***The following resources related to this article are available online at
www.sciencemag.org (this information is current as of May 4, 2014):***

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/344/6183/519.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/04/30/344.6183.519.DC1.html

This article **cites 31 articles**, 9 of which can be accessed free:
http://www.sciencemag.org/content/344/6183/519.full.html#ref-list-1

This article appears in the following **subject collections:**
Genetics
http://www.sciencemag.org/cgi/collection/genetics

mechanism more directly, we simulated maize yields using the Agricultural Production Systems Simulator (APSIM) model for a representative site under sowing densities of 24,000 and 35,000 plants per acre, or 6.0 and 7.5 plants per m², respectively. Simulations under higher density exhibited higher average yields, but significantly higher sensitivity to VPD than simulations for lower density (fig. S11). The estimated yield response to high VPD 61 to 90 days after sowing in the simulations (–15% and –28% per kPa, respectively) agreed with the cross-sectional analysis of the USDA data (Fig. 4A). The importance of density is also suggested by the fact that soybeans, which show weaker evidence of increased drought sensitivity, have exhibited relatively constant sowing densities since 1995 in this region (24, 25).

Our results agree with the general notion that as farmers become more adept at removing all nonwater constraints to crop production, the sensitivity to drought generally increases (26). Given the dominant role of temperature (via VPD) in driving water stress in this region, our results are also consistent with the finding that heat sensitivity of maize yields in Indiana increased in recent decades (27). One implication is that climate change effects may be more severe than predicted by models that assume current crop genetics and management. Climate model projections indicate that July VPD for this region will become more severe, with an expected increase in average VPD of roughly 20% over the next 50 years (Fig. 4B), driven both by higher temperatures and reduced relative humidity. At current VPD sensitivity, these VPD trends would reduce yields by about 15% over the next 50 years. If maize yields continue to become increasingly sensitive to VPD, then yield losses from VPD trends could be as much as 30% (Fig. 4C).

Overall, we find no evidence to support the notion that farmers' yields are becoming less sensitive to drought in the main maize- and soybean-growing states. Instead, we find evidence that drought sensitivity in maize, in particular sensitivity to high VPD, has steadily increased over the past 18 years. Whether the recent push by seed companies to develop and market drought-tolerant seeds will reverse this trend remains to be seen. It is clear that cultivar changes are not the only relevant factor for changes in drought sensitivity, because agronomic trends—often facilitated by the cultivar changes—can be as or more important. It is also clear that new, field-level data sets on farm performance can complement field-based experiments in efforts to understand changes in drought sensitivity in this important production region.

References and Notes

1. H. Campos, M. Cooper, J. E. Habben, G. O. Edmeades, J. R. Schussler, *Field Crops Res.* **90**, 19–34 (2004).
2. C. M. Löffler et al., *Crop Sci.* **45**, 1708 (2005).
3. K. Chenu, R. Deihimfard, S. C. Chapman, *New Phytol.* **198**, 801–820 (2013).
4. F. Tardieu, *J. Exp. Bot.* **63**, 25–31 (2012).
5. Food and Agriculture Organization of the United Nations (FAO), FAO Statistical Databases (2013); http://faostat.fao.org.
6. J. McGrath, D. B. Lobell, *Glob. Change Biol.* **17**, 2689–2696 (2011).
7. R. J. Donohue, M. L. Roderick, T. R. McVicar, G. D. Farquhar, *Geophys. Res. Lett.* **40**, 3031–3035 (2013).
8. D. N. Duvick, *Adv. Agron.* **86**, 83–145 (2005).
9. M. Tollenaar, J. Wu, *Crop Sci.* **39**, 1597 (1999).
10. G. Hammer et al., *Crop Sci.* **49**, 299 (2009).
11. NRC, *The Impact of Genetically Engineered Crops on Farm Sustainability in the United States* (National Research Council, Washington, DC, 2010).
12. D. J. Lyon, G. L. Hammer, G. B. McLean, J. M. Blumenthal, *Agron. J.* **95**, 884 (2003).
13. M. Tollenaar, E. Lee, *Field Crops Res.* **75**, 161–169 (2002).
14. M. Bänziger, P. S. Setimela, D. Hodson, B. Vivek, *Agric. Water Manage.* **80**, 212–224 (2006).
15. J. H. Friedman, *Ann. Stat.* **19**, 1–67 (1991).
16. Materials and methods are available as supplementary materials on *Science* Online.
17. D. B. Lobell et al., *Nat. Clim. Change* **3**, 497–501 (2013).
18. J. D. Ray, R. W. Gesch, T. R. Sinclair, L. Hartwell Allen, *Plant Soil* **239**, 113–121 (2002).
19. B. Barnabás, K. Jäger, A. Fehér, *Plant Cell Environ.* **31**, 11–38 (2008).
20. A. E. Fonseca, M. E. Westgate, *Field Crops Res.* **94**, 114–125 (2005).
21. P. Woli, J. W. Jones, K. T. Ingram, C. W. Fraisse, *Agron. J.* **104**, 287 (2012).
22. C. D. Messina, D. Podlich, Z. Dong, M. Samples, M. Cooper, *J. Exp. Bot.* **62**, 855–868 (2011).
23. National Agriculture Statistics Service, Corn Objective Yield Survey Data, 1992-2006 (NASS, U.S. Department of Agriculture, Washington, DC, 2007); http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1749.
24. National Agriculture Statistics Service, Soybean Objective Yield Survey Data, 1992-2006 (NASS, U.S. Department of Agriculture, Washington, DC, 2007); http://usda01.library.cornell.edu/usda/current/SoyObjYield/SoyObjYield-07-27-2007.pdf.
25. National Agriculture Statistics Service, 2013 Illinois Annual Statistics Bulletin (NASS, U.S. Department of Agriculture, Washington, DC, 2013); www.nass.usda.gov/Statistics_by_State/Illinois/Publications/Annual_Statistical_Bulletin/2013/toc.html#fieldcrops.
26. D. J. Connor, R. S. Loomis, K. G. Cassman, *Crop Ecology: Productivity and Management in Agricultural Systems* (Cambridge Univ. Press, Cambridge, 2011).
27. M. J. Roberts, W. Schlenker, in *The Economics of Climate Change: Adaptations Past and Present*, G. D. Libecap, R. H. Steckel, Eds. (Univ. of Chicago Press, Chicago, 2011), pp. 225–251.

