A genome-wide scan for genes involved in primary vesicoureteric reflux

H Kelly, C M Molony, J M Darlow, M E Pirker, A Yoneda, A J Green, P Puri, D E Barton

Background: Vesicoureteric reflux (VUR) is the retrograde flow of urine from the bladder along the ureter towards the kidneys. VUR is common, occurring in an estimated 1–2% of Caucasian newborns and is a major cause of end-stage renal failure and hypertension in both children and adults. An estimated 30–50% of children presenting with urinary tract infections (UTI) have VUR, and the reflux of infected urine can cause pylonephritis and resultant permanent renal damage. Anatomically, primary vesicoureteric reflux arises from a submucosal ureter that is of insufficient length, or a ureteric orifice that is too wide to allow for the operation of a proper valvular mechanism. VUR occurs frequently in families; approximately half of the offspring will be affected but the mode of inheritance is unknown. Micturating cystourethrography (MCUG) in children may be performed because of UTI or suspicion of renal-tract anomalies from prenatal ultrasonography, and in either case both VUR and renal tract anomalies, including duplex renal systems, are commonly found. Furthermore, VUR is more common in those with duplex than with single renal tracts, and duplex kidneys are commonly found in those with VUR. In both mice and humans, individuals with various renal tract defects, including VUR, occur in the same sibships, indicating that mutations of the same genes can cause both renal developmental anomalies and VUR, with considerable phenotypic variation, for reasons yet to be identified.

To date, the only genes identified as mutated in VUR have been in cases with other anomalies, or in families in which other anomalies were segregating. Hu et al observed VUR among other features in mice in which the gene for uropakin III (locus UPK3A) had been homozygously inactivated. Linkage and mutation screening studies of UPK3A and other members of the uropakin family (Kelly et al, unpublished data) in humans did not show evidence for their involvement in VUR, but mutations in UPK3A have now been found in some cases of renal aplasia, hypoplasia and dysplasia, including some with VUR. The PAX2 gene has been shown to be mutated in renal-coloboma syndrome, which includes VUR as part of a complex phenotype, and in renal hypoplasia and VUR without eye defects, but has not been found to be mutated in uncomplicated VUR. Investigation of a patient with multiple congenital anomalies, including severe bilateral VUR, found a Y;3 translocation that disrupted ROBO2, and investigation of this gene in 124 families with VUR with potential autosomal dominance revealed mutations in two families. Both index cases had bilateral VUR (one having hypoplastic kidneys and nephropathy, the other having a unilateral duplex system), and most affected relatives in both families had renal anomalies. Inheritance in all of the above cases appeared to be autosomal dominant, but mutations in these genes account for a very small proportion of VUR in general.

Attempts at mapping other genes for VUR have included genome-wide scans of small numbers of multi-generation families (the latter yet to be reported in full), or individual or small collections of cases with chromosomal rearrangements, although most patients in the latter group also had other renal tract anomalies. These studies identified several possible loci, and indicated that the disease is genetically heterogeneous. Several candidate-gene studies to confirm some of these and other loci possibly involved in VUR have been unsuccessful.

In this paper, we present the results of a genome-wide linkage scan with 4710 SNP markers and 129 families, each with ≥2 individuals who meet the criteria for a diagnosis of primary VUR.

METHODS

Subjects

The samples for this study were collected at Our Lady’s Children’s Hospital Crumlin, and the National Children’s Hospital.

Abbreviations: HLOD, heterogeneity logarithm of odds; OMIM, Online Mendelian Inheritance in Man; MCUG, micturating cystourethrography; NPL, nonparametric linkage; SNP, single-nucleotide polymorphism; TDT, transmission disequilibrium test; UTI, urinary tract infections; VUR, vesicoureteric reflux

