



The Challenges of Chromosome Y Analysis and the Implications for Chronic Kidney Disease

Kerry Anderson^{1*}, Marisa Cañadas-Garre¹, Robyn Chambers¹,
Alexander Peter Maxwell^{1,2} and Amy Jayne McKnight¹

¹ Epidemiology and Public Health Research Group, Centre for Public Health, Queen's University of Belfast, c/o Regional Genetics Centre, Belfast City Hospital, Belfast, United Kingdom, ² Regional Nephrology Unit, Belfast City Hospital, Belfast, United Kingdom

OPEN ACCESS

Edited by:

Veronique Vitart,
University of Edinburgh,
United Kingdom

Reviewed by:

Samuel Lukowski,
University of Queensland,
Australia

Audrey Qiuyan Fu,
University of Idaho,
United States

*Correspondence:

Kerry Anderson
kanderson26@qub.ac.uk

Specialty section:

This article was submitted to
Human Genomics,
a section of the journal
Frontiers in Genetics

Received: 12 April 2019

Accepted: 24 July 2019

Published: 04 September 2019

Citation:

Anderson K, Cañadas-Garre M,
Chambers R, Maxwell AP and
McKnight AJ (2019) The Challenges
of Chromosome Y Analysis and
the Implications for Chronic Kidney
Disease.
Front. Genet. 10:781.
doi: 10.3389/fgene.2019.00781

The role of chromosome Y in chronic kidney disease (CKD) remains unknown, as chromosome Y is typically excluded from genetic analysis in CKD. The complex, sex-specific presentation of CKD could be influenced by chromosome Y genetic variation, but there is limited published research available to confirm or reject this hypothesis. Although traditionally thought to be associated with male-specific disease, evidence linking chromosome Y genetic variation to common complex disorders highlights a potential gap in CKD research. Chromosome Y variation has been associated with cardiovascular disease, a condition closely linked to CKD and one with a very similar sexual dimorphism. Relatively few sources of genetic variation in chromosome Y have been examined in CKD. The association between chromosome Y aneuploidy and CKD has never been explored comprehensively, while analyses of microdeletions, copy number variation, and single-nucleotide polymorphisms in CKD have been largely limited to the autosomes or chromosome X. In many studies, it is unclear whether the analyses excluded chromosome Y or simply did not report negative results. Lack of imputation, poor cross-study comparability, and requirement for separate or additional analyses in comparison with autosomal chromosomes means that chromosome Y is under-investigated in the context of CKD. Limitations in genotyping arrays could be overcome through use of whole-chromosome sequencing of chromosome Y that may allow analysis of many different types of genetic variation across the chromosome to determine if chromosome Y genetic variation is associated with CKD.

Keywords: chromosome Y, chronic kidney disease, genome-wide association, genotyping arrays, haplogroup, imputation, LOY, microdeletion

INTRODUCTION

To date, the contribution of chromosome Y to the development and progression of chronic kidney disease (CKD) has remained largely unexplored. Over 50 genome-wide association studies (GWASs) have been conducted in renal diseases during the last 10 years (MacArthur et al., 2017), yet only one has reported details of chromosome Y analysis (Nanayakkara et al., 2014). For example, one of the most comprehensive meta-analysis GWASs conducted in renal disease included over 2.5 million single-nucleotide polymorphisms (SNPs), genotyped in 110,517 individuals; however, no chromosome

Y SNPs were included (Gorski et al., 2017). The exclusion of chromosome Y from genomic analyses may previously have been justifiable, based on the assumption that chromosome Y was a “genetic wasteland” (Maan et al., 2017); but as more research is published, it is becoming clear that chromosome Y variation may be useful for identifying individuals with increased susceptibility to the disease. It was traditionally thought that chromosome Y only carried genes important for male-specific traits. However, pseudoautosomal regions (PARs) of sequence homology with chromosome X are found on the tips of chromosome Y. Gene expression levels between chromosome X and chromosome Y PAR homologs can be subject to male expression bias, whereby chromosome Y PAR genes are more highly expressed than their chromosome X counterparts, which could account for sex differences in disease (Snell and Turner, 2018). Additionally, upon the complete sequencing and characterization of the chromosome by Skaletsky et al. (2003), it was revealed that approximately 50% of protein-coding genes present on the male-specific region (MSY) expressed in non-gonadal tissues (Skaletsky et al., 2003), and it is, therefore, likely that they could play a role in common complex disease. For example, Charchar et al. (2012) performed such analysis on chromosome Y and identified an increased risk of coronary artery disease (CAD) in men from haplogroup I, but it is unclear whether this finding is relevant to CKD. The major rationale for including chromosome Y in studies of disease risk stems from the goal of identifying genetic features that may contribute to sex-specific presentations of disease. For example, cardiovascular disease (CVD) incidence is similar between men and women, but progression of the disease differs, with age of onset approximately 10 years later for women (De Smedt et al., 2016). Chromosome Y analysis to detect variation contributing to disease risk is a logical step. A very similar sexual dimorphism exists in CKD; prevalence is greater in women, but kidney disease in men progresses more rapidly to end-stage renal disease (ESRD), the most severe form of CKD (Hill et al., 2016). In this case, a male-specific marker of accelerated CKD progression could prove useful in identifying which patients are at greater risk of rapid loss of renal function. The issue of having better markers for CKD progression is relevant when examining current biomarkers for the diagnosis of CKD. Renal function is assessed by measuring either serum creatinine, or less commonly cystatin C, and an equation is used to determine an estimated glomerular filtration rate (eGFR). However, these eGFR equations are less accurate for certain individuals, such as those with low muscle mass, extreme body mass indexes, and early-stage CKD, a group whose identification is key to allow implementation of preventative

Abbreviations: AZF, Azoospermia factor; CAD, Coronary artery disease; CAKUT, Congenital anomalies of the kidney and urinary tract; CKD, Chronic kidney disease; CNV, Copy number variation; CVD, Cardiovascular disease; DAZ, Deleted in azoospermia; eGFR, Estimated glomerular filtration rate; ESRD, End-stage renal disease; GWAS, Genome-wide association study; HDL, High-density lipoprotein; HIV, Human immunodeficiency virus; IgAN, Immunoglobulin A nephropathy; LDL, Low-density lipoprotein; LOY, Loss of chromosome Y; MI, Myocardial infarction; MSY, Male-specific region of chromosome Y; PAR, Pseudoautosomal region; PCa, Prostate cancer; SNP, Single-nucleotide polymorphism; WAGR, Wilms tumor, aniridia, genitourinary anomalies, and mental retardation; WES, Whole-exome sequencing; WGS, Whole-genome sequencing; YAP, Chromosome Y Alu insertion polymorphism.

measures (Gentile and Remuzzi, 2016). Extensive reviews of the literature have highlighted that there are relatively few alternative kidney function biomarkers, and none have improved upon the limitations of serum creatinine or cystatin C (Cañadas-Garre et al., 2018; Cañadas-Garre et al., 2019b). Therefore, while exclusion of chromosome Y in genomic analysis of renal disease may previously have been justifiable, it does highlight a distinct gap in our knowledge of how chromosome Y genetic variation may play a role in renal disease pathogenesis.

CHROMOSOME Y IN DISEASE

While chromosome Y variation has been linked to a number of male-specific conditions such as prostate cancer (PCa), it has also been shown to influence the risk profile of men for common complex disease such as CAD and influence the progression of HIV (Sezgin et al., 2009). The clearest link thus far between chromosome Y and disease is related to infertility.

