The American Journal of Human Genetics, Volume *109*

Supplemental information

A genome sequencing system for universal newborn screening, diagnosis, and precision medicine for severe genetic diseases

Stephen F. Kingsmore, Laurie D. Smith, Chris M. Kunard, Matthew Bainbridge, Sergey Batalov, Wendy Benson, Eric Blincow, Sara Caylor, Christina Chambers, Guillermo Del Angel, David P. Dimmock, Yan Ding, Katarzyna Ellsworth, Annette Feigenbaum, Erwin Frise, Robert C. Green, Lucia Guidugli, Kevin P. Hall, Christian Hansen, Charlotte A. Hobbs, Scott D. Kahn, Mark Kiel, Lucita Van Der Kraan, Chad Krilow, Yong H. Kwon, Lakshminarasimha Madhavrao, Jennie Le, Sebastien Lefebvre, Rebecca Mardach, William R. Mowrey, Danny Oh, Mallory J. Owen, George Powley, Gunter Scharer, Seth Shelnutt, Mari Tokita, Shyamal S. Mehtalia, Albert Oriol, Stavros Papadopoulos, James Perry, Edwin Rosales, Erica Sanford, Steve Schwartz, Duke Tran, Martin G. Reese, Meredith Wright, Narayanan Veeraraghavan, Kristen Wigby, Mary J. Willis, Aaron R. Wolen, and Thomas Defay.

