Supplementary Material

1. Variant calling, quality control and variant validations

1.1 Variant calling

Alignment

The sequencing reads of fastq files were found for 20,937 samples. Alignments were performed to the human genome reference UCSC build 37 (hg19) using BWA (v0.5.9) (Li and Durbin 2009). Basic summary of mapping results was shown in **Supplemental Table S3**.

Variant calling

We performed SNP and Indel calling using The Genome Analysis Toolkit (GATK) (3.7-0-gcfedb67) (DePristo et al. 2011) with the following steps: read duplicate marking using Picard MarkDuplicates, base quality recalibration using GATK BaseRecalibrator and PrintReads, single sample variant discovery using GATK HaplotypeCaller and joint calling using GATK GenotypeGVCFs.

1.2 Variant recalibration and filtering

GATK Variant Quality Score Recalibration (VQSR) model was applied to filter variants. With a sensitivity threshold of 99.6%, HapMap3.3 and 1KG Omni 2.5 SNP were utilized as training sites to filter SNPs. Axiom Exome Plus and 1KG gold standard sites were applied to filter indels and a 95.0% sensitivity threshold was used.

After generation of initial variant calls set, we re-checked VCF files and did filters regarding read depth, mapping quality, inbreeding coefficients, Hardy-Weinberg equilibrium and the number of individuals with coverage at the site. The call set then comprised variants that were *MQ* >20 (called with posterior probability >99%), were targeted in at least 80% individuals, and had a total depth across samples between 20,937 to 20,937,000 (about 1-1000X per sample on average). We also excluded sites where inbreeding coefficient (*InbreedingCoeff*) was <-0.2 or sites violating Hardy-Weinberg equilibrium (Chi-square $P<10^{-6}$). To assure genotypes with high confidence and accuracy, we further set individual genotypes

with *GQ*<20 or *DP*<10 to missing data. After such filtering, variants with more than 15% of missing genotypes across individuals were eliminated for downstream analyses.

After above filtering, Ti/Tv ratios for synonymous, missense, stop-gained, and splice variants were 5.61, 2.25, 2.08, and 1.68, respectively, which is consistent with Tennessen et al. 2012.

1.3 Sample quality control and selection

We further removed 603 individuals because of mismatches between self-declared and genetically determined sex, and 199 samples due to outliers of heterozygosity using PLINK (Purcell et al. 2007). Besides, 176 individuals with a call rate less than 10% were also removed.

We performed a KING (Manichaikul et al. 2010) analysis on the filtered data to remove related individuals. Specifically, we performed LD pruning with PLINK (--indep-pairwise 50 5 0.5) to the variants with minor allele frequency (MAF) >5%. We then extracted a list of 19,973 individuals that contains no pairs of individuals with a 1st-, 2nd- or 3rd-degree relationship with KING (--degree 3 --unrelated). After excluding these samples, we repeated the KING investigations and found no kinships in the remaining data set. Thus, our final datasets consist of 19,973 individuals, including 10,539 healthy control subjects and 9,434 patients with psoriasis, within 1,320 protein coding genes (**Supplemental Table S4**). Locus-based summary of ~200,000 filtered variant calls are provided in CNGB Sequence Archive (CNSA, https://db.cngb.org/cnsa/, Guo et al. 2020) under accession number CNP0001423**.**

1.4 Analyses of phenotype effects

Phenotype effects were assessed. Psoriasis cases and controls showed similar properties of Ti/Tv, heterozygosity and genotype missingness (**Supplemental Fig. S1**). To confirm that patterns of population structure were not significantly biased by phenotype effects, we carried out a principal component analysis (PCA) on the merged dataset of samples from the 1000 Genomes Project Consortium et al. 2015 (1KGP) and our study (n=20,376, **Supplemental Fig. S2**) using EIGENSOFT (Patterson et al. 2006). PCA was performed with biallelic SNPs with a MAF > 1% and pruned by LD.