Infertility

While up to 7% of men are infertile, only 15–30% of these cases have a known genetic cause (Neto et al., 2016). Numerical and structural defects in chromosome Y have been linked to male infertility. An extra copy of chromosome Y (47, XYY) is the second most-frequent aneuploidy of the sex chromosomes, present in 1/1,000 men (Bardsley et al., 2013) and can result in a complete lack of spermatozoa production (azoospermia) or a severely low sperm count (oligospermia) (McLachlan and O’Bryan, 2010). Chromosome Y microdeletions (deletions less than 5 megabases in size) (Halder et al., 2013) have long been associated with infertility (Stuppia et al., 1997), and three azoospermia factor regions, AZFa, AZFb, and AZFc, have been identified on the long arm of chromosome Y (Vogt et al., 1996). The most clinically significant recurrent microdeletions see the complete loss of each AZF region, or the combined loss of AZFb and AZFc, with approximately 80% of all microdeletions being a complete loss of AZFc (Krausz et al., 2014). The AZFc region, which contains the deleted in azoospermia (DAZ) gene family, is completely deleted in 5–10% of azoospermia/severe oligospermia cases, making it the most frequent genetic cause of infertility in men (Ferlin et al., 2005). Partial AZF deletions (Lu et al., 2009) and *gr/gr* deletions (Bansal et al., 2016) are associated with spermatogenic failure and have also been associated with different chromosome Y haplogroups (Lu et al., 2009; Ran et al., 2013; Xue et al., 2013). Copy number variation (CNV) in certain chromosome Y genes, such as *GOLGA2P3Y* and *RBMY1*, has been associated with reduced sperm count (Sen et al., 2016) and motility (Yan et al., 2017). However, while certain genes and deletions have been associated with male infertility, further research is required to establish a complete pathogenic mechanism.

Prostate Cancer

Another male-specific condition with which chromosome Y has been linked is PCa. Loss of chromosome Y (LOY) has

been observed in PCa (Noveski et al., 2016; Zhou et al., 2016). Specific deletions in chromosome Y genes have been associated with PCa, several of which, including the sex-determining factor *SRY*, were found to increase in frequency with increasing PCa stage (Perinchery et al., 2000). A similar study detected loss of the region containing *SRY* at a similar rate, and also observed this loss in surrounding benign prostate hyperplasia tissue, perhaps indicating that loss of *SRY* is a precursor for PCa (Jordan et al., 2001). Loss of *SRY* may prevent the negative regulation of the androgen receptor AR, leading to increased androgen receptor activity and thus PCa growth. An additional chromosome Y gene, *KDM5D*, is also thought to interact with the androgen receptor, altering the sensitivity of docetaxel, a drug commonly used in androgen deprivation therapy (Komura et al., 2016). Chromosome Y haplogroups (Paracchini et al., 2003), short tandem repeats (Carvalho et al., 2010; Nargesi et al., 2011), and a number of different genes, many identified through co-expression networks, have been associated with PCa (Khosravi et al., 2014). As the exact mechanism of PCa pathogenesis has yet to be elucidated, a further role of chromosome Y in PCa may yet emerge.

Cardiovascular Disease

CVD and its associated conditions are an example of some of the strongest evidence linking chromosome Y to common complex disease. CVD is a prime example of a condition that exhibits a complex sexual dimorphism; incidence of CVD is higher in men than in age-matched women, but the relative risk of mortality is higher in women with CVD (Möller-Leimkühler, 2007). Chromosome Y genetic variation has previously been linked to CVD; carriers of haplogroup I-defining SNP rs2032597 (also known as M170) had a ~50% higher age-adjusted risk of CAD than men with other chromosome Y lineages in two independent cohorts, and the joint analysis of both cohorts (Charchar et al., 2012). The presence of the A form of two SNPs, rs768983 (A/G) in *TBL1Y* and rs3212292 (A/T) in *USP9Y*, was associated with lower levels of triglycerides and higher levels of high-density lipoprotein (HDL)-cholesterol compared with the other haplotypes in Black individuals of African origin (Russo et al., 2008). These SNPs were almost entirely monomorphic in the other ethnic groups included in the analysis, highlighting the differences in risk profiles between different ethnic groups and may explain the lower risk of CVD in Black individuals. However, conflicting studies in this area make the connection between chromosome Y and CVD less clear. While the YAP polymorphism, caused by an Alu insertion, was associated with an increased risk of atherosclerotic plaque formation at a particular bifurcations (Voskarides et al., 2014), other studies failed to find any association between this polymorphism and low-density lipoprotein (LDL)-cholesterol (Shoji et al., 2002; Hiura et al., 2008), hypertension (Hiura et al., 2008; Kostrzewa et al., 2013), or myocardial infarction (MI) (Hiura et al., 2008). Other chromosome Y haplogroups have been investigated in CVD, and haplogroup K was found to be associated with a 2.5× increased risk of atherosclerotic plaque development (Hiura et al., 2008), but not with hypertension (Kostrzewa et al., 2013).

Additional haplogroup analyses failed to identify any association between haplogroups and either hypertension (Kostrzewa et al., 2013) or venous thrombosis (de Haan et al., 2016). Conflicting evidence of association between the *HindIII*(±) polymorphism has also been presented, where *HindIII*(+) has been associated with increased systolic and diastolic blood pressure (Charchar et al., 2002) and MI in hypertensive patients (García et al., 2003). However, *HindIII*(-) has also been reported as associated with increased blood pressure, although this study was conducted in pre-pubescent boys, which may account for the conflicting results (Shankar et al., 2007). In other studies, no significant association at all was found between the *HindIII*(±) polymorphism and blood pressure (Rodriguez et al., 2005; Russo et al., 2006; Kostrzewa et al., 2013).

CHROMOSOME Y IN CHRONIC KIDNEY DISEASE

The potential links between chromosome Y genetic variation and CKD have not been systematically explored. Several types of genetic variation in chromosome Y have been discussed above in relation to disease: loss of chromosome Y (LOY), an extra copy of chromosome Y (47, XYY), chromosome Y microdeletions, CNVs, SNPs, and haplogroups. While many of these genetic variations, particularly microdeletions, CNVs, and SNPs, have been explored in autosomes in individuals with CKD, studies of these variants in chromosome Y in CKD are limited. LOY has been detected in other renal conditions, such as in renal cell carcinoma tumors (Dagher et al., 2013) but not in any condition falling under the umbrella of CKD. As a minimally invasive biomarker, LOY in CKD could be detected through traditional karyotyping or by using SNP arrays (Forsberg, 2017). No study to date has tested for association between mosaic LOY in peripheral blood and CKD; and as mosaic LOY in blood is also associated with smoking (Dumanski et al., 2015), higher cancer risk (Forsberg et al., 2014), aging (Forsberg et al., 2017), and age-related macular degeneration (Grassmann et al., 2019), any association study would require adjustment for these factors. An additional copy of chromosome Y results in a mild syndrome known as 47, XYY. CKD is not typically associated with 47, XYY, although a single participant with posterior urethral valves carrying an additional copy of chromosome Y was identified in a study of congenital anomalies of the kidney and urinary tract (CAKUT) (Caruana et al., 2014). Although 47, XYY does not appear to be associated with adult-onset CKD, its relevance in CKD has not really been explored, so the evidence is limited. As with LOY, XYY can be detected using karyotyping or SNP arrays. There are currently no known associations between chromosome Y microdeletions and CKD, as studies in CKD to date have generally focused only on the specific microdeletions described in autosomes. Microdeletions in the *HNF1B* gene on chromosome 17q12 have been identified in both children and adults with CKD (Musetti et al., 2014; Verbitsky et al., 2015). A study of idiopathic CKD found a 1.3Mb deletion in *HNF1B* in 9% of participants tested (Clissold et al., 2018).

