Current Biology, Vol. 13, 2118–2123, December 2, 2003, ©2003 Elsevier Science Ltd. All rights reserved. DOI 10.1016/j.cub.2003.11.025

Positive Selection on a Human-Specific Transcription Factor Binding Site Regulating *IL4* Expression

Matthew V. Rockman,[1,*] Matthew W. Hahn,[1] Nicole Soranzo,[2] David B. Goldstein,[2] and Gregory A. Wray[1]
[1]Department of Biology
Duke University
Box 90338
Durham, North Carolina 27708
[2]Department of Biology
University College London
Gower St., London WC1E 6BT
United Kingdom

Summary

A single nucleotide polymorphism in the promoter of the multifunctional cytokine *Interleukin 4* (*IL4*) affects the binding of NFAT, a key transcriptional activator of *IL4* in T cells [1, 2]. This regulatory polymorphism influences the balance of cytokine signaling in the immune system, with important consequences—positive and negative—for human health [3–12]. We determined that the NFAT binding site is unique to humans; it arose by point mutation along the lineage separating humans from other great apes. We show that its frequency distribution among human subpopulations has been shaped by the balance of selective forces on IL4's diverse roles. New statistical approaches, based on parametric and nonparametric comparisons to neutral variants typed in the same individuals, indicate that differentiation among subpopulations at the *IL4* promoter polymorphism is too great to be attributed to neutral drift. The allele frequencies of this binding site represent local adaptation to diverse pathogenic challenges; disease states associated with the common derived allele are side-effects of positive selection on other *IL4* functions.

Results and Discussion

The Origin of the Functional Promoter Polymorphism

The cytokine IL4 mediates a variety of interactions among components of the immune system. IL4 induces immature effector T cells to assume a Th2 phenotype and also represses Th1-inducing signals [13]; it induces B-cells to undergo immunoglobulin type-switching and secretion of IgE [14]; and it downregulates expression of one of the major coreceptors for HIV, CCR5, while upregulating the other, CXCR4 [15]. These diverse roles place conflicting demands on *IL4* expression, which is regulated at the level of transcription.

When Th2 cells are stimulated by antigen-presenting cells, signal transduction cascades lead to the dephosphorylation and nuclearization of the nuclear factor for activated T cells (NFAT) and to the transcriptional upregulation of the AP1 group of transcription factors [2]. Cooperative binding of NFAT and AP1 to the *IL4* promoter activates its transcription. The human *IL4* promoter contains six invariant binding sites for NFAT [16], but because of a polymorphism at position −524 from the transcriptional start site (dbSNP rs2243250), some promoter alleles contain an experimentally verified seventh binding site [1]. Transfection of allelic reporter constructs into cultured T cells has shown that the presence of this seventh NFAT binding site has a synergistic affect on transcription rate: −524T drives more than 3-fold greater expression than −524C [1]. The influence of the site on *IL4* expression has also been confirmed in vivo [17].

The allele bearing the binding site (−524T) has been associated by case-control studies, linkage mapping, and transmission-disequilibrium tests to atopy, asthma, and atopic dermatitis, perhaps via IL4's role in IgE expression [3–7]; to subacute sclerosing panencephalitis [8] and severe respiratory syncitium virus disease [9, 10] through its effect on the Th1/Th2 ratio; and to susceptibility to syncytium-inducing variants of HIV, mediated by its upregulation of *CXCR4* [11]. Despite these associations with disease risks, the −524T allele occurs at high frequencies in human populations. If −524T is the ancestral allele, its high frequencies may indicate that the deleterious effects of this binding site are a new feature of human evolution. If the allele is derived, the high frequencies may result from positive selection on other aspects of IL4 biology—for example, its role in increasing the inducibility of the Th2 response to certain immune challenges. In addition, the −524T allele is associated with overall increased survivorship in HIV+ cohorts [12], due perhaps to its roles in downregulating *CCR5* and in suppression of HIV transcription in monocytes [18]. Although HIV is too recent to have played a role in the evolution of the *IL4* promoter, earlier retroviruses may have exerted similar selective pressures.