Polarization of the Effects of Autoimmune and Neurodegenerative Risk Alleles in Leukocytes

Towfique Raj,[1,2,3,4] Katie Rothamel,[5] Sara Mostafavi,[6] Chun Ye,[4] Mark N. Lee,[3,4] Joseph M. Replogle,[1,4] Ting Feng,[5] Michelle Lee,[1] Natasha Asinovski,[5] Irene Frohlich,[1] Selina Imboywa,[1] Alina Von Korff,[1] Yukinori Okada,[2,3,4,7,8] Nikolaos A. Patsopoulos,[1,2,3,4] Scott Davis,[5] Cristin McCabe,[1,4] Hyun-il Paik,[5] Gyan P. Srivastava,[1,2,3,4] Soumya Raychaudhuri,[2,3,4,9] David A. Hafler,[4,10] Daphne Koller,[6] Aviv Regev,[4,11] Nir Hacohen,[4,12] Diane Mathis,[5] Christophe Benoist,[13,5]* Barbara E. Stranger,[13,14]* Philip L. De Jager[1,2,3,4]*

To extend our understanding of the genetic basis of human immune function and dysfunction, we performed an expression quantitative trait locus (eQTL) study of purified CD4⁺ T cells and monocytes, representing adaptive and innate immunity, in a multi-ethnic cohort of 461 healthy individuals. Context-specific cis- and trans-eQTLs were identified, and cross-population mapping allowed, in some cases, putative functional assignment of candidate causal regulatory variants for disease-associated loci. We note an over-representation of T cell–specific eQTLs among susceptibility alleles for autoimmune diseases and of monocyte-specific eQTLs among Alzheimer's and Parkinson's disease variants. This polarization implicates specific immune cell types in these diseases and points to the need to identify the cell-autonomous effects of disease susceptibility variants.

It is necessary to understand the functional consequences of disease-associated genetic variation as we strive to unravel the causal chain of events linking genetic risk factors to clinical syndromes. Because most disease-associated variants localize to noncoding regions of the genome, an initial role of their function can be inferred from their proximity to a gene and association to gene expression levels (1–6) or to epigenomic features derived from representative tissues and cell types (7, 8). Given the central role of inflammation in many diseases, we leveraged differences in the

genetic structure of human populations to systematically interrogate the role of peripheral blood cell types that represent the lymphoid (CD4+ T lymphocyte) and myeloid (CD14+CD16− monocyte) arms of the immune system. These distinct programs of gene expression were profiled as part of the of Immune Variation (ImmVar) Project.