This paper is freely available online under the BMJ Journals unlocked scheme, see http://jmg.bmj.com/info/unlocked.dtl
Hospital, Tallaght, both in Dublin, Ireland. Ethics approval was granted, and informed consent was obtained from all subjects and/or their parents. Families with ≥2 affected siblings with primary VUR of any grade were collected. All families are Caucasian and the majority considered to be of homogeneous Irish ancestry. All index cases were referred because of recurrent UTIs and diagnosed by MCUG. Siblings of index cases were screened by MCUG. Three parents and one grandparent of affected children were classed as affected because they had been diagnosed with primary VUR in the past; others with merely a history of renal problems were classified as of unknown affectedness status. Patients with VUR secondary to other conditions were not collected. Note: most authors agree that secondary VUR can result from bladder dysfunction (neurogenic bladder and non-neurogenic neurogenic bladder (Hinman–Allen syndrome)), posterior urethral valves, or ureterocoele, but some authors include complete duplex renal systems as a cause, and some count some cases with duplex systems as primary and some as secondary. All families included in our study were considered to exhibit primary VUR. It was noted that in one of the excluded families, only one of the affected members had a posterior urethral valve, and in one of the included families, one of the affected members had a ureterocoele. From a developmental genetics point of view, both types of condition could probably safely be included (see Introduction and Discussion).

Genotyping
Samples were sent to Illumina (San Diego, California, USA) for genotyping. The panel (linkage panel III) consisted of 4753 SNP markers, with an average distance between markers of 600 Kb. In total, 626 DNA samples from 133 families were genotyped, but 4 families were subsequently removed because exclusion of members giving non-mendelian results left only one affected member. The following figures for the remaining 129 families refer only to genotyped individuals. They included 255 parents (3 affected, 5 unaffected, and 247 with unknown affectedness status), 283 affected siblings (117 male and 166 female), 37 unaffected siblings (17 male and 20 female), 32 siblings with unknown affectedness status (14 male and 18 female) and 2 grandparents (one affected and one of unknown status). There were 107 families with 2 genotyped affected children, 20 with 3, 1 with 4 and 1 with 5, giving 183 affected sibling pairs in total. Six samples gave no genotype results, thus 638 samples (620 samples + 18 replicates) were genotyped in total. After cleaning the data for mendelian errors and departures from Hardy–Weinberg equilibrium (ie removing obvious genotyping errors) with Pedstats software, 4710 SNPs were used in the analysis. Of the possible 3 004 980 genotypes, 3 002 877 were reported, giving a success rate of 99.93%. The mean heterozygosity over all markers was 46.1%.

Statistical analysis
Both non-parametric and parametric linkage analysis for the scan was performed using MERLIN V.1.0.1.10.3 (http://www.sph.umich.edu/csg/abecasis/Merlin). NPL scores were based on a procedure for converting scores for individual inheritance vectors into Z scores for single or multiple pedigrees, which are used to construct a likelihood ratio test for linkage and define a (NPL) LOD score statistic. The dataset was prepared for analysis using MEGA2 software (http://www.megasoftware.net/) and markers in tight LD (D<0.7) were removed from the dataset prior to analysis using the SNPLINK program (http://www.icr.ac.uk/research/research_sections/cancer_genetics/cancer_genetics_teams/molecular_and_population_genetics/software_and_databases/index.shtml). The resultant panel consisted of 3861 markers. Although all patients included in the study had primary VUR, 25 families included patients with additional phenotypic features. These features were mainly duplex kidneys (20 families, in most of which only one of

![Figure 1](http://jmg.bmj.com/)

Figure 1 Layout plots of nonparametric linkage (NPL) scores from the genome scan. (A) Full dataset; (B) uncomplicated dataset. The vertical axis shows NPL scores. Plots are scaled vertically to show the maximum peak (4.10) in the uncomplicated dataset.
the patients with VUR had a duplex system) but also included fused kidneys, renal hypoplasia, large ureterocoele, branchial cleft remnants, or mental retardation with abnormal facies. After the initial scan, these families were removed to create a subset for further analysis. The reduced set consisted of 104 families, comprising 480 members.

