Towards understanding the inherited susceptibility for nephropathy in diabetes

Merlin C. Thomas\textsuperscript{a}, Per-Henrik Groop\textsuperscript{a,b,c}, and Karl Tryggvason\textsuperscript{d}

Purpose of review
The burden of nephropathy is unequally shared across patients with diabetes. The majority of the variability in incident nephropathy remains unaccounted for by conventional risk factors. There appears to be an inherited predisposition for diabetic nephropathy, but this does not follow simple Mendelian rules. Any inherited predisposition for nephropathy is far more complicated. This article reviews the recent advances in understanding of the genetics and epigenetics of diabetic nephropathy.

Recent findings
A few candidate genes have been reproducibly associated with diabetic nephropathy, and recent genome-wide linkage studies have also identified chromosomal loci for susceptibility genes, including 3q, 7q, 10p, 14q and 18q. Unbiased, genome-wide linkage studies have identified specific loci and genome-wide association studies a number of new loci. However, any roles of those genes in the molecular pathobiology remain to be established. Moreover, their individual contribution to the variability in incident nephropathy in diabetes appears to be small.

Summary
New genome-wide approaches offer new opportunities to identify genes associated with diabetic nephropathy. However, such approaches have key limitations. Upto the present time, genetic testing has failed to identify a gene or combination of genes that will substantially identify those patients most at risk for diabetic nephropathy. It may be that epigenetic regulation of gene expression may represent a more important contributor to an inherited predisposition to diabetic nephropathy. Nonetheless, genetic studies may provide valuable information regarding the pathobiology of nephropathy and potential targets for its treatment.

Keywords
chronic kidney disease, diabetic complications, diabetic nephropathy, genome-wide association studies

INTRODUCTION
Diabetes currently affects between 3 and 4% of the world’s population \cite{1}. Although diabetes deleteriously affects many organ systems, its impact on the kidney is among the most significant. Some degree of renal dysfunction is observed in at least half of all patients with type 2 diabetes \cite{2,3} and approximately one in three patients with type 1 diabetes \cite{4–7}. Diabetic kidney disease is the leading single cause of end-stage kidney disease. However, the clinical and socioeconomic burden imparted by diabetic kidney disease is much broader than the requirement for dialysis or transplantation. In individuals with diabetes, the presence and severity of kidney disease adversely affects their well being, significantly contributes to burden of morbidity and increases their risk of a premature death \cite{8,9}. Indeed, in adults with type 1 diabetes, excess mortality associated with diabetes is almost entirely confined to those with nephropathy \cite{10,11}.

The burden of nephropathy is unequally shared across patients with diabetes. A prolonged duration, or inadequate metabolic and/or blood pressure control may explain some cases. But even diabetic
individuals with excellent control may still develop renal complications of their disease. Moreover, the long term survival of some of Banting and Best’s original patients stand as testament to the fact that some patients appear to be ‘protected’ despite many decades of marked hyperglycaemia, even in the absence of effective treatments. The vast majority of the variability in incident nephropathy remains unaccounted for by conventional risk factors. One possible explanation may be that there is an inherited predisposition for diabetic nephropathy. This article reviews some of the recent advances in the search for its basis.

HERITABILITY OF DIABETIC NEPHROPATHY

Strong evidence for an inherited susceptibility to diabetic nephropathy comes from family aggregation studies, showing that some families have an increased risk of renal disease whereas others do not, despite a shared predisposition for diabetes [12,13]. For example, in the Diabetes Control and Complications Trial (DCCT), there was an increased risk of nephropathy (defined as a urinary albumin excretion rate >40 mg/day) in relatives of nephropathy-positive patients when compared with relatives of nephropathy-negative patients [odds ratio (OR) = 5.4; 95% CI 2.2–13.7; P < 0.001] [14]. Although it may be argued that shared environmental factors may also be important, a family history of nephropathy (or hypertension or premature cardiovascular disease as its surrogate [15]) is an independent risk factor for the development of nephropathy in patients with diabetes. For example, the diabetic offspring of parents with diabetes and proteinuria have a cumulative risk of developing renal disease of over 70%, and a three to four times increased incidence of renal disease compared with children of diabetic parents without renal disease [12,16,17]. The risk appears to be further increased if both parents have diabetic nephropathy. This has led to the suggestion that an inherited predisposition to nephropathy is a dominant trait [12,16,18–20].