We used a rigorous sampling and cell purification protocol, minimizing circadian and experimental noise, to generate gene expression profiles from blood cells of 461 healthy individuals sampled in Boston, Massachusetts, including African Americans (AA), European Americans (EU), and East Asian Americans (EA) (Table 1, fig. S1, and table S1) (see also supplementary materials and methods). We purified naïve CD4+ CD62L^hi T cells, chosen to minimize experiential or environment-linked variation in effector/memory pools. Similarly, CD14+CD16− monocytes are a recent and short-lived cell population that is less likely to reflect an individual's life history. We quantified the extent of differential gene expression between each pairwise combination of populations using the V_{ST} statistic for each of the 19,114 genes passing quality control (fig. S2). V_{ST} (range: 0 to 1) is similar to the fixation index (F_{ST}): It quantifies the proportion of expression-level variance explained by population-level differences (1). The low median pairwise V_{ST} estimates (ranging from 0.005 to 0.009) suggest minimal gene expression variance between populations. Nonetheless, several genes are highly differentiated ($V_{ST} > 0.2$; top 1% of all V_{ST} scores) in expression between pairs of human populations (Fig. 1, A and B, and tables S2 and S3).

For each human population and cell type, we associated residual RNA expression phenotypes with genotyped or imputed (using reference haplotypes from The 1000 Genomes Project) (9) SNPs in a ±1-Mb window around the transcription start

[1]Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences, Departments of Neurology and Psychiatry, Brigham and Women's Hospital, Boston, MA 02115, USA. [2]Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA. [3]Harvard Medical School, Boston, MA 02115, USA. [4]The Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [5]Department of Microbiology and Immunobiology, Division of Immunology, Harvard Medical School, Boston, MA 02115, USA. [6]Department of Computer Science, Stanford University, Stanford, CA 94305, USA. [7]Department of Human Genetics and Disease Diversity, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan. [8]Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan. [9]Division of Rheumatology, Immunology and Allergy, Department of Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA. [10]Departments of Neurology and Immunobiology, Yale School of Medicine, New Haven, CT 06520, USA. [11]Department of Biology, Massachusetts Institute of Technology, and Howard Hughes Medical Institute, Cambridge, MA 02139, USA. [12]Center for Immunology and Inflammatory Diseases, Massachusetts General Hospital, Charlestown, MA 02129, USA. [13]Section of Genetic Medicine, Department of Medicine, University of Chicago, IL 60637, USA. [14]Institute for Genomics and Systems Biology, University of Chicago, Chicago, IL 60637, USA.

*Corresponding author. E-mail: christophe_benoist@hms.harvard.edu (C.B.); bstranger@medicine.bsd.uchicago.edu (B.E.S.); pdejager@partners.org (P.L.D.J.)

Table 1. Significant cis-eQTLs detected at FDR = 0.05. N/A, not applicable.

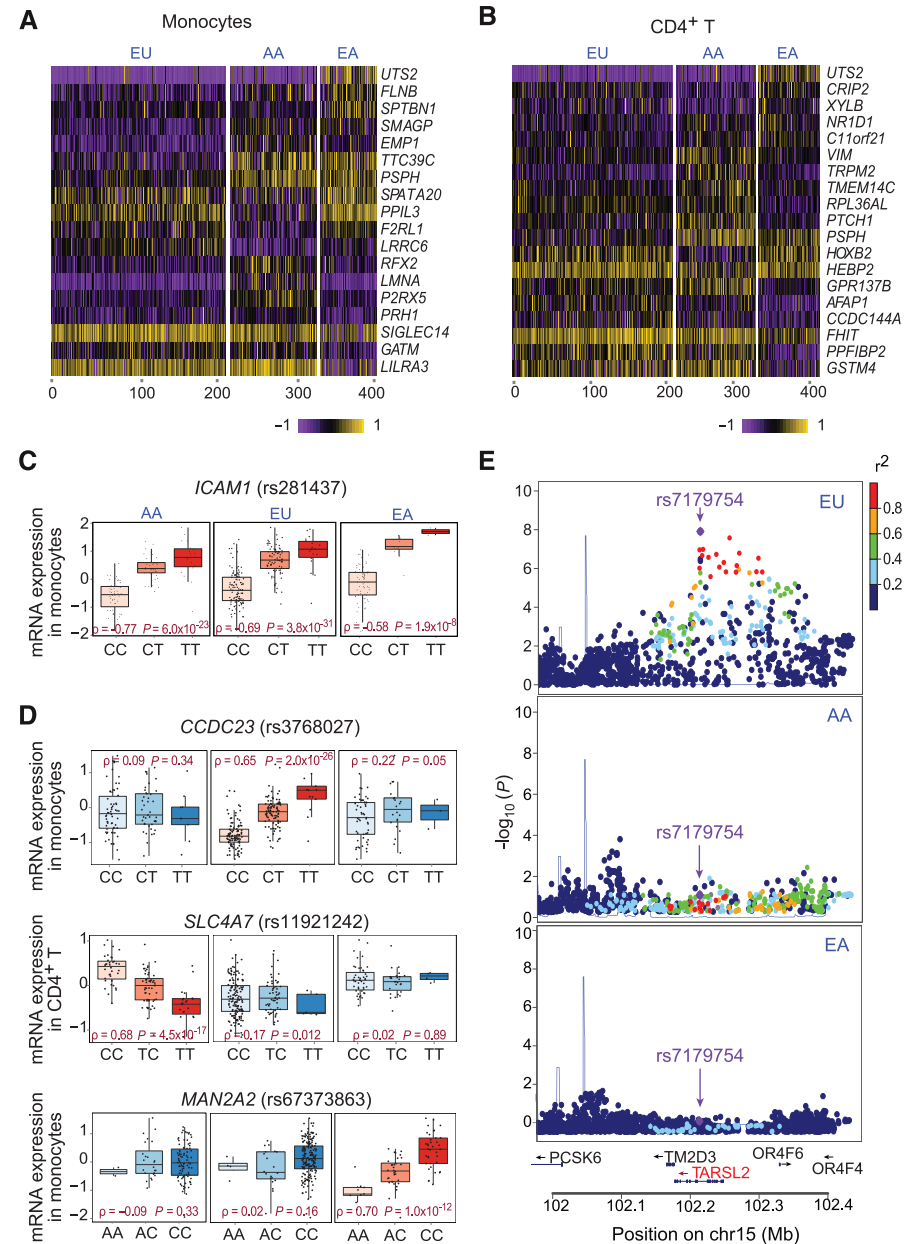| Population	Monocyte		CD4+ T cell		Shared
	No. of participants	No. of genes	No. of participants	No. of genes	No. of genes
European American	211	3090	213	2352	1178
African American	112	1318	112	722	259
East Asian	78	1181	82	592	215
≥Two populations	N/A	1352	N/A	739	328
Three populations	401	537	407	255	102
Nonredundant	401	3703	407	2672	1372
Meta-analysis	401	6210	407	4975	2789

