

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

Running title: VUR genome scan

Helena Kelly, Cliona M Molony, John M Darlow, Martina E Pirker, Akihiro Yoneda, Andrew J Green, Prem Puri, David E Barton

Correspondence to:

David E Barton

The National Centre for Medical Genetics, Our Lady's Children's Hospital Crumlin, Dublin 12, Ireland

david.barton@olhsc.ie

Tel +353 1 409 6749

Fax +353 1 456 0953

Authors' affiliations

Helena Kelly, John Darlow, The Children's Research Centre and The National Centre for Medical Genetics, Our Lady's Children's Hospital Crumlin, Dublin 12, Ireland

Cliona Molony, Rosetta Inpharmatics, 401 Terry Avenue N., Seattle, Washington 98109, U.S.A.

Martina Pirker, The Children's Research Centre, Our Lady's Children's Hospital Crumlin, Dublin 12, Ireland

Akihiro Yoneda, The Children's Research Centre, Our Lady's Children's Hospital Crumlin, Dublin 12, Ireland, and Department of Paediatric Surgery, Osaka University Medical School, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan

Andrew J Green, David E Barton, The National Centre for Medical Genetics, Our Lady's Children's Hospital Crumlin, Dublin 12, Ireland, and Department of Medical Genetics and UCD Conway Institute for Biomolecular and Biomedical Research, National University of Ireland, Dublin

Prem Puri, The Children's Research Centre, Our Lady's Children's Hospital Crumlin, Dublin 12, Ireland, and National Children's Hospital, Tallaght, Dublin 24, Ireland

Key words: vesico-ureteral reflux (= the MeSH term), genome scan (No MeSH term exists for this.)

Word count: 3,876

Background: Vesicoureteric Reflux (VUR) is the retrograde flow of urine from the bladder into the ureters. It is the most common urological anomaly in children and a major cause of end stage renal failure and hypertension in both children and adults. VUR is seen in approximately 1-2% of newborn Caucasians and is frequently familial.

Aim and methods: In order to search for genetic loci involved in VUR, we have performed a genome-wide linkage scan using 4710 single nucleotide polymorphisms (SNPs) in 609 individuals from 129 Irish families with more than one member affected.

Results: Nonparametric analysis of the data set yielded moderately suggestive linkage at chromosome 2q37 ($NPL_{max} = 2.67$, $p \sim 0.0002$), while analysis of a subset without any additional features, such as duplex kidneys, yielded a maximum NPL score of 4.1 ($p = 0.00001$), reaching levels of genome-wide statistical significance. Suggestive linkage was also seen at 10q26 and 6q27, and there were several smaller peaks.

Conclusion: Our results confirm the previous conclusion that VUR is genetically heterogeneous, and lend support to several disease-associated regions indicated by smaller studies, as well as indicating new regions of interest for investigation.

Vesicoureteric reflux (vesicoureteral reflux, MIM 193000) 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 newborn Caucasians¹ and is a major cause of end stage renal failure and hypertension in both children and adults.^{1, 2, 3} 30-50% of children presenting with urinary tract infections (UTI) have VUR,⁴ and the reflux of infected urine can cause pyelonephritis 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 fulfillment of a proper valvular mechanism.^{1, 3} VUR occurs frequently in families; approximately half of the offspring will be affected but the mode of inheritance is unknown.¹ Micturating cystourethrography (MCUG) in children may be performed because of UTI or suspicion of renal tract anomalies from prenatal ultrasound, and in either case both VUR and renal tract anomalies, including duplex renal systems, are commonly found^{5, 6}. Furthermore, VUR is more common in those with duplex than with single renal tracts, and duplex kidneys are commonly found in those with VUR^{6, 7}. In both Mouse and Man, individuals with different 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^{3, 8, 9}.

The only genes so far identified to be mutated in VUR have been in cases with other anomalies, or in families in which other anomalies were segregating. Hu *et al.* observed VUR amongst other features in mice in which the gene for uroplakin III (locus *UPK3A*) had been homozygously inactivated.¹⁰ Linkage¹¹ and mutation screening studies¹² of *UPK3A* and other members of the uroplakin family^{13, 14} (Kelly, H., Puri, P., Barton, D.E., 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 a/hypo/dysplasia, including some with VUR.^{15, 16} The *PAX2* gene has been shown to be mutated in the renal-coloboma syndrome, which includes VUR as part of a complex phenotype,^{17, 18} and in renal hypoplasia and VUR without eye defects¹⁹, but has not been found to be mutated in uncomplicated VUR.²⁰ Investigation of a patient with multiple congenital anomalies, including severe bilateral VUR, found a Y;3 translocation which disrupted *ROBO2*, and investigation of this gene in 124 families with VUR with potential autosomal dominance (AD) revealed mutations in 2 families.²¹ 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 AD, but mutations in these genes account for a very small proportion of VUR in general.