Supplemental Information

Supplemental Figures

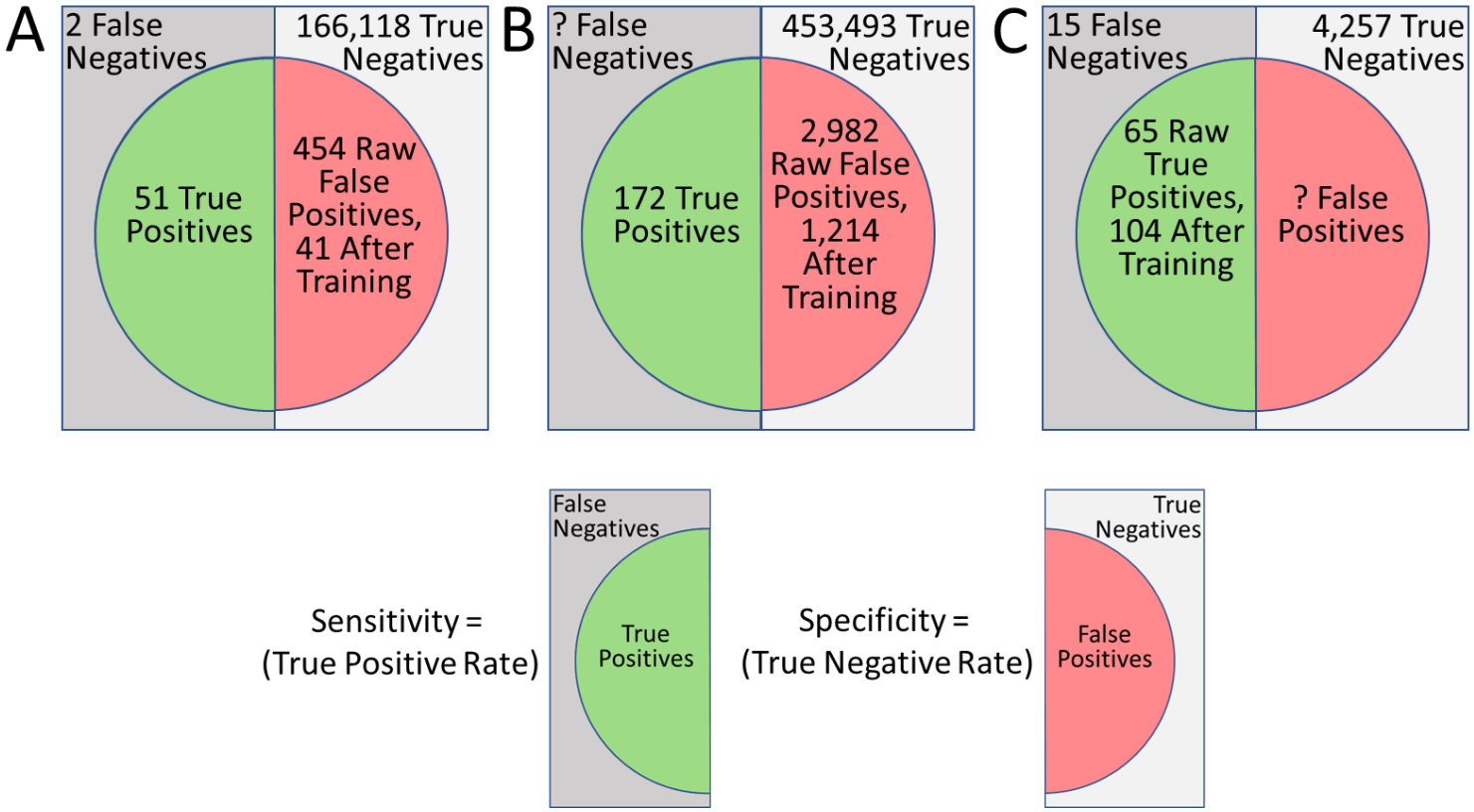

Figure S1: Impact of training on the sensitivity (true positive rate) and specificity (true negative rate) of NBS-MS and NBS-rWGS. A. Postanalytical tools reduced false positives from NBS-MS of 48 disorders from 454 to 41, improving specificity from 99.7% to 99.98%.[47] Of note, NBS-MS false positives excluded newborns with birth weight <1.8 kg and DBS obtained at <24 hours or >7 days.[47] **B.** Root cause analysis reduced false positives from NBS-rWGS of 388 disorders from 2,982 to 1,214, improving specificity from 99.3% to 99.7%. **C.** Addition of positive individuals by GEM and inclusion of ClinVar 3712323 increased NBS-rWGS true positives from 65 to 104 of 119, improving sensitivity from 55% to 87%. Of note, these results included NBS-rWGS of newborns with birth weight <1.8 kg and DBS obtained at >7 days.

