Approach to the evaluation of heritable diseases and update on familial focal segmental glomerulosclerosis

Michelle P. Winn

Duke Medical Center, Division of Nephrology, Durham, North Carolina, USA

Abstract
Focal segmental glomerulosclerosis (FSGS) is a pathological entity that is a significant cause of morbidity and mortality throughout the world. It is also a significant cause of end-stage renal disease (ESRD). Glomerular disease is the third leading cause of ESRD, and FSGS comprises a significant proportion of this subgroup. Up to 20% of individuals with ESRD have FSGS. It has been reported in patients from varied ethnic backgrounds including individuals who are of Spanish, North American, North European and African descent. The diagnosis of FSGS is based on renal pathology and requires the presence of areas of glomerular sclerosis and tuft collapse that are both focal and segmental. The clinical hallmarks of FSGS include proteinuria, nephrotic syndrome and, frequently, the progressive loss of renal function. At present, there are no consistently reliable treatments for FSGS and response rates to available treatments have been estimated at <30–50%. FSGS has been characterized previously as having primary (idiopathic), secondary and familial forms. In the latter category, both autosomal recessive and dominant inheritance patterns have been reported. Advances in molecular genetics technology and mapping, including high-throughput genotyping for genomic screening, provide powerful tools for the analysis of renal diseases. Genes associated with many familial renal disorders that lead to ESRD have been isolated; these include Alport’s nephropathy, familial juvenile nephronophthisis and adult polycystic disease. Recently, the genetic mutation (ACTN4) causing a form of autosomal dominant FSGS (ACTN4) and congenital nephrotic syndromes (NPHS2) have been described. The existence of hereditary forms of FSGS permits the use of molecular genetics techniques to study the pathogenesis of this disorder.

Keywords: FSGS; familial disease

Introduction
Focal segmental glomerulosclerosis (FSGS) is a pathological diagnosis that is seen in ~5% of adults and 20% of children with end-stage renal disease (ESRD) [1]. It is observed in varied ethnic groups and in all ages [2]. Up to 64% of glomerular lesions in Blacks are FSGS as demonstrated on renal biopsy, the incidence remains between 15 and 20% in other ethnic groups [3]. The incidence of this diagnosis appears to be increasing. It is unclear if this is due to an increase in biopsy patterns or a true increase in disease frequency. The diagnosis is based on areas of glomerular sclerosis and tuft collapse that are both focal and segmental. Common clinical findings are proteinuria, nephrotic syndrome, occasional haematuria and frequent progression to ESRD. Because FSGS has become an important cause of ESRD, it is essential to understand the molecular basis and pathogenesis of this disease process.

There are various subtypes of FSGS, which include primary (idiopathic or sporadic), secondary, familial and FSGS associated with congenital syndromes. Primary FSGS is the most common form of FSGS. Secondary FSGS is associated with another disease process. Examples of secondary FSGS include that which is associated with sickle cell disease, human immunodeficiency virus, heroin nephropathy, reflux nephropathy and others. FSGS has been associated with various congenital syndromes such as Charcot-Marie-Tooth, Laurence-Moon-Biedl syndrome and craniomandibular dermatodysostosis [4–6]. Autosomal dominant and recessive forms of familial FSGS have been reported [7–9]. The renal pathology and clinical characteristics of familial FSGS are virtually identical to that of the more common sporadic forms. While familial FSGS is extremely rare, by using molecular genetic techniques such as linkage analysis, candidate
gene analysis and positional cloning, we can gain beneficial insight into the pathogenesis of this disorder. To elucidate one approach in which to study heritable disease and the current level of understanding of the genetics of FSGS, the Duke familial FSGS study will be used as a paradigm and recently discovered genes, which are thought to cause FSGS will be discussed.

**Patient selection**

Stringent inclusion and exclusion criteria, as well as declaration of affection status are important in minimizing misclassification. Misclassification of families or individuals can lead to false positive and negative results, making linkage analysis elusive. The inclusion criteria in our evaluation of familial FSGS as well as affection status are given below in Tables 1 and 2 [10].