CANDIDATE GENES

The constraints of genetic sequencing technology have, until very recently, hampered the exploration and identification of specific genetic traits that may predispose to diabetic nephropathy. For the most part, genetic research has previously been restricted to association studies of so-called ‘candidate genes’ that are selected on the basis of their postulated role in diabetic nephropathy. However, this approach may be problematic as the pathophysiology of diabetic nephropathy is poorly understood, so it is difficult to make ‘educated guesses’ about candidate genes. Many hypothetical candidate genes have been explored in association studies in case and control cohorts. Some of those that are listed below are reviewed in detail elsewhere [21]:

1. Adiponectin
2. Aldose reductase
3. Angiotensin-converting enzyme (ACE)
4. Angiotensin II receptor (AGTR1)
5. Angiotensinogen
6. Apolipoprotein E
7. Atrial natriuretic peptide (ANP)
8. Carnosinase
9. Cysteinyl-tRNA synthetase (CARS)
10. Decorin
11. Engulfment and cell motility 1 gene (ELMO-1)
12. FRMD3 (4.1 protein ezrin, radixin, moesin [FERM] domain containing 3)
13. Heterotrimeric G-protein
14. Glucokinase (MODY 2)
15. GLUT-1 transporter
16. Heparan sulphate
17. Hepatocyte nuclear factor-1 beta (MODY 5)
18. 11B hydroxysteroid dehydrogenase (11B HSD)
19. Interacellular adhesion molecule-1 (ICAM-1)
20. Leukocyte-endothelial cell adhesion molecule 1 (LECAM-1)

KEY POINTS

- Despite association of a large number of ‘candidate genes’ and chromosomal loci with diabetic nephropathy, the findings of most of the associations have not been replicated, and their possible contributions to the disease process remain to be established.
- New genome-wide approaches offer new opportunities to identify genes associated with nephropathy. However, such approaches have key limitations.
- Epigenetic regulation of gene expression may represent an important contributor to an inherited predisposition to diabetic nephropathy.
- Although the development of a genetic test for diabetic nephropathy is not likely to occur in the near future, genetic studies may provide valuable information regarding the pathobiology of nephropathy and potential targets for its treatment.
- Identification of novel diabetic nephropathy susceptibility genes may facilitate the development of badly needed animal models that can be used to study disease mechanisms and for testing new drugs.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.
(21) Lipoprotein lipase (LPL)
(22) Lin domain kinase 2 (LIMK2)
(23) Matrix metalloproteinase-9 (MMP-9)
(24) Methylene-tetrahydrofolate reductase
(25) Na/H exchanger
(26) Nitric oxide synthase (especially endothelial NOS)
(27) Plasminogen activator inhibitor-1 (PAI-1)
(28) Peroxisome proliferator-activated receptor (PPAR)
(29) Ribosomal protein S12 (RPS12)
(30) Solute carrier protein S12 (SLC12A3)
(31) Set7 histone lysine methyltransferase
(32) T-cell receptor β-chain (TCRBC)
(33) Transforming growth factor β (TGFβ)
(34) Vitamin D receptor (VDR)
(35) Vascular endothelial growth factor (VEGF)

For example, the bi-allelic ACE insertion/deletion (I/D) polymorphism is associated with at least half of the phenotypic variation in serum ACE activity [22]. Patients homozygous for the deletion allele have the highest ACE activity, whereas those with the insertion polymorphism have the lowest. Initial studies demonstrated that the D allele was associated with nephropathy and premature death in patients with type 1 diabetes. For example, the Genétique de la Nephropathie Diabétique (GENEDIAB) study found that the D allele was associated with the prevalence and severity of nephropathy in 499 patients with type 1 diabetes [23]. This association has been confirmed in some, but not all populations. However, when pooled in a meta-analysis of published data, this association remains significant, though modest in its effect size [24**]. Whether this is mediated via differential responsiveness to blockade of the renin–angiotensin system (RAS) [25], direct actions on renal RAS activity or its linkage to another pathogenic locus [26] remains to be established.