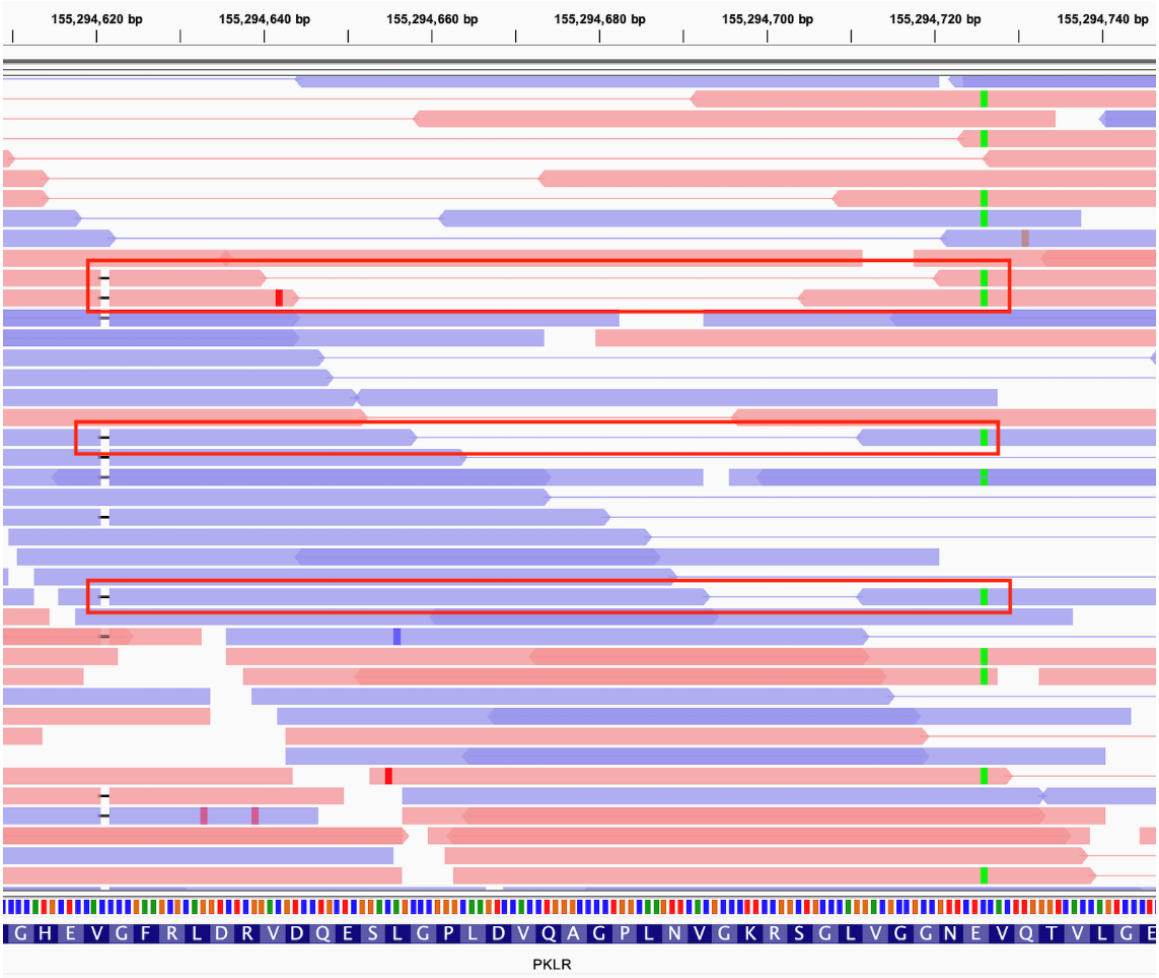

Figure S2: Visualization of paired sequence reads on a 120 nt region of Chr 1 demonstrating that ClinVar variants 280113 (*PKLR* g.155,294,726G>T, NM_000298.6: c.721G>T p.Glu241Ter), shown in green, and 1163645 (*PKLR* g.155294621del, NM_000298.6: c.826del p.Val276fs*45), shown as a black hash, occurred in the same read in a positive UKBB subject (red boxes).

Supplemental Tables

Table S1: Correction of gene association for ClinVar variants with more than one associated gene.

Table S2: Quality control (QC) metrics for gDNA extraction and WGS from 5 x 3 mm^2 punches from 260 archived CA dried blood spots (DBS). Abbreviations: GB, Gigabases; nt, nucleotides; Std Dev, standard deviation; Hom, homozygous; Het, heterozygous; Ti, transition; Tv, transversion; GC, guanine cytosine; CCDS CNV, Copy number variants overlapping consensus coding domains; MIM, Mendelian Inheritance in Man.

Table S3: Results of Delphi adjudication for suitability for NBS by rapid WGS (NBS-rWGS) of 457 disease-gene dyads (446 disorders associated with 346 genes) that are included in Genome to Treatment (GTRx). 388 (85%) gene-disorder dyads (317 [92%] genes associated with 381 [85%] disorders) were retained for evaluation in retrospective datasets. Group A (295, 76%) were conditions for which there were not major gaps in the evidence, high likelihood of benefit, and low risk of harm. Group B (93, 24%) were conditions for which there were gaps in the evidence or uncertainty regarding net benefit that required further assessment by NBS-rWGS implementation research. Abbreviations: Inh, Inheritance; AR, Autosomal recessive; AD, Autosomal dominant; XR, X-linked recessive; XD, X-linked dominant. Note that carrier females may exhibit symptoms in some XR disorders due to effects of Lyonization.

