CLINICAL RESEARCH ARTICLE

Genetic associations of hemoglobin in children with chronic kidney disease in the PediGFR Consortium

Meredith A. Atkinson¹, Rui Xiao², Anna Köttgen³, Elke Wühl⁴, Craig S. Wong⁵, Matthias Wuttke³, Aysun K. Bayazit⁶, Salim Çalışkan⁷, Bradley A. Warady⁸, Franz Schaefer⁴ and Susan L. Furth²

BACKGROUND: Genome-wide association studies (GWAS) in healthy populations have identified variants associated with erythrocyte traits, but genetic causes of hemoglobin variation in children with CKD are incompletely understood.

METHODS: The Pediatric Investigation of Genetic Factors Linked with Renal Progression (PediGFR) Consortium comprises three pediatric CKD cohorts: Chronic Kidney Disease in Children (CKiD), Effect of Strict Blood Pressure Control and ACE Inhibition on the Progression of CRF in Pediatric Patients (ESCAPE), and Cardiovascular Comorbidity in Children with CKD (4C). We performed cross-sectional and longitudinal association studies of single-nucleotide polymorphisms (SNPs) in 1125 patients.

RESULTS: Children of European ($n = 725$) or Turkish ($n = 400$) ancestry (EA or TA) were included. In cross-sectional analysis, two SNPs (rs10758658 and rs12718597) previously associated with RBC traits were significantly associated with hemoglobin levels in children of EA and TA. In longitudinal analysis, SNP rs2540917 was nominally associated with hemoglobin in EA and TA children.

CONCLUSIONS: SNPs associated with erythrocyte traits in healthy populations were marginally significant for an association with hemoglobin. Further analyses/replication studies are needed in larger CKD cohorts to investigate SNPs of unknown significance associated with hemoglobin. Functional studies will be required to confirm that the observed associations between SNPs and clinical phenotype are causal.

Pediatric Research (2019) 85:324–328; <https://doi.org/10.1038/s41390-018-0148-z>

INTRODUCTION

Anemia is a prevalent comorbidity in children with chronic kidney disease (CKD), attributable to a number of physiologic factors including impaired erythropoiesis, iron deficiency, and inflammation/iron restriction.¹ Despite the high prevalence of anemia in this population, there is substantial variation in individual hemoglobin levels at any given glomerular filtration rate (GFR), as well as by race/ethnicity and nationality.^{2–4} Well-established differences in normative hemoglobin levels in individuals of European vs. African ancestry (both healthy and with CKD),^{4,5} along with recent studies showing that specific genetic variants are associated with a variety of red blood cell traits,^{6–8} suggest that genetic variability is also a significant contributor to the risk for anemia, but currently the genetic causes of hemoglobin variation in children with CKD are incompletely understood.

Most children with the anemia of CKD receive standard treatment with erythropoiesis-stimulating agents (ESA) and iron supplementation, but the prevalence of treatment-resistant anemia is high.⁹ In light of adult clinic trial data regarding the adverse cardiovascular outcomes associated with the use of escalating ESA doses and observational evidence in children that higher ESA doses are associated with increased risk for mortality,^{3,10–12} it would be useful to determine how genetic

variation contributes to hemoglobin variability, decline, and resistance to anemia treatment in patients with kidney disease, and if the effects are magnified in this population.

Genome-wide association studies (GWAS) allow for the identification of genomic regions in individuals where genetic variations, or single-nucleotide polymorphisms (SNPs), are associated with unique phenotypes.^{13,14} GWAS provide a unique opportunity to delineate the genetic epidemiology and molecular basis of a variety of CKD comorbidities including anemia.¹⁴ A 2009 cross-sectional GWAS in 24,167 European-ancestry individuals showed that lower hemoglobin levels were associated with SNPs in genes known to regulate erythropoiesis and red cell membrane function, iron homeostasis, and hepcidin expression.⁶ In addition, a GWAS in healthy children identified genetic variants cross-sectionally associated with a variety of erythrocyte traits,⁷ but these associations have not been previously studied in children with CKD, in whom the effects may be more pronounced due to kidney disease-associated erythropoietin deficiency/dysregulation and iron restriction. In turn, the goal of this analysis was to investigate the cross-sectional association of hemoglobin in children with CKD with SNPs previously identified as associated with hemoglobin levels in community-based adult and pediatric populations.^{6,7} In addition, we sought to determine whether any