A microdeletion at 11p13, the region containing the *PAX6* and *WT1* genes, results in Wilms tumor, aniridia, genitourinary anomalies, and mental retardation (WAGR) syndrome, patients with which have, among other symptoms, significant ESRD (van Heyningen et al., 2007). Larger deletions and duplications, known as CNVs, have been studied in CKD but largely in pediatric cohorts of CAKUT. A study of two distinct cohorts of adult Han Chinese individuals found several CNVs in the *DEFA1A3* locus associated with renal dysfunction in immunoglobulin A nephropathy (IgAN) patients (Ai et al., 2016). *NPHP1* gene deletions resulting in nephronophthisis, one of the most prevalent causes of ESRD in children, have also been detected in patients with adult-onset ESRD (Snoek et al., 2018). Among these studies, different CNV detection methods were used, including SNP arrays and whole-exome sequencing (Caruana et al., 2014; Bekheirnia et al., 2017). While some studies actively did not analyze any of the known chromosome Y CNVs (Verbitsky et al., 2015; Li et al., 2017), whether chromosome Y was included in the analysis in other studies was unclear. Therefore, due to either lack of reported analysis or indeed lack of association, there are currently no known associations between chromosome Y CNVs and CKD.

As outlined above, more than 50 studies have tried to unravel the genetic variation of CKD explained by SNPs (MacArthur et al., 2017). To date, 140 autosomal and X chromosome SNPs have been associated with CKD (Cañadas-Garre et al., 2019a). However, no studies have reported significant associations with chromosome Y SNPs, probably because they were methodologically excluded. Indeed, only one GWAS of renal dysfunction appears to have included chromosome Y SNPs in their analysis (Nanayakkara et al., 2014). No significant associations were detected between CKD and chromosome Y SNPs in this study, but due to differences in prevalence of chromosome Y haplogroups between different populations, this does not necessarily mean chromosome Y variation is not linked to CKD. For example, the most common haplogroup in European populations, R1b, is not present in the Sinhalese participants of the study by Nanayakkara and colleagues, whose dominant haplogroup is the R2 haplogroup (Kivisild et al., 2003), therefore demonstrating the need for analysis in a range of populations. Given that chromosome Y SNPs have demonstrated to play a role in other diseases, particularly CVD, a condition with strong links to CKD, are we missing associations between chromosome Y SNPs and CKD?

GWAS EXCLUSION

Chromosome Y SNPs make up approximately 0.07% (60,505/84,387,209) of all recorded biallelic SNPs within the genome (Gibbs et al., 2015). Therefore, a possible explanation for the lack of significant findings on chromosome Y in relation to CKD may be the underrepresentation of chromosome Y on commonly used genotyping arrays. **Figure 1** shows that, although chromosome Y is completely excluded from some arrays, its representation on other platforms is actually greater than the percentage of chromosome Y SNPs in the genome.

However, although chromosome Y may be represented almost proportionally on genotyping arrays, only 4% of chromosome Y SNPs in the genome (60,555) are analyzed on the largest array (2,445 on the Illumina Omni-5.4 Array). Therefore, while representation of chromosome Y on traditional genotyping arrays may be *proportional*, it is far from comprehensive. Insufficient gene coverage may also explain the lack of significant findings in chromosome Y. However, as shown in **Figure 2**, the SNPs offered on commonly used genotyping arrays provide sufficient coverage of chromosome Y genes (Zerbino et al., 2018).

Although only a fraction of chromosome Y SNPs are present on available genotyping platforms, an often greater-than-proportional number of SNPs is dedicated to chromosome Y on common genotyping platforms. So why is chromosome Y excluded from analyses? Chromosome Y content between genotyping arrays is variable, which in turn greatly limits the number of SNPs available for cross-study meta-analysis. For example, for a meta-GWAS of cohorts genotyped using the arrays outlined in **Figure 1**, no chromosome Y SNPs would be eligible for inclusion, as zero chromosome Y SNPs are found on the Affymetrix 500K arrays. This may explain why chromosome Y is excluded from some larger meta-GWAS. Even if arrays do contain chromosome Y SNPs, there is a distinct lack of overlap between arrays, and as more arrays are considered, the number of SNPs that are common to all platforms is reduced. For example, **Figure 1** shows that only 19 SNPs are in common between the other seven included arrays.

However, the same can be said for the rest of the SNPs on commonly used genotyping platforms. Whole-genome SNPs (including chromosome Y) that are common between the same eight genotyping arrays are outlined in **Figure 1**. Only 7,851 SNPs feature on all eight arrays. Yet in larger renal meta-GWAS, by using imputation, as many as 2.5 million autosomal and chromosome X SNPs can be included (Pattaro et al., 2016; Gorski et al., 2017). Imputation is a process that allows inference of ungenotyped SNPs in a sample, based on panels of haploid reference sequences (Marchini et al., 2007), meaning that the number of loci for which information is available can be dramatically increased from the number obtained from direct genotyping alone. For example, less than a million directly typed markers were imputed to approximately 96 million variants using the Human Reference Consortium and UK10K haplotype resources in UK Biobank samples genotyped using either the UK BiLEVE or UK Biobank Axiom array (Bycroft et al., 2018). In this case, imputation allowed more than a hundred times the number of directly genotyped SNPs to be available for analysis. Even after applying quality control thresholds, such as for minor allele frequency and imputation quality, as many as 12 million SNPs could be available for association analysis (Haas et al., 2018). The same imputed genotypes were used in a recent meta-GWAS, in which five million SNPs were common between the three studies and available for the meta-analysis (Xue et al., 2018).

However, as chromosome Y is haploid and a large portion of the chromosome does not undergo recombination, accurate and reliable chromosome Y imputation is, despite recent efforts (Zhang et al., 2013), not widely implemented.