Table S4: Approximate US population incidence of the 388 disorders retained for NBS-rWGS. A single value is given for rare disorders with multiple causative genes. Note that disorders may be significantly more common in specific racial,

ethnic, and ancestral groups. The published incidence for many disorders is an underestimate since it is based on definitive diagnosis in symptomatic individuals, which is frequently not pursued to comprehensive methods such as WES or WGS. Abbreviations: MIM, Mendelian Inheritance in Man; 100k, 100,000 individuals.

Table S5: UK Biobank subject counts by zygosity underpinning 2,982 positive diplotypes in 454,707 UK Biobank subjects in 388 NBS-rWGS disorders. Abbreviations: Chr, Chromosome.

Table S6: Results of root cause analysis of 2,982 positive diplotypes in 454,707 UK Biobank subjects in 388 NBS-rWGS disorders. 94 variants were block-listed as a result of root cause analysis. Abbreviations: Inh, inheritance; AR, autosomal recessive; XR, X-linked recessive, XD, X-linked dominant; AD, autosomal dominant; LB, likely benign; VUS, variant of uncertain significance; FP, false positive. *ClinVar variant 431443 (*INS* NC_000011.10: g.2161302G>C, NM_000207.2: c.-152C>G) is associated with AR permanent neonatal diabetes mellitus, rather than AD (Demiral M, Demirbilek H, Çelik K, Okur N, Hussain K, Ozbek MN. Neonatal diabetes due to homozygous INS gene promoter mutations: Highly variable phenotype, remission and early relapse during the first 3 years of life. Pediatr Diabetes. 2020 21:1169-1175). **ClinVar variant 132994 (*RYR1* NM_000540.3: c.10348-6C>G) is associated with AR autosomal recessive myopathy and ophthalmoplegia (MIM: 255320), rather than AD (Wilmshurst, J. M., Lillis, S., Zhou, H., Pillay, K., Henderson, H., Kress, W., Muller, C. R., Ndondo, A., Cloke, V., Cullup, T., Bertini, E., Boennemann, C., et al. (2010). RYR1 mutations are a common cause of congenital myopathies with central nuclei. Ann. Neurol. 68, 717-726).

Table S7: Root cause analysis of 112 Hemophilia A-positive, ICD10 code D66 negative UKBB subjects by genomic NBS. The severity of Hemophilia A disease associated with *F8* variants was from the CDC Hemophilia Mutation Project Database (https://www.cdc.gov/ncbddd/hemophilia/champs.html). The accession number for F8 was NM_000132.4. Abbreviations: VUS, variant of uncertain significance; LP, likely benign; TP, true positive.

Table S8: Comparison of results of prospective Dx-rWGS with retrospective NBS-rWGS for 388 disorders in 4,376 critically ill children with suspected genetic disorders, and their parents. Abbreviations: ID, subject ID; Inh, inheritance; Class, variant pathogenicity classification; Zyg, Variant zygosity; Chr, Chromosome; AD, autosomal dominant; AR, autosomal recessive; XR, X-linked recessive, XD, X-linked dominant; P, Pathogenic; LP, Likely pathogenic; CHT, compound heterozygous; Het, heterozygous; Hom, Homozygous; Hemi, Hemizygous.

Table S9: Delphi panel counterfactual analysis of clinical utility of earlier diagnosis by NBS-rWGS compared with actual age at diagnosis by rWGS in 43 children. Abbreviations: ID, subject ID; DOL, Day of life. *Henter, J.I., Horne, A., Aricó, M., Egeler, R.M., Filipovich, A.H., Imashuku, S., Ladisch, S., McClain, K., Webb, D., Winiarski, J., et al. (2007). HLH-2004: Diagnostic and therapeutic guidelines for hemophagocytic lymphohistiocytosis. Pediatr. Blood Cancer. 48, 124-31.

Table S10: Comparison of P and LP variants identified by ClinVar and Mastermind (Genomenon) in 28 disorders associated with 16 genes.