Attempts at mapping other genes for VUR have been genome-wide scans of small numbers of multi-generation families,^{22, 23} the latter yet to be reported in full, or individual or small collections of cases with chromosomal rearrangements^{24, 25, 26, 27} though 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.^{28, 29, 30}

In this paper we present the results of a genome-wide linkage scan with 4710 SNP markers and 129 families each with two or more 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, Tallaght, both in Dublin, Ireland. Ethical approval was granted, and informed consent was obtained from all subjects and/or their parents. Families

with 2 or more affected sibs 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. Sibs 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 affection 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 ureterocele³¹, but some authors include complete duplex renal systems as a cause³², and some count some cases with duplex systems as primary and some as secondary.³³ 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 ureterocele. 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 600Kb.

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 of unknown affection status), 283 affected siblings (117 male and 166 female), 37 unaffected siblings (17 male and 20 female), 32 siblings with unknown affection status (14 male and 18 female) and 2 grandparents (one affected and one of unknown status). There were 107 families with two genotyped affected children, 20 with three, one with four and one with five, giving 183 affected sib-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 (*i.e.* removing obvious genotyping errors) with Pedstats³⁴, 4,710 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%.

Analysis

Both non-parametric and parametric linkage analysis for the scan was performed using MERLIN v.1.0.1.10.3.^{34, 35} NPL scores were based on a procedure for converting scores for individual inheritance vectors into Z-scores for a single or multiple pedigrees³⁶ which are used to construct a likelihood ratio test for linkage and define a (NPL) LOD score statistic.³⁷ The data set was prepared for analysis using MEGA2³⁸ and markers in tight LD ($D' > 0.7$) were removed from the data set prior to analysis using SNPLINK.³⁹ The resultant panel consisted of 3,861 markers. Though 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 the VUR cases had a duplex system) but, in other families, fused kidneys, renal hypoplasia, large ureterocele, 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 and 480 individuals.

A genetic model of inheritance for VUR has been proposed both in segregation studies⁴⁰ and in studies of large multi-generational families.^{22, 29} 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.⁴⁰ Power for the complete cohort is estimated to be ~79% ($\alpha = 0.05$), given the genetic model and a common test marker allele frequency and $\theta = 0.1$. This is likely to be an over-estimate in general, as power will have fluctuated substantially based on marker- and multimarker-informativeness, θ increasing, or the sample size reduction, as in the Uncomplicated cohort setting. For parametric analysis, heterogeneity was investigated using HLOD statistic. Abreu *et al*⁴¹ 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 guidelines⁴²: $p < 0.0007$ and a LOD score of 2.2 (NPL 3.18) was suggestive of linkage, and $p < 0.00002$ and a LOD score of 3.6 (NPL 4.08) was considered significant linkage.⁴² All regions with a p-value of ≤ 0.02 (NPL ~0.9), in the nonparametric analysis are reported and used to define linkage peak intervals. Regions with NPL > 2.00 are 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 QTDT.⁴³

RESULTS

Tables 1 and 2 show the results of the genome-wide linkage analysis based on the full-set of 129 families (FS). 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).

Table 1 Linkage Results: Regions showing suggestive evidence of linkage in the full data set.

Location	SNP markers		Interval (cM)	Interval (Mbp) ^a	Width (Mbp)	LOD	max NPL	p-value
1q23.2-q25.2	rs1053074	rs1923626	152-175	158.3-175.3	17.05	-18.10	2.10	0.0001
2q37.2-q37.3	rs869214	rs16747	246-261	235.6-242.4	6.79	-15.33	2.67	0.0002
3p13-q21.2	rs7064	rs713161	96-131	126.3-127.4	1.1	-17.36	1.36	0.006
3q26.31	rs2046718	rs753293	177-178.	175.0-176.6	1.6	-20.8	0.98	0.02
4p16.1 – p15.31	rs1981635	rs729918	18-30	10.2-20.1	9.9	-19.77	1.33	0.007
6q24.1-q27	rs1931992	rs756519	146-188	142.6-170.7	28.1	-9.16	2.55	0.0003
7q36.1-q36.3	rs1547958	rs1343750	165-179	150.3-155.2	4.94	-14.94	2.19	0.0007
10q25.2-q26.3	rs1327551	rs880340	120-165	112.8-134.9	22.1	-11.38	2.35	0.0005
13q33.2-q33.3	rs981900	rs1876723	100-108	105.3-107.4	2.12	-15.17	1.47	0.005
16q24.1-q24.3	rs454087	rs8577	107-123	83.1-88.6	5.5	-4.02	1.36	0.007
20p12.2-p11.23	rs742920	rs761461	30-41	11.0-18.2	7.25	-17.73	2.09	0.001
21q22.3	rs234705	rs235310	53-56	42.8-45.1	2.3	-25.69	1.35	0.006

^a Estimated by closest reported marker.

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 ten regions across the genome (table 2). The proportion of linked families (α) was estimated as

Table 2 HLOD scores for complete data set: All areas giving HLOD scores over 1.2 are reported.

Chromosome	cM	HLOD	α
1	169	1.45	0.29
2	260	2.90	0.40
3	129	1.55	0.29
4	21	1.64	0.29
5	145	1.31	0.25
6	163	3.55	0.42
7	172	1.84	0.33
10	154	1.78	0.36
13	104	1.37	0.29
20	38	1.25	0.29

approximately 1/3 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 noted above. Nonparametric linkage analysis yielded six regions that are potentially interesting candidate regions based on their NPL scores under standard criteria⁴² (table 1 and fig 1A). All of these regions were coincident with regions that exhibited evidence of heterogeneity (HLOD > 1.2) under the parametric analysis (see table 2). In addition to the six regions mentioned above, six additional regions in the

nonparametric scan provided NPL scores which surpassed the less stringent threshold of p-value < 0.02, three of which were also coincident with regions showing heterogeneity (tables 1 and 2).