Fig. 1. Transcriptome variation among human populations. Heat map of genes that are differentially expressed ($V_{ST} > 0.2$) in monocytes (**A**) and T cells (**B**). (**C**) Example of a cis-eQTL shared across populations. (**D**) Examples of population-specific cis-eQTLs. (**E**) Regional association plots of an EU-specific cis-eQTL in the *TARSL2* locus.

site of each gene. Of the 19,114 genes tested in T cells, we detected 2352, 592, and 722 genes with evidence of a cis expression quantitative trait locus (eQTL) effect in EU, EA, and AA participants, respectively [false discovery rate (FDR) = 0.05] (Table 1 and tables S4 to S6). In monocytes, we detected more genes with cis-eQTLs: 3090, 1181,

and 1318 in EU, EA, and AA participants, respectively (Table 1 and tables S7 to S9). Up to 70% of the genes with cis-eQTLs in monocytes replicate associations noted in a previous study (4) (fig. S3).

We used a Bayesian model average (BMA) and a hierarchical model (HM) (10) to jointly

analyze the three human populations in a single framework. This method was applied separately to each cell type. With BMA-HM, we observed a high degree of cross-population sharing, with 96% [95% confidence interval (CI): 89 to 100%] and 94% (95% CI: 83 to 100%) of cis-eQTLs being shared by all three populations in monocytes
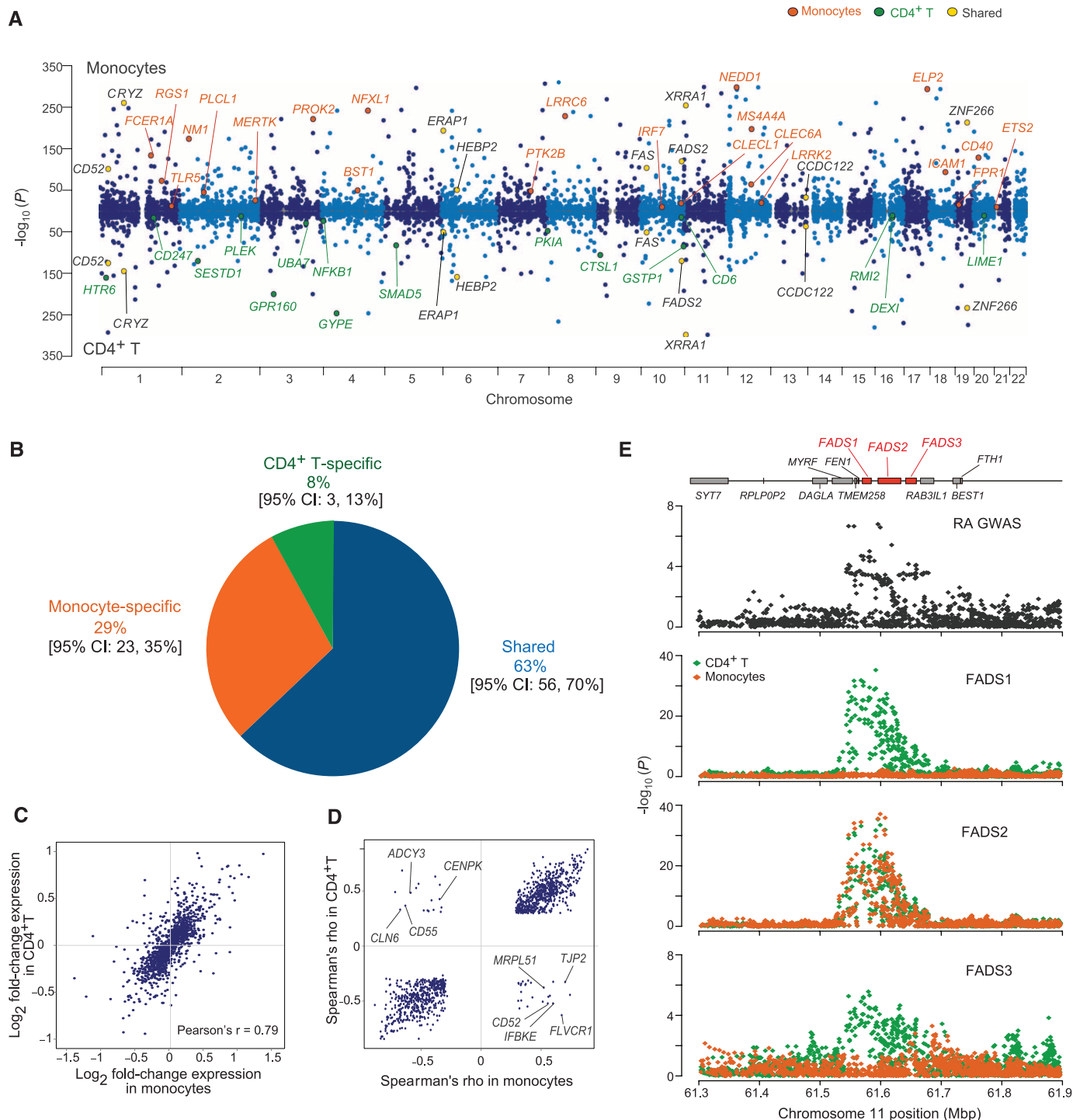

Fig. 2. Comparison of meta-analysis results from each cell type. (**A**) Manhattan plots of the most significant cis-eQTL per gene in monocytes (top) and T cells (bottom). The highlighted genes are representative of genes that (i) are shared, (ii) are cell-specific, or (iii) overlap with a disease association. (**B**) Inference of the proportion of cis-eQTL sharing between the two cell types using a Bayesian

hierarchical model. (**C**) Expression-level fold-change for the most significant cis-eQTLs in T cells and monocytes. (**D**) Spearman's ρ (direction of allelic effect on gene expression) for the most significant cis-eQTLs shared by the two cell types. We highlight genes that have discrepant directions of effect. (**E**) Example of a context-specific cis-eQTL in a rheumatoid arthritis locus. Mbp, megabase pairs.