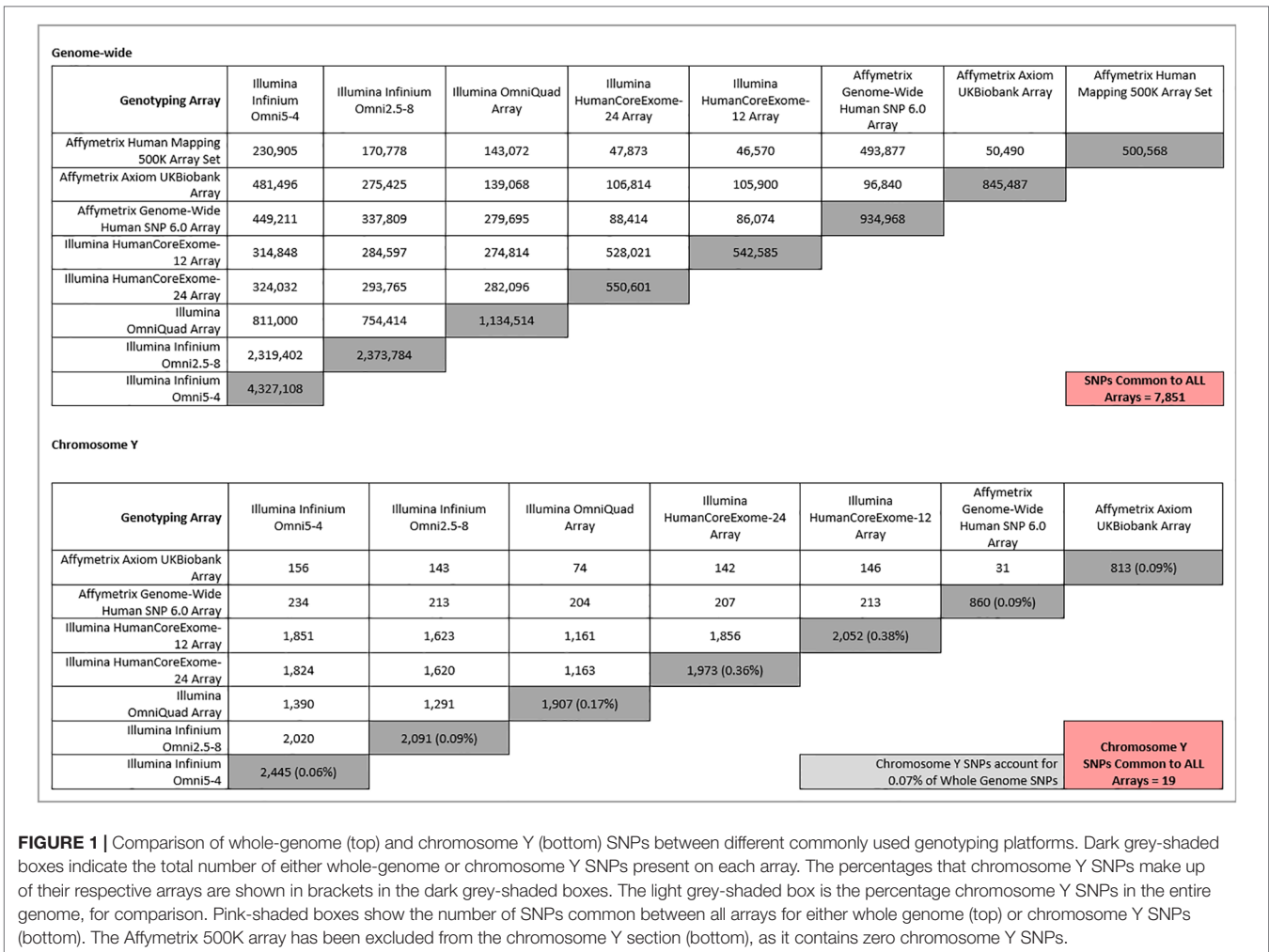


FIGURE 1 | Comparison of whole-genome (top) and chromosome Y (bottom) SNPs between different commonly used genotyping platforms. Dark grey-shaded boxes indicate the total number of either whole-genome or chromosome Y SNPs present on each array. The percentages that chromosome Y SNPs make up of their respective arrays are shown in brackets in the dark grey-shaded boxes. The light grey-shaded box is the percentage chromosome Y SNPs in the entire genome, for comparison. Pink-shaded boxes show the number of SNPs common between all arrays for either whole genome (top) or chromosome Y SNPs (bottom). The Affymetrix 500K array has been excluded from the chromosome Y section (bottom), as it contains zero chromosome Y SNPs.

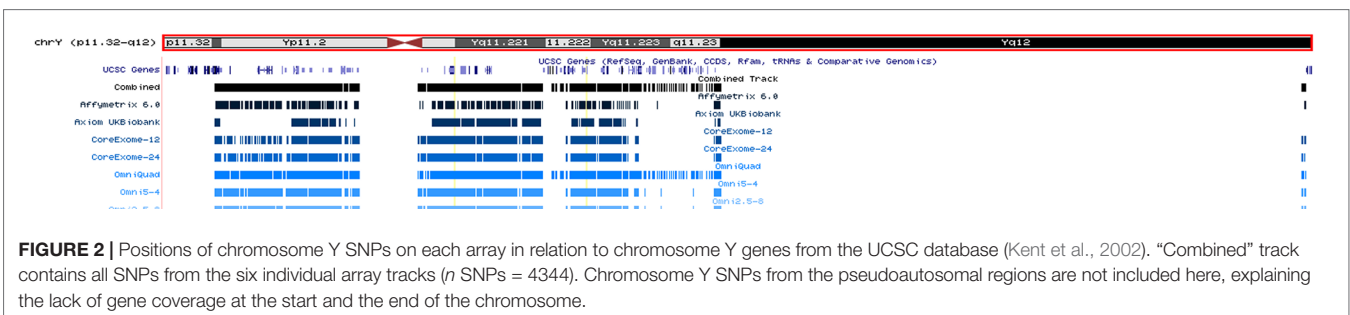


FIGURE 2 | Positions of chromosome Y SNPs on each array in relation to chromosome Y genes from the UCSC database (Kent et al., 2002). “Combined” track contains all SNPs from the six individual array tracks (n SNPs = 4344). Chromosome Y SNPs from the pseudoautosomal regions are not included here, explaining the lack of gene coverage at the start and the end of the chromosome.

Even chromosome X, whose imputation has been achieved (Marchini and Howie, 2010), may be excluded due to the need to impute it separately, so exclusion of both sex chromosomes in GWAS is common. While this lack of recombination in the majority of chromosome Y should actually aid imputation, chromosome Y reference panels for imputation are not widely available, and therefore, chromosome Y is often excluded from the analysis at this stage. For example, the Sanger Imputation service offers five different reference panels to impute data to,

none of which include chromosome Y (McCarthy et al., 2016). However, the lack of recombination across this section of chromosome Y means that any genetic variations in the MSY pass directly from father to son and means that certain genetic variations are often inherited together. Patterns in these genetic variations are known as haplotypes and can be used to group individuals into haplogroups. These groups or “clades” are defined by single markers that differentiate one clade from another, and genotyping of these markers can be used to sort individuals into

different haplogroups. It has also been shown that some of these haplogroups are associated with certain phenotypes. For example, haplogroup I, one of the most frequently occurring haplogroup in the UK (Winney et al., 2012), is associated with an increased risk of CAD (Charchar et al., 2012). In many cases, haplotyping negates the need for large numbers of SNPs to be genotyped in genetic association analyses; for example, only 11 SNPs need to be genotyped to cover 95% of the haplogroups present within the UK (Charchar et al., 2012). Haplotyping can be performed using any SNPs genotyping method, including SNP arrays (Kim and Misra, 2007). All arrays included in **Figure 1** provide coverage of major haplogroup-defining SNPs and would therefore be suitable for use in haplogroup association studies. However, to date, no such analysis has been carried out in CKD. This glaring lack of investigation offers the opportunity to perform a complete analysis of chromosome Y genetic variation in CKD. A full analysis of SNPs/haplogroups, CNVs, microdeletions, LOY, and XYY could be performed using SNP arrays, but the limitations of these arrays in chromosome Y have been outlined above. The decreasing cost of and increased coverage offered by whole-genome (WGS) and whole-exome sequencing (WES) may offer a solution for improving investigations in chromosome Y (Levy and Myers, 2016).