1.5 Validation

As part of quality control and validation procedure, we sequenced 24 CHB samples from 1KGP with the same targeted sequencing strategy (Tang et al. 2014). Such 24 CHB samples were also sequenced and performed with variant-calling independently by Lan et al. 2017 using whole-genome sequencing at high

depths (~80X), which could be utilized for genotype comparisons for our current study. The average genotype concordance for SNPs and Indels was 99.54% and 96.03% (**Supplemental Table S1**), respectively.

Rare Indels variants were validated by Sanger sequencing (Zhen et al. 2019). We observed that 98.4% of individuals (123/125) had genotype concordance of 31 rare Indels from 125 samples of targeted sequence (**Supplemental Table S2**), which showed that the data quality of these Indels was highly reliable. Of 15 Indel PTVs assessed, 15 (100%) were confirmed.

1.6 Data annotation

After decomposition and normalization with vt v0.5772-60f436c3 (Tan et al. 2015), variants were annotated using Variant Effect Predictor version 92 (McLaren et al. 2016) and the LOFTEE v0.3-beta plugin (Karczewski 2020) to identify protein-truncating variants (PTVs), resulting in 8,720 stop gain, splice-site, or frameshift variants that are significantly predicted to disrupt gene function with "high confidence".

Combined Annotation Dependent Depletion (CADD) (Kircher et al. 2014) scores were annotated via online website https://cadd.gs.washington.edu/. Pathogenic variants were extracted from ClinVar (Landrum et al. 2018) (pathogenic or likely pathogenic, with at least 1 entry with 1 star) and HGMD (Stenson et al. 2017) ("DM" tags).

We used the following approaches to define the gene-set:

1. ExAC constrained genes: we downloaded information on the probability of being loss-of-function intolerant (*pLI*) from ftp://ftp.broadinstitute.org/pub/ExAC_release/release0.3.1/functional_gene_constraint/fordist_cleaned_exac_r03_march16_z_pli_rec_null_data.txt. We defined ExAC constrained genes as those having a $pLI \geq 0.9$.
2. ClinGen haploinsufficient, OMIM dominant and recessive, essential in culture, essential in mice, GWAS hits, FMRP interactors, and olfactory receptors: such gene lists were obtained from the ExAC (Exome Aggregation Consortium et al. 2016) via the website https://github.com/macarthur-lab/gene_lists.

2. Eight genes that were previously identified as targets of selection (besides *FUT2* and *EFCAB13)*

A total of 18 PTVs in 14 genes (**Fig. 5A; Supplemental Table S10**) were identified with significant population allele frequency differences ($Z(F_{ST})$ > 4.3; Bonferroni corrected P < 0.05): *FUT2*, *FMO2*, *TMPRSS3*, *EFCAB13*, *DMKN*, *FCGR2A*, *SLC5A9*, *OR8D4*, *WDR27*, *POLN*, *CST2*, *CLECL1*, *OR10X1* and *SLC22A1*. Of these 18 loci, 14 PTVs in 10 genes were previously reported to be under positive selection, harboring genes linked to immune response (*FUT2* (Karlsson et al. 2014; Ferrer-Admetlla et al. 2009a; Galinsky et al. 2016), *CLECL1* (Fu and Akey 2013a)), DNA repair (*POLN* (Arbiza et al. 2006a; DeGiorgio et al. 2014a)), protein binding (*WDR27* (Voight et al. 2006a), *CST2* (Tang et al. 2007; Clark et al. 2003a)), calcium ion binding (*EFCAB13* (Pickrell et al. 2009; Palamara et al. 2018)), drug transporter (*SLC22A1* (Li et al. 2011a, 2014a)), nutrition (*SLC5A9* (Prost et al. 2018a)) and sense of smell (*OR10X1* (Clark et al. 2003a), *OR8D4*).

CLECL1 (C-Type Lectin Like 1) encodes a type II transmembrane, which is highly expressed on dendritic and B cells and may act as a T-cell costimulatory molecule (Ryan et al. 2002, 4). *CLECL1* may function in mediating immune cell-cell interactions and thus is involved in immune response (Ishigaki et al. 2017). *CLECL1* is associated with white blood cell count, lymphocyte counts, multiple sclerosis and type 1 diabetes (MacArthur et al. 2017; Schadt et al. 2008). *Fu et al*. (Fu and Akey 2013b) suggested *CLECL1* was under selection.

