-1a 31:MIP-1b 32:Serpin E1
33:RANTES 34:ADF-1 35:TNF-a 36sTREM-1

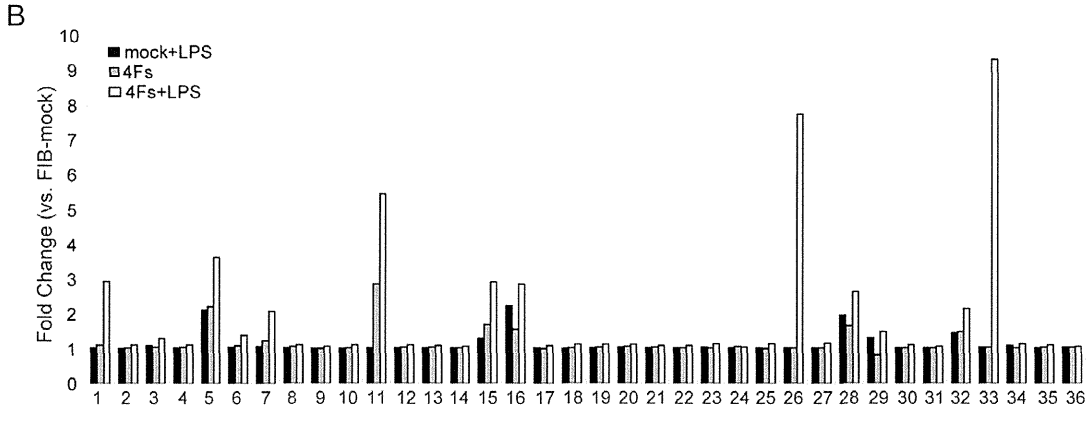


Figure 4. Four TRN-inducers adopted inflammatory cytokine secretion in response to LPS. FIB-mock and FIB-4Fs were treated or untreated with 10 μg/ml LPS. The cytokine levels of supernatant medium were assessed using the Proteome Profiler Human Cytokine Array, Pannel A. Array images were collected by LAS-3000 imaging system (A). The intensity of each spot was determined by Multi Gauge softwere (B).
doi:10.1371/journal.pone.0033474.g004

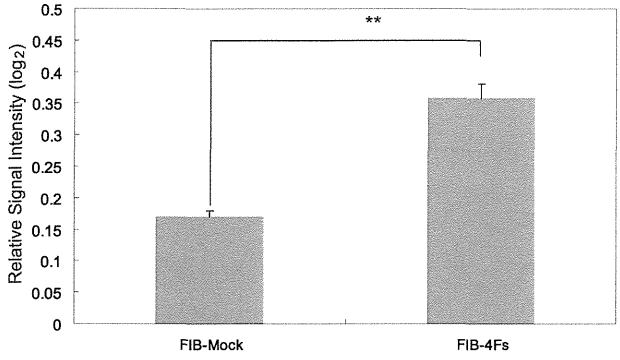


Figure 5. Four TRN-inducers adopted chemotaxis activity towards CCL2. Chemotaxis activity was measured by performing a transwell assay. FIB-mock and FIB-4Fs were cultured in transwells and incubated in the lower-chamber containing 5 μM of CCL2. The cells that migrated to the bottom of transwell were stained with Calcein-AM. Relative signal intensities were calculated by comparing fluorescent intensities of CCL2 treated to untreated cells (n = 3). ** represents *P-value*≤0.01 (t-test).
doi:10.1371/journal.pone.0033474.g005

SPI1 and FIB-4Fs were not statistically significant, suggesting that SPI1 also regulated monocytic genes although SPI1 alone was not sufficient to induce genes that are indispensable for the monocytic functions. In addition, we also tested CD115, CD16 and CCR2 monocyte surface markers but did not detect any positive cells (data not shown), suggesting that even FIB-4Fs is a partial conversion towards monocytes.