To determine whether the binding site is unique to humans or is characteristic of a wider range of primates, we sequenced the proximal *IL4* promoter in all five extant great ape species (human, chimpanzee, bonobo, gorilla, and orangutan) and one old-world monkey (baboon). The *IL4* promoter exhibits remarkably conservative evolution (Figure 1). The 350 bp upstream from the start of transcription, which contains the six invariant NFAT binding sites, has experienced only two nucleotide substitutions among these primates, corresponding to a substitution rate of only 8% of the average rate for unconstrained DNA in this clade [19]. This region, including all six NFAT binding sites, is also strongly conserved in mouse, rat, and cow. In contrast, the region around −524 is not conserved in cow or rodents, but in the primates −524C is near the center of a 43 bp invariant region (Figure 1B). The phylogeny of promoter sequences implies that the seventh NFAT binding site arose during human ancestry as the result of a C→T point mutation.

Positive Selection on IL4 −524

To address the role of natural selection in shaping allele frequencies at this locus, we genotyped −524 in six

*Correspondence: matthew.rockman@duke.edu

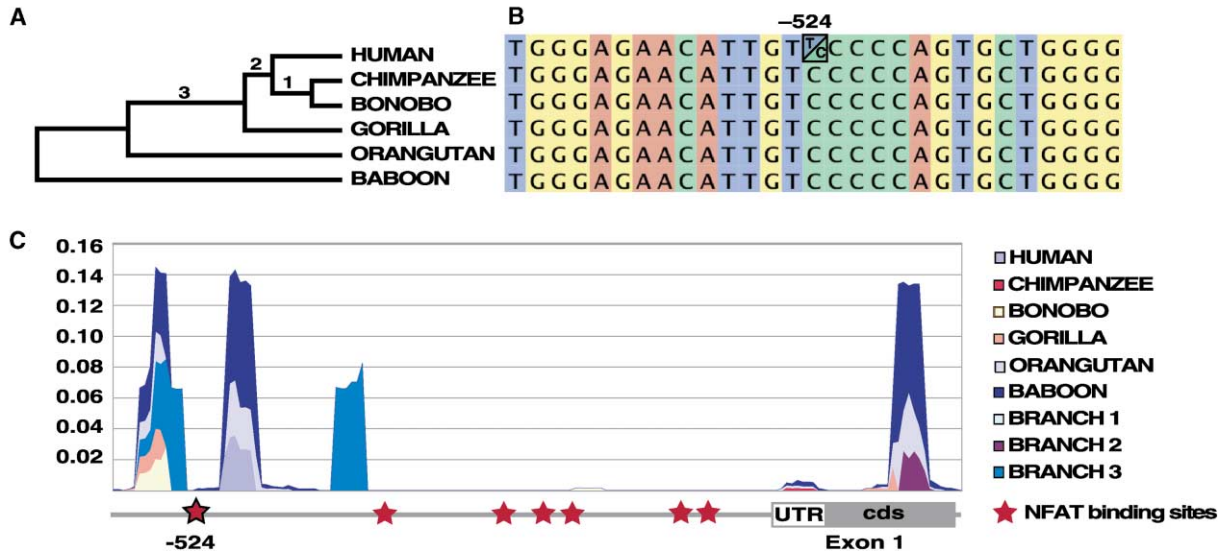

Figure 1. Evolution of the *IL4* Promoter

(A) Phylogenetic relationships among the extant great apes, with baboon as the outgroup. The branch lengths are roughly proportional to time [19].

(B) Alignment of *IL4* promoter sequence around the human −524C/T polymorphism. The −524T allele is derived.

(C) Sliding-window plot of evolutionary divergence in the *IL4* proximal promoter. The histogram shows the number of substitutions per site along each branch of the phylogeny in part (A), estimated for 30-base windows every 5 bases. Substitutions were estimated from the best-fit maximum likelihood model, which incorporates unequal equilibrium nucleotide frequencies and unequal rates of transitional and transversional substitutions. The histograms are stacked so that the total height represents the density of substitutions over the entire phylogeny. The proximal promoter exhibits dramatic conservation, and the polymorphism falls in an island of conservation in the more rapidly evolving distal region.

human subpopulations (Table 1) and compared the pattern of differentiation among subpopulations to that observed for 18 mutually unlinked SNPs, each more than 200 kb from any known gene and thus unlikely to be affected by selection. In the absence of selection, allele frequencies differ among subpopulations as the result of genetic drift, and the differentiation can be described by Wright's F_{ST} [20], which ranges from 0 (no differentiation) to 1 (fixed differences between subpopulations). Alleles at each unlinked locus drift in frequency independently, but the expected drift variance is a function of population size, allele frequency, and time. There is currently no robust theoretical model from which to derive expected F_{ST}s, but expectations can be derived empirically from unlinked neutral loci. Deviations from the neutral distribution—both high and low—suggest a role for natural selection [21–24]. The use of candidate neutral loci, rather than random loci, is necessary because the