WHOLE GENOME/EXOME SEQUENCING AND CHROMOSOME Y

WGS has already been utilized in CKD. WES detected diagnostic variants (Lata et al., 2018; Groopman et al., 2019) and CNVs in individuals with CKD (Bekheirnia et al., 2017); chromosome Y was not included in either of these studies, as they only targeted variants with known links to CKD. WGS can also detect LOY/XYY, microdeletions, SNPs and therefore, haplogroups (Muzzey et al., 2015). It offers the added benefit of multiple long reads, meaning there is a reduced risk of genotypes being lost to poor genotyping, as there is with arrays. WGS would generate a complete profile of chromosome Y to be analyzed, rather than needing multiple alternative methods (karyotyping, chromosomal arrays, and SNP arrays) to provide the same data. As knowledge of chromosome Y haplogroups grows, it is expected that more haplogroup markers will be added to the phylogeny (van Oven et al., 2014) and that WGS prevents re-genotyping of samples for certain markers as the whole sequence will be available. WGS also overcomes the issue of imputation; genotypes do not need to be inferred if the whole sequence is available. Lack of common SNPs between different studies is also improved by access to the entire sequence. However, due to sequence homology between the sex chromosomes and the large regions of repetitive sequences within chromosome Y, sequencing presents some challenges in itself. Repetitive sequences make sequencing more difficult. However, tools are being developed to try and combat some of these issues (Webster et al., 2019), and as sequencing technologies develop and costs fall, longer reads of greater read depth may aid in mapping complex repeated sequences. For example, Oxford Nanopore technology has recently achieved read lengths of up to 2.2Mb (Payne et al., 2018), and this has since been used

to sequence and assemble the first chromosome Y of African Origin, where sequence continuity increased by almost 800% than did previous methods and amounted to 21.5Mb of total sequence (Kuderna et al., 2019). In short, although presenting some challenges of its own, WGS/WES resolves many of the major issues of analyzing chromosome Y with SNP arrays and allows multiple types of variation to be considered using a single test, offering the most comprehensive analysis of chromosome Y possible.

CONCLUSIONS

Chromosome Y analysis remains challenging due to lack of common coverage by genotyping arrays, the need to process chromosome Y data separately from autosomal data, and the current inability to accurately impute chromosome Y to the same standards that have been achieved in autosomal imputation. The inclusion of chromosome Y in GWAS and other genetic analyses is inconsistent, and in many cases, it is not clear if the analysis was not performed, or if negative results have simply not been reported. For this reason, the known contribution of chromosome Y genetic variation to disease remains limited, particularly in renal disease. The sexual dimorphism in CKD provides a rationale for further investigations of factors influencing sex-related progression to ESRD, perhaps using methods such as targeted next-generation sequencing to analyze chromosome Y specifically. While other factors, such as hormone profiles, may influence disease progression, the current lack of chromosome Y analysis in renal disease means that the contribution of genetic variation in chromosome Y to renal disease progression remains unknown.

AUTHOR CONTRIBUTIONS

KA: Conception and design of the work; acquisition, analysis, and interpretation of the data for the article; and drafting, critical revision, and final approval of the manuscript. MCG: Interpretation of the data for the article and drafting, critical revision, and final approval of the manuscript. RC: Acquisition, analysis, and interpretation of the data for the article and drafting and final approval of the manuscript. APM: Conception and design of the work; interpretation of the data for the article; and drafting, critical revision, and final approval of the manuscript. AJM: Conception and design of the work; interpretation of the data for the article; and drafting, critical revision, and final approval of the manuscript.

FUNDING

This work has been partly funded by the Medical Research Council (Award reference MC_PC_15025) and the Public Health Agency R&D Division (Award Reference STL/4760/13). KA and MG are funded by a Science Foundation Ireland-Department for the Economy (SFI-DfE) Investigator Program Partnership Award (15/IA/3152).