The subset of samples, the uncomplicated set (UCS), defined by excluding the 25 families in which some affected members had additional phenotypic features, was analyzed 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. Nonparametric analysis yielded a maximum NPL score of NPL 4.1 (p = 0.00001) at 2q37, reaching levels of genome-wide statistical significance,⁴² others

potentially interesting at 10q26 and 6q27, as well as nine additional candidate regions on eight different chromosomes with results having p-values < 0.02 (table 3 and fig 1B), the most significant of these, however (on 17q) was neither supported by the parametric analysis (table 4) nor seen at all in the full set.

Table 3 Regions showing evidence or suggestive evidence of linkage in the uncomplicated data set.

Location	SNP markers		Interval (cM)	Interval (Mbp) ^a	Width (Mbp)	LOD	max NPL	p-value
1q23.3-q32.2	rs968853	rs946125	161-203	163.4-208.6	41.4	-14.17	1.96	0.001
2q37.1-q37.3	rs887062	rs16747	243-261	234.4-242.4	8.0	-3.99	4.10	<0.00001
3p26.3-p26.2	rs1499260	rs902982	3-12	1.6-4.6	3.1	-14.28	1.67	0.003
3q26.31-q26.32	rs1563436	rs1489630	174-181	173.3-179.3	6.0	-10.39	1.60	0.0031
6q23.3-q27	rs1041480	rs756519	137-188	135.4-170.7	35.3	-2.88	2.65	0.0002
7q36.1-q36.3	rs1547958	rs1343750	164-179	150.3-155.2	4.9	-10.88	1.81	0.002
10q25.2-q26.3	rs1050755	rs880340	119-165	112.0-134.9	22.9	-3.89	2.89	0.00013
13q33.2-q33.3	rs981900	rs1876723	100-107	105.3-107.4	2.1	-12.02	1.55	0.004
16q23.3-q24.3	rs967955	rs1006547	102-122	80.7-88.4	7.7	-1.38	1.63	0.003
17q24.2-q25.3	rs755424	rs733342	91-126	64.3-78.4	14.1	-3.31	2.56	0.0003
20p12.2-p12.1	rs803880	rs727472	32-41	11.8-17.3	5.5	-12.58	1.66	0.003
21q22.3	rs876498	rs2256207	52-58	42.7-46.9	4.2	-18.16	1.73	0.002

^a See Table 1

Under the proposed genetic model, parametric analysis of the UCS provided a number of regions with just moderately negative LOD scores (not shown), in contrast to the strongly negative LOD scores obtained across the entire genome observed in the FS parametric analysis. Two of the regions, where the degree of 'rejection' decreased in the UCS versus the FS parametric analysis, overlapped with two of those highlighted in the nonparametric analysis, 2q37 and 6q27.

Table 4 HLOD scores for reduced data set

Chromosome	cM	HLOD	α
1	168	1.30	0.31
2	260	3.37	0.50
3	177	1.40	0.34
6	163	3.28	0.45
7	172	1.37	0.34
10	154	2.19	0.48
13	23.2	1.32	0.30
20	38	1.36	0.32

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), while there were more modest changes elsewhere. Individual familial LOD scores were obtained and it appeared that there was no significant enrichment or association for uncomplicated status and individual family contributions to linkage (familial LOD score > 0), Fisher's exact test $p > 0.05$ in both of these regions. Thus, any increase in linkage signal or change in support for a particular region was likely to be a reflection of the changes in study population composition, combined with the removal of families that, by chance, may have previously rejected linkage strongly in these regions. It appears 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.

While the evidence for linkage in the UCS appeared to be stronger, the regions highlighted covered greater segments of the genome. The mean interval of the significant nonparametric regions in the UCS 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.

DISCUSSION

This was by far the largest such study yet undertaken, both in terms 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 (UB), an outgrowth of the mesonephric (Wolffian) duct.^{8,9} Reciprocal signaling between the UB and the metanephrogenic mesenchyme stimulates the UB to grow to form the ureter and branch to form the collecting ducts, and stimulates the metanephrogenic mesenchyme to form the kidney. The part of the mesonephric 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 a large number of genes is 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 from the above that studies of the genetics of VUR in humans have indicated heterogeneity^{22, 23, 25, 26, 29, 30}, and that there has been disagreement between them. Even in a study with a large number of families, such as ours, genes whose mutations only account for 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 fifteen are listed in our 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 artifacts. At least those that show agreement with results of earlier studies should be given further consideration.