A genetic model of inheritance for VUR has been proposed both in segregation studies and in studies of large multi-generational families. A single model was tested in the parametric analysis, which corresponded to an autosomal dominant mode of inheritance with a disease frequency of 0.01 and incomplete penetrance according to the previously proposed genetic model. Power for the complete cohort is estimated to be 79% (\(p = 0.05\)), given the genetic model, a common test marker allele frequency and \(\theta = 0.1\). This is likely to be an overestimate in general, as power will have fluctuated substantially based on marker and multimarker informativeness, increasing \(\theta\), or the sample size reduction, as in our uncomplicated cohort setting. For parametric analysis, heterogeneity was investigated using the heterogeneity logarithm of odds (HLOD) statistic. Abreu et al suggest that an HLOD of approximately 1.2 for a fully penetrant autosomal dominant genetic model corresponds to a \(p\) value of 0.01 and results exceeding this level are reported for exploratory purposes, as these parameters are rough approximations of a genetic model of VUR. Stringent criteria for statistical significance of genome-wide scans were used, based on previously published guidelines: a \(p\) value of 0.001 and a LOD score of 2.2 (NPL = 3.18) was suggestive of linkage, and \(p < 0.001\) and a LOD score of 3.6 (NPL = 4.08) was considered significant linkage. All regions with \(p \leq 0.02\) (NPL = \(\sim 0.9\)) in the nonparametric analysis were reported and used to define linkage peak intervals. Regions with NPL >2.00 were also highlighted in the results as potentially interesting. Exploratory analysis of association in the linkage panel was performed using the transmission disequilibrium test (TDT) using the discrete trait testing option for general pedigrees in the QTDT program (http://www.sph.umich.edu/csg/abecasis/QTDT).41

RESULTS

Tables 1 and 2 show the results of the genome-wide linkage analysis based on the full set of 129 families. Under genetic homogeneity, all parametric multipoint LOD scores were negative, thus rejecting support for linkage under the model used for the complete sample (table 1). However, positive HLOD scores allowing for heterogeneity indicated that a subset of families might be informative under this model and showed linkage that was suggestive in 10 regions across the genome (table 2). The proportion of linked families \((\pi)\) was estimated as approximately one-third of the sample for each of these regions, and the remaining families must reject linkage of these regions based on the strong negative parametric LOD scores as noted above. NPL analysis yielded six regions that are potentially interesting candidate regions based on their NPL scores under standard criteria (fig 1A, table 1). All of these regions were coincident with regions that exhibited evidence of heterogeneity (HLOD ≥1.2) under the parametric analysis (table 2). In addition to these six regions, six additional regions in the nonparametric scan provided NPL scores that exceeded the less stringent threshold of \(p < 0.02\), three of which were also coincident with regions showing heterogeneity (tables 1 and 2).

The subset of samples we termed the uncomplicated set (UCS), defined by excluding the 25 families in which some affected members had additional phenotypic features, was also analysed to see whether this resulted in stronger linkage support. Greater support for the existence of linkage was provided from the UCS in both the parametric analysis under heterogeneity and nonparametric analyses. NPL analysis yielded a maximum NPL score of 4.1 (\(p < 0.001\)) at 2q37, reaching levels of genome-wide statistical significance. It also identified other potentially interesting loci at 10q26 and 6q27, and nine additional candidate regions on eight different chromosomes with \(p < 0.02\) (table 3, fig 1B); however, the most significant of these, (on 17q) was neither supported by the parametric analysis (table 4) nor seen at all in the full set.

Under the proposed genetic model, parametric analysis of the UCS provided a number of regions with only moderately negative LOD scores, in contrast to the strongly negative LOD