REFERENCES

- Ai, Z., Li, M., Liu, W., Foo, J. N., Mansouri, O., Yin, P., et al. (2016). Low α -defensin gene copy number increases the risk for IgA nephropathy and renal dysfunction. *Sci. Transl. Med.* 8, 345–388. doi: 10.1126/scitranslmed.aaf2106
- Bansal, S. K., Jaiswal, D., Gupta, N., Singh, K., Dada, R., Sankhwar, S. N., et al. (2016). Gr/gr deletions on Y-chromosome correlate with male infertility: an original study, meta-analyses, and trial sequential analyses. *Sci. Rep.* 6, 19798. doi: 10.1038/srep19798
- Bardsley, M. Z., Kowal, K., Levy, C., Gosek, A., Ayari, N., Tartaglia, N., et al. (2013). 47, XYY syndrome: clinical phenotype and timing of ascertainment. *J. Pediatr.* 163, 1085–1094. doi: 10.1016/j.jpeds.2013.05.037
- Bekheirnia, M. R., Bekheirnia, N., Bainbridge, M. N., Gu, S., Akdemir, Z. H. C., Gambin, T., et al. (2017). Whole-exome sequencing in the molecular diagnosis of individuals with congenital anomalies of the kidney and urinary tract and identification of a new causative gene. *Genet. Med.* 19, 412–420. doi: 10.1038/gim.2016.131
- Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature* 562, 203–209. doi: 10.1038/s41586-018-0579-z
- Cañadas-Garre, M., Anderson, K., Cappa, R., Skelly, R., Smyth, L. J., McKnight, A. J., et al. (2019a). Genetic susceptibility to chronic kidney disease—some more pieces for the heritability puzzle. *Front. Genet.* 10, 453. doi: 10.3389/fgene.2019.00453
- Cañadas-Garre, M., Anderson, K., McGoldrick, J., Maxwell, A. P., and McKnight, A. J. (2018). Genomic approaches in the search for molecular biomarkers in chronic kidney disease. *J. Transl. Med.* 16, 292. doi: 10.1186/s12967-018-1664-7
- Cañadas-Garre, M., Anderson, K., McGoldrick, J., Maxwell, A. P., and McKnight, A. J. (2019b). Proteomic and metabolomic approaches in the search for biomarkers in chronic kidney disease. *J. Proteomics* 193, 93–122. doi: 10.1016/j.jprot.2018.09.020
- Caruana, G., Wong, M. N., Walker, A., Heloury, Y., Webb, N., Johnstone, L., et al. (2014). Copy-number variation associated with congenital anomalies of the kidney and urinary tract. *Pediatr. Nephrol.* 30, 487–495. doi: 10.1007/s00467-014-2962-9
- Carvalho, R., Pinheiro, M. F., and Medeiros, R. (2010). Localization of candidate genes in a region of high frequency of microvariant alleles for prostate cancer susceptibility: the chromosome region Yp11.2 genetic variation. *DNA Cell Biol.* 29, 3–7. doi: 10.1089/dna.2009.0905
- Charchar, F. J., Bloomer, L. D. S., Barnes, T. A., Cowley, M. J., Nelson, C. P., Wang, Y., et al. (2012). Inheritance of coronary artery disease in men: an analysis of the role of the Y chromosome. *Lancet* 379, 915–922. doi: 10.1016/S0140-6736(11)61453-0
- Charchar, F. J., Tomaszewski, M., Padmanabhan, S., Lacka, B., Upton, M. N., Inglis, G. C., et al. (2002). The Y chromosome effect on blood pressure in two European populations. *Hypertension* 39, 353–356. doi: 10.1161/hy0202.103413
- Clissold, R. L., Ashfield, B., Burrage, J., Hannon, E., Bingham, C., Mill, J., et al. (2018). Genome-wide methylomic analysis in individuals with HNF1B intragenic mutation and 17q12 microdeletion. *Clin. Epigenetics* 10, 97. doi: 10.1186/s13148-018-0530-z
- Dagher, J., Dugay, F., Verhoest, G., Cabillic, F., Jaillard, S., Henry, C., et al. (2013). Histologic prognostic factors associated with chromosomal imbalances in a contemporary series of 89 clear cell renal cell carcinomas. *Hum. Pathol.* 44, 2106–2115. doi: 10.1016/j.humpath.2013.03.018
- de Haan, H. G., van Hylckama Vlieg, A., van der Gaag, K. J., de Knijff, P., and Rosendaal, F. R. (2016). Male-specific risk of first and recurrent venous thrombosis: a phylogenetic analysis of the Y chromosome. *J. Thromb. Haemost.* 14, 1971–1977. doi: 10.1111/jth.13437
- De Smedt, D., De Bacquer, D., De Sutter, J., Dallongeville, J., Gevaert, S., De Backer, G., et al. (2016). The gender gap in risk factor control: effects of age and education on the control of cardiovascular risk factors in male and female coronary patients. The EUROASPIRE IV study by the European Society of Cardiology. *Int. J. Cardiol.* 209, 284–290. doi: 10.1016/j.ijcard.2016.02.015
- Dumanski, J. P., Rasi, C., Lonn, M., Davies, H., Ingelsson, M., Giedraitis, V., et al. (2015). Smoking is associated with mosaic loss of chromosome Y. *Science (80-)* 347, 81–83. doi: 10.1126/science.1262092
- Ferlin, A., Tessari, A., Ganz, F., Marchina, E., Barlati, S., Garolla, A., et al. (2005). Association of partial AZFc region deletions with spermatogenic impairment and male infertility. *J. Med. Genet.* 42, 209–213. doi: 10.1136/jmg.2004.025833
- Forsberg, L. A. (2017). Loss of chromosome Y (LOY) in blood cells is associated with increased risk for disease and mortality in aging men. *Hum. Genet.* 136, 657–663. doi: 10.1007/s00439-017-1799-2
- Forsberg, L. A., Gisselsson, D., and Dumanski, J. P. (2017). Mosaicism in health and disease—clones picking up speed. *Nat. Rev. Genet.* 18, 128–142. doi: 10.1038/nrg.2016.145
- Forsberg, L. A., Rasi, C., Malmqvist, N., Davies, H., Pasupulati, S., Pakalapati, G., et al. (2014). Mosaic loss of chromosome Y in peripheral blood is associated with shorter survival and higher risk of cancer. *Nat. Genet.* 46, 624–628. doi: 10.1038/ng.2966
- García, E. C., González, P., Castro, M. G., Alvarez, R., Reguero, J. R., Batalla, A., et al. (2003). Association between genetic variation in the Y chromosome and hypertension in myocardial infarction patients. *Am. J. Med. Genet. A* 122A, 234–237. doi: 10.1002/ajmg.a.20376
- Gentile, G., and Remuzzi, G. (2016). Novel biomarkers for renal diseases? None for the moment (but one). *J. Biomol. Screen.* 21, 655–670. doi: 10.1177/1087057116629916
- Gibbs, R. A., Boerwinkle, E., Doddapaneni, H., Han, Y., Korchina, V., Kovar, C., et al. (2015). A global reference for human genetic variation. *Nature* 526, 68–74. doi: 10.1038/nature15393
- Gorski, M., van der Most, P. J., Teumer, A., Chu, A. Y., Li, M., Mijatovic, V., et al. (2017). 1000 Genomes-based meta-analysis identifies 10 novel loci for kidney function. *Sci. Rep.* 7, 45040. doi: 10.1038/srep45040
- Grassmann, F., Kiel, C., den Hollander, A. I., Weeks, D. E., Lotery, A., Cipriani, V., et al. (2019). Y chromosome mosaicism is associated with age-related macular degeneration. *Eur. J. Hum. Genet.* 27, 36–41. doi: 10.1038/s41431-018-0238-8
- Groopman, E. E., Marasa, M., Cameron-Christie, S., Petrovski, S., Aggarwal, V. S., Milo-Rasouly, H., et al. (2019). Diagnostic utility of exome sequencing for kidney disease. *N. Engl. J. Med.* 380, 142–151. doi: 10.1056/NEJMoa1806891
- Haas, M. E., Aragam, K. G., Emdin, C. A., Bick, A. G., International Consortium for Blood Pressure, I. C. for B, Hemani, G., et al. (2018). Genetic association of albuminuria with cardiometabolic disease and blood pressure. *Am. J. Hum. Genet.* 103, 461–473. doi: 10.1016/j.ajhg.2018.08.004
- Halder, A., Jain, M., Chaudhary, I., Gupta, N., and Kabra, M. (2013). Fluorescence in situ hybridization (FISH) using non-commercial probes in the diagnosis of clinically suspected microdeletion syndromes. *Indian J. Med. Res.* 138, 135–142. doi: 10.4103/0971-6866.112877
- Hill, N. R., Fatoba, S. T., Oke, J. L., Hirst, J. A., O'Callaghan, C. A., Lasserson, D. S., et al. (2016). Global prevalence of chronic kidney disease—a systematic review and meta-analysis. *PLoS One* 11, e0158765. doi: 10.1371/journal.pone.0158765
- Hiura, Y., Fukushima, Y., Kokubo, Y., Okamura, T., Goto, Y., Nonogi, H., et al. (2008). Effects of the Y chromosome on cardiovascular risk factors in Japanese men. *Hypertens. Res.* 31, 1687–1694. doi: 10.1291/hyres.31.1687
- Jordan, J. J., Hanlon, A. L., Al-Saleem, T. I., Greenberg, R. E., and Tricoli, J. V. (2001). Loss of the short arm of the Y chromosome in human prostate carcinoma. *Cancer Genet. Cytogenet.* 124, 122–126. doi: 10.1016/S0165-4608(00)00340-X
- Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., et al. (2002). The human genome browser at UCSC. *Genome Res.* 12, 996–1006. doi: 10.1101/gr.229102
- Khosravi, P., Gazestani, V. H., Asgari, Y., Law, B., Sadeghi, M., and Goliaei, B. (2014). Network-based approach reveals Y chromosome influences prostate cancer susceptibility. *Comput. Biol. Med.* 54, 24–31. doi: 10.1016/j.compbiomed.2014.08.020
- Kim, S., and Misra, A. (2007). SNP genotyping: technologies and biomedical applications. *Annu. Rev. Biomed. Eng.* 9, 289–320. doi: 10.1146/annurev.bioeng.9.060906.152037
- Kivisild, T., Rootsi, S., Metspalu, M., Mastana, S., Kaldma, K., Parik, J., et al. (2003). The genetic heritage of the earliest settlers persists both in Indian tribal and caste populations. *Am. J. Hum. Genet.* 72, 313–332. doi: 10.1086/346608
- Komura, K., Jeong, S. H., Hinohara, K., Qu, F., Wang, X., Hiraki, M., et al. (2016). Resistance to docetaxel in prostate cancer is associated with androgen receptor activation and loss of KDM5D expression. *Proc. Natl. Acad. Sci. U. S. A.* 113, 6259–6264. doi: 10.1073/pnas.1600420113
- Kostrzewa, G., Broda, G., Konarzewska, M., Krajewski, P., and Płoski, R. (2013). Genetic polymorphism of human Y chromosome and risk factors for