**Clinical spectrum of disease**

Over 26 single-generation and 34 multi-generation kindreds with familial FSGS were obtained. The single-generation families had more aggressive disease and lower mean age of presentation. The male to female ratio was equal. Fifty per cent of our families progressed to ESRD by 30 years of age. Over 540 blood samples were obtained with more than 200 affected individuals. All families had multiple affected individuals and various ethnicities were represented (including Black, White and Native American kindreds). There was decreased renal survival in single-generation vs multi-generation families, Black vs White families, and those presenting with >3 g/24 h of proteinuria. Forty-one patients underwent 48 renal transplant procedures. There was a 62% graft survival at 10 years. Only one patient had recurrence of FSGS in the renal allograft. This is important data as it argues against there being a circulating factor involved in familial FSGS. One would presume that this one person with recurrence may have actually had idiopathic FSGS and coincidentally was a member of a family with familial FSGS [1]. Of interest, of four siblings from a kindred with familial FSGS, two of the siblings donated kidneys to their two siblings. One of the recipients lost his graft due to non-compliance with immunosuppressives, the other recipient has an allograft, which continues to function well at present; however, both donors have progressed to ESRD [11]. This is an important object lesson in carefully screening potential donors for familial FSGS and that potent immunosuppressives may delay the progression of this disorder.

**Linkage analysis**

Currently, highly polymorphic markers called microsatellites are utilized when performing genome-wide scans for linkage analysis. Microsatellites are short tandem repeats with unique sequences flanking each end (Figure 1). They typically repeat every 2–4 bases (di-, tri- and tetra-nucleotide repeats) with highly informative repeats usually being greater than 15. Primers are prepared which are complementary to the unique sequence surrounding the repeat and a polymerase chain reaction is performed to amplify the area. If there is an equal number of repeats on each chromosome, an individual is said to be homozygous at that locus, a person would be heterozygous if there are an unequal number. The more polymorphic a microsatellite is (e.g. different number of alleles at a given locus), the more informative it is for linkage analysis. Microsatellite markers are located throughout the human genome. Although they are located every 30–4000 nt, useful microsatellites are located every 300–50000 nt. Typically, a map with a 5–10 cM spacing is used for the screen. A centiMorgan is a unit of genetic distance that is equivalent to a 1% chance of recombination during meiosis. It is roughly equivalent to a physical distance of 1 Mb, which is equal to 1 million nt [12] (Figure 2).

After genotyping is completed, linkage analysis is performed. ‘Linkage’ refers to the tendency of certain genes to be inherited together. Two genes are said to be ‘linked’ if they are often inherited together, due to their close proximity on a chromosome. Figure 3 demonstrates the concept of linkage. The completely shaded symbols denote an affected individual and the open symbols are unaffected. Notice that the affected father in this pedigree has the 1,2 haplotype, therefore, one of those alleles has to be from the affected chromosome. The mother has the 3,4 haplotype and is not affected. The affected children both have the one allele from the father, the unaffected children have the two from the father. Therefore, the disease has to be associated with the one allele (i.e. is located close to the one allele) and therefore segregates (or travels) with this allele, so they are inherited together.

A logarithm of the odds (LOD) score is a statistical analysis of the likelihood of having found a linkage of the disease of interest to a certain place on a
The LOD score is the base-10 log of the ratio of the likelihood of the data under linkage to the likelihood of the data under no linkage [13]. LOD scores are a function of the recombination fraction (θ) and are calculated over a range of θ values to evaluate the likelihood of the disease being completely linked to the locus (θ = 0) or completely unlinked (θ = 0.5) [14,15]. They also provide an estimate of location of the disease gene. LOD scores of ≥3.0 are indicative of linkage (1000:1 odds in favour of linkage) with a 5% threshold for error. LOD scores of less than −2.0 exclude linkage. Values between −2.0 and 3.0 are inconclusive and require additional family data. The maximum LOD score is used as evidence for linkage. Prior to starting an analysis, the power of the available data set must be determined. If the power is insufficient to detect a LOD score of three, it will be difficult to determine if linkage exists. Table 3 illustrates the available power in the Duke data set.

The two largest families were utilized for initial analysis of familial FSGS. Family 6530, a seven generation, 399 member kindred from New Zealand and Family 6524, a five generation, 65 member family from North Carolina [10] were originally studied. Each family exhibits an autosomal dominant mode of inheritance (e.g. there is male-to-male transmission and affected individuals in each generation).