Table 1. Differentiation at *IL4* −524

	IL4 −524 F_{ST}						
frequency[a]	**0.76**	0	0.19	0.46***	0.52**	0.21	Cameroon
neutral F_{ST}	0.13	**0.74**	0.17	0.43***	0.49***	0.19	China
	0.05	0.05	**0.43**	0.10*	0.15	0	Ethiopia
	0.10	0.03	0.02	**0.21**	0	0.09	India
	0.12	0.08	0.04	0.03	**0.17**	0.13	Italy
	0.33	0.18	0.25	0.22	0.30	**0.42**	New Guinea
	Cameroon	China	Ethiopia	India	Italy	New Guinea	

One asterisk indicates $p < 0.05$, two asterisks indicate $p < 0.01$, and three asterisks indicate $p < 0.0001$.

[a]Values on the diagonal are frequencies of −524T.

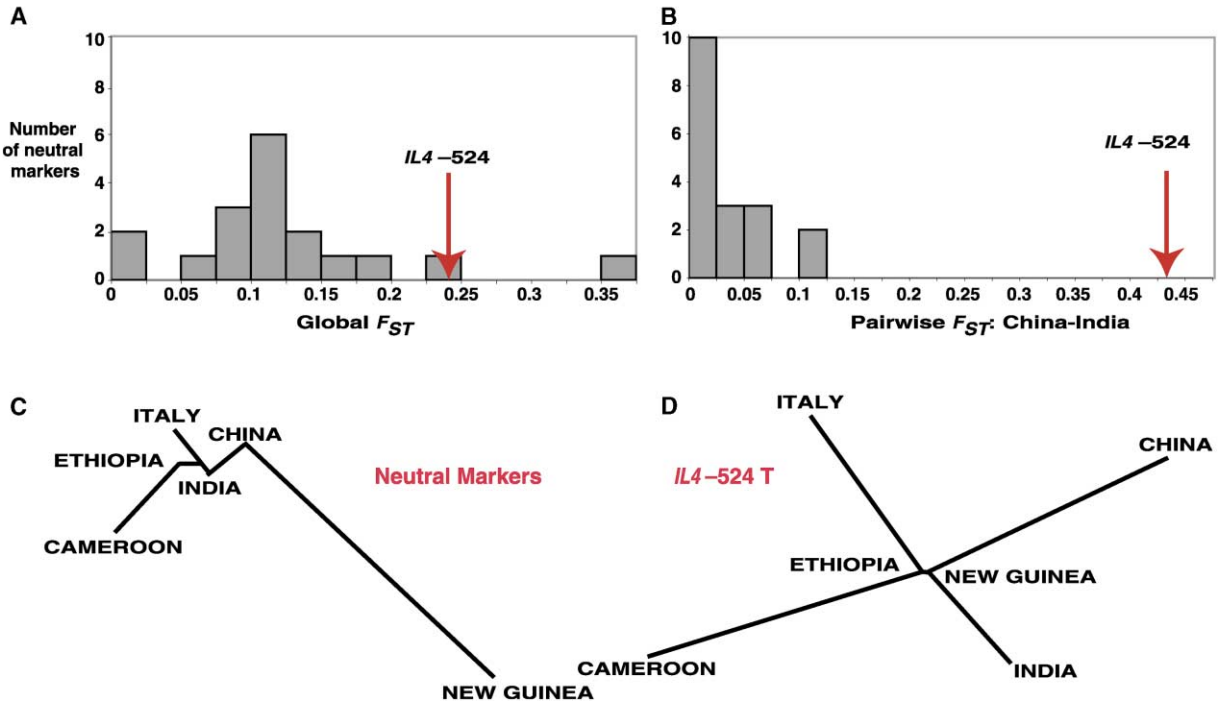

Figure 2. Nonneutral Pattern of Differentiation at *IL4* −524

(A) Global F_{ST} at −524 is higher than 17 of 18 neutral markers.

(B) Pairwise F_{ST} between China and India is dramatically elevated at −524.

(C) The maximum likelihood phylogeny of subpopulations based on allele frequencies at the neutral markers. Branch lengths are proportional to expected drift variance, the amount of divergence expected among the subpopulations due to genetic drift. India, China, and Ethiopia are connected to the phylogeny by branches with length zero.