Another potential candidate gene that has been explored in diabetic nephropathy is ELMO1 (engulfment and cell motility 1). ELMO-1 is known to be involved in fibrogenesis, inducing the key regulator TGFβ, and matrix synthesis. Genetic variants at the ELMO1 locus have been found to be associated with diabetic nephropathy in a number of different cohorts and racially distinct populations [27].

A third example of a recent candidate gene to be studied is SET7, a protein that has recently been identified to be important in epigenetic regulation of genetic transcription. Indeed, in-vitro studies in human endothelial cells have demonstrated that SET7 is a key element for hyperglycaemia-mediated changes in histone methylation that contribute to metabolic memory [28]. Such data make SET7 an appealing candidate gene for an inherited predisposition to diabetic nephropathy. This has now been studied in patients with type 1 diabetes, demonstrating that indeed there appears to be an association between SET7 polymorphisms and the development of diabetic nephropathy [29**]. This is a good example of how new science can rapidly be translated into new genetics, when interesting new targets are identified.

Another approach has been to explore in detail those genetic regions that appear to be associated with the development and progression of diabetic nephropathy using genome-wide association studies (GWAS) in case–control cohorts. For example, chromosome 3q has been identified as a major locus for diabetic nephropathy susceptibility genes [30–33]. Within this locus, Vionnet et al. [34] examined association of 14 candidate genes at the 3q locus in Danish, Finnish and French cohorts, and found the strongest association of diabetic nephropathy with polymorphisms in the adiponectin (ADIPOQ) gene.

Although the candidate gene approach may appear appealing, it is ultimately limited by our limited understanding of molecular disease mechanisms (i.e. of what pathways drive diabetic complications). Moreover, even the most logical candidates, when analysed together, account for even as little as 5% of the genetic variability in nephropathy [18]. Another problem with these candidate studies is their limited power in detecting common gene variants that confer only a modest increase in the risk of disease, although such polymorphisms may account for much more variability in the incidence of nephropathy between patients. However, one of the most important problems with ‘candidate’ genetic studies, thus far, has been that most of the studies performed have been small and underpowered to detect a significant difference in outcomes, given the wide metabolic and clinical management variability across and between patients. This means that, although nominally significant, falsely positive findings have tended to be over emphasized, whereas falsely negative studies have not been published or presented.

Furthermore, at least initially, some studies made limited or no attempts to replicate findings or they failed to adjust for multiple comparisons. This has led to a consensus whereby all genetic association studies are now required to include documented replication of findings as a prerequisite for publication.

Yet another problem coming from small study size is that, to make up for reduced power, some
studies have also pooled outcomes, like microalbuminuria and end-stage renal disease (ESRD). Rather than enhancing the conclusions, this has potentially made them more confusing. For example, the effects of RAS blockade are less apparent for the prevention of microalbuminuria in patients with type 1 diabetes [35], than the prevention of ESRD [36]. So if RAS blockade works differentially, it is conceivable that a genetic predisposition may work similarly. Equally, pooling macroalbuminuria and ESRD cases may be problematic, if more patients with macroalbuminuria die than develop ESRD, meaning that ESRD patients are ‘survivors’ [37].

One recent approach to try to overcome the deficiencies of individual studies has been to look at the pooled effect of genetic variants across a number of studies and ethnic groups, in the same way that meta-regression has been applied with effect to clinical trials. Such studies potentially suffer from the publication bias and heterogeneity introduced in different study designs in different populations, but also have advantages in reflecting larger patients numbers across a range of settings. The most recent attempt in diabetic kidney disease has pooled data from 671 genetic association studies (most of which were case–control studies), which identified 34 reproduced genetic variants associated with diabetic nephropathy [24**]. Of these, only three genetic variants met stringent criteria for significance, the ACE I/D polymorphism (ACE rs179975), the lipoprotein polymorphisms (APOE E2/3/4) and the polyol pathway polymorphism (AKRB1 CA repeat Z–2). Together these genes accounted for only a fraction of variability in diabetic nephropathy.