POLN, Polymerase (DNA Directed) Nu, encodes a DNA polymerase type-A family member, which is indispensable for DNA cross-links repair and homologous recombination (Moldovan et al. 2010). *POLN* is associated with response to paliperidone in schizophrenia and Alzheimer's disease (MacArthur et al. 2017). *POLN* was reported to be under positive selection in the ancestral lineage of hominids (Arbiza et al. 2006b) while *DeGiorgio et al.* suggested *POLN* was under long-term balancing selection in in Africans and Europeans (DeGiorgio et al. 2014b).

WDR27 encodes a protein with multiple WD repeats, which may form scaffolds for protein-protein interaction and play key roles in cell signaling (Safran et al. 2010). *WDR27* is associated with eye morphology, type 1 diabetes, and insomnia (MacArthur et al. 2017), in which rs3736712 with significantly high F_{ST} was associated with eye tail length (Cha et al. 2018). *Voight et al*. showed evidence that *WDR27* was under recent positive selection in Europeans (Voight et al. 2006b).

CST2 (Cystatin SA) belongs to type 2 cystatin gene family and encodes a secreted thiol protease inhibitor, which is abundant in saliva, tears and seminal plasma (Safran et al. 2010). *CST2* is associated with blood protein levels, response to TNF antagonist treatment, body mass index (change over time) in cancer or chronic obstructive pulmonary disease (COPD) (MacArthur et al. 2017). *CST2* shows lineage-specific selection in humans compared with the chimpanzee and mouse lineages (Clark et al. 2003b) and recent positive selection in Europeans (K et al. 2007).

SLC22A1 (Solute Carrier Family 22 Member 1) encodes Organic Cation Transporter (OCT) 1 protein. *SLC22A1* functions as a cellular exporter of acylcarnitines in hepatocytes and is primarily expressed in the liver. The drug transporter *SLC22A1* appears to be under recent positive selection in the American (Li et al. 2011b), Chinese Uyghur(Li et al. 2014b) and Chinese Hui populations (Li et al. 2014b).

SLC5A9 (Solute Carrier Family 5 Member 9) encodes solute carrier family 5 (sodium/glucose cotransporter), member 9. *SLC5A9* is enriched in the small intestine in the GTEx data (Carithers and Moore 2015). *SLC5A9* is crucial for pathways of transport of glucose and other sugars, bile salts and organic acids, metal ions and amine compounds and metabolism (Safran et al. 2010; Tazawa et al. 2005, 4). *SLC5A9* is under positive selection in birds-of-paradise (Prost et al. 2018b).

The olfactory receptor (OR) genes make up the largest gene family in mammalian. While *Gimelbrant et al.* observed weak purifying selection affecting over half of OR repertoire (Gimelbrant et al. 2004), particular cases of positive selection in human OR genes have also been reported (Ferrer-Admetlla et al. 2009b; Clark et al. 2003b; Gilad et al. 2003; Nielsen et al. 2007), in which *OR10X1* (Olfactory Receptor Family 10 Subfamily X Member 1) with high F_{ST} in rs863362 has been reported to show lineage-specific selection in humans compared with the chimpanzee and mouse lineages (Clark et al. 2003b). *OR8D4* is also a member of the OR gene family.

3. Four genes without previous evidence of selection

TMPRSS3 (Transmembrane protease serine 3) encodes a transmembrane serine protease. Mutations in *TMPRSS3* were identified in congenital and childhood onset autosomal recessive deafness (Lee et al. 2003, 3). *TMPRSS3* is overexpressed in pancreatic cancer (Wallrapp et al. 2000). Rs34966432 is a benign splice acceptor variant.