and T cells, respectively (Fig. 1C). A small number of population-specific cis-eQTLs were found (Fig. 1, D and E). Using a different approach, a measure of the proportion of overlap estimated from the enrichment of low P values (π_1) (11),

we observe a similarly high degree of pairwise population sharing of cis-eQTLs (70 to 90%) (table S10).

The direction of allelic effect [that is, whether a given cis-eQTL is associated with an allele in

the same (or opposite) direction across pairs of populations] is highly conserved: When the top single-nucleotide polymorphism (SNP)–gene pair is shared across pairs of populations, all cis-eQTLs are concordant, suggesting that causal regulatory variation generally affects gene expression in the same direction in all populations (figs. S4 and S5). Similarly, the effect size of shared cis-eQTLs is significantly correlated (Pearson's $r = 0.70$ to 0.87) in each pairwise combination of populations for the same cell type (figs. S6 and S7), suggesting little population-specific modulation of eQTL effect.

A meta-analysis maximized our power to detect cis-eQTLs in 401 participants with monocyte data and 407 with T cell data; we used a random effects model optimized to detect associations under heterogeneity (12). We detected 50% more genes with cis-eQTLs in the meta-analysis: 6210 genes in monocytes and 4975 genes in T cells (Fig. 2A and tables S11 and S12). Up to 17% of genes in monocytes and 14% of genes in T cells demonstrate multiple independent eQTL effects (figs. S8 and S9). In addition, among the genes with shared effects across populations, the number of candidate causal variants in strong linkage disequilibrium (LD) with the lead SNP was reduced by more than 50% to an average of 21.2 SNPs when an LD threshold of $r^2 > 0.8$ with the lead SNP is considered in the meta-analysis (fig. S10).