Peaks in our study which coincided or overlapped with areas identified in other studies are illustrated in fig 2. A peak (NPL 1.47 in the FS, 1.55 UCS) 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, where we have identified a peak NPL of 2.35 (FS), 2.89 (UCS), all cases had either urinary tract anomalies (including VUR), genital anomalies or both.²⁵

The first genome scan for VUR²² studied seven extended families, whereas our study used 129 families, but they were only nuclear families, and it was essentially an affected sib-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 more than ten 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 information content, provide better resolution and may be less biased due to genotyping error rates.^{44, 45} Nonetheless there is good agreement 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 VUR²⁷ and the principal peak of the genome scan of Conte et al.²³ is on 3q, though 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 cases of terminal deletions of 2q (mainly with breakpoints at 2q37, but a few more proximal), plus four unkaryotyped close relatives with similar phenotypes.^{46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57 (and refs in 57)} Of these 81 cases, seven are reported to have had urinary tract anomalies,^{48, 57, 58, 59, 60, 61, 62} one of whom was recorded to have bilateral VUR,⁵⁹ and three cases without recorded anomalies were reported to have recurrent UTIs and therefore may also have had VUR.^{59, 63} 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/threonine phosphatase, ILKAP, which may participate in the wnt signaling 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 before in genetic studies of VUR in humans. (VUR has been reported to be linked to the HLA region on 6p⁶⁴, but we had no peak in this region.)

Yeung et al.,⁶⁵ concluded that mild reflux affects mostly females and a small proportion of males, and 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 probands. Several studies of VUR have noted a preponderance of males (⁶⁵ and refs therein) whereas our series had more affected females. Our index cases were ascertained through UTIs, which, after the first year of life, are more common in females, and may therefore give more emphasis to genes causing mild VUR without renal anomalies than ascertainment by other methods, such as detection of hydronephrosis on prenatal ultrasound.

The former conclusion that VUR is a homogeneous disorder^{3, 40} is no longer tenable; it is clearly heterogeneous. The sub-group of 25 families with additional phenotypic features was too small to analyze on its own. It was to be expected that the removal of this sub-group 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. 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. It may therefore be that many of the genes represented by our peaks may be associated with VUR, with or without renal and ureteric anomalies, depending upon the mutation, and that the removal of the sub-group had its effect rather randomly, 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 data set, or due to random chance alone. Without corroborating or even suggestive evidence from additional studies, results such as this 17q peak, would 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 FS and UCS respectively) which, 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 hard to decide which genes to investigate, so that fine-mapping is necessary in some regions. Similar genome scans on other groups of VUR patients should provide valuable replication of these results and help eliminate false positive peaks. In the meantime we shall be investigating the candidates that are already likely. Our total number of families is now over 200, and we have available 600 DNA samples from healthy Irish controls. These will be useful for investigation of which mutations are causative and which are normal polymorphisms, and possibly which common variants are associated 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 a simple diagnostic test is not a possibility, 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 counseling, 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.

Competing interests: None.

Figure legends

Figure 1 Layout plots of NPL scores from the genome scan. A. Full data set. B. Uncomplicated data set. The vertical axis shows NPL scores. Plots are scaled vertically to show the maximum peak on the full data set.

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.

REFERENCES

- 1 **Eccles MR**, Jacobs GH. The genetics of primary vesico-ureteric reflux. *Ann Acad Med Singapore* 2000;**29**(3):337-45.
- 2 **Eccles MR**, Bailey RR, Abbott GD, Sullivan MJ. Unravelling the genetics of vesicoureteric reflux: a common familial disorder. *Hum Mol Genet* 1996;**5 Spec No**:1425-9.
- 3 **Devriendt K**, Groenen P, Van Esch H, van Dijck M, Van de Ven W, Fryns JP, Proesmans W. Vesico-ureteral reflux: a genetic condition? *Eur J Pediatr* 1998;**157**(4):265-71.
- 4 **Chertin B**, Puri P. Familial vesicoureteral reflux. *J Urol* 2003;**169**(5):1804-8.
- 5 **Krzemien G**, Roszkowska-Blaim M, Kostro I, Szmigielska A, Madzik J. [Vesicoureteral reflux in children with prenatal suspicion of urinary tract abnormalities]. *Przegl Lek* 2006;**63 Suppl 3**:127-30.
- 6 **Siomou E**, Papadopoulou F, Kollios KD, Photopoulos A, Evagelidou E, Androulakis P, Siamopoulou A. Duplex collecting system diagnosed during the first 6 years of life after a first urinary tract infection: a study of 63 children. *J Urol* 2006;**175**(2):678-81; discussion 81-2.
- 7 **Seseke F**, Strauss A, Seseke S, Zappel H, Ringert RH, Zoller G. [Long-term experience with Cohen ureteral reimplantation in bilateral vesicoureteral reflux in childhood]. *Urologe A* 2006;**45**(7):852-7.
- 8 **Ichikawa I**, Kuwayama F, Pope JCT, Stephens FD, Miyazaki Y. Paradigm shift from classic anatomic theories to contemporary cell biological views of CAKUT. *Kidney Int* 2002;**61**(3):889-98.
- 9 **Murawski IJ**, Gupta IR. Vesicoureteric reflux and renal malformations: a developmental problem. *Clin Genet* 2006;**69**(2):105-17.
- 10 **Hu P**, Deng FM, Liang FX, Hu CM, Auerbach AB, Shapiro E, Wu XR, Kachar B, Sun TT. Ablation of uroplakin III gene results in small urothelial plaques, urothelial leakage, and vesicoureteral reflux. *J Cell Biol* 2000;**151**(5):961-72.
- 11 **Kelly H**, Ennis S, Yoneda A, Bermingham C, Shields DC, Molony C, Green AJ, Puri P, Barton DE. Uroplakin III is not a major candidate gene for primary vesicoureteral reflux. *Eur J Hum Genet* 2005;**13**(4):500-2.
- 12 **Giltay JC**, van de Meerakker J, van Amstel HK, de Jong TP. No pathogenic mutations in the uroplakin III gene of 25 patients with primary vesicoureteral reflux. *J Urol* 2004;**171**(2 Pt 1):931-2.
- 13 **Jiang S**, Gitlin J, Deng FM, Liang FX, Lee A, Atala A, Bauer SB, Ehrlich GD, Feather SA, Goldberg JD, Goodship JA, Goodship TH, Hermanns M, Hu FZ, Jones KE, Malcolm S, Mendelsohn C, Preston RA, Retik AB, Schneck FX, Wright V, Ye XY, Woolf AS, Wu XR, Ostrer H, Shapiro E, Yu J, Sun TT. Lack of major involvement of human uroplakin genes in vesicoureteral reflux: implications for disease heterogeneity. *Kidney Int* 2004;**66**(1):10-9.
- 14 **Jenkins D**, Bitner-Glindzicz M, Malcolm S, Allison J, de Bruyn R, Flanagan S, Thomas DF, Belk RA, Feather SA, Bingham C, Southgate J, Woolf AS. Mutation analyses of Uroplakin II in children with renal tract malformations. *Nephrol Dial Transplant* 2006;**21**(12):3415-21.
- 15 **Jenkins D**, Bitner-Glindzicz M, Malcolm S, Hu CC, Allison J, Winyard PJ, Gullett AM, Thomas DF, Belk RA, Feather SA, Sun TT, Woolf AS. De novo Uroplakin IIIa heterozygous mutations cause human renal adysplasia leading to severe kidney failure. *J Am Soc Nephrol* 2005;**16**(7):2141-9.