FMO2 (Flavin-containing monooxygenase 2) encodes an enzyme called dimethylaniline monooxygenase [N-oxide-forming] 2. *FMO2* is engaged in the xenobiotics metabolism, including

therapeutic drugs (Phillips et al. 2007). *FMO2* is associated with blood protein levels (MacArthur et al. 2017) and tuberculosis (Mekonnen and Bekele 2017). A distinctive ethnic-geographic difference was reported in the expression and distribution of *FMO2* polymorphisms between African and non-African populations (Veeramah et al. 2008; Whetstine et al. 2000; Krueger et al. 2004). Rs28369860 is a benign frameshift as well as an eQTL for expression in artery, lung, adipose, muscle, thyroid, adrenal gland, esophagus, heart, skins, cells and stomach (Carithers and Moore 2015). This variant has a higher frequency in Africa, consistent with higher rates of functional *FMO2*1* in Africans than in other populations. However, previous studies showed no evidence of selection in *FMO2* (Veeramah et al. 2008) and suggested the most likely explanation for absence of functional *FMO2*1* allele outside Africa is due to its loss in a bottleneck during out-of-Africa (OOA) migration (Veeramah et al. 2008).

DMKN (Dermokine) encodes encoding dermokine. Expression of the *DMKN* results in production of four dermokine isoform families with different expression patterns and cellular locations (Leclerc et al. 2014). *DMKN* is overexpressed in inflammatory diseases and it was expressed in epithelial tissues in a transcriptionally and post-transcriptionally complex way (Naso et al. 2007) .

FCGR2A (Fc fragment of IgG receptor IIa) encodes low affinity immunoglobulin gamma Fc region receptor II-a, which is found on the surface of phagocytic cells such as macrophages and neutrophils (Safran et al. 2010; Guilliams et al. 2014). *FCGR2A* is overexpressed on whole blood (Carithers and Moore 2015). *FCFR2A* is associated with Kawasaki disease (Khor et al. 2011) and response to Anti-TNF Therapy in Rheumatoid Arthritis (Avila-Pedretti et al. 2015).

4. References for Supplemental Materials

Arbiza L, Dopazo J, Dopazo H. 2006a. Positive selection, relaxation, and acceleration in the evolution of the human and chimp genome. *PLoS Comput Biol* **2**: e38.

Arbiza L, Dopazo J, Dopazo H. 2006b. Positive selection, relaxation, and acceleration in the evolution of the human and chimp genome. *PLoS Comput Biol* **2**: e38.

Avila-Pedretti G, Tornero J, Fernández-Nebro A, Blanco F, González-Alvaro I, Cañete JD, Maymó J, Alperiz M, Fernández-Gutiérrez B, Olivé A, et al. 2015. Variation at FCGR2A and functionally related genes is associated with the response to anti-TNF therapy in rheumatoid arthritis. *PLoS ONE* **10**: e0122088.

Carithers LJ, Moore HM. 2015. The Genotype-Tissue Expression (GTEx) Project. *Biopreserv Biobank* **13**: 307–308.

Cha S, Lim JE, Park AY, Do J-H, Lee SW, Shin C, Cho NH, Kang J-O, Nam JM, Kim J-S, et al. 2018. Identification of five novel genetic loci related to facial morphology by genome-wide association studies. *BMC Genomics* **19**: 481.

Clark AG, Glanowski S, Nielsen R, Thomas PD, Kejariwal A, Todd MA, Tanenbaum DM, Civello D, Lu F, Murphy B, et al. 2003a. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios. *Science* **302**: 1960–1963.

Clark AG, Glanowski S, Nielsen R, Thomas PD, Kejariwal A, Todd MA, Tanenbaum DM, Civello D, Lu F, Murphy B, et al. 2003b. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios. *Science* **302**: 1960–1963.

DeGiorgio M, Lohmueller KE, Nielsen R. 2014a. A model-based approach for identifying signatures of ancient balancing selection in genetic data. *PLoS Genet* **10**: e1004561.

DeGiorgio M, Lohmueller KE, Nielsen R. 2014b. A model-based approach for identifying signatures of ancient balancing selection in genetic data. *PLoS Genet* **10**: e1004561.

DePristo MA, Banks E, Poplin RE, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, et al. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* **43**: 491–498.

Exome Aggregation Consortium, Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, et al. 2016. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**: 285–291.