To identify trans-eQTLs in which a SNP is associated with expression of a distal gene (>1 Mb from the SNP or on a different chromosome), we performed a trans-eQTL meta-analysis in our data. At a conservative Bonferroni-corrected threshold of $P < 3 \times 10^{-12}$, we identified 482 trans-eQTL associations (427 are interchromosomal) involving 55 genes specific to monocytes, 31 genes specific to T cells, and 4 genes shared by both cell types (figs. S11 and S12 and tables S13 and S14). Our results support the existence of trans-eQTLs, but their mechanism remains largely unknown.

We assessed the proportion of cell type–specific cis-eQTLs with BMA-HM (10) and estimated that 29% of cis-eQTLs are specific to monocytes, whereas 8% are specific to T cells, with 63% shared across cell types (Fig. 2B). Using a different approach that estimates the likelihood of eQTL rediscovery (11), we obtain similar estimates of sharing: $\pi_1 = 69$, 68, and 62% in the EU, AA, and EA participants (table S15). With reference epigenomic data from the ENCODE project (13), the cell type–specific cis-eQTLs were relatively enriched (quantified by relative risk) in the histone marks and deoxyribonuclease (DNase) I sites derived from the pertinent cell type (fig. S13). For instance, T cell–specific cis-eQTLs were more likely to coincide with DNase I sites derived from naïve CD4+ T cells than from monocytes. This confluence of epigenomic annotations and functional analysis results strengthens the interpretation of such epigenomic annotations, which are inferred to relate to functional effects of given variants.
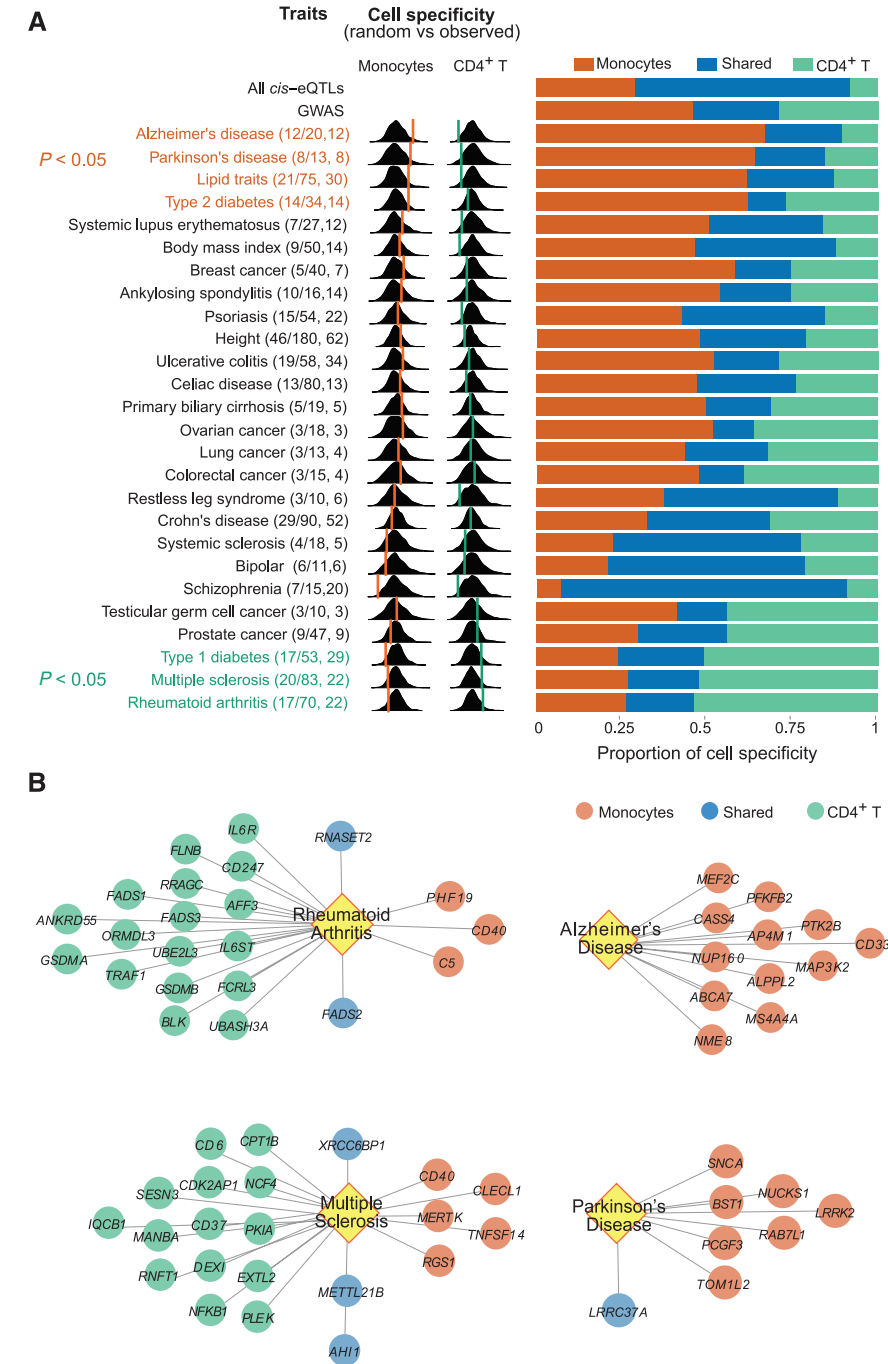