(D) Branch lengths estimated for *IL4* −524, constrained by the topology in part (C). Ethiopia and New Guinea are connected to the phylogeny by branches with length zero. The likelihood ratio test [25] compares the probability of the *IL4* −524 data given the branch lengths from part (C) to that given the branch lengths from part (D). Together, these analyses indicate that the distribution of allele frequencies at *IL4* –524 is inconsistent with neutrality.

fraction of random loci under selection is unknown and the distribution F_{ST}s for random loci is therefore uninformative with respect to the action of selection.

Global F_{ST} among all six subpopulations at −524 is higher than that observed for 17 of the 18 neutral markers (Figure 2A) but is not significantly different from the estimated neutral F_{ST} (0.239 versus 0.127; two-tailed $p = 0.13$, based on bootstrap resampling of the difference between single-locus and 18-locus neutral F_{ST}). However, if the mode of selection is local adaptation, populations with similar selective regimes will acquire similar allele frequencies (yielding low F_{ST}), whereas in others selection will drive allele frequencies apart (yielding high F_{ST}). Consequently, pairwise comparisons of populations are more informative than the global F_{ST} about patterns of local selection. At *IL4* –524, four of 15 pairwise F_{ST}s are higher than any observed for the neutral markers, and five of the 15 are significantly elevated (Table 1); in particular, the F_{ST} at −524 for the China-India comparison exceeds the neutral F_{ST} by more than 10 standard deviations (Figure 2B; $p < 0.0001$). In the global F_{ST} value, these elevated pairwise F_{ST}s are partially masked by several very low pairwise F_{ST}s. Three pairwise values are zero, which can be indicative of balancing selection (i.e., allele frequencies held constant by selection), although statistical power to detect such selection among human subpopulations is low. Overall, China and Cameroon share high frequencies of *IL4* −524T, India and Italy share low frequencies, and the difference between the high and low frequencies is greater than expected under neutrality, again with strong statistical support.

To better understand the pattern of local selection, we applied a model that explicitly incorporates differing degrees of shared coancestry among populations. We estimated the phylogenetic relationships among the six subpopulations based on the 18 neutral markers by using a maximum likelihood model of independent neutral drift of gene frequencies along the branches of the phylogeny; the branch lengths on the resulting tree are proportional to the expected drift variance, a function of the effective population size and number of generations along the branch [25]. The maximum likelihood estimate based on neutral markers is shown in Figure 2C; the branch lengths for *IL4* −524 on the same topology are shown in Figure 2D. A likelihood ratio test [25] rejects the neutral branch lengths for *IL4* −524 ($\chi^2 = 20.01$, df = 9, $p = 0.018$). The distribution of allele frequencies at

−524 is therefore inconsistent with a neutral history at this site. This result points to a role for natural selection in the structuring of *IL4* −524 frequencies.

Cis-Regulation Is the Target of Selection

The distinct pattern of differentiation at −524 may be due to selection acting directly on this site, or it may be due to selection on a site in linkage disequilibrium (LD). Resequencing studies of African-American, European-American, and Korean population samples [9, 26] have identified no nonsynonymous polymorphisms in *IL4*, but −524 could be in LD with selected sites in other linked genes within the chromosome 5q cytokine cluster. To test this possibility, we identified sites in LD with −524 in African-American (AA) and European-American (EA) samples based on the resequencing data collected by the SeattleSNPs project [26]; the frequency of −524T in these samples (0.71 and 0.17) is similar to that in our Cameroon and Italy samples (0.76 and 0.17). In both AA and EA samples, −524 is not in linkage disequilibrium with any sites in the cytokine cluster genes centromeric (5′) to *IL4*, including *IL3*, *CSF2*, *IL5*, and *IL13*, except for a single SNP in the 3′UTR of *IL13*. This SNP exhibits weak LD with −524 in the EA sample only. Within the *IL4* locus, −524 exhibits significant LD with many noncoding sites of unknown function, including sites in the *IL13-IL4* intergenic region and in the introns of *IL4*. The 5′UTR SNP, +33 C/T, which is in perfect LD with −524 in European Americans, Japanese [27], and Koreans [9], is not in perfect LD with −524 in African Americans and does not show an elevated F_{ST} in a pairwise comparison of the AA and EA samples. Although SeattleSNPs data are not available for nearby genes telomeric (3′) to *IL4*, no nonsynonymous variants are known for any genes within a megabase in that direction.