Ferrer-Admetlla A, Sikora M, Laayouni H, Esteve A, Roubinet F, Blancher A, Calafell F, Bertranpetit J, Casals F. 2009a. A natural history of FUT2 polymorphism in humans. *Mol Biol Evol* **26**: 1993–2003.

Ferrer-Admetlla A, Sikora M, Laayouni H, Esteve A, Roubinet F, Blancher A, Calafell F, Bertranpetit J, Casals F. 2009b. A Natural History of FUT2 Polymorphism in Humans. *Molecular Biology and Evolution* **26**: 1993–2003.

Fu W, Akey JM. 2013a. Selection and Adaptation in the Human Genome. *Annu Rev Genom Hum Genet* **14**: 467–489.

Fu W, Akey JM. 2013b. Selection and Adaptation in the Human Genome. *Annu Rev Genom Hum Genet* **14**: 467–489.

Galinsky KJ, Loh P-R, Mallick S, Patterson NJ, Price AL. 2016. Population Structure of UK Biobank and Ancient Eurasians Reveals Adaptation at Genes Influencing Blood Pressure. *The American Journal of Human Genetics* **99**: 1130–1139.

Gilad Y, Bustamante CD, Lancet D, Pääbo S. 2003. Natural selection on the olfactory receptor gene family in humans and chimpanzees. *Am J Hum Genet* **73**: 489–501.

Gimelbrant AA, Skaletsky H, Chess A. 2004. Selective pressures on the olfactory receptor repertoire since the human-chimpanzee divergence. *Proc Natl Acad Sci USA* **101**: 9019–9022.

Guilliams M, Bruhns P, Saeys Y, Hammad H, Lambrecht BN. 2014. The function of Fcγ receptors in dendritic cells and macrophages. *Nat Rev Immunol* **14**: 94–108.

Guo X, Chen F, Gao F, Li L, Liu K, You L, Hua C, Yang F, Liu W, Peng C, et al. 2020. CNSA: a data repository for archiving omics data. *Database (Oxford)* **2020**.

Ishigaki K, Kochi Y, Suzuki A, Tsuchida Y, Tsuchiya H, Sumitomo S, Yamaguchi K, Nagafuchi Y, Nakachi S, Kato R, et al. 2017. Polygenic burdens on cell-specific pathways underlie the risk of rheumatoid arthritis. *Nat Genet* **49**: 1120–1125.

K T, Kr T, M S. 2007. A New Approach for Using Genome Scans to Detect Recent Positive Selection in the Human Genome. *PLoS biology* **5**. https://pubmed.ncbi.nlm.nih.gov/17579516/ (Accessed June 21, 2020).

Karczewski K. 2020. *konradjk/loftee*. https://github.com/konradjk/loftee (Accessed June 21, 2020).

Karlsson EK, Kwiatkowski DP, Sabeti PC. 2014. Natural selection and infectious disease in human populations. *Nat Rev Genet* **15**: 379–393.

Khor CC, Davila S, Breunis WB, Lee Y-C, Shimizu C, Wright VJ, Yeung RSM, Tan DEK, Sim KS, Wang JJ, et al. 2011. Genome-wide association study identifies FCGR2A as a susceptibility locus for Kawasaki disease. *Nat Genet* **43**: 1241–1246.

Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. 2014. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* **46**: 310–315.

Krueger SK, Siddens LK, Martin SR, Yu Z, Pereira CB, Cabacungan ET, Hines RN, Ardlie KG, Raucy JL, Williams DE. 2004. Differences in FMO2*1 allelic frequency between Hispanics of Puerto Rican and Mexican descent. *Drug Metab Dispos* **32**: 1337–1340.

Lan T, Lin H, Zhu W, Laurent TCAM, Yang M, Liu X, Wang J, Wang J, Yang H, Xu X, et al. 2017. Deep whole-genome sequencing of 90 Han Chinese genomes. *Gigascience* **6**: 1–7.

Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Jang W, et al. 2018. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res* **46**: D1062–D1067.

Leclerc EA, Huchenq A, Kezic S, Serre G, Jonca N. 2014. Mice deficient for the epidermal dermokine β and γ isoforms display transient cornification defects. *J Cell Sci* **127**: 2862–2872.