Candidate genes and loci that were linked previously to FSGS were evaluated for linkage. The PAX2 gene is associated with coloboma, vesicoureteric reflux and renal abnormalities as well as FSGS [16]. Familial steroid-resistant nephrotic syndrome is an autosomal recessive disorder found in European families [17]. Families with chromosome 1 and 17 linked Charcot-Marie-Tooth have been described with FSGS [6,18]. A family identified by Mathis et al., with autosomal dominant familial FSGS was linked to chromosome 19q [19,20]. These genes were excluded in our families as the cause of their disease (see Table 4).

Multipoint LOD scores combine two or more polymorphic microsatellite markers to make them more informative. When a multipoint analysis was done for chromosome 19q autosomal dominant FSGS, the initial two-point LOD score excluded a 15cM area. However, when a multipoint LOD score was performed using the flanking markers D19S213 and D19S223 with D19S191, the area of exclusion was more than 60cM (see Figure 4). The genetic heterogeneity seen between the Mathis et al. patients and the Duke families explains the phenotypic heterogeneity. The Duke cohort had higher-grade proteinuria, lower average age of presentation and more progression to ESRD than the Mathis cohort.

A genome-wide scan was undertaken with Family 6530 and Family 6524. In addition, three smaller New Zealand families were utilized for haplotype analysis as it was postulated that one gene might cause familial FSGS in New Zealand due to the closed environment. The genome scan revealed a linkage of the disease in Family 6530 to chromosome 11q with a maximum LOD score of 9.89 in the affecteds-only analysis and 6.88 in the full pedigree analysis. An affecteds-only analysis is a more stringent test and minimizes mis-classification that can lead to erroneous results. The disease of Family 6524 was excluded from linkage to this area. The three smaller New Zealand families shared no common alleles among affecteds in this area [10].

Fine-mapping using polymorphic microsatellite markers expose recombination events. Recombinations in
affected individuals allow a region to be narrowed for
gene mutation analysis (Figure 5). Subsequent fine-
mapping using this strategy identified recombination
events at D11S1343 and D11S1394 in key affected
individuals which has reduced the candidate region to
/C24
3.5cM. Candidate genes in the region include a family
of matrix metalloproteinases (which cleave different
types of collagen), inhibitors of apoptosis genes 1 and 2
(apoptotic suppressors) and radixin (binds the barbed
end of actin filaments to the plasma membrane).
Preliminary analysis of several metalloproteinases
using a combination of Southern blot analysis, SSCP,
HPLC and direct sequencing of exons as well as
employed. Table 4.

<table>
<thead>
<tr>
<th>Candidate gene region</th>
<th>Markers genotyped</th>
<th>Approximate exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAX2 (10q)</td>
<td>PAX2</td>
<td>22cM</td>
</tr>
<tr>
<td>Chr-1 linked CMT-type 1</td>
<td>APOA2</td>
<td>12cM</td>
</tr>
<tr>
<td>Chr-17 linked CMT-type 1</td>
<td>D1S194</td>
<td>12cM</td>
</tr>
<tr>
<td></td>
<td>DJ7S953</td>
<td>14cM</td>
</tr>
<tr>
<td></td>
<td>DJ7S921</td>
<td>12cM</td>
</tr>
<tr>
<td>Familial steroid-resistant nephrotic syndrome (1q)</td>
<td>DJ5480</td>
<td>12cM</td>
</tr>
<tr>
<td></td>
<td>DJ52883</td>
<td>7cM</td>
</tr>
<tr>
<td>Autosomal dominant FSGS (19q)</td>
<td>DJ9S191</td>
<td>&gt;15cM</td>
</tr>
</tbody>
</table>

The table does not include families with less than three affected individuals or idiopathic cases ascertained.