In order to examine the expression profiles of reconstructed monocyte TRN, we performed a global transcriptome analysis of FIB-4Fs and FIB-SPI1 using the Illumina microarrays. Based on the aforementioned dataset, we first defined the monocytes- and fibroblasts-specific genes by selecting the transcripts which expressed 2-fold or greater and *p-value*<0.03 (t-test) in monocytes (477 genes) and genes which expressed 2-fold or less and *p-value*<0.03 (t-test) in fibroblasts (488 genes), respectively. To evaluate the genes regulated in FIB-4Fs and FIB-SPI1 samples, we selected genes which were differentially expressed 2-fold or greater and *p-value*<0.05 (t-test) as compared to fibroblasts. Gene-association analysis revealed that FIB-SPI1 and FIB-4Fs induced 36.1% and 56.6% of the monocyte-specific genes, respectively (Figure 6B). The most upregulated monocyte-specific genes included well-known monocyte markers such as *CD14* (Rank:1), *IL1B* (Rank:2), *MX1* (Rank:3), *MAFB* (Rank:5) or *CD163* (Rank:12) and expression profile of well-known monocyte markers were shown in Table 2. These results suggested that TRN-inducers reconstructed the monocyte TRN to a greater degree than *SPI1* alone, although the monocyte expression profile was not completely mimicked. The analysis also showed that TRN-inducers repressed nearly 60% of fibroblast-specific genes whereas SPI1 alone repressed close to 40% (Figure 6C), suggesting an antagonistic interplay between reconstructed monocyte TRN and original fibroblast TRN. To further investigate the differential gene regulation by TRN-inducers, we categorized the genes either into up-regulated, down-regulated or unchanged based on a 2-fold cutoff (Figure 6D). Interestingly, expression of genes categorized in the up-regulated group was significantly higher in monocyte. On the other hand, expression of genes categorized in the down-regulated group were significantly lower in monocyte (*p-value*<0.001, Wilcoxon rank sum test), suggesting that differential regulation by the TRN-inducers were not random but clearly biased towards monocyte-like profile. To investigate the effect of additional factors to the *SPI1* target genes, we observed the expression distribution of SPI1 target genes. Interestingly, additional three factors (*CEBPA*, *MNDA* and *IRF8*) further up-regulated the expression of the SPI1 target genes (*p-value*<0.001, Wilcoxon signed-rank test) (Figure 6E).

Functional Relevancy of Differentially Expressed Genes

We next asked whether the TRN-induced cells had acquired functional expression properties of monocytes. To achieve this we performed Gene Ontology (GO) analysis for biological process category using DAVID [21] and KEGG pathway database. The DAVID analysis revealed that both FIB-4Fs and FIB-SPI1 significantly induced genes categorized to immune system process (GO: 0002376), a major function of monocytes (Table 3). Interestingly, statistical analysis indicated that genes in this category were additionally enhanced in FIB-4Fs as compared to FIB-SPI1 ($p\text{-value} = 6.20\times10^{-3}$; Fisher exact test; Table 2). Furthermore, KEGG pathway analysis revealed that 13 genes induced by *SPI1* were mapped to Toll-like receptor signaling pathway involved in LPS response as compared to 21 genes by FIB-4Fs. In addition, these 13 genes induced by *SPI1* were also mapped to the endocytosis pathway in which phagocytosis process is one of the sub-categories, as compared to 20 by FIB-4Fs. Further comparison between FIB-SPI1 and FIB-4Fs revealed that lysosome pathway genes, which are involved in the post-phagocytosis process, were significantly enriched in FIB-4Fs as compared to FIB-SPI1 (*p-value*<0.05, Fisher exact test). Taken together, these results suggested that the three additional factors further enhanced the biological functions typical for monocytes in fibroblasts.

Motif Activity Analysis Confirms Reconstruction of Monocyte Network

The average regulatory effects of each TF can be represented by computing the TF binding motif activities in a given cell type or a condition [1]. To investigate the propagation of gene expression regulated by TRN-inducers, top ten monocyte-associated motifs showed increased representations in FIB-SPI1 and more so in FIB-4Fs (Figure 7A). This result suggested that the three factors could additionally induce or enhance the monocyte TRN as compared to *SPI1* alone, possibly due to a more complete reconstruction of monocyte TRN. Interestingly, the activity of the SPI1 motif was significantly higher in FIB-4F than that in FIB-SPI1 (Figure 7B), confirming that *SPI1* works in concert with other TFs to elicit transcriptional activities at the promoter regions of target genes.

Discussion

We have found that key biological functions of monocytes can be accompanied in human dermal fibroblasts. This was achieved by reconstructing the monocyte TRN using the four TRN-inducers and the induced cells displayed marker profile, morphology, LPS response and chemotaxis activities similar to that of monocytes. The expression profile analysis of induced cells also revealed the synergistic role of 4 TRN inducers by greatly enhancing the expression and the motif activity involved in monocyte functional pathways as compared to *SPI1* alone.

Since only a limited number of the upstream elements can induce a substantial number of the downstream genes in a hierarchical–gene model, we aimed to identify a broader set of the most upstream genes by relying not only on differential expression analysis but also on text mining for selecting TRN elements. In general, the transcript expressions of TRN-elements are not necessarily specific to a single cell-type [22,23], and the expression levels do not always correlate strongly with the actual activity of