GENOME-WIDE LINKAGE AND ASSOCIATION STUDIES

The development of new genetic sequencing technologies and computing resources to analyse their results has significantly changed the approach to genetic disease. Instead of small studies exploring familial clustering across a limited number of genomic variations, focusing on a few plausible linkages, genetic epidemiology has provided a means to examine common variations across the entire genome repeated across medium size and large populations. The GWAS have a number of advantages as well as key limitations, including requirement for large sample sizes, the potential for erroneous results with the massive number of statistical tests performed and for biases related to phenotyping of study participants, and selection of cases and controls. Although larger and more expensive, cohort studies are potentially more representative than selected case–control studies, and have particular advantages for the study of diabetic nephropathy (see below). In theory, by sequencing all or most of the genomes such studies are unconstrained by prior hypotheses regarding genetic associations with disease. As such, linkage and GWAS are considered ‘unbiased’ and ‘hypothesis generating’, rather than ‘hypothesis driven’. The utility of GWAS is based on the ‘common disease, common variant’ hypothesis, which proposes that genetic influences on common diseases may be attributable to a limited number of common allelic variants, present in a more than 1–5% of the population [38], each individually conferring a modest increase in relative risk, but which together confer a phenotypic predisposition. Although such an approach has advantages in detecting prevalent genetic associations, this approach may not identify rare variants and structural variants that may be more strongly associated with disease or information regarding disease pathobiology. However, in selected cases subsequent application of more detailed deep sequencing to extend the results of successful GWAS can be used to identifying additional, rare and/or highly functional variants.

Over the last decade, there have been a number of attempts to identify key loci for diabetic nephropathy using genome-wide linkage approaches. Together, they have identified several chromosomal regions that appear to be associated with the development and progression of diabetic nephropathy, and contain genes that are plausibly associated with diabetic pathophysiology. For example, four independent genome-wide linkage studies [30–33] have identified chromosome 3q as a major locus for diabetic nephropathy susceptibility genes. We have recently demonstrated a significant association of diabetic nephropathy to a cluster of three genes on that chromosomal region [39], including NCK1 and TMEM22. Both these proteins are expressed in glomeruli. In particular, Nck1 appears to be an important link between phosphorylated nephrin and the actin cytoskeleton during the development of podocyte foot processes, as well as in their regeneration during repair of effaced foot processes following glomerular injury [40,41].

Other chromosomal regions have also been associated with diabetic nephropathy in different populations. For example, the Family Investigation of Nephropathy and Diabetes (FIND), looking at the predisposition for nephropathy in diabetic sibling pairs concordant and discordant for diabetic nephropathy, found strongest evidence of linkage to the diabetic nephropathy trait on chromosomes 7q, 10p, 14q and 18q [42]. These have been subsequently replicated by genome-wide linkage studies.
studies in other populations. For example, the locus on chromosome 18q demonstrated in the FIND study has also been linked to diabetic nephropathy in African-Americans with diabetes [32] and Pima Indians [43]. This locus harbours two carnosinase genes, CNDP1 and CNDP2, and carnosinase degrades carnosine (β-alanyl-l-histidine) [44], which is thought to be a renal antioxidant and inhibitor of advanced glycation. Certainly, a single nucleotide polymorphism (SNP) in the 3′ untranslated region of CNDP2 (rs7577) appears to be associated with increased risk of diabetic nephropathy in Swedish patients with type 2 diabetes [44].

In the last 5 years, GWAS have further advanced understanding of the genetics of common complex diseases, including type 1 and 2 diabetes. More recently attention has turned to the study of diabetic complications, including retinopathy, neuropathy and nephropathy. Nephropathy lends itself to study as its incidence clearly segregates across diabetic populations. Only 25–40% of patients with type 1 diabetes will ever develop overt nephropathy [4–6,45], whereas almost all patients with type 1 diabetes eventually develop retinopathy.

In the first study to be published in patients with type 1 diabetes, GWAS were implemented in 820 case participants (284 with proteinuria and 536 with prevalent end-stage renal disease) and 885 control individuals with type 1 diabetes without nephropathy from the Genetics of Kidneys in Diabetes (GoKinD) collection [26]. Subsequent confirmation of implicated SNPs was sought in 1304 participants of the DCCT/Epidemiology of Diabetes Interventions and Complications (EDIC) study. Despite its size, this study identified a total of only 13 SNPs located in four genomic loci, which were associated with diabetic nephropathy with \( P < 1 \times 10^{-5} \). The strongest demonstrated association was at the FRMD3 [4.1 protein ezrin, radixin, moesin (FERM) domain containing 3] locus (OR=1.45, \( P = 5.0 \times 10^{-7} \)). Whether this association is causal remains to be established, although previous studies have demonstrated that the FERM domain at least on ezrin, radixin, moesin is able to interact with advanced glycated end product-modified proteins [46,47].