The elevated F_{ST} at −524 and the absence of LD between −524 and any SNP with a known function both point to the −524 SNP as the site directly subject to selection. However, it remains possible that selection is acting on an unknown regulatory site in LD with −524. Another possibility is that some component of selection on *IL4* haplotypes involves *cis*-epistatic interactions among regulatory sites [28, 29]. However, selection on linked coding variants may be excluded as a cause of the pattern of allele frequencies at −524.

Selection on Immunoregulatory Balance in Health and Disease

The additional NFAT binding site at −524T creates a hair-trigger for *IL4* transcription. The resulting sensitivity is advantageous when a Th2 response is required, as when individuals are confronted with extracellular pathogens such as intestinal helminths, or when CCR5 expression poses risks, as in the face of HIV infection. On the other hand, −524C is favored when a Th1 response is required, as when individuals are confronted with intracellular pathogens, such as tuberculosis or leprosy, or when they face CXCR4 binding viruses. This balance of selective forces happens in a background of selection against the pleiotropic effects of *IL4* expression, including asthma and atopy. The role of positive selection in driving the derived binding site to high frequencies raises the possibility that many common diseases will be side-effects of common, positively selected variants. Although common variants will be a boon for disease mapping, the influence of positive selection complicates models of their population dynamics; models have assumed that the simple allelic spectrum of common diseases results from population expansion and weak negative selection [30, 31].

Previously documented instances of natural selection on components of the human immune system have involved molecules that directly interact with antigens (e.g., MHC molecules [32], immunoglobulins [33], CCR5 [24, 34], and FY [35]). Selection on the *IL4* promoter, on the other hand, acts to alter the whole balance of interactions within the immune system. Whereas diversifying selection of antigenic proteins may result in a coevolutionary race with the antigen binding proteins, selection on IL4 signaling offers an evolutionary response that pathogen proteins cannot simply evolve around. The *IL4* promoter illustrates the importance of regulatory variation in immune system evolution [36] and demonstrates the ease with which new beneficial regulatory interactions can arise by point mutation [37].

Experimental Procedures

We PCR-amplified an 828 bp fragment of the *IL4* locus, including the first exon and proximal promoter, from chimpanzee (*Pan troglodytes*), bonobo (*Pan paniscus*), gorilla (*Gorilla gorilla*), orangutan (*Pongo pygmaeus*), and baboon (*Papio papio*) by using primers IL4f-CTTGCCAAGGGCTTCCTTAT and IL4r-CTGGAGAGATGGTGCCA GAT. Fragments were cloned into Topo-TA vectors (Invitrogen), and multiple clones were sequenced in each direction. All base-calls were confirmed in multiple identical clones. The promoter was also sequenced directly from PCR products from two additional chimpanzees; both were homozygous −524C. The sliding-window analysis of substitution rates was performed in HYPHY v0.95 (http://www.hyphy.org/). We also used the public rat, mouse, and cow sequences, accessions X53087, AL645741, and U14131, respectively.

We PCR-amplified DNA from unrelated individuals from six locales: Southern Italy (chromosomes sampled = 90), Cameroon (88), Ethiopia (88), China (Singaporean Chinese) (80), India (Uttar Pradesh) (90), and Papua New Guinea (Madang Coastal) (84). *IL4* −524T was genotyped by PCR-RFLP [4]. We also sequenced 10% of the samples to verify genotypes. All genotype frequencies were consistent with Hardy-Weinberg proportions ($p > 0.05$). Human DNA samples were collected in the Goldstein lab with informed consent or were anonymized legacy collections provided to the Goldstein laboratory by collaborators from other academic research universities.

We genotyped 18 markers in the same individuals as above by using Taqman assays (ABI). The markers were selected based on data from the SNP Consortium according to three criteria. First, each is >200 kb from the nearest known gene and >50 kb from the nearest mapped EST (UCSC April 2002 freeze, NCBI Build29; some annotations have subsequently changed). Second, each SNP is unlinked to any other (at least 50 cM apart if on the same chromosome). Third, each SNP is polymorphic (minor allele frequency > 0.05) in the four ethnic groups typed by the SNP Consortium (http://snp.cshl.org), and the minor allele frequency in the European-derived CEPH population is at least 0.3. The markers are therefore enriched in highly heterozygous loci. Data are available in the Supplemental Data published with this article online.