¹Division of Pediatric Nephrology, Johns Hopkins University School of Medicine, Baltimore, MD, USA; ²Department of Biostatistics, Epidemiology & Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; ³Division of Genetic Epidemiology, Institute for Medical Biometry and Statistics, Medical Center and Faculty of Medicine, University of Freiburg, Freiburg, Germany; ⁴Pediatric Nephrology Division, University of Heidelberg, Heidelberg, Germany; ⁵Department of Pediatrics, Division of Nephrology, University of New Mexico Children's Hospital, Albuquerque, NM, USA; ⁶Department of Pediatric Nephrology, Cukurova University, Adana, Turkey; ⁷Department of Pediatric Nephrology, Istanbul University, Cerrahpasa Medical Faculty, Istanbul, Turkey and ⁸Division of Nephrology, Children's Mercy Hospital, Kansas City, MO, USA
Correspondence: Meredith A. Atkinson (matkins3@jhmi.edu)

These authors contributed equally: Meredith A. Atkinson, Rui Xiao.

Received: 30 March 2018 Revised: 30 May 2018 Accepted: 4 June 2018

Published online: 15 August 2018

Table 1. Demographic and clinical characteristics at baseline and follow-up time in 1125 children

	European ancestry			Turkish ancestry	
	CKiD (<i>n</i> = 262)	4C (<i>n</i> = 251)	Escape (<i>n</i> = 212)	4C (<i>n</i> = 339)	Escape (<i>n</i> = 61)
Baseline characteristics					
Age, years – median (IQR)	11.1 (7.4–14.4)	12.1 (8.9–15.0)	11.8 (8.6–15.1)	12.4 (9.5–15.3)	12.0 (10.1–14.3)
Female, <i>n</i> (%)	95 (36.3%)	76 (30.3%)	88 (41.5%)	127 (37.5%)	28 (45.9%)
eGFR, ml/min/1.73 m ² – median (IQR)	43.2 (33.1–54.7)	25.1 (17.1–34.4)	36.1 (25.1–45.1)	25.9 (17.6–34.3)	27.5 (19.5–46.3)
ESA treated – <i>n</i> (%)	30 (11.5%)	89 (35.5%)	16 (7.5%)	55 (16.2%)	8 (13.2%)
ACEi treated – <i>n</i> (%)	118 (45%)	119 (47.4%)	0 (0%)	113 (33.3%)	0 (0%)
Iron treated – <i>n</i> (%)	72 (27.5%)	81 (32.3%)	23 (10.8%)	158 (46.6%)	24 (39.3%)
Non-glomerular CKD – <i>n</i> (%)	230 (87.8%)	198 (78.9%)	185 (87.3%)	315 (92.9%)	53 (86.9%)
Longitudinal characteristics					
Number of visits, mean (SD)	5.1 ± 1.9	4.0 ± 2.0	22.6 ± 12.0	4.2 ± 2.0	15.7 ± 10.3
Follow-up period, years – mean (SD)	4.19 ± 2.02	1.78 ± 1.06	4.46 ± 2.21	1.70 ± 1.04	2.82 ± 2.06

of these SNPs are associated with hemoglobin decline over time in the setting of progressive CKD.

MATERIALS AND METHODS

Sample description

We utilized the genotype and phenotypic data on hemoglobin in the Pediatric Investigation for Genetic Factors Linked to Renal Progression (PediGFR) Consortium. PediGFR is composed of three studies of CKD in children: the Chronic Kidney Disease in Children (CKiD) cohort study (*n* = 262), the Effect of Strict Blood Pressure Control and ACE Inhibition on Progression of Chronic Renal Failure in Pediatric Patients (ESCAPE) trial (*n* = 273), and the Cardiovascular Comorbidity in Children with CKD (4C) study (*n* = 590).^{15–17} Each participating study was approved by its respective Institutional Review Board, and all study subjects or parents/legal guardians provided informed consent for participation in the study and genetic research. In this analysis, participants were excluded if they self-reported non-European or non-Turkish ancestry.