Lee YJ, Park D, Kim SY, Park WJ. 2003. Pathogenic mutations but not polymorphisms in congenital and childhood onset autosomal recessive deafness disrupt the proteolytic activity of TMPRSS3. *J Med Genet* **40**: 629–631.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754–1760.

Li J, Lou H, Yang X, Lu D, Li S, Jin L, Pan X, Yang W, Song M, Mamatyusupu D, et al. 2014a. Genetic architectures of ADME genes in five Eurasian admixed populations and implications for drug safety and efficacy. *J Med Genet* **51**: 614–622.

Li J, Lou H, Yang X, Lu D, Li S, Jin L, Pan X, Yang W, Song M, Mamatyusupu D, et al. 2014b. Genetic architectures of ADME genes in five Eurasian admixed populations and implications for drug safety and efficacy. *J Med Genet* **51**: 614–622.

Li J, Zhang L, Zhou H, Stoneking M, Tang K. 2011a. Global patterns of genetic diversity and signals of natural selection for human ADME genes. *Hum Mol Genet* **20**: 528–540.

Li J, Zhang L, Zhou H, Stoneking M, Tang K. 2011b. Global patterns of genetic diversity and signals of natural selection for human ADME genes. *Hum Mol Genet* **20**: 528–540.

MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, Junkins H, McMahon A, Milano A, Morales J, et al. 2017. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res* **45**: D896–D901.

Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W-M. 2010. Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**: 2867–2873.

McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, Flicek P, Cunningham F. 2016. The Ensembl Variant Effect Predictor. *Genome Biol* **17**: 122.

Mekonnen E, Bekele E. 2017. An ancestral human genetic variant linked to an ancient disease: A novel association of FMO2 polymorphisms with tuberculosis (TB) in Ethiopian populations provides new insight into the differential ethno-geographic distribution of FMO2*1. *PLoS ONE* **12**: e0184931.

Moldovan G-L, Madhavan MV, Mirchandani KD, McCaffrey RM, Vinciguerra P, D'Andrea AD. 2010. DNA polymerase POLN participates in cross-link repair and homologous recombination. *Mol Cell Biol* **30**: 1088–1096.

Naso MF, Liang B, Huang CC, Song X-Y, Shahied-Arruda L, Belkowski SM, D'Andrea MR, Polkovitch DA, Lawrence DR, Griswold DE, et al. 2007. Dermokine: an extensively differentially spliced gene expressed in epithelial cells. *J Invest Dermatol* **127**: 1622–1631.

Nielsen R, Hellmann I, Hubisz M, Bustamante C, Clark AG. 2007. Recent and ongoing selection in the human genome. *Nat Rev Genet* **8**: 857–868.

Palamara PF, Terhorst J, Song YS, Price AL. 2018. High-throughput inference of pairwise coalescence times identifies signals of selection and enriched disease heritability. *Nat Genet* **50**: 1311–1317.

Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis. *PLoS Genet* **2**: e190.

Phillips IR, Francois AA, Shephard EA. 2007. The flavin-containing monoooxygenases (FMOs): Genetic variation and its consequences for the metabolism of therapeutic drugs.

Pickrell JK, Coop G, Novembre J, Kudaravalli S, Li JZ, Absher D, Srinivasan BS, Barsh GS, Myers RM, Feldman MW, et al. 2009. Signals of recent positive selection in a worldwide sample of human populations. *Genome Research* **19**: 826–837.

Prost S, Armstrong EE, Nylander J, Thomas GWC, Suh A, Petersen B, Dalen L, Benz BW, Blom MPK, Palkopoulou E, et al. 2018a. Genome analyses show strong selection on coloration, morphological and behavioral phenotypes in birds-of-paradise. *bioRxiv* 287086.

Prost S, Armstrong EE, Nylander J, Thomas GWC, Suh A, Petersen B, Dalen L, Benz BW, Blom MPK, Palkopoulou E, et al. 2018b. Genome analyses show strong selection on coloration, morphological and behavioral phenotypes in birds-of-paradise. *bioRxiv* 287086.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**: 559–575.