---

### Table 1

<table>
<thead>
<tr>
<th>Location</th>
<th>SNP markers</th>
<th>Interval (cM)</th>
<th>Interval (Mbp)*</th>
<th>Width (Mbp)</th>
<th>LOD</th>
<th>Maximum NPL</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1q23.2-q25.2</td>
<td>rs1053074 rs1932626</td>
<td>152–175</td>
<td>158.3–175.3</td>
<td>17.05</td>
<td>-18.10</td>
<td>2.10</td>
<td>1×10⁻⁴</td>
</tr>
<tr>
<td>2q37.2-q37.3</td>
<td>rs869214 rs16747</td>
<td>246–261</td>
<td>235.6–242.4</td>
<td>6.79</td>
<td>-15.33</td>
<td>2.67</td>
<td>2×10⁻⁴</td>
</tr>
<tr>
<td>3p13-q21.2</td>
<td>rs7064 rs713161</td>
<td>96–131</td>
<td>126.3–127.4</td>
<td>1.1</td>
<td>-17.36</td>
<td>1.36</td>
<td>6×10⁻³</td>
</tr>
<tr>
<td>3q26.31</td>
<td>rs2046718 rs732939</td>
<td>177–178</td>
<td>175.0–176.6</td>
<td>1.6</td>
<td>-20.8</td>
<td>0.98</td>
<td>0.02</td>
</tr>
<tr>
<td>4p1.61-p15.31</td>
<td>rs1981635 rs729918</td>
<td>18–30</td>
<td>10.2–20.1</td>
<td>9.9</td>
<td>-19.77</td>
<td>1.33</td>
<td>7×10⁻⁵</td>
</tr>
<tr>
<td>6q24.1-q27</td>
<td>rs1931992 rs756519</td>
<td>146–188</td>
<td>142.6–170.7</td>
<td>28.1</td>
<td>-9.16</td>
<td>2.55</td>
<td>3×10⁻⁴</td>
</tr>
<tr>
<td>7q26.1-q26.3</td>
<td>rs1547958 rs1343750</td>
<td>165–179</td>
<td>150.3–155.2</td>
<td>4.94</td>
<td>-14.94</td>
<td>2.19</td>
<td>7×10⁻⁴</td>
</tr>
<tr>
<td>10q25.2-q26.3</td>
<td>rs1327551 rs880340</td>
<td>120–165</td>
<td>112.8–134.9</td>
<td>22.1</td>
<td>-11.38</td>
<td>2.35</td>
<td>5×10⁻⁴</td>
</tr>
<tr>
<td>13q33.2-q33.3</td>
<td>rs981900 rs1876723</td>
<td>100–108</td>
<td>105.3–107.4</td>
<td>2.12</td>
<td>-15.17</td>
<td>1.47</td>
<td>5×10⁻³</td>
</tr>
<tr>
<td>16q24.1-q24.3</td>
<td>rs454087 rs8577</td>
<td>107–123</td>
<td>83.1–88.6</td>
<td>5.5</td>
<td>-4.02</td>
<td>1.36</td>
<td>7×10⁻³</td>
</tr>
<tr>
<td>20p12.2–p11.23</td>
<td>rs742920 rs761461</td>
<td>30–41</td>
<td>11.0–18.2</td>
<td>7.25</td>
<td>-17.73</td>
<td>2.09</td>
<td>1×10⁻³</td>
</tr>
<tr>
<td>21q22.3</td>
<td>rs234705 rs235310</td>
<td>53–56</td>
<td>42.8–45.1</td>
<td>2.3</td>
<td>-25.69</td>
<td>1.35</td>
<td>6×10⁻³</td>
</tr>
</tbody>
</table>

HLOD, logarithm of odds; NPL, nonparametric linkage.

*Estimated by closest reported marker.

---

### Table 2

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>cM</th>
<th>HLOD</th>
<th>(\pi)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>169</td>
<td>1.45</td>
<td>0.29</td>
</tr>
<tr>
<td>2</td>
<td>260</td>
<td>2.90</td>
<td>0.40</td>
</tr>
<tr>
<td>3</td>
<td>129</td>
<td>1.55</td>
<td>0.29</td>
</tr>
<tr>
<td>4</td>
<td>21</td>
<td>1.64</td>
<td>0.29</td>
</tr>
<tr>
<td>5</td>
<td>145</td>
<td>1.31</td>
<td>0.25</td>
</tr>
<tr>
<td>6</td>
<td>163</td>
<td>3.55</td>
<td>0.42</td>
</tr>
<tr>
<td>7</td>
<td>172</td>
<td>1.84</td>
<td>0.33</td>
</tr>
<tr>
<td>10</td>
<td>154</td>
<td>1.78</td>
<td>0.36</td>
</tr>
<tr>
<td>13</td>
<td>104</td>
<td>1.37</td>
<td>0.29</td>
</tr>
<tr>
<td>20</td>
<td>38</td>
<td>1.25</td>
<td>0.29</td>
</tr>
</tbody>
</table>