SNP genotyping and imputation

Using a candidate marker approach, we sought to determine the association between genetic variation and hemoglobin levels in these pediatric CKD cohorts via the analysis of the most promising SNP markers identified previously. Hemoglobin was measured locally at individual study sites. DNA was genotyped using the Illumina® Infinium 2.5M-8 microarray, and genotype imputation was performed according to standard protocols¹⁸ separately for each group using SHAPEIT v2.r644¹⁹ and IMPUTE2 v2.3.0.²⁰ Phased haplotypes from the 1000 Genomes project²¹ were used as the reference panel. Details of genotyping, imputation, and quality control have been previously described.¹³ Twenty-four previously reported SNPs associated with red blood cell traits were included for the analysis.^{6,7} All candidate SNPs and subjects included in the analysis passed the quality control.

Statistical analysis

Standard descriptive statistics such as mean, standard deviation, and frequency were used to summarize the patient characteristics. The distribution of the outcome variable was assessed graphically in order to meet the normality assumption for modeling. For each candidate SNP, linear mixed-effects models were used to assess the association of hemoglobin measured repeatedly over time and allelic dosage of the SNP obtained from imputation. This

model accounts for within-subject correlation due to repeated measures by including subject-specific random effect. Models were adjusted for age at baseline, cohort indicator (CKiD, ESCAPE, or 4C), and the first three components from the principle components analysis (PCA) to account for potential population stratification. Medication use (iron, ESA, angiotensin converting enzyme inhibitors) was included in models as the percentage of time (months prescribed medication/total months of follow-up) on each of the medications. Estimated glomerular filtration rate (eGFR) based on height and serum creatinine²² as well as its interaction with SNP dosage was included as time-varying covariates. Separate analyses were conducted for different ethnicity groups. Cross-sectional analysis was performed on baseline hemoglobin by linear regression models for each candidate SNP. Similarly, cohort indicator, age and eGFR at baseline, medication use at baseline, and the first three components from PCA, were included in the linear regression models. False discovery rate (FDR) was applied to correct for multiple testing. SNP annotation was performed using SnpPA²³ (<http://snipa.helmholtz-muenchen.de/snipa/>), the UCSC Genome Browser²⁴ (<https://genome.ucsc.edu/>), and the data from the Sankaran lab²⁵ (http://www.bloodgenes.org/RBC_MPRA/view/index.html).

RESULTS

We conducted a candidate SNP association study in our pediatric population of 1125 children of European (*n* = 725) or Turkish (*n* = 400) ancestry. Baseline characteristics and follow-up time for subjects by ethnicity and study cohort are presented in Table 1. Median GFR at baseline was noted to be higher in the children in the CKiD cohort compared to the other study cohorts. In terms of medication use at baseline, a higher proportion of children of European ancestry enrolled in the 4C cohort were treated with an ESA at baseline compared to the other groups. No ESCAPE subjects were treated with ACE-inhibitors at baseline, given that it was a trial in which an ACE-inhibitor was the intervention of interest. Children of Turkish ancestry enrolled in the 4C cohort had the highest prevalence of iron use at baseline. Children of European ancestry enrolled in CKiD demonstrated the longest median follow-up time at 4.89 years, with 4C subjects of both ethnicities followed for just under 2 years.

Table 2. Significant results of multivariate^a cross-sectional analysis of SNP and hemoglobin

SNP	Unadjusted <i>p</i> -value	Adjusted <i>p</i> -value	Coefficient estimate ± SD	Minor allele frequency	Chromosome	Reference allele and assoc. trait in CHARGE ^a	Coefficient estimate in CHARGE ^a
European ancestry							
rs10758658	0.012	0.29	0.53 ± .21	0.19	9	A, MCH	−0.0048
rs12718597	0.044	0.052	0.36 ± .18	0.37	7	A, MCV	0.0032
Turkish ancestry							
rs10758658	0.051	0.19	−0.68 ± .35	0.15	9	A, MCH	−0.0048
rs12718597	0.056	0.19	0.49 ± .25	0.39	7	A, MCV	0.0032
rs172629	0.038	0.19	0.74 ± .35	0.20	4	G, MCV	−0.0043
rs2075671	0.051	0.19	0.69 ± .35	0.16	7	A, RBC	0.0086
rs2540917	0.031	0.19	0.56 ± .26	0.29	2	C, MCV	−0.0031
rs7255045	0.036	0.19	0.71 ± .34	0.23	19	A, MCV	−0.0037
rs857721	0.030	0.19	0.62 ± .28	0.27	1	A, MCHC	−0.0022