- 16 **Schonfelder EM**, Knuppel T, Tasic V, Miljkovic P, Konrad M, Wuhl E, Antignac C, Bakkaloglu A, Schaefer F, Weber S. Mutations in Uroplakin IIIA are a rare cause of renal hypodysplasia in humans. *Am J Kidney Dis* 2006;**47**(6):1004-12.
- 17 **Sanyanusin P**, McNoe LA, Sullivan MJ, Weaver RG, Eccles MR. Mutation of PAX2 in two siblings with renal-coloboma syndrome. *Hum Mol Genet* 1995;**4**(11):2183-4.
- 18 **Sanyanusin P**, Schimmenti LA, McNoe LA, Ward TA, Pierpont ME, Sullivan MJ, Dobyns WB, Eccles MR. Mutation of the PAX2 gene in a family with optic nerve colobomas, renal anomalies and vesicoureteral reflux. *Nat Genet* 1995;**9**(4):358-64.
- 19 **Nishimoto K**, Iijima K, Shirakawa T, Kitagawa K, Satomura K, Nakamura H, Yoshikawa N. PAX2 gene mutation in a family with isolated renal hypoplasia. *J Am Soc Nephrol* 2001;**12**(8):1769-72.
- 20 **Choi KL**, McNoe LA, French MC, Guilford PJ, Eccles MR. Absence of PAX2 gene mutations in patients with primary familial vesicoureteric reflux. *J Med Genet* 1998;**35**(4):338-9.
- 21 **Lu W**, van Eerde AM, Fan X, Quintero-Rivera F, Kulkarni S, Ferguson H, Kim HG, Fan Y, Xi Q, Li QG, Sanlaville D, Andrews W, Sundaresan V, Bi W, Yan J, Giltay JC, Wijmenga C, de Jong TP, Feather SA, Woolf AS, Rao Y, Lupski JR, Eccles MR, Quade BJ, Gusella JF, Morton CC, Maas RL. Disruption of ROBO2 Is Associated with Urinary Tract Anomalies and Confers Risk of Vesicoureteral Reflux. *Am J Hum Genet* 2007;**80**(4):616-32.
- 22 **Feather SA**, Malcolm S, Woolf AS, Wright V, Blaydon D, Reid CJ, Flinter FA, Proesmans W, Devriendt K, Carter J, Warwicker P, Goodship TH, Goodship JA. Primary, nonsyndromic vesicoureteric reflux and its nephropathy is genetically heterogeneous, with a locus on chromosome 1. *Am J Hum Genet* 2000;**66**(4):1420-5.
- 23 **Conte ML**, Bertoli-Avella AM, de Graaf BM, Lama G, La Manna A, Rambaldi PF, Oostra BA, Perrotta S. A novel locus for primary familial vesicoureteral reflux maps to chromosome 3q. *Eur J Hum Genet* 2006;**14 Suppl 1**:335.
- 24 **Groenen PM**, Vanderlinden G, Devriendt K, Fryns JP, Van de Ven WJ. Rearrangement of the human CDC5L gene by a t(6;19)(p21;q13.1) in a patient with multicystic renal dysplasia. *Genomics* 1998;**49**(2):218-29.
- 25 **Ogata T**, Muroya K, Sasagawa I, Kosho T, Wakui K, Sakazume S, Ito K, Matsuo N, Ohashi H, Nagai T. Genetic evidence for a novel gene(s) involved in urogenital development on 10q26. *Kidney Int* 2000;**58**(6):2281-90.
- 26 **Vats KR**, Ishwad C, Singla I, Vats A, Ferrell R, Ellis D, Moritz M, Surti U, Jayakar P, Frederick DR, Vats AN. A locus for renal malformations including vesico-ureteric reflux on chromosome 13q33-34. *J Am Soc Nephrol* 2006;**17**(4):1158-67.
- 27 **Stefanou EG**, Crocker M, Boon A, Stewart H. Cryptic mosaicism for monosomy 20 identified in renal tract cells. *Clin Genet* 2006;**70**(3):228-32.
- 28 **Yoneda A**, Cascio S, Green A, Barton D, Puri P. Angiotensin II type 2 receptor gene is not responsible for familial vesicoureteral reflux. *J Urol* 2002;**168**(3):1138-41.
- 29 **Sanna-Cherchi S**, Reese A, Hensle T, Caridi G, Izzi C, Kim YY, Konka A, Murer L, Scolari F, Ravazzolo R, Ghiggeri GM, Gharavi AG. Familial Vesicoureteral Reflux: Testing Replication of Linkage in Seven New Multigenerational Kindreds. *J Am Soc Nephrol* 2005:ASN.2004121034.
- 30 **van Eerde AM**, Koelman BPC, van de Kamp JM, de Jong TPVM, Wijmenga C, Giltay JC. Linkage study in four large Dutch families with VUR: exclusion of a previously reported locus and (separate) candidate genes. *Eur J Hum Genet* 2006;**14 Suppl 1**:334.
- 31 **Decter RM**. Update on vesicoureteral reflux: pathogenesis, nephropathy, and management. *Rev Urol* 2001;**3**(4):172-8.