All areas giving heterogeneity logarithm of odds (HLOD) scores >1.2 are reported.
scores obtained across the entire genome observed in the parametric analysis of the full set. Some of the regions, where the degree of “rejection” decreased in the UCS versus the FS parametric analysis, overlapped with some of those highlighted in the nonparametric analysis (tables 1 and 3).

Again, significant HLOD scores were coincident with significant or suggestive NPL scores, with HLODs of 3.37 and 3.28 for 2q37 and 6q27, respectively. As the overall number of families decreased, the proportion of linked families increased to as much as 0.5 in these regions (table 4); elsewhere, there were more modest changes. Individual familial LOD scores were obtained; it seemed from these that there was no significant enrichment or association for uncomplicated status, though the mutations may differ from those causing additional anomalies. It is also not surprising that studies of VUR in humans have indicated heterogeneity, and that there remains significant genetic heterogeneity in this VUR population, and that the inclusion of cases with additional phenotypic features cannot explain fully the observed statistics. Although the evidence for linkage in the uncomplicated set appeared to be stronger, the regions highlighted covered greater segments of the genome (the mean interval of the significant nonparametric regions was 12.9 Mb).

Additional exploratory association testing using the TDT gave no significant results after correction for multiple testing. None of the most significant association results (p<0.01) before correction were located in the regions supported by the linkage analysis.

**Table 3** Regions showing evidence or suggestive evidence of linkage in the uncomplicated dataset

<table>
<thead>
<tr>
<th>Location</th>
<th>SNP markers</th>
<th>Interval (cm)</th>
<th>Interval (Mbp)*</th>
<th>Width (Mbp)</th>
<th>LOD</th>
<th>Maximum NPL</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1q23.3–q32.2</td>
<td>rs968853 rs646125</td>
<td>161–203</td>
<td>163.4–208.6</td>
<td>41.4</td>
<td>−14.17</td>
<td>1.96</td>
<td>1×10⁻³</td>
</tr>
<tr>
<td>2q37.1–q37.3</td>
<td>rs687062 rs16747</td>
<td>243–261</td>
<td>234.4–242.4</td>
<td>8.0</td>
<td>−3.99</td>
<td>4.10</td>
<td>1×10⁻⁵</td>
</tr>
<tr>
<td>3p26.3–p26.2</td>
<td>rs1490926 rs920982</td>
<td>3–12</td>
<td>1.6–4.6</td>
<td>3.1</td>
<td>−14.28</td>
<td>1.67</td>
<td>3×10⁻⁴</td>
</tr>
<tr>
<td>3q26.31–q26.32</td>
<td>rs1565436 rs1489630</td>
<td>174–181</td>
<td>173.3–179.3</td>
<td>6.0</td>
<td>−10.39</td>
<td>1.60</td>
<td>3×10⁻⁳</td>
</tr>
<tr>
<td>6q23.3–q27</td>
<td>rs1014180 rs756519</td>
<td>137–188</td>
<td>135.4–170.7</td>
<td>35.3</td>
<td>−2.88</td>
<td>2.65</td>
<td>2×10⁻⁴</td>
</tr>
<tr>
<td>7q26.1–q36.3</td>
<td>rs1547958 rs1343750</td>
<td>164–179</td>
<td>150.3–155.2</td>
<td>4.9</td>
<td>−10.88</td>
<td>1.81</td>
<td>2×10⁻⁴</td>
</tr>
<tr>
<td>10q25.2–q26.3</td>
<td>rs1050755 rs880340</td>
<td>119–165</td>
<td>112.0–134.9</td>
<td>22.9</td>
<td>−3.89</td>
<td>2.89</td>
<td>1×10⁻⁴</td>
</tr>
<tr>
<td>13q33.2–q33.3</td>
<td>rs981900 rs1876723</td>
<td>100–107</td>
<td>105.3–107.4</td>
<td>2.1</td>
<td>−12.02</td>
<td>1.55</td>
<td>4×10⁻³</td>
</tr>
<tr>
<td>16q23.3–q24.3</td>
<td>rs967955 rs1005647</td>
<td>102–122</td>
<td>80.7–88.4</td>
<td>7.7</td>
<td>−1.38</td>
<td>1.63</td>
<td>3×10⁻³</td>
</tr>
<tr>
<td>17q24.2–q25.3</td>
<td>rs755424 rs733342</td>
<td>91–126</td>
<td>64.3–78.4</td>
<td>14.1</td>
<td>−3.31</td>
<td>2.56</td>
<td>3×10⁻⁴</td>
</tr>
<tr>
<td>20p12.2–p12.1</td>
<td>rs803980 rs727472</td>
<td>32–41</td>
<td>11.8–17.3</td>
<td>5.5</td>
<td>−12.58</td>
<td>1.66</td>
<td>3×10⁻³</td>
</tr>
<tr>
<td>21q22.3</td>
<td>rs876498 rs2256207</td>
<td>52–58</td>
<td>42.7–46.9</td>
<td>4.2</td>
<td>−18.16</td>
<td>1.73</td>
<td>2×10⁻³</td>
</tr>
</tbody>
</table>