MCH mean corpuscular hemoglobin, *MCV* mean corpuscular volume, *MCHC* mean corpuscular hemoglobin concentration, *RBC* red blood cell count
^aModel adjusted for age at baseline, eGFR, ESA use, iron use, ACEi use, cohort indicator, and the first three PCA components

Table 3. Significant results of multivariate^a longitudinal analysis of SNP and hemoglobin

SNP	Unadjusted <i>p</i> -value	Adjusted <i>p</i> -value	Coefficient estimate ± SD	Minor allele frequency	Chromosome	Reference allele and assoc. trait in CHARGE ^a	Coefficient estimate
European ancestry							
rs10758658	0.074	0.40	0.27 ± .15	0.19	9	A, MCH	−0.0048
rs12718597	0.083	0.40	0.20 ± .11	0.37	7	A, MCV	0.0032
rs2540917	<0.001	<0.002	−0.57 ± .12	0.44	2	C, MCV	−0.0031
rs7255045	0.0047	0.06	−0.37 ± .13	0.31	19	A, MCV	−0.0037
rs7385804	0.013	0.10	0.30 ± .12	0.39	7	C, Hct	−0.1592
Turkish ancestry							
rs10758658	0.016	0.19	−0.52 ± .007	0.15	9	A, MCH	−0.0048
rs2540917	0.004	0.096	0.45 ± 0.005	0.29	2	C, MCV	−0.0031
rs9483788	0.047	0.37	−0.37 ± 0.17	0.22	6	C, Hct	0.2172

MCH mean corpuscular hemoglobin, *MCV* mean corpuscular volume, *HCT* hematocrit
^aModel adjusted for age at baseline, eGFR, ESA use, iron use, ACEi use, cohort indicator, and the first three PCA components

Cross-sectional analysis

Table 2 displays results of a cross-sectional analysis of the association of SNPs with hemoglobin at baseline. Among children of European ancestry, two SNPs (rs10758658 and rs12718597) were significantly associated with hemoglobin ($p = .012$ and $.044$, respectively). These two SNPs also showed marginal significance in children of Turkish ancestry, as did rs2075671. In addition, in the Turkish cohort, four SNPs (rs17218597, rs2540917, rs7255045, and rs857721) were identified to be significantly associated with hemoglobin ($p < .05$). Effects on hemoglobin were all of a similar magnitude in the model adjusted for other covariates. However, none of the SNPs passed the threshold after FDR correction for multiple testing.

Longitudinal analysis

Table 3 summarizes results of longitudinal analysis of the association between SNPs and hemoglobin measured over time. In children of European ancestry, the two SNPs (rs10758658 and rs12718597) associated (marginally) with baseline hemoglobin in both European and Turkish cohorts were also marginally significant in longitudinal analysis ($p = .074$ and $.083$). Among the two SNPs, rs10758658 was also observed to be significant in Turkish children ($p = .016$). Three additional SNPs (rs2540917,

rs7255045, and rs7385804) were identified in longitudinal analysis ($p < .0001$, $p = .0047$, and $p = .013$, respectively) in children of European ancestry, among which the SNP rs2540917 also showed significant association in Turkish children ($p = .004$). There was an additional SNP (rs9483788) identified to be marginally associated with hemoglobin in Turkish children ($p = .047$). Effect size of association with hemoglobin was of a similar magnitude for all the identified SNPs adjusted for other covariates.