Ryan EJ, Marshall AJ, Magaletti D, Floyd H, Draves KE, Olson NE, Clark EA. 2002. Dendritic cell-associated lectin-1: a novel dendritic cell-associated, C-type lectin-like molecule enhances T cell secretion of IL-4. *J Immunol* **169**: 5638–5648.

Safran M, Dalah I, Alexander J, Rosen N, Iny Stein T, Shmoish M, Nativ N, Bahir I, Doniger T, Krug H, et al. 2010. GeneCards Version 3: the human gene integrator. *Database (Oxford)* **2010**: baq020.

Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY, Kasarskis A, Zhang B, Wang S, Suver C, et al. 2008. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol* **6**: e107.

Stenson PD, Mort M, Ball EV, Evans K, Hayden M, Heywood S, Hussain M, Phillips AD, Cooper DN. 2017. The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Hum Genet* **136**: 665–677.

Tan A, Abecasis GR, Kang HM. 2015. Unified representation of genetic variants. *Bioinformatics* **31**: 2202–2204.

Tang H, Jin X, Li Y, Jiang H, Tang X, Yang X, Cheng H, Qiu Y, Chen G, Mei J, et al. 2014. A large-scale screen for coding variants predisposing to psoriasis. *Nat Genet* **46**: 45–50.

Tang K, Thornton KR, Stoneking M. 2007. A New Approach for Using Genome Scans to Detect Recent Positive Selection in the Human Genome ed. K.H. Wolfe. *PLoS Biol* **5**: e171.

Tazawa S, Yamato T, Fujikura H, Hiratochi M, Itoh F, Tomae M, Takemura Y, Maruyama H, Sugiyama T, Wakamatsu A, et al. 2005. SLC5A9/SGLT4, a new Na+-dependent glucose transporter, is an essential transporter for mannose, 1,5-anhydro-D-glucitol, and fructose. *Life Sci* **76**: 1039–1050.

Tennessen JA, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, McGee S, Do R, Liu X, Jun G, et al. 2012. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* **337**: 64–69.

The 1000 Genomes Project Consortium, Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, Zhang Y, Ye K, Jun G, et al. 2015. An integrated map of structural variation in 2,504 human genomes. *Nature* **526**: 75–81.

Veeramah KR, Thomas MG, Weale ME, Zeitlyn D, Tarekegn A, Bekele E, Mendell NR, Shephard EA, Bradman N, Phillips IR. 2008. The potentially deleterious functional variant flavin-containing monooxygenase 2*1 is at high frequency throughout sub-Saharan Africa. *Pharmacogenet Genomics* **18**: 877–886.

Voight BF, Kudaravalli S, Wen X, Pritchard JK. 2006a. A Map of Recent Positive Selection in the Human Genome ed. L. Hurst. *PLoS Biol* **4**: e72.

Voight BF, Kudaravalli S, Wen X, Pritchard JK. 2006b. A Map of Recent Positive Selection in the Human Genome ed. L. Hurst. *PLoS Biol* **4**: e72.

Wallrapp C, Hähnel S, Müller-Pillasch F, Burghardt B, Iwamura T, Ruthenbürger M, Lerch MM, Adler G, Gress TM. 2000. A novel transmembrane serine protease (TMPRSS3) overexpressed in pancreatic cancer. *Cancer Res* **60**: 2602–2606.

Whetstine JR, Yueh MF, McCarver DG, Williams DE, Park CS, Kang JH, Cha YN, Dolphin CT, Shephard EA, Phillips IR, et al. 2000. Ethnic differences in human flavin-containing monooxygenase 2 (FMO2) polymorphisms: detection of expressed protein in African-Americans. *Toxicol Appl Pharmacol* **168**: 216–224.

Zhen Q, Yang Z, Wang W, Li B, Bai M, Wu J, Ge H, Dong Z, Shen J, Tang H, et al. 2019. Genetic Study on Small Insertions and Deletions in Psoriasis Reveals a Role in Complex Human Diseases. *Journal of Investigative Dermatology* S0022202X19315519.