F_{ST} was estimated by the diploid method of Weir ([38], page 178) implemented in a perl program (M.W.H. and J. Stajich) available through the Bioperl project [39] (http://www.bioperl.org). All F_{ST} values less than 0 were set equal to 0. Significance of the difference between F_{ST} at *IL4* −524 and the neutral F_{ST} was estimated by comparison to a bootstrap-resampled distribution; the difference be-

tween the single-locus F_{ST} of a randomly sampled neutral marker and the multi-locus F_{ST} of 18 randomly resampled neutral markers was calculated 10,000 times. P values are two-tailed. With a Bonferroni correction for multiple comparisons, which is extremely conservative in this case because the pairwise comparisons are not independent, three of 15 pairwise comparisons remain highly significant. The phylogenetic likelihood ratio test also accommodates the lack of independence among population pairs. Subpopulation phylogenetic trees were estimated with the CONTML gene frequency model [25] in Phylip v3.573 (http://evolution.genetics.washington.edu/phylip.html). The model assumes a bifurcating population history with independent genetic drift (i.e., loci are unlinked and neutral) along each branch of the tree. Modest constant gene flow along the branches of the phylogeny will mimic the effect of large population size (and result in short branches); however, the effect of other deviations from a bifurcating population history is unknown. The neutral markers strongly support a bifurcating history over one with no shared coancestry among populations according to a likelihood ratio test (star phylogeny versus bifurcating phylogeny, $\chi^2 = 15.5$, $df = 3$, $p < 0.002$). We estimated the effect of phylogenetic uncertainty by recalculating the likelihood ratio test statistic for each of the 105 possible trees and weighting the results by the likelihoods of the trees; the *IL4* results remain significantly nonneutral ($\chi^2 = 18.31$, $df = 9$, $p = 0.0318$).

Linkage disequilibrium was tested via the EM likelihood method for unphased data and implemented in Arlequin 2.000 [40] with $p = 0.05$ as a significance criterion. Similar results were found via the composite LD method ([38], page 127), implemented in Bioperl [39]. Genes telomeric to *IL4* were searched for nonsynonymous SNPs via the LocusLink interface to dbSNP build 116 (http://www.ncbi.nlm.nih.gov/SNP/).

Supplemental Data

Supplemental Data include a table showing genotypes for neutral marker SNPs and are available with this article online at http://www.current-biology.com/cgi/content/full/13/23/2118/DC1/.

Acknowledgments

Thanks to Anne Stone and Dagan Loisel for primate DNA samples. We gratefully acknowledge the support of the NSF (M.V.R., M.W.H., G.A.W.) and NASA (G.A.W.) and the Leverhulme Trust (D.B.G. and N.S.). D.B.G. is a Royal Society/Wolfson Research Merit Award Holder. We thank Ziheng Yang, Mark Rausher, and Jim Balhoff for valuable comments.

Received: September 1, 2003
Revised: October 20, 2003
Accepted: October 20, 2003
Published: December 2, 2003