**DISCUSSION**

This is by far the largest linkage study of VUR yet undertaken, in terms both of families and markers tested. The results from the scan revealed 10–15 regions of varying sizes and levels of support or significance across the genome. It is likely that some of these regions will not be replicated, but the smaller ones should not necessarily be rejected out of hand.

The genetics of VUR reflect the complexity of the development of the urinary tract. VUR results from poor function of the vesicoureteric valve. The development of the renal tract in the embryo begins with the formation of the ureteric bud, an outgrowth of the mesonephric (Wolffian) duct. Reciprocal signalling between the ureteric bud and the metanephrogenic mesenchyme stimulates the ureteric bud to grow to form the ureter and to branch to form the collecting ducts, and stimulates the metanephrogenic mesenchyme to form the kidney. The part of the mesenchymal duct between the urogenital sinus and the newly developed ureter is removed by apoptosis, whereupon the freed end of the ureter inserts into the bladder wall and the vesicoureteric valve is formed. The precise position at which the UB grows out from the mesonephric duct is critical for the development of the whole urinary tract and many genes are involved in the control of this process. Aberrant or multiple ureteric budding can lead not only to VUR but to a range of abnormalities of the kidneys and urinary tract. This is at least partly because some of the same genes that are involved in ureteric budding are also involved in later stages of the process. Thus, mutation of any one of many genes can cause anomalies of the ureter and/or kidney, and VUR may be found in combination with such anomalies. It is possible that mutation of several of these genes may be responsible for isolated VUR, though the mutations may differ from those causing additional anomalies.

In view of this, it is not surprising that there is not very much difference between our two sets of results (with and without the families with duplex kidneys and other anomalies), confirming our opinion that these families do not represent a different genetic entity. It is also not surprising that studies of the genetics of VUR in humans have indicated heterogeneity, and that there has been disagreement between them. Even in a study such as ours, with a large number of families, genes in which mutations are only found in a small proportion of families may not be detected at traditionally accepted significance levels. Our results yielded a number of peaks in the HLOD/NPL topography, of which 15 are listed in the tables. However, the fact that many of these peaks are well below accepted levels of credence as potential loci of causative mutations does not indicate that they should immediately be
ignored as artefacts; at least those that show agreement with results of earlier studies should be given further consideration.