Another strong association was also identified at the CARS (cysteinyl-tRNA synthetase) locus (OR=1.36, \( P = 3.1 \times 10^{-6} \)). CARS is a cytosolic regulator of cysteine metabolism and its incorporation into proteins. Cysteine metabolism appears to be important for renal health as the lysosomal storage disorder cystinosis ultimately leads to renal impairment caused by the accumulation of free cysteine in cellular lysosomes. Cysteamine, which depletes intracellular cysteine, is also a potent renal antioxidant [48]. Although circumstantial, such data suggest the association with CARS may be more than a coincidence.

Although interesting, this analysis has a number of potential limitations, including the dominance of patients with prevalent ESRD used as cases. As many patients with overt nephropathy die before developing ESRD, these may represent the survivors rather than the high-risk patients with nephropathy in whom genetic testing may have particular advantages [37]. Moreover, the relatively small numbers of patients in this study led to the use of a lower threshold for significance \( (1 \times 10^{-5}) \) than is ideal (usually \( <10^{-7} \)), raising the possibility for both false positive and false negative results. It also means that this study was underpowered to detect genetic effects of a smaller magnitude than has been observed with much larger analyses of other complex disorders, including type 2 diabetes.

**WHY DOES IT NOT WORK LIKE IN THE MOVIES?**

Although the myth of genetic determinism (or fate) still pervades much of modern society, it remains a ‘science fiction’. It is now clear that even with comprehensive genetic sequencing only a small fraction of the variability in complex phenotypic traits may be explained by genetic variations. With the exception of rare monogenetic disorders, the pathogenesis of disease appears to be the result of complex interactions between environmental factors and inherited predisposition. There may be a number of reasons why it has not been possible to tease out the genetic component from the mixture. These include that the variability in response is often polygenic; haplotype structure may be a more important determinant of phenotype than individual SNPs; there is difficulty in defining a quantitative trait, for example, in the case of diabetic nephropathy individuals with established complications (cases) are by definition survivors, although nephropathy is also the leading risk factor for premature mortality; and populations are not individuals. If genes simply facilitate the actions of environmental stimuli, then the same genes may be associated with either protection or disease depending on the environment, meaning that the net effect may appear nonsignificant overall, although they may be contextually very important.

**EPIGENETICS OF DIABETIC NEPHROPATHY**

Epigenetics has emerged as an increasingly interesting paradigm to understand complex
non-Mendelian diseases. In essence, epigenetics represents ‘the structural adaptation of chromosomal regions so as to store, retain and recall past experiences in a way to shape present and future behaviour’ [49]. Such adaptations are made possible because DNA does not exist naked within a eukaryotic cell, but rather extraordinary lengths of DNA are carefully packaged as a DNA–protein complex (chromatin) into the tight confines of the cell nucleus. The structure and accessibility of chromatin significantly influences the regulation of gene expression, potentially even more so than the DNA sequence itself. There are many different epigenetic modifications. The most well recognized of which include DNA methylation, histone posttranslational modifications and RNA-based mechanisms, which have differing effects on gene expression depending on the type of modification, their location and combination. For example, covalent modifications to the unstructured tails of histones with acetylation, methylation, phosphorylation, ubiquitinylation, sumoylation, and/or other changes result in a histone ‘code’ that is read and translated into signals for activation or repression of associated genes. Certain histone modifications are most often associated with repressed genes, like H3K9 and H3K27 methylation, which function as repressive marks. Other modifications like methylation of histones H3K4 and H3K36 are associated with transcriptionally active genes. For example, transiently exposing endothelial cells to high glucose triggers increased methylation of H3K4, a transcription activating mark adjacent to the proximal promoter region of the nuclear factor kappa B (NFκB) subunit p65 gene [28]. This results in increased expression of p65 and its downstream effects on inflammatory pathways. Importantly this ‘epigenetic mark’ is quite specific, and other histone modifications are not seen adjacent to the p65 promoter. Moreover, inhibition of this specific methylation event prevents glucose-mediated upregulation of p65 and its associated inflammasome.