- 32 **Greenbaum LA**, Mesrobian HG. Vesicoureteral reflux. *Pediatr Clin North Am* 2006;**53**(3):413-27, vi.
- 33 **Orellana P**, Pizarro ME, Garcia F, Baquedano P. Daño renal en reflujo vesicoureteral asociado a doble sistema pieloureteral [Renal damage in vesicoureteric reflux associated with duplex pyeloureteric systems]. *Rev Esp Med Nucl* 2005;**24**(6):387-91.
- 34 **Abecasis GR**, Wigginton JE. Handling marker-marker linkage disequilibrium: pedigree analysis with clustered markers. *Am J Hum Genet* 2005;**77**(5):754-67.
- 35 **Abecasis GR**, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002;**30**(1):97-101.
- 36 **Whittemore AS**, Halpern J. A class of tests for linkage using affected pedigree members. *Biometrics* 1994;**50**(1):118-27.
- 37 **Kong A**, Cox NJ. Allele-sharing models: LOD scores and accurate linkage tests. *Am J Hum Genet* 1997;**61**(5):1179-88.
- 38 **Mukhopadhyay N**, Almasy L, Schroeder M, Mulvihill WP, Weeks DE. Mega2: data-handling for facilitating genetic linkage and association analyses. *Bioinformatics* 2005;**21**(10):2556-7.
- 39 **Webb EL**, Sellick GS, Houlston RS. SNPLINK: multipoint linkage analysis of densely distributed SNP data incorporating automated linkage disequilibrium removal. *Bioinformatics* 2005;**21**(13):3060-1.
- 40 **Chapman CJ**, Bailey RR, Janus ED, Abbott GD, Lynn KL. Vesicoureteric reflux: segregation analysis. *Am J Med Genet* 1985;**20**(4):577-84.
- 41 **Abreu PC**, Hodge SE, Greenberg DA. Quantification of type I error probabilities for heterogeneity LOD scores. *Genet Epidemiol* 2002;**22**(2):156-69.
- 42 **Lander E**, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature Genetics* 1995;**11**:241-7.
- 43 **Abecasis GR**, Cookson WO, Cardon LR. Pedigree tests of transmission disequilibrium. *Eur J Hum Genet* 2000;**8**(7):545-51.
- 44 **Nsengimana J**, Renard H, Goldgar D. Linkage analysis of complex diseases using microsatellites and single-nucleotide polymorphisms: application to alcoholism. *BMC Genet* 2005;**6 Suppl 1**:S10.
- 45 **Schaid DJ**, Guenther JC, Christensen GB, Hebring S, Rosenow C, Hilker CA, McDonnell SK, Cunningham JM, Slager SL, Blute ML, Thibodeau SN. Comparison of microsatellites versus single-nucleotide polymorphisms in a genome linkage screen for prostate cancer-susceptibility loci. *Am J Hum Genet* 2004;**75**(6):948-65.
- 46 **Alkuraya FS**, Kimonis VE, Holt L, Murata-Collins JL. A patient with a ring chromosome 2 and microdeletion of 2q detected using FISH: Further support for "ring chromosome 2 syndrome". *Am J Med Genet A* 2005;**132**(4):447-9.
- 47 **Armstrong L**, Allanson JE, Weaver DD, Bevan CJ, Hobart HH. Unrelated patients with a rearrangement of chromosome 2 causing duplication of 2p23 and deletion of 2q37. *Am J Med Genet A* 2005;**134**(3):299-304.
- 48 **Bacino CA**, Kashork CD, Davino NA, Shaffer LG. Detection of a cryptic translocation in a family with mental retardation using FISH and telomere region-specific probes. *Am J Med Genet* 2000;**92**(4):250-5.
- 49 **Bruyere H**, Rajcan-Separovic E, Doyle J, Pantzar T, Langlois S. Familial cryptic translocation (2;17) ascertained through recurrent spontaneous abortions. *Am J Med Genet A* 2003;**123**(3):285-9.
- 50 **Caliskan MO**, Karauzum SB, Mihci E, Tacoy S, Luleci G. Subtelomeric chromosomal rearrangements detected in patients with idiopathic mental retardation and dysmorphic features. *Genet Couns* 2005;**16**(2):129-38.

