Am. J. Hum. Genet. 68:1119–1129, 2001

Identification of a New Candidate Locus for Uric Acid Nephrolithiasis

Maria Neve Ombra,[1,*] Paola Forabosco,[1,*] Stefania Casula,[1] Andrea Angius,[1] Gianbattista Maestrale,[1] Enrico Petretto,[1] Giuseppina Casu,[1] Giacomo Colussi,[2] Enzo Usai,[3] Paola Melis,[1] and Mario Pirastu[1]

[1]Institute of Molecular Genetics, CNR, Alghero, Italy; [2]Niguarda-Cà Granda Hospital, Milan; [3]Clinic of Urology, SS. Trinità Hospital, Cagliari University, Cagliari, Italy

Renal stone formation is a common multifactorial disorder, of unknown etiology, with an established genetic contribution. Lifetime risk for nephrolithiasis is ~10% in Western populations, and uric acid stones account for 5%–10% of all stones, depending on climatic, dietary, and ethnic differences. We studied a small, isolated founder population in Sardinia, characterized by an increased prevalence of uric acid stones, and performed a genomewide search in a deep-rooted pedigree comprising many members who formed uric acid renal stones. The pedigree was created by tracing common ancestors of affected individuals through a genealogical database based on archival records kept by the parish church since 1640. This genealogical information was used as the basis for the study strategy, involving screening for alleles shared among affected individuals, originating from common ancestors, and utilization of large pedigrees to obtain greater power for linkage detection. We performed multistep linkage and allele-sharing analyses. In the initial stage, 382 markers were typed in 14 closely related affected subjects; interesting regions were subsequently investigated in the whole sample. We identified two chromosomal regions that may harbor loci with susceptibility genes for uric acid stones. The strongest evidence was observed on 10q21-q22, where a LOD score of 3.07 was obtained for D10S1652 under an affected-only dominant model, and a LOD score of 3.90 was obtained using a dominant pseudomarker assignment. The localization was supported also by multipoint allele-sharing statistics and by haplotype analysis of familial cases and of unrelated affected subjects collected from the isolate. In the second region on 20q13.1-13.3, multipoint nonparametric scores yielded suggestive evidence in a ~20-cM region, and further analysis is needed to confirm and fine-map this putative locus. Replication studies are required to investigate the involvement of these regions in the genetic contribution to uric acid stone formation.

Introduction

Nephrolithiasis is a common disease with multifactorial etiopathogenesis (Jaeger 1996; Baggio 1999). Its prevalence in Western populations is ~10% (Serio and Fraioli 1999; Rivers et al. 2000), although geographical and ethnic differences among populations have also been reported (Soucie et al. 1996). The major classes of formed stones are calcium oxalate, calcium phosphate, uric acid, struvite, and cystine. The majority of stone formers have disturbances either in the metabolism and excretion of stone constituents or in promoters (Siddiqui et al. 1998) and inhibitors of crystallization (Dussol and Berland 1998).

Clinical and epidemiological studies have documented that several types of risk factors are involved in disease etiology, such as dietary habits, warm climate, and familial occurrence (Jaeger 1996; Curhan et al. 1997). It is well recognized that nephrolithiasis has a genetic component (Curhan et al. 1997; Scheinman 1999) and specific genes that cause the cystine lithiasis have been identified (Pras et al. 1994; Bisceglia et al. 1997) (CSNU [MIM 220100]). Evidence for linkage on chromosome 12q12-q14 has been reported for idiopathic calcium stone formation (Scott et al. 1999) (oxalate [MIM 167030]), whereas absorptive hypercalciuria, a common cause of kidney stones, has been mapped to 1q23.3-q24 (Reed 1999) (hypercalciuria [MIM 143870]). However, genes that predispose to the most common forms of kidney stones remain unknown, and, therefore, important genetic factors involved in disease etiology remain to be identified.

Gene mapping for complex traits is a major challenge of current research studies, since it is hampered by several factors, such as genetic heterogeneity, small effects of disease alleles on risks, and confounding effects caused by gene-environment and gene-gene interactions. Numerous susceptibility genes are likely to be involved in common diseases such as nephrolithiasis, and most of these genes may be common in populations, sug-

Received December 27, 2000; accepted for publication March 7, 2001; electronically published April 17, 2001.

Address for correspondence and reprints: Dr. Maria Neve Ombra, Istituto di Genetica Molecolare, CNR, Casella Postale, 07040 Santa Maria La Palma (SS), Italy. E-mail: ombra@igm.ss.cnr.it