Peaks in our study that coincided or overlapped with areas identified in other studies are illustrated in fig 2. A peak (NPL 1.47 in the full set, 1.55 in the uncomplicated set) in 13q33 coincides with a region identified by interstitial deletion mapping in children with renal anomalies, some of whom had VUR.26 Similarly, in a study of patients with terminal deletions of 10q26, a location at which we identified a peak NPL of 2.35 (full set) and 2.89 (uncomplicated set), all cases had

Figure 2  Comparison of regions suggestive of linkage in this study with those identified in other studies. (A) chromosome 1; (B) chromosome 2; (C) chromosome 3; (D) chromosome 10. The vertical line indicates the position of the peak in the uncomplicated set. (E) chromosome 13; (F) chromosome 20. Regions of suggestive evidence for linkage from the full set of families used in this study are shown in red; results from the uncomplicated set are shown in light green. Results from other studies are coded as follows: Feather et al, blue; Casas et al, brown; Ogata et al (renal development region), purple; Ogata et al (genital development region), maroon; Vats et al, aqua. The chromosomes are drawn to the scale of nucleotides rather than of their cytogenetic lengths.
either urinary-tract anomalies (including VUR), genital anoma-
lies or both.25

The first genome scan for VUR22 studied seven extended
families, whereas our study used 129 families, but ours were
only nuclear families, thus it was essentially an affected sibling-
pair study. The other difference was that the first study used
375 microsatellite markers, which are highly variable, so that
nearly all individuals are heterozygotes and therefore nearly all
markers are informative in every family, whereas our study
used >10 times as many markers, but they were SNPs with an
average heterozygosity of 46%, meaning that each marker will
be uninformative in some families. However, SNPs in high-
density, such as those used here, can provide greater informa-
tion content and better resolution, and may be less biased by
genotyping error rates.44 45 Nonetheless, there is good agree-
ment between the two studies, with coincident or overlapping
peaks on chromosomes 1, 3, 13 and 20 (fig 2). Furthermore,
there has been a recent report of mosaic monosomy 20 in the
renal tract of a child with bilateral VUR23 and the principal peak
of the genome scan of Conte et al42 is on 3q, although the exact
region has yet to be published.

Our region of greatest significance is in the terminal band of
chromosome 2, q37. We have searched the case reports of 77
terms of anuric deletions of 2q (mainly with breakpoints at
2q37, but a few more proximal), plus four unknarly typed close
relatives with similar phenotypes66 67 (see also Casas et al77 and
references therein). Of these 81 cases, 7 were reported to have
urinary tract anomalies,66 67 1 of whom was recorded to have
bilateral VUR,77 and 3 cases without recorded anomalies
were reported to have recurrent UTIs and therefore may also
have had VUR.68 69 As this region is gene-rich, even small
deletions tend to cause multiple anomalies, and VUR may have
been under-reported. One of the candidates at the centre of our
region is the transcription factor TWIST2, but the sequence of
this human gene is unfortunately incomplete, as there are gaps
in both the reference and Celera genome sequences in this
region. Other possible candidates are the transcription factor
HES6 and the integrin-linked kinase-associated serine/threo-
nine phosphatase, ILKAP, which might participate in the wnt
signalling pathway.

Thus we have detected in our families six regions supported
by earlier studies, on 1q, 2q 3q, 10q, 13q and 20p (fig 2).
However, others of similar significance, on chromosomes 6q,
7q, 16q and 21q, have not been reported previously in genetic
studies of VUR in humans (VUR has been reported to be linked
to the HLA region on 6p,64 but we had no peak in this region).
Yeung et al65 concluded that mild reflux affects mostly females
and a small proportion of males, whereas severe reflux with
congenital renal malformation affects almost exclusively
males, and our own data are in broad agreement.66 67 It is
therefore possible that the relative significances of different
regions might be affected by the method of ascertainment of
proband. Several studies of VUR have noted a preponderance
of males (Yeung et al65 and references therein) whereas our
series had more affected female patients. Our index cases were
ascertained through UTIs, which, after the first year of life, are
more common in female patients, and may therefore give more
emphasis to genes causing mild VUR without renal anomalies
than does ascertainment by other methods, such as detection of
hydronephrosis on prenatal ultrasound.