Fig. 3. Polarization of cis-regulatory effects for disease-associated variants. (**A**) (Left) For each evaluated trait, we report, in parentheses, the number of trait-associated (GWAS) SNPs with cis-eQTL effects over the total number of SNPs, and the number of genes influenced by one or more of these SNPs. For each trait, we present the distribution of the proportion of cell specificity (estimated using a Bayesian hierarchical model) observed for each of 1000 random samplings of matched SNP sets. The proportion of cell-specific cis-eQTL effects observed for a given trait is shown over this distribution using an orange line for monocytes and a green line for T cells. (Right) We report the proportion of cis-eQTLs that are monocyte-specific (orange), shared (blue), and T cell–specific (green). (**B**) Four diseases with significant polarization (FDR < 0.05) in the distribution of cis-eQTLs: Multiple sclerosis and rheumatoid arthritis are enriched in T cell effects (green), whereas Alzheimer's and Parkinson's diseases are enriched in monocyte effects (orange).

We went on to examine whether cis-eQTLs shared across cell types have similar direction and effect sizes. The comparison of the effect size for an eQTL (most significant SNP per gene) in the two cell types (Fig. 2C) showed concordance in effect size for most eQTLs (Pearson's $r = 0.79$). However, in 42 genes, the most significant SNP was associated with discordant directional effects across cell types: The allele associated with higher expression in one cell type is associated with lower expression in the other (Fig. 2D and fig. S14). We identified another dimension of cell type specificity: 6% of genes with cis-eQTLs shared across cell types are influenced by different, independent SNPs in each of the two cell types. Furthermore, 7% of independent SNPs in T cells and 11% in monocytes were associated with expression levels of more than one gene, with several SNPs affecting different genes in a cell type–dependent manner (Fig. 2E).

Leveraging 6341 SNPs (LD-pruned to $r^2 < 0.4$) associated with one or more of 415 human diseases or traits [National Institutes of Health genome-wide association study (GWAS) catalog] (14), we note 887 independent SNPs that influence expression of 1088 genes in cis (tables S16 and S17). To distinguish coincidental co-localizations of GWAS and eQTL associations, we used regulatory trait concordance (RTC) scores (15), which assesses whether a cis-eQTL and a trait association are tagging the same functional variant. Applying a stringent RTC threshold of 0.9, we identified 106 GWAS SNPs in T cells for which the trait and expression associations may be tagging the same effect; in monocytes, 123 GWAS SNPs met this criterion. Given the cell types profiled in our study, we found the expected significant enrichment (permutation $P < 1 \times 10^{-4}$) for cis-eQTLs among autoimmune disease–associated variants: Of the 425 GWAS SNPs associated with one or more autoimmune disease, 143 have cis-regulatory effects on 164 genes in monocytes and/or T cells (fig. S15 and tables S18 and S19).

Disease and trait-associated cis-eQTLs show more cell type specificity than average cis-eQTLs. Using BMA-HM, we find that 46% of all trait-associated variants are monocyte-specific, 29% are T cell–specific, and 25% are shared (Fig. 3A). In addition, variants associated with some diseases display a preponderance of T cell–specific cis-regulatory effects: These include susceptibility alleles for multiple sclerosis, rheumatoid arthritis, and type 1 diabetes (Fig. 3, A and B). On the other hand, Alzheimer's disease (AD), Parkinson's disease (PD), lipid traits, and type 2 diabetes susceptibility alleles are enriched in monocyte-specific effects (Fig. 3, A and B). In fact, the AD susceptibility alleles are significant eQTLs only in monocytes (Fig. 3B and figs. S16 and S17). Because the targeted genes are mostly expressed in both monocytes and T cells, this observed polarization is not driven by the lack of expression of the implicated genes in T cells (fig. S18).