- 51 **Chassaing N**, De Mas P, Tauber M, Vincent MC, Julia S, Bourrouillou G, Calvas P, Bieth E. Molecular characterization of a cryptic 2q37 deletion in a patient with Albright hereditary osteodystrophy-like phenotype. *Am J Med Genet A* 2004;**128**(4):410-3.
- 52 **Giardino D**, Finelli P, Gottardi G, Clerici D, Mosca F, Briscioli V, Larizza L. Cryptic subtelomeric translocation t(2;16)(q37;q24) segregating in a family with unexplained stillbirths and a dysmorphic, slightly retarded child. *Eur J Hum Genet* 2001;**9**(12):881-6.
- 53 **Giardino D**, Finelli P, Gottardi G, De Canal G, Della Monica M, Lonardo F, Scarano G, Larizza L. Narrowing the candidate region of Albright hereditary osteodystrophy-like syndrome by deletion mapping in a patient with an unbalanced cryptic translocation t(2;6)(q37.3;q26). *Am J Med Genet A* 2003;**122**(3):261-5.
- 54 **Lukusa T**, Smeets E, Vogels A, Vermeesch JR, Fryns JP. Terminal 2q37 deletion and autistic behaviour. *Genet Couns* 2005;**16**(2):179-80.
- 55 **Lukusa T**, Vermeesch JR, Holvoet M, Fryns JP, Devriendt K. Deletion 2q37.3 and autism: molecular cytogenetic mapping of the candidate region for autistic disorder. *Genet Couns* 2004;**15**(3):293-301.
- 56 **Shrimpton AE**, Braddock BR, Thomson LL, Stein CK, Hoo JJ. Molecular delineation of deletions on 2q37.3 in three cases with an Albright hereditary osteodystrophy-like phenotype. *Clin Genet* 2004;**66**(6):537-44.
- 57 **Casas KA**, Mononen TK, Mikail CN, Hased SJ, Li S, Mulvihill JJ, Lin HJ, Falk RE. Chromosome 2q terminal deletion: report of 6 new patients and review of phenotype-breakpoint correlations in 66 individuals. *Am J Med Genet A* 2004;**130**(4):331-9.
- 58 **Wang TH**, Johnston K, Hsieh CL, Denney PA. Terminal deletion of the long arm of chromosome 2 in a premature infant with karyotype: 46,XY,del(2)(q37). *Am J Med Genet* 1994;**49**(4):399-401.
- 59 **Batstone PJ**, Simpson S, Bonthron DT, Keng WT, Hamilton D, Forsyth L, Sales M, Pratt N, Goudie D. Effective monosomy or trisomy of chromosome band 2q37.3 due to the unbalanced segregation of a 2;11 translocation. *Am J Med Genet A* 2003;**118**(3):241-6.
- 60 **Viot-Szoboszlai G**, Amiel J, Doz F, Prieur M, Couturier J, Zucker JN, Henry I, Munnich A, Vekemans M, Lyonnet S. Wilms' tumor and gonadal dysgenesis in a child with the 2q37.1 deletion syndrome. *Clin Genet* 1998;**53**(4):278-80.
- 61 **Reddy KS**, Flannery D, Farrer RJ. Microdeletion of chromosome sub-band 2q37.3 in two patients with abnormal situs viscerum. *Am J Med Genet* 1999;**84**(5):460-8.
- 62 **Waters BL**, Allen EF, Gibson PC, Johnston T. Autopsy findings in a severely affected infant with a 2q terminal deletion. *Am J Med Genet* 1993;**47**(7):1099-103.
- 63 **Bijlsma EK**, Aalfs CM, Sluitjer S, Oude Luttikhuis ME, Trembath RC, Hoovers JM, Hennekam RC. Familial cryptic translocation between chromosomes 2qter and 8qter: further delineation of the Albright hereditary osteodystrophy-like phenotype. *J Med Genet* 1999;**36**(8):604-9.
- 64 **Macintosh M**, Almarhoos G, Heath DA. HLA linkage with familial vesicoureteral reflux and pelviureteric junction obstruction. *Tissue Antigens* 1989;**34**:185-9.
- 65 **Yeung CK**, Godley ML, Dhillon HK, Gordon I, Duffy PG, Ransley PG. The characteristics of primary vesico-ureteric reflux in male and female infants with pre-natal hydronephrosis. *Br J Urol* 1997;**80**(2):319-27.
- 66 **Pirker ME**, Colhoun E, Puri P. Renal scarring in familial vesicoureteral reflux: Is prevention possible? *J Urol* 2006;**176**(4S):1842-6.
- 67 **Pirker ME**, Mohanan N, Colhoun E, Barton D, Green A, Puri P. Familial vesicoureteral reflux: Influence of sex on prevalence and expression. *J Urol* 2006;**176**(4S):1776-80.
- 68 **Redon R**, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R,

Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Zhang J, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME. Global variation in copy number in the human genome. *Nature* 2006;**444**(7118):444-54.

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence on a worldwide basis to the BMJ Publishing Group Ltd to permit this article (if accepted) to be published in Journal of Medical Genetics and any other BMJ PGL products and sublicences such use and exploit all subsidiary rights, as set out in the Journal's licence (<http://JMG.bmjournals.com/misc/ifora/licenceform.shtml>).

Figure 1

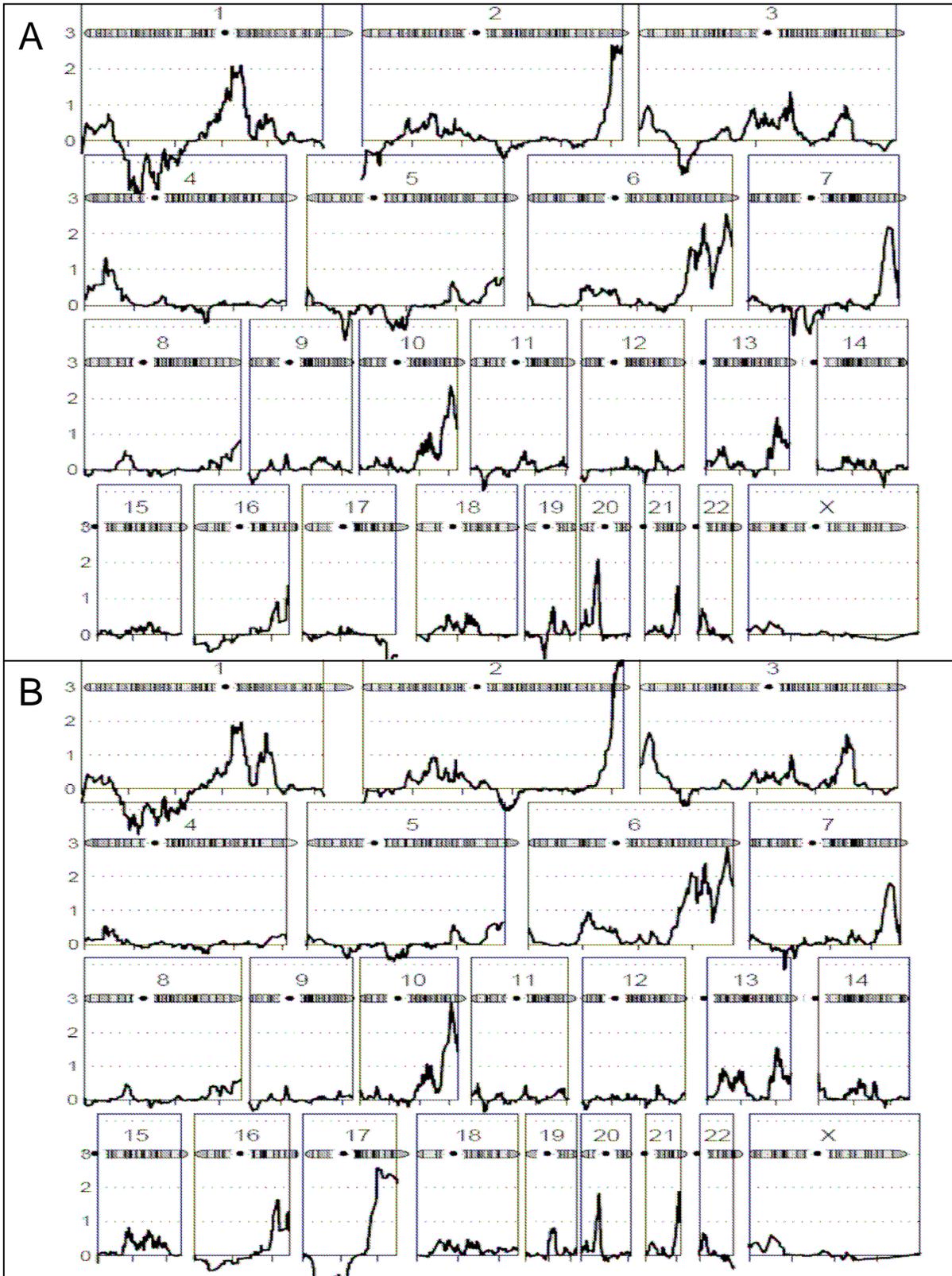


Figure 2