DISCUSSION

We performed cross-sectional and longitudinal analyses to examine the associations between hemoglobin and 24 SNP genotypes previously identified in cross-sectional analyses to be associated with hematological traits in both adults and children,^{6,7} in a cohort of children with CKD of European and Turkish ancestry. Among the previously reported loci, 7 reached significance for an association with Hgb at baseline (2 in both Europeans and Turks and an additional 5 in Turks only), while 6 reached significance for an association with longitudinal change in hemoglobin (2 in both Europeans and Turks, 3 in Europeans only and 1 in Turks only). SNPs rs12718597 and rs10758658, intron variants within the IKZF1 and RCL1 genes, respectively, were associated with hemoglobin

both cross-sectionally and longitudinally among children of European ancestry, while the SNP on the RCL1 gene was also associated with hemoglobin both at baseline and longitudinally among children of Turkish ancestry.^{23,24} Both the IKZF1 and RCL1 genes are associated with mean corpuscular volume (MCV),^{23,24} a measure of the average volume of red cells that is often positively correlated with hemoglobin.²⁶ rs2540917, which was associated with hemoglobin both cross-sectionally and longitudinally in children of Turkish ancestry, has in previously published studies been suggested to be on the BCL11A gene,⁶ which has also been associated with MCV.²⁷ The transcription factor BCL11A is one of the most potent known regulators of the fetal-to-adult hemoglobin switch/fetal hemoglobin silencing that occurs shortly after birth, and is considered a promising therapeutic target in the setting of sickle cell disease and β -thalassemia.²⁸ However, in our annotation we were unable to confirm a specific gene or erythrocyte trait associated with rs2540917. Similarly rs172629, which we found to be cross-sectionally associated with hemoglobin in Turks, has been previously reported to be on the KIT gene and associated with MCV;⁶ our annotation did not confirm a specific gene or red cell trait association for this intergenic variant.

Other genes on which SNPs were associated with hemoglobin in our analysis have been associated with MCV (rs7255045, RTDBN), hematocrit (rs7385804, TFR2; and rs9483788, HBS1L/MYB), or other erythrocyte phenotypes (rs2075671, ZAN; and rs857721, SPTA1).^{23,24} The KIT and RTBDN genes have roles in hematopoiesis and retinoid binding, respectively.⁶ TFR2 (transferrin receptor 2) mediates cellular uptake of transferrin-bound iron and is involved in iron metabolism, while HBS1L/MYB is a quantitative trait locus controlling fetal hemoglobin and erythrocyte volume and hemoglobin content.^{6,27} Although not all of the SNPs most strongly associated with hemoglobin in this analysis are intron variants which may limit their biologic plausibility as regulators of red cell phenotype, some may be linked SNPs which do not reside within a specific gene or affect protein function, but nevertheless may correspond to a particular red cell response.²⁹ An estimated ~80% of the phenotypic heritability in common traits may be explained by non-coding regulatory variants, which may alter the regulation of gene transcription.²⁵ However, establishing the causality of GWAS-identified regulatory variants entails identification of target genes and demonstration of molecular function and connection to phenotype, which are beyond the scope of the present study.²⁵

Our results also highlight loci which demonstrate a differential association with hemoglobin between different ancestral groups. Both at baseline and longitudinally, a SNP on the RCL1 gene was associated with a higher or increasing hemoglobin in Europeans compared to a lower or decreasing hemoglobin in Turks. Similarly, rs2540917 was associated with a longitudinal decrease in hemoglobin in Europeans while the opposite association was observed in Turks. While the observation of opposite associations for a causal variant between studies sampled from one population may be unexpected, opposite associations across different ethnic groups do occur and have been reported in other studies,^{30,31} and may be explained by population differences. For example, opposite associations may indicate heterogeneous effects of the genetic variant resulting from different genetic background. It is known that linkage disequilibrium (LD) patterns of the same genetic region can differ between ethnic populations,^{32,33} leading to inconsistent association patterns when non-causal variants are tested.³⁴ Therefore, if different LD structures are present in our samples across the region and the SNP being tested is not causal, it is possible to see the opposite associations between the two populations. In addition, this phenomenon is also likely the result of the complex interplay of multiple loci and environmental factors which may confound the association with hemoglobin levels. Examining the association between a single locus and the outcome of interest

fails to account for multi-locus effect or take into account correlations with other causal variants. Differences in the minor allele frequency for each of these SNPs by ethnicity may also explain the observed differential association with hemoglobin. Despite adjustment for principal components, there may be additional, ancestry-specific confounders of the association between SNPs of interest and hemoglobin, such as differences in the prevalence of thalassemias between subjects of European and Turkish ancestry, to account for differential findings.