**Table 4. Candidate gene analysis**

<table>
<thead>
<tr>
<th>Number of families with ≥3 affected individuals</th>
<th>Total number of affected individuals</th>
<th>Number of affected individuals sampled</th>
<th>Total number of individuals sampled</th>
<th>Average LOD score range (SE)</th>
<th>Maximum LOD score range</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>117</td>
<td>68</td>
<td>362</td>
<td>0.2–8.97 (0.01–0.08)</td>
<td>0.56–33.7</td>
</tr>
</tbody>
</table>

*Arrow on vertical axis denotes LOD score of -2.00*

Fig. 4. Multipoint LOD scores combine two or more polymorphic microsatellite markers to make them more informative.

affected individuals allow a region to be narrowed for
gene mutation analysis (Figure 5). Subsequent fine-
mapping using this strategy identified recombination
events at D11S1343 and D11S1394 in key affected
individuals which has reduced the candidate region to
~3.5cM. Candidate genes in the region include a family
of matrix metalloproteinases (which cleave different
types of collagen), inhibitors of apoptosis genes 1 and 2
(apoptotic suppressors) and radixin (binds the barbed
end of actin filaments to the plasma membrane).
Preliminary analysis of several metalloproteinases
using a combination of Southern blot analysis, SSCP,
HPLC and direct sequencing of exons as well as
intron–exon boundaries has revealed evidence of polymorphisms that are common to affected and control individuals.

**New discoveries**

Recently, two gene mutations have been described in
connection with familial FSGS. Previously published
reports of linkage of familial FSGS to chromosome 19q
[19] and autosomal recessive steroid-resistant nephrotic
syndrome to chromosome 1q [17] are the basis for
these new findings. Both mutations were identified by
narrowing the area of interest, using recombination events, to the smallest possible area (the minimal candidate region or MCR); however, subsequent approaches to mutation discovery diverged.

**ACTN4**

After linkage of familial FSGS to chromosome 19q to a large family from Oklahoma, the disease in two additional families was also linked to this area. Using recombination events in affected and unaffected individuals, the area of interest was narrowed to a ~3.5 Mb region [21]. After examining the NPHS1 gene (whose gene product nephrin causes congenital nephrotic syndrome of the Finnish type) which is located in this interval [22,23] and finding no evidence of mutation in the families linked to this area, bioinformatic techniques were utilized to identify candidate genes in the region. ACTN4 was identified via BLAST analysis. BLAST is a program, which searches the National Center for Biotechnology Information (NCBI) databases for sequence similarities. ACTN4 was an excellent candidate gene based on previously published data and was evaluated for mutations in affected individuals of the linked families.

Alpha-actinin-4 (which is encoded by ACTN4) is expressed in glomerular podocytes and involved in cytoskeletal cellular functions. It had been reported previously to be induced prior to foot process effacement in experimental nephrosis [24,25]. Upon sequence analysis of ACTN4 in affected family members, missense mutations were identified in each family. Further evaluation revealed restriction digest changes in affected individuals and not in controls. Alpha-actinin-4 binds actin. When the identified mutations were introduced into cDNA, there was greater binding of F-actin. The actin binding experiments led the authors to postulate that perhaps the mutations alter the mechanical characteristics of the glomerular podocyte, alter the interaction of alpha-actinin-4 with other proteins or alters glomerular hemodynamics due to the location of alpha-actinin-4 in the walls of blood vessels [21].

**NPHS2**

Autosomal recessive, steroid-resistant nephrotic syndrome was linked to chromosome 1q via linkage analysis of several families with the disorder [17]. The approach for identification of the mutations causing this disease utilized traditional positional cloning techniques. The area of interest was refined to a 2.7 cM region. There was a known YAC (yeast artificial chromosome) contig, which spanned this region. A YAC is an artificial chromosome, which is created by combining fragments of foreign DNA with the sequence elements necessary for chromosome function in yeast cells. A PAC (P1 artificial chromosome) contig was constructed using information from the YAC contig. Several known genes and expressed-sequence tags (ESTs are short sequences of cDNA) were assigned