The former conclusion that VUR is a homogeneous dis-
order4 40 is no longer tenable; it is clearly heterogeneous. The
subgroup of 25 families with additional phenotypic features
was too small to analyse on its own. We expected that the
removal of this subgroup would cause a reduction in the NPL
scores of those regions that we had identified as containing loci
associated with renal anomalies, on 2q37, 10q26, and 13q33,
but in fact, all scores rose, particularly that on 2q, and there
were modest changes, up or down, to the NPL scores of other
regions, with the exception of a new peak appearing on 17q.
Thus, many of the genes represented by our peaks may be
associated with VUR, with or without renal and ureteric
anomalies, depending upon the mutation, hence the removal
of the subgroup had a rather random effect, by changing the
relative proportions of mutated loci left in the sample. Drastic
changes in results, such as the appearance of the 17q peak, may
have arisen due to fluctuations in the amount of evidence
supporting and/or rejecting linkage in the underlying dataset,
or due to random chance alone. Without corroborating or even
suggestive evidence from additional studies, results such as this
17q peak require replication before any further investigation is
warranted.

Our genome scan has revealed a number of large regions
(mean intervals of 9.1 and 12.9 MB in the full and uncom-
licated sets, respectively) that, when combined together, provide
a very large region for follow-up studies. In some regions, there
are good candidate genes, but in others, it is much less obvious
which gene might be the relevant one. There is almost nothing
known about many of the genes revealed by the Human
Genome Project, making it difficult to decide which genes to
investigate, so that fine mapping is necessary in some regions.
Similar genome scans on other groups of patients with VUR
should provide valuable replication of these results and help
eliminate false positive peaks. In the meantime, we shall be
investigating the likely candidates. Our total number of families
is now >200 and we have 600 DNA samples available from
healthy Irish controls. These will be useful for investigation of
which mutations are causative and which are normal poly-
morphisms, and possibly which common variants are asso-
ciated with an increased risk of VUR or associated
developmental anomalies. The recently published data on
copy-number variation in the human genome68 will also be
taken into account.

As we have said earlier, VUR results from developmental
anomaly and can have serious consequences of hypertension
and renal failure, yet in many cases it resolves spontaneously
without any damage. VUR is clearly genetically heterogeneous,
so genotyping will not provide a simple diagnostic test, but
finding the genes will be the first step in the process of
understanding the aetiology of the condition. It is to be hoped
that it may be possible to identify which mutations are
compatible with spontaneous resolution and whether any are
related to inevitable renal damage, which will help with
prognosis and genetic counselling, and in due course may lead
to new forms of treatment.

ACKNOWLEDGEMENTS
This work was funded by grants from The Children's Medical and
Research Foundation, Our Lady's Children's Hospital, Crumlin, Dublin
12, Ireland. We wish to thank Josephine Mulligan for her help with the
clinical information and Sean Ennis for helpful advice.

Authors' affiliations
H Kelly, A J Green, D E Barton, The National Centre for Medical Genetics
and University College Dublin Department of Medical Genetics, Our Lady's
Children's Hospital, Crumlin, Dublin, Ireland
C Molony, Rosetta Inpharmatics, Seattle, Washington, USA
J Darlow, A Yonedo, The National Centre for Medical Genetics and
The Children's Research Centre, Our Lady's Children's Hospital, Crumlin,
Dublin, Ireland
P Puri, M Pirker, The Children's Research Centre, Our Lady's Children's
Hospital, Crumlin, Dublin, Ireland
A Yonedo, The Children's Research Centre, Our Lady's Children's Hospital,
Crumlin, Dublin, Ireland and Department of Paediatric Surgery, Osaka
University Medical School, Osaka, Japan

www.jmedgenet.com
REFERENCES


A genome-wide scan for genes involved in primary vesicoureteric reflux

H Kelly, C M Molony, J M Darlow, M E Pirker, A Yoneda, A J Green, P Puri and D E Barton

doi: 10.1136/jmg.2007.051086

Updated information and services can be found at:
http://jmg.bmj.com/content/44/11/710

These include:

References
This article cites 67 articles, 6 of which you can access for free at:
http://jmg.bmj.com/content/44/11/710#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections
Hypertension (60)
Open access (198)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/