Our study does have limitations. Given the relatively small sample size of our cohort, we were limited to a candidate gene approach to identifying relevant SNPs, as the cohort would be underpowered for a GWAS; this precluded the identification of novel SNPs associated with hemoglobin in this cohort.³⁵ Small sample size also limits interpretation of nominal vs. genome-wide significance of results, but the presence of an association between specific SNPs and hemoglobin in both ancestry cohorts is suggestive of an association despite the lack of significance after multiple correction. Given the low prevalence of CKD in children, there is currently no larger group of subjects in which to explore this association. Due to limitations of the available hematologic data in PediGFR subjects, we were restricted to hemoglobin as the single marker of erythrocyte phenotype in this analysis, rather than examining a wide variety of erythrocyte traits as in previously reported studies. However, many of the previously examined phenotypes including MCV are positively correlated with hemoglobin. Anemia-targeted therapies including ESA and iron supplements were included in the analysis as percentage of observed time subjects were prescribed the interventions, but data on specific ESA and iron dose were not available for all subjects; thus, we were not able to additionally account for the effects of ESA or iron dose changes on hemoglobin. There was variation in subject follow up time by study cohort, which may have limited our ability to observe a change in hemoglobin in subjects followed for shorter periods of time. The previously identified SNPs we examined were noted to be cross-sectionally associated with hemoglobin in community based cohorts;^{6,7} a risk allele for lower hemoglobin may not necessarily be associated with faster hemoglobin decline, limiting the interpretation of our longitudinal analysis. Finally, given the restricted ethnicities of children included in this analysis, these findings may not be generalizable to all children with CKD and in particular to children of African ancestry. However, to our knowledge this is the first study to explore the genetic contribution to red cell phenotype in children with CKD.

Our results confirming the association between a subset of previously identified SNPs and hemoglobin provide additional evidence supporting these associations in children. Among some of the variants associated with differences in red blood cell traits, it is possible that the effects on hemoglobin are more pronounced among children with the conditions of erythropoietin dysregulation, iron restriction, and inflammation that commonly emerge in advancing CKD. Thus it is a crucial step toward the application of precision medicine to test these results among children with CKD, in whom the biologic implications and potential treatment options are most relevant.

In conclusion, we demonstrate for the first time that SNPs previously reported to be associated with erythrocyte traits were nominally significant for an association with hemoglobin in children of European and Turkish ancestry with CKD. Differences in single SNPs are unlikely to account for most of the variability in hemoglobin levels observed in children with CKD. However, in the case of larger observed effect sizes, there is potential for these findings to yield prognostic information that may, in the future, guide individualized pharmacologic decisions in the treatment of the anemia of CKD. Moving forward, we will plan for further analyses/replication studies in larger adult CKD cohorts (such as the Chronic Renal Insufficiency Cohort Study [CRIC]) to investigate

SNPs of unknown significance that appear to be associated with hemoglobin. Functional studies will be required to confirm that the observed associations between SNPs of interest and clinical phenotype are causal.

ACKNOWLEDGEMENTS

The CKiD Study is supported by grants from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), with additional funding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, and the National Heart, Lung, and Blood Institute (U01-DK-66143, U01-DK-66174, U01DK-082194, and U01-DK-66116). The CKiD website is located at <http://www.statepijhsph.edu/ckid>. The ESCAPE trial and the 4C Study were conducted by the ESCAPE Clinical Research Network. The ESCAPE Trial was supported by grants from the European Commission Fifth Framework Programme (QLRT-2001-00908), Boehringer Ingelheim Foundation, Kuratorium für Dialyse und Nierentransplantation e.V., and the Baxter Extramural Grant Program. Support for the 4C Study was received from the ERA-EDTA Research Programme, the KfH Foundation for Preventive Medicine and the German Federal Ministry of Education and Research (reference number: 01EO0802). The genotyping data utilized for the PediGFR Consortium studies (CKiD, 4C, and ESCAPE) were supported by NIDDK Grant No. R01-DK-082394. Additional support was provided by the EU 7th Framework Programme (EURENomics, Grant 2012-305608). M.A.A. was supported by the National Institutes of Health (NIH)/NIDDK (K23-DK-084116). The work of M.W. and A.K. was supported by the CRC 1140 of the German Research Foundation. M.A.A. and S.L.F. were supported by an unrestricted grant from Amgen, Inc.