References

1. Rosenwasser, L.J., and Borish, L. (1997). Genetics of atopy and asthma: the rationale behind promoter-based candidate gene studies (IL-4 and IL-10). Am. J. Respir. Crit. Care Med. *156*, S152–S155.
2. Weirenga, E.A., and Messer, G. (2000). Regulation of interleukin 4 gene transcription. Am. J. Respir. Crit. Care Med. *162*, S81–S85.
3. Kawashima, T., Noguchi, E., Arinami, T., Yamakawa-Kobayashi, K., Nakagawa, H., Otsuka, F., and Hamaguchi, H. (1998). Linkage and association of an interleukin 4 gene polymorphism with atopic dermatitis in Japanese families. J. Med. Genet. *35*, 502–504.
4. Burchard, E.G., Silverman, E.K., Rosenwasser, L.J., Borish, L., Yandava, C., Pillari, A., Weiss, S.T., Hasday, J., Lilly, C.M., Ford, J.G., et al. (1999). Association between a sequence variant in the IL-4 gene promoter and FEV1 in asthma. Am. J. Respir. Crit. Care Med. *160*, 919–922.
5. Sandford, A.J., Chagani, T., Zhu, S., Weir, T.D., Bai, T.R., Spinelli, J.J., Fitzgerald, J.M., Behbehani, N.A., Tan, W.C., and Pare, P.D. (2000). Polymorphisms in the IL4, IL4RA, and FCERIB genes in asthma severity. J. Allergy Clin. Immunol. *106*, 135–140.
6. Noguchi, E., Nukaga-Nishio, Y., Jian, Z., Yokouchi, Y., Kamioka, M., Yamakawa-Kobayashi, K., Hamaguchi, H., Matsui, A., Shibasaki, M., and Arinami, T. (2001). Haplotypes of the 5′ region of the IL-4 gene and SNPs in the intergene sequence between IL-4 and IL-13 genes are associated with atopic asthma. Hum. Immunol. *62*, 1251–1257.
7. Beghé, B., Barton, S., Rorke, S., Peng, Q., Sayers, I., Gaunt, T., Keith, T.P., Clough, J.B., Holgate, S.T., and Holloway, J.W. (2003). Polymorphisms in the interleukin-4 and interleukin-4 receptor a chain genes confer susceptibility to asthma and atopy in a Caucasian population. Clin. Exp. Allergy *33*, 1111–1117.
8. Inoue, T., Kira, R., Nakao, F., Ihara, K., Bassuny, W.M., Kusuhara, K., Nihei, K., Takeshita, K., and Hara, T. (2002). Contribution of the interleukin-4 gene to susceptibility to subacute sclerosing panencephalitis. Arch. Neurol. *59*, 822–827.
9. Choi, E.H., Lee, H.J., Yoo, T., and Chanock, S.J. (2002). A common haplotype of interleukin-4 gene is associated with severe respiratory syncytial virus disease in Korean children. J. Infect. Dis. *186*, 1207–1211.
10. Hoebee, B., Rietveld, E., Bont, L., Oosten, M., Hodemaekers, H.M., Nagelkerke, N.J., Neijens, H.J., Kimpen, J.L., and Kimman, T.G. (2003). Association of severe respiratory syncytial virus bronchiolitis with interleukin-4 and interleukin-4 receptor alpha polymorphisms. J. Infect. Dis. *187*, 2–11.
11. Nakayama, E.E., Hoshino, Y., Xin, X., Liu, H., Goto, M., Watanabe, N., Taguchi, H., Hitani, A., Kawana-Tachikawa, A., Fukushima, M., et al. (2000). Polymorphism in the interleukin-4 promoter affects acquisition of human immunodeficiency virus type 1 syncytium-inducing phenotype. J. Virol. *74*, 5452–5459.
12. Nakayama, E.E., Meyer, L., Iwamoto, A., Persoz, A., Nagai, Y., Rouzioux, C., Delfraissy, J.F., Debre, P., McIlroy, D., Theodorou, I., et al. (2002). Protective effect of interleukin-4 –589T polymorphism on human immunodeficiency virus type 1 disease progression: relationship with virus load. J. Infect. Dis. *185*, 1183–1186.
13. Murphy, K.M., and Reiner, S.L. (2002). The lineage decisions of helper T cells. Nat. Rev. Immunol. *2*, 933–944.
14. Finkelman, F.D., Holmes, J., Katona, I.M., Urban, J.F., Jr., Beckmann, M.P., Park, L.S., Schooley, K.A., Coffman, R.L., Mosmann, T.R., and Paul, W.E. (1990). Lymphokine control of in vivo immunoglobulin isotype selection. Annu. Rev. Immunol. *8*, 303–333.
15. Valentin, A., Lu, W., Rosati, M., Schneider, R., Albert, J., Karlsson, A., and Pavlakis, G.N. (1998). Dual effect of interleukin 4 on HIV-1 expression: implications for viral phenotypic switch and disease progression. Proc. Natl. Acad. Sci. USA *95*, 8886–8891.
16. Burke, T.F., Casolaro, V., and Georas, S.N. (2000). Characterization of P5, a novel NFAT/AP-1 site in the human IL-4 promoter. Biochem. Biophys. Res. Commun. *270*, 1016–1023.
17. Nakashima, H., Miyake, K., Inoue, Y., Shimizu, S., Akahoshi, M., Tanaka, Y., Otsuka, T., and Harada, M. (2002). Association between IL-4 genotype and IL-4 production in the Japanese population. Genes Immun. *3*, 107–109.
18. Goletti, D., Kinter, A.L., Coccia, E.M., Battistini, A., Petrosillo, N., Ippolito, G., and Poli, G. (2002). Interleukin (IL)-4 inhibits phorbol-ester induced HIV-1 expression in chronically infected U1 cells independently from the autocrine effect of endogenous tumour necrosis factor-a, IL-1b, and IL-1 receptor antagonist. Cytokine *17*, 28–35.
19. Yi, S., Ellsworth, D.L., and Li, W.-H. (2002). Slow molecular clocks in old world monkeys, apes, and humans. Mol. Biol. Evol. *19*, 2191–2198.
20. Wright, S. (1951). The genetical structure of populations. Ann. Eugen. *15*, 323–354.
21. Lewontin, R.C., and Krakauer, J. (1973). Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. Genetics *74*, 175–195.
22. Bowcock, A.M., Kidd, J.R., Mountain, J.L., Hebert, J.M., Carotenuto, L., Kidd, K.K., and Cavalli-Sforza, L.L. (1991). Drift, admixture, and selection in human evolution: a study with DNA polymorphisms. Proc. Natl. Acad. Sci. USA *88*, 839–843.