On the basis of these results, we suggest that the inflammatory component of susceptibility to neurodegenerative diseases may be more strongly driven by the myeloid compartment of the immune system. This is consistent with the altered phagocytic function that we have reported in monocytes of individuals carrying the CD33 AD susceptibility variant (16). The putative role of inflammation has previously been invoked in aging-related neurodegenerative diseases such as AD and PD (17), but our results bring a genetic underpinning and candidate pathways to these observations. Because our study examines young and healthy individuals, we provide support for a role of myeloid cells in the prodromal phase of neurodegenerative diseases. Functionally, we cannot say that blood-derived monocytes themselves are the key cell type; they are likely to be proxies for the infiltrating macrophages (which differentiate from monocytes) and/or resident microglia found at the sites of neuropathology. Both of these cell types share many of the transcriptome patterns of monocytes.

Despite its inherent logistic challenges, the profiling of purified primary cells from multiple human populations was clearly beneficial in several respects, including the reduction of the credible set of variants associated with a given trait. Though the majority of eQTLs are shared, the ImmVar study provides clear examples of context specificity of eQTLs and has yielded insights into the relative role of two representative cell types in exerting the functional consequences of disease-associated variants. Further, the depletion of disease-associated eQTLs with effects in both cell types is intriguing and suggests that the evolutionary history of cell-specific and pleiotropic variants is an important area of future investigation. Overall, this component of the ImmVar Project represents

a strong foundation for future explorations of the role of disease susceptibility variants in an increasingly diverse matrix of cell types, stimuli, and in vivo contexts.

References and Notes

1. B. E. Stranger et al., PLOS Genet. 8, e1002639 (2012).
2. T. Lappalainen et al., Nature 501, 506–511 (2013).
3. A. Battle et al., Genome Res. 24, 14–24 (2014).
4. B. P. Fairfax et al., Nat. Genet. 44, 502–510 (2012).
5. H. J. Westra et al., Nat. Genet. 45, 1238–1243 (2013).
6. D. L. Nicolae et al., PLOS Genet. 6, e1000888 (2010).
7. G. Trynka et al., Nat. Genet. 45, 124–130 (2013).
8. M. T. Maurano et al., Science 337, 1190–1195 (2012).
9. G. R. Abecasis et al., Nature 491, 56–65 (2012).
10. T. Flutre, X. Wen, J. Pritchard, M. Stephens, PLOS Genet. 9, e1003486 (2013).
11. J. D. Storey, R. Tibshirani, Proc. Natl. Acad. Sci. U.S.A. 100, 9440–9445 (2003).
12. B. Han, E. Eskin, Am. J. Hum. Genet. 88, 586–598 (2011).
13. B. E. Bernstein et al., Nature 489, 57–74 (2012).
14. L. A. Hindorff et al., Proc. Natl. Acad. Sci. U.S.A. 106, 9362–9367 (2009).
15. A. C. Nica et al., PLOS Genet. 6, e1000895 (2010).
16. E. M. Bradshaw et al., Nat. Neurosci. 16, 848–850 (2013).
17. C. K. Glass, K. Saijo, B. Winner, M. C. Marchetto, F. H. Gage, Cell 140, 918–934 (2010).

Reconstructing the DNA Methylation Maps of the Neandertal and the Denisovan

David Gokhman,[1] Eitan Lavi,[1] Kay Prüfer,[2] Mario F. Fraga,[3] José A. Riancho,[4] Janet Kelso,[2] Svante Pääbo,[2] Eran Meshorer,[1,5]* Liran Carmel[1]*

Ancient DNA sequencing has recently provided high-coverage archaic human genomes. However, the evolution of epigenetic regulation along the human lineage remains largely unexplored. We reconstructed the full DNA methylation maps of the Neandertal and the Denisovan by harnessing the natural degradation processes of methylated and unmethylated cytosines. Comparing these ancient methylation maps to those of present-day humans, we identified ~2000 differentially methylated regions (DMRs). Particularly, we found substantial methylation changes in the HOXD cluster that may explain anatomical differences between archaic and present-day humans. Additionally, we found that DMRs are significantly more likely to be associated with diseases. This study provides insight into the epigenetic landscape of our closest evolutionary relatives and opens a window to explore the epigenomes of extinct species.

The high-coverage genomes of the Denisovan (1) and the Neandertal (2), from whom present-day humans split 550,000 to 765,000 years ago (2), allowed the comparison of our

DNA to that of our closest extinct relatives. However, differences could be sought exclusively at the sequence level and consequently did not provide direct insights regarding epigenetic