AUTHOR CONTRIBUTIONS

M.A.A. participated in the analysis and interpretation of data, drafted and revised the article providing intellectual content of critical importance to the work, and approved the final draft. R.X. participated in the analysis and interpretation of the data, revised the article providing intellectual content of critical importance to the work, and approved the final draft. A.K. participated in the conception and design of the analysis and the interpretation of data, revised the article and provided intellectual content of critical importance to the work, and approved the final draft. E.W. participated in the conception and design of the analysis and in the analysis and interpretation of data, revised the article and provided intellectual content of critical importance to the work, and approved the final draft. C.S.W. participated in the conception and design of the analysis and in the analysis and interpretation of data, revised the article and provided intellectual content of critical importance to the work, and approved the final draft. M.W. participated in the conception and design of the analysis and in the analysis and interpretation of data, revised the article and provided intellectual content of critical importance to the work, and approved the final draft. A.K.B. participated in the design of the analysis, revised the article for critical intellectual content, and approved the final draft. S.Ç. participated in the design of the analysis, revised the article for critical intellectual content, and approved the final draft. B.A.W. participated in the conception and design of the analysis and in the interpretation of data, revised the article and provided intellectual content of critical importance to the work, and approved the final draft. F.S. participated in the conception and design of the analysis and in the interpretation of data, revised the article and provided intellectual content of critical importance to the work, and approved the final draft. S.L.F. participated in the conception and design of the analysis and in the interpretation of data, revised the article and provided intellectual content of critical importance to the work, and approved the final draft.

ADDITIONAL INFORMATION

Competing interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Atkinson, M. A. & Furth, S. L. Anemia in children with chronic kidney disease. *Nat. Rev. Nephrol.* **7**, 635–641 (2011).
- Fadrowski, J. J. et al. Hemoglobin decline in children with chronic kidney disease: baseline results from the chronic kidney disease in children prospective cohort study. *Clin. J. Am. Soc. Nephrol.* **3**, 457–462 (2008).
- Borzycz-Duzalka, D. et al. Management of anemia in children receiving chronic peritoneal dialysis. *J. Am. Soc. Nephrol.* **24**, 665–676 (2013).