23. Beaumont, M.A., and Nichols, R.A. (1996). Evaluating loci for use in the genetic analysis of population structure. Proc. R. Soc. Lond. B. Biol. Sci. *263*, 1619–1626.
24. Bamshad, M.J., Mummidi, S., Gonzalez, E., Ahuja, S.S., Dunn, D.M., Watkins, W.S., Wooding, S., Stone, A.C., Jorde, L.B., Weiss, R.B., et al. (2002). A strong signature of balancing selection in the 5′ cis-regulatory region of CCR5. Proc. Natl. Acad. Sci. USA *99*, 10539–10544.
25. Felsenstein, J. (1981). Evolutionary trees from gene frequencies and quantitative characters: finding maximum likelihood estimates. Evolution *35*, 1229–1242.
26. SeattleSNPs. (2003). NHBLI Program for Genomic Applications. (Seattle, WA: University of Washington, Fred Hutchinson Cancer Research Center) (http://pga.gs.washington.edu).
27. Suzuki, I., Yamaguchi, E., Hizawa, N., Itoh, A., and Kawakami, Y. (1999). A new polymorphism in the 5′ flanking region of the human interleukin (IL)-4 gene. Immunogenetics *49*, 738–739.
28. Cantagrel, A., Navaux, F., Loubet-Lescoulie, P., Nourhashemi, F., Enault, G., Abbal, M., Constantin, A., Laroche, M., and Mazieres, B. (1999). Interleukin-1β, interleukin-1 receptor antagonist, interleukin-4, and interleukin-10 gene polymorphisms: relationship to occurrence and severity of rheumatoid arthritis. Arthritis Rheum. *42*, 1093–1100.
29. Rockman, M.V., and Wray, G.A. (2002). Abundant raw material for cis-regulatory evolution in humans. Mol. Biol. Evol. *19*, 1991–2004.
30. Reich, D.E., and Lander, E.S. (2001). On the allelic spectrum of human disease. Trends Genet. *17*, 502–510.
31. Pritchard, J.K. (2001). Are rare variants responsible for susceptibility to complex diseases? Am. J. Hum. Genet. *69*, 124–137.
32. Hughes, A.L., and Nei, M. (1988). Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature *335*, 167–170.
33. Tanaka, T., and Nei, M. (1989). Positive selection observed at the variable-region genes of immunoglobulins. Mol. Biol. Evol. *6*, 447–459.
34. Dean, M., Carrington, M., and O'Brien, S.J. (2002). Balanced polymorphism selected by genetic versus infectious human disease. Annu. Rev. Genomics Hum. Genet. *3*, 263–292.
35. Hamblin, M.T., Thompson, E.E., and DiRienzo, A. (2002). Complex signatures of natural selection at the Duffy blood group locus. Am. J. Hum. Genet. *70*, 369–383.
36. Mitchison, N.A. (1997). Partitioning of genetic variation between regulatory and coding gene segments: the predominance of software variation in genes encoding introvert proteins. Immunogenetics *46*, 46–52.
37. Stone, J., and Wray, G.A. (2001). Rapid evolution of cis-regulatory sequences via local point mutations. Mol. Biol. Evol. *18*, 1764–1770.
38. Weir, B.S. (1996). Genetic data analysis II. (Sunderland, MA: Sinauer Associates).
39. Stajich, J.E., Block, D., Boulez, K., Brenner, S.E., Chervitz, S.A., Dagdigian, C., Fuellen, G., Gilbert, J.G., Korf, I., Lapp, H., et al. (2002). The Bioperl toolkit: Perl modules for the life sciences. Genome Res. *12*, 1611–1618.
40. Schneider, S., Roessli, D., and Excoffier, L. (2000). Arlequin version 2.000: A Software for Population Genetics Data Analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Accession Numbers

Sequences were deposited in Genbank with accession numbers AY339644-48.