Some epigenetic modifications are also heritable, consistent with the Lamarckian ‘Theory of Inheritance of Acquired Characteristics’, which states that ‘If an organism changes during life in order to adapt to its environment, those changes may be passed on to its offspring’. The most well known example of trans-generational persistence of epigenetic changes is observed in the agouti viable yellow (Avy) gene. If the gene has little or no methylation, then it is active in all cells, and the mouse is yellow. But if Avy is highly methylated, its expression is switched off and the mouse is a sooty-brown colour. In between these two extremes, Avy can be methylated to varying degrees, resulting in a spectrum of mottled mice, in which Avy activity even varies from cell to cell. The more pregnant mice are fed methyl donors, the browner and more heavily mottled are the baby mice [50]. But more importantly, this environmental stimulus also affects the next generation in the same way [51].

In theory, epigenetic imprinting may also be important for an inherited predisposition for complex diseases [52], like diabetic nephropathy. For example, as stated above, the expression of ACE is modified by genetic polymorphisms, which are associated with incident nephropathy in patients with type 1 diabetes. However, somatic ACE contains two CpG islands in its proximal promoter region. When this DNA is de-methylated and associated histones are acetylated, ACE expression is markedly upregulated. But when these CpG islands are methylated, ACE expression is silenced [53]. Such data make it possible that genetic and epigenetic phenotypes may be identical, and that genetic association studies may be informing us about what sequences epigenetic changes must affect in order to result in disease. The application of high-throughput sequencing technologies beyond merely DNA sequencing, to include the methylation of cytosines and histone modifications on a genome-wide scale, represents the likely new course towards the realisation of a genetic predisposition to nephropathy. As an early taste of things to come, researchers at Temple University Medical School have recently identified 187 genes that are differentially methylated between African–American and Hispanic diabetes patients with end-stage renal disease and diabetes patients without nephropathy [54]. Interestingly, almost one in five of these genes (21%) had been previously implicated in diabetic nephropathy by genome association studies. Such data suggest that we may have identified the right genes for diabetic nephropathy all along, but that we failed to appreciate that variability in their expression was only weakly determined by genetic polymorphisms.

**CONCLUSION**

Although it is clear that there is an inherited predisposition for nephropathy in diabetes, and a number of potential contributors have been identified, none can, as yet, substantially explain why some individuals and some families seem to have an inordinate burden of disease. There will be no simple genetic test for diabetic nephropathy. However, genetic studies potentially have a more important role, informing scientists as to which specific pathways, when modified, alter the pathobiology of diabetic nephropathy. Certainly, genetic
studies have the potential to identify variants in genes or regulatory elements outside genes that may modulate the risk for diabetic nephropathy, although ‘functionalization’ of the variants identified remains poorly established for most variants, thus far. Identification of diabetic nephropathy-associated DNA variants may also facilitate the development of appropriately ‘humanized’ animal models that mimic human diabetic nephropathy, as adequate experimental models for human diabetic nephropathy do not currently exist [55]. Finally, the genetic studies can also inform about mechanistic pathways that, when modified, alter the pathobiology of diabetic nephropathy. Subsequent targeting of these pathways with selective pharmacological therapies may offer real opportunities for disease prevention. For example, ACE gene polymorphisms have been associated with diabetic nephropathy, although their overall effect is modest. But pharmaco- logical blockade of the RAS may achieve changes in the RAS that are more complete, and potentially more renoprotective than a ‘protective’ ACE polymorphism. Essentially if we can only find what works a little, maybe this can be magnified, and in combination with other interventions, achieve the clinical effects we desperately want for our patients.

Acknowledgements
The authors’ work on diabetic nephropathy has been supported in part by grants from the Australian National Health and Medical Research Council (NHMRC), the Juvenile Diabetes Research Foundation and Kidney Health Australia, Folkhälsan Research Foundation, Wilhelm and Else Stockmann Foundation, Academy of Finland, the Knut and Alice Wallenberg Foundation, the Novo Nordisk Foundation, Swedish Research Council and Sweden’s Foundation for Strategic Research.

Conflicts of interest
There are no conflicts of interest.

REFERENCES AND RECOMMENDED READING
Papers of particular interest, published within the annual period of review, have been highlighted as: of outstanding interest
Additional references related to this topic can also be found in the Current World Literature section in this issue (p. 226).