- Atkinson, M. A. et al. Hemoglobin differences by race in children with CKD. *Am. J. Kidney Dis.* **55**, 1009–1017 (2010).
- Robins, E. B. & Blum, S. Hematologic reference values for African American children and adolescents. *Am. J. Hematol.* **82**, 611–614 (2007).
- Ganesh, S. K. et al. Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium. *Nat. Genet.* **41**, 1191–1198 (2009).
- Li, J. et al. GWAS of blood cell traits identifies novel associated loci and epistatic interactions in Caucasian and African-American Children. *Hum. Mol. Genet.* **22**, 1457–1464 (2013).
- Ding, K. et al. Genetic variants that confer resistance to malaria are associated with red blood cell traits in African-Americans: an electronic medical record-based genome-wide association study. *G3 (Bethesda)* **3**, 1061–1068 (2013).
- Atkinson, M. A., Martz, K., Warady, B. A. & Neu, A. M. Risk for anemia in pediatric chronic kidney disease patients: a report of NAPRTCS. *Pediatr. Nephrol.* **25**, 1699–1706 (2010).
- Pfeffer, M. A. et al. A trial of darbepoetin alfa in type 2 diabetes and chronic kidney disease. *N. Engl. J. Med.* **361**, 2019–2032 (2009).
- Singh, A. K. et al. Correction of anemia with epoetin alfa in chronic kidney disease. *N. Engl. J. Med.* **355**, 2085–2098 (2006).
- Lestz, R. M., Fivush, B. A. & Atkinson, M. A. Association of higher erythropoiesis stimulating agent dose and mortality in children on dialysis. *Pediatr. Nephrol.* **29**, 2021–2028 (2014).
- Wuttke, M. et al. Genetic loci associated with renal function measures and chronic kidney disease in children: the Pediatric Investigation for Genetic Factors Linked with Renal Progression Consortium. *Nephrol. Dial. Transplant.* **31**, 262–269 (2016).
- Gupta, J. et al. Genome-wide association studies in pediatric chronic kidney disease. *Pediatr. Nephrol.* **31**, 1241–1252 (2016).
- Furth, S. L. et al. Design and methods of the Chronic Kidney Disease in Children (CKiD) prospective cohort study. *Clin. J. Am. Soc. Nephrol.* **1**, 1006–1015 (2006).
- ESCAPE Trial Group, Wühl, E. et al. Strict blood-pressure control and progression of renal failure in children. *N. Engl. J. Med.* **361**, 1639–1650 (2009).
- Querfeld, U. et al. The Cardiovascular Comorbidity in Children with Chronic Kidney Disease (4C) study: objectives, design, and methodology. *Clin. J. Am. Soc. Nephrol.* **5**, 1642–1648 (2010).
- Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. *Nat. Rev. Genet.* **11**, 499–511 (2010).
- Delaneau, O., Zagury, J. F. & Marchini, J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat. Methods* **10**, 5–6 (2013).
- Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
- 1000 Genomes Project Consortium, Abecasis et al. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
- Schwartz, G. J. et al. New equations to estimate GFR in children with CKD. *J. Am. Soc. Nephrol.* **20**, 629–637 (2009).
- Arnold, M., Raffler, J., Pfeufer, A., Suhre, K. & Kastenmüller, G. SNIpA: an interactive, genetic variant-centered annotation browser. *Bioinformatics* **31**, 1334–1336 (2015).
- Kent, W. J. et al. The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
- Ulirsch, J. C. et al. Systematic functional dissection of common genetic variation affecting red blood cell traits. *Cell* **165**, 1530–1545 (2016).
- Sarma, P. R. in *Clinical Methods: The History, Physical, and Laboratory Examinations* 3rd edn (eds Walker, H. K., Hall, W. D. & Hurst, J. W.) Ch. 152 (Butterworths, Boston, 1990).
- Fanis, P., Kousiappa, I., Phylactides, M. & Kleanthous, M. Genotyping of BCL11A and HB51L-MYB SNPs associated with fetal haemoglobin levels: a SNaPshot minisequencing approach. *BMC Genom.* **15**, 108 (2014).
- Basak, A. & Sankaran, V. G. Regulation of the fetal hemoglobin silencing factor BCL11A. *Ann. NY Acad. Sci.* **1368**, 25–30 (2016).
- Marian, A. J. The enigma of genetics etiology of atherosclerosis in the post-GWAS era. *Curr. Atheroscler. Rep.* **14**, 295–299 (2012).
- Singleton, A. B. et al. alpha-Synuclein locus triplication causes Parkinson's disease. *Science* **302**, 841 (2003).
- Tan, E. K. et al. Alpha synuclein promoter and risk of Parkinson's disease: microsatellite and allelic size variability. *Neurosci. Lett.* **36**, 70–72 (2003).
- Schneider, J. A. et al. DNA variability of human genes. *Mech. Ageing Dev.* **124**, 17–25 (2003).
- Nielsen, R., Hubisz, M. J. & Clark, A. G. Reconstituting the frequency spectrum of ascertained single-nucleotide polymorphism data. *Genetics* **168**, 2373–2382 (2004).
- Lin, P.-I., Vance, J. M., Pericak-Vance, M. A. & Martin, E. R. No gene is an island: the flip-flop phenomenon. *Am. J. Hum. Genet.* **80**, 531–538 (2007).
- Amos, W., Driscoll, E. & Hoffman, J. I. Candidate genes versus genome-wide associations: which are better for detecting genetic susceptibility to infectious disease? *Proc. Biol. Sci.* **278**, 1183–1188 (2011).