- cardiovascular diseases: a study in WOBASZ cohort. *PLoS One* 8, e68155. doi: 10.1371/journal.pone.0068155
- Krausz, C., Hoefsloot, L., Simoni, M., Tüttelmann, F., and European Academy of Andrology, and European Molecular Genetics Quality Network (2014). EAA/EMQN best practice guidelines for molecular diagnosis of Y-chromosomal microdeletions: state-of-the-art 2013. *Andrology* 2, 5–19. doi: 10.1111/j.2047-2927.2013.00173.x
- Kuderna, L. F. K., Lizano, E., Julià, E., Gomez-Garrido, J., Serres-Armero, A., Kuhlwil, M., et al. (2019). Selective single molecule sequencing and assembly of a human Y chromosome of African origin. *Nat. Commun.* 10, 4. doi: 10.1038/s41467-018-07885-5
- Lata, S., Marasa, M., Li, Y., Fasel, D. A., Groopman, E., Jobanputra, V., et al. (2018). Whole-exome sequencing in adults with chronic kidney disease: a pilot study. *Ann. Intern. Med.* 168, 100–109. doi: 10.7326/M17-1319
- Levy, S. E., and Myers, R. M. (2016). Advancements in next-generation sequencing. *Annu. Rev. Genomics Hum. Genet.* 17, 95–115. doi: 10.1146/annurev-genom-083115-022413
- Li, M., Carey, J., Cristiano, S., Susztak, K., Coresh, J., Boerwinkle, E., et al. (2017). Genome-wide association of copy number polymorphisms and kidney function. *PLoS One* 12, e0170815. doi: 10.1371/journal.pone.0170815
- Lu, C., Zhang, J., Li, Y., Xia, Y., Zhang, F., Wu, B., et al. (2009). The b2/b3 subdeletion shows higher risk of spermatogenic failure and higher frequency of complete AZFc deletion than the gr/gr subdeletion in a Chinese population. *Hum. Mol. Genet.* 18, 1122–1130. doi: 10.1093/hmg/ddn427
- Maan, A. A., Eales, J., Akbarov, A., Rowland, J., Xu, X., Jobling, M. A., et al. (2017). The Y chromosome: a blueprint for men's health? *Eur. J. Hum. Genet.* 25, 1181–1188. doi: 10.1038/ejhg.2017.128
- MacArthur, J., Bowler, E., Cerezo, M., Gil, L., Hall, P., Hastings, E., et al. (2017). The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.* 45, D896–D901. doi: 10.1093/nar/gkx1133
- Marchini, J., and Howie, B. (2010). Genotype imputation for genome-wide association studies. *Nat. Rev. Genet.* 11, 499–511. doi: 10.1038/nrg2796
- Marchini, J., Howie, B., Myers, S., McVean, G., and Donnelly, P. (2007). A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* 39, 906–913. doi: 10.1038/ng2088
- McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A. R., Teumer, A., et al. (2016). A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* 48, 1279–1283. doi: 10.1038/ng.3643
- McLachlan, R. I., and O'Bryan, M. K. (2010). State of the art for genetic testing of infertile men. *J. Clin. Endocrinol. Metab.* 95, 1013–1024. doi: 10.1210/jc.2009-1925
- Möller-Leimkühler, A. M. (2007). Gender differences in cardiovascular disease and comorbid depression. *Dialogues Clin. Neurosci.* 9, 71–83.
- Musetti, C., Quaglia, M., Mellone, S., Pagani, A., Fusco, I., Monzani, A., et al. (2014). Chronic renal failure of unknown origin is caused by *HNF1B* mutations in 9% of adult patients: a single centre cohort analysis. *Nephrology* 19, 202–209. doi: 10.1111/nep.12199
- Muzzev, D., Evans, E. A., and Lieber, C. (2015). Understanding the basics of NGS: from mechanism to variant calling. *Curr. Genet. Med. Rep.* 3, 158–165. doi: 10.1007/s40142-015-0076-8
- Nanayakkara, S., Senevirathna, S. T. M. L. D., Abeysekera, T., Chandrajith, R., Ratnatunga, N., Gunarathne, E. D. L., et al. (2014). An integrative study of the genetic, social and environmental determinants of chronic kidney disease characterized by tubulointerstitial damages in the North Central Region of Sri Lanka. *J. Occup. Health* 56, 28–38. doi: 10.1539/joh.13-0172-OA
- Nargesi, M. M., Ismail, P., Razack, A. H. A., Pasalar, P., Nazemi, A., Oshkoor, S. A., et al. (2011). Linkage between prostate cancer occurrence and Y-chromosomal DYS loci in Malaysian subjects. *Asian Pac. J. Cancer Prev.* 12, 1265–1268.
- Neto, F. T. L., Bach, P. V., Najari, B. B., Li, P. S., and Goldstein, M. (2016). Genetics of male infertility. *Curr. Urol. Rep.* 17, 70. doi: 10.1007/s11934-016-0627-x
- Noveski, P., Madjunkova, S., Sukarova Stefanovska, E., Matevska Geshkovska, N., Kuzmanovska, M., Dimovski, A., et al. (2016). Loss of Y chromosome in peripheral blood of colorectal and prostate cancer patients. *PLoS One* 11, e0146264. doi: 10.1371/journal.pone.0146264
- Paracchini, S., Pearce, C. L., Kolonel, L. N., Altshuler, D., Henderson, B. E., and Tyler-Smith, C. (2003). A Y chromosomal influence on prostate cancer risk: the multi-ethnic cohort study. *J. Med. Genet.* 40, 815–819. doi: 10.1136/jmg.40.11.815
- Pattaro, C., Teumer, A., Gorski, M., Chu, A. Y., Li, M., Mijatovic, V., et al. (2016). Genetic associations at 53 loci highlight cell types and biological pathways relevant for kidney function. *Nat. Commun.* 7, 10023. doi: 10.1038/ncomms10023
- Payne, A., Holmes, N., Rakyan, V., and Loose, M. (2018). BulkVis: a graphical viewer for Oxford nanopore bulk FAST5 files. *Bioinformatics* 35, 2193. doi: 10.1093/bioinformatics/bty841
- Perinchery, G., Sasaki, M., Angan, A., Kumar, V., Carroll, P., and Dahiya, R. (2000). Deletion of Y-chromosome specific genes in human prostate cancer. *J. Urol.* 163, 1339–1342. doi: 10.1016/S0022-5347(05)67774-9
- Ran, J., Han, T., Ding, X., Wei, X., Zhang, L., Zhang, Y., et al. (2013). Association study between Y-chromosome haplogroups and susceptibility to spermatogenic impairment in Han People from southwest China. *Genet. Mol. Res.* 12, 59–66. doi: 10.4238/2012.January.22.4
- Rodriguez, S., Chen, X., Miller, G. J., and Day, I. N. M. (2005). Non-recombining chromosome Y haplogroups and centromeric *HindIII* RFLP in relation to blood pressure in 2,743 middle-aged Caucasian men from the UK. *Hum. Genet.* 116, 311–318. doi: 10.1007/s00439-004-1221-8
- Russo, P., Siani, A., Miller, M. A., Karanam, S., Esposito, T., Gianfrancesco, F., et al. (2008). Genetic variants of Y chromosome are associated with a protective lipid profile in Black men. *Arterioscler. Thromb. Vasc. Biol.* 28, 1569–1574. doi: 10.1161/ATVBAHA.108.168641
- Russo, P., Venezia, A., Lauria, F., Strazzullo, P., Cappuccio, F., Iacoviello, L., et al. (2006). *HindIII*(+/-) polymorphism of the Y chromosome, blood pressure, and serum lipids: no evidence of association in three white populations. *Am. J. Hypertens.* 19, 331–338. doi: 10.1016/j.amjhyper.2005.10.003
- Sen, S., Agarwal, R., Ambulkar, P., Hinduja, I., Zaveri, K., Gokral, J., et al. (2016). Deletion of GOLGA2P3Y but not GOLGA2P2Y is a risk factor for oligozoospermia. *Reprod. Biomed. Online* 32, 218–224. doi: 10.1016/j.rbmo.2015.11.001
- Sezgin, E., Lind, J. M., Shrestha, S., Hendrickson, S., Goedert, J. J., Donfield, S., et al. (2009). Association of Y chromosome haplogroup I with HIV progression, and HAART outcome. *Hum. Genet.* 125, 281–294. doi: 10.1007/s00439-008-0620-7
- Shankar, R., Charchar, F., Eckert, G., Saha, C., Tu, W., Dominiczak, A., et al. (2007). Studies of an association in boys of blood pressure and the Y chromosome. *Am. J. Hypertens.* 20, 27–31. doi: 10.1016/j.amjhyper.2006.06.013
- Shoji, M., Tsutaya, S., Shimada, J., Kojima, K., Kasai, T., and Yasujima, M. (2002). Lack of association between Y chromosome Alu insertion polymorphism and hypertension. *Hypertens. Res.* 25, 1–3. doi: 10.1291/hyres.25.1
- Skaletsky, H., Kuroda-Kawaguchi, T., Minx, P. J., Cordum, H. S., Hillier, L., Brown, L. G., et al. (2003). The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 423, 825–837. doi: 10.1038/nature01722
- Snell, D. M., and Turner, J. M. A. (2018). Current biology review sex chromosome effects on male–female differences in mammals. *Curr. Biol.* 28, R1313–R1324. doi: 10.1016/j.cub.2018.09.018
- Snoek, R., van Setten, J., Keating, B. J., Israni, A. K., Jacobson, P. A., Oetting, W. S., et al. (2018). NPHP1 (Nephrocystin-1) gene deletions cause adult-onset ESRD. *J. Am. Soc. Nephrol.* 29, 1772–1779. doi: 10.1681/ASN.2017111200
- Stuppia, L., Gatta, V., Mastroprimiano, G., Pompetti, F., Calabrese, G., Guanciali Franchi, P., et al. (1997). Clustering of Y chromosome deletions in subinterval E of interval 6 supports the existence of an oligozoospermia critical region outside the DAZ gene. *J. Med. Genet.* 34, 881–883. doi: 10.1136/jmg.34.11.881
- van Heyningen, V., Hoovers, J. M. N., de Kraker, J., and Crolla, J. A. (2007). Raised risk of Wilms tumour in patients with aniridia and submicroscopic WT1 deletion. *J. Med. Genet.* 44, 787–790. doi: 10.1136/jmg.2007.051318
- van Oven, M., Van Geystelen, A., Kayser, M., Decorte, R., and Larmuseau, M. H. (2014). Seeing the wood for the trees: a minimal reference phylogeny for the human Y chromosome. *Hum. Mutat.* 35, 187–191. doi: 10.1002/humu.22468
- Verbitsky, M., Sanna-Cherchi, S., Fasel, D. A., Levy, B., Kiryluk, K., Wuttke, M., et al. (2015). Genomic imbalances in pediatric patients with chronic kidney disease. *J. Clin. Invest.* 125, 2171–2178. doi: 10.1172/JCI80877
- Vogt, P. H., Edelmann, A., Kirsch, S., Henegariu, O., Hirschmann, P., Kiesewetter, E., et al. (1996). Human Y chromosome azoospermia factors (AZF) mapped to different subregions in Yq11. *Hum. Mol. Genet.* 5, 933–943. doi: 10.1093/hmg/5.7.933