---

Table 5. Chromosome 11 summary LOD scores; Family 6530

<table>
<thead>
<tr>
<th>Thetas</th>
<th>0.00</th>
<th>0.05</th>
<th>0.10</th>
<th>0.15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affecteds only analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D11S2002</td>
<td>−3.92</td>
<td>−0.41</td>
<td>0.11</td>
<td>0.29</td>
</tr>
<tr>
<td>D11S2000</td>
<td>9.89</td>
<td>9.06</td>
<td>8.19</td>
<td>7.29</td>
</tr>
<tr>
<td>D11S1391</td>
<td>4.92</td>
<td>4.56</td>
<td>3.98</td>
<td>3.40</td>
</tr>
<tr>
<td>D11S1986</td>
<td>3.36</td>
<td>5.05</td>
<td>4.85</td>
<td>4.39</td>
</tr>
<tr>
<td>D11S1998</td>
<td>−6.58</td>
<td>0.09</td>
<td>0.63</td>
<td>0.80</td>
</tr>
<tr>
<td>Full pedigree analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thetas</td>
<td>0.00</td>
<td>0.05</td>
<td>0.10</td>
<td>0.15</td>
</tr>
<tr>
<td>D11S2002</td>
<td>−8.22</td>
<td>−3.89</td>
<td>−2.41</td>
<td>−1.55</td>
</tr>
<tr>
<td>D11S2000</td>
<td>2.78</td>
<td>6.11</td>
<td>6.86</td>
<td>6.88</td>
</tr>
<tr>
<td>D11S1391</td>
<td>−5.54</td>
<td>−2.82</td>
<td>−1.65</td>
<td>−0.97</td>
</tr>
<tr>
<td>D11S1986</td>
<td>4.11</td>
<td>4.32</td>
<td>4.18</td>
<td>3.86</td>
</tr>
<tr>
<td>D11S1998</td>
<td>−1.98</td>
<td>3.13</td>
<td>4.28</td>
<td>4.58</td>
</tr>
</tbody>
</table>

---

Fig. 5. Fine-mapping using polymorphic microsatellite markers expose recombination events. Recombinant events are seen when offspring have a different genotype from that of either parent. Recombinations in affected individuals allow the region of interest to be narrowed and subsequent analysis of the area for mutations.
to the PAC contig. After mutation screening of the known candidate genes, one of the ESTs was found to be expressed in kidney tissue. The transcript of interest was amplified from kidney tissue of an affected individual and found to have barely detectable product. The gene structure was then determined and sequenced in affected individuals from different families. Ten different mutations were found in 14 families. This novel gene was named NPHS2 with a protein product designated podocin [26]. The protein product was found to have similarity to the band-7-stomatin protein family, the greatest being to stomatin and Caenorhabditis elegans MEC-2. Analysis of podocin with computer modelling suggests it is an integral membrane protein with a transmembrane domain. In situ hybridization revealed transcript expression restricted to the podocyte. The function of podocin is not yet known, however, C.elegans MEC-2 links the mechanosensory channel and the microtubule cytoskeleton of the touch receptor neurons [27].

Conclusion

FSGS is a major cause of morbidity and mortality. It is seen in all ethnicities and ages and the incidence appears to be increasing, especially in Blacks. Our data have shown relatively early presentation, high-grade proteinuria and rapid progression of ESRD. There was only one recurrence of FSGS in 48 allograft transplants, suggesting a circulating factor does not play a role in familial FSGS. The pathogenesis of disorders such as idiopathic FSGS is very difficult to study due to the numerous etiologies of the disorder. Familial FSGS has already proven to be heterogeneous with disease genes on three different chromosomes thus far.

Important insights have already been gained from the two identified mutated genes, which cause inherited FSGS. It has long been known that there is podocyte foot process fusion and effacement on electron microscopy, however, there is now molecular genetic evidence that specific integral components of the podocyte, which are essential to function, have mutations in this disorder—specifically, cytoskeletal components. Nephrin, the culprit in congenital nephrotic syndrome of the Finnish type has been found to localize to the glomerular podocyte slit diaphragm [28]. The molecular mechanisms remain to be fully elucidated. However, there are new specific targets to evaluate.

Molecular genetic research continues to progress rapidly. Access to entire genomic maps of polymorphic microsatellite markers has made linkage analysis less difficult if suitable cohorts of individuals are available for study. The Internet has advanced genetic research with on-line databases of genomic sequences, tools to compare sequences as well as the modelling of novel proteins and more. As the International Human Genome Sequencing Consortium finishes the Human Genome Project, genetic research will require even more bioinformatics knowledge and the pace of novel findings will undoubtedly increase. The overall goal remains the same, to apply new discoveries to common forms of kidney disease.

References