- Voskarides, K., Hadjipanagi, D., Papazachariou, L., Griffin, M., and Panayiotou, A. G. (2014). Evidence for contribution of the Y chromosome in atherosclerotic plaque occurrence in men. *Genet. Test. Mol. Biomarkers* 18, 552–556. doi: 10.1089/gtmb.2014.0020
- Webster, T. H., Couse, M., Grande, B. M., Karlins, E., Phung, T. N., Richmond, P. A., et al. (2019). Identifying, understanding, and correcting technical artifacts on the sex chromosomes in next-generation sequencing data. *GigaScience* 8. doi: 10.1093/gigascience/giz074
- Winney, B., Boumertit, A., Day, T., Davison, D., Echeta, C., Evseeva, I., et al. (2012). People of the British Isles: preliminary analysis of genotypes and surnames in a UK-control population. *Eur. J. Hum. Genet.* 20, 203–210. doi: 10.1038/ejhg.2011.127
- Xue, A., Wu, Y., Zhu, Z., Zhang, F., Kemper, K. E., Zheng, Z., et al. (2018). Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. *Nat. Commun.* 9, 2941. doi: 10.1038/s41467-018-04951-w
- Xue, C., Zhao, J., Liu, F., Lu, C., Yang, S.-T., and Bai, F.-W. (2013). Two-stage *in situ* gas stripping for enhanced butanol fermentation and energy-saving product recovery. *Bioresour. Technol.* 135, 396–402. doi: 10.1016/j.biortech.2012.07.062
- Yan, Y., Yang, X., Liu, Y., Shen, Y., Tu, W., Dong, Q., et al. (2017). Copy number variation of functional RBMY1 is associated with sperm motility: an azoospermia factor-linked candidate for asthenozoospermia. *Hum. Reprod.* 32, 1521–1531. doi: 10.1093/humrep/dex100
- Zerbino, D. R., Achuthan, P., Akanni, W., Amode, M. R., Barrell, D., Bhai, J., et al. (2018). Ensembl 2018. *Nucleic Acids Res.* 46, D754–D761. doi: 10.1093/nar/gkx1098
- Zhang, F., Chen, R., Liu, D., Yao, X., Li, G., Jin, Y., et al. (2013). YHap: a population model for probabilistic assignment of Y haplogroups from re-sequencing data. *BMC Bioinformatics* 14, 331. doi: 10.1186/1471-2105-14-331
- Zhou, W., Machiela, M. J., Freedman, N. D., Rothman, N., Malats, N., Dagnall, C., et al. (2016). Mosaic loss of chromosome Y is associated with common variation near TCL1A. *Nat. Genet.* 48, 563–568. doi: 10.1038/ng.3545

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Anderson, Cañadas-Garre, Chambers, Maxwell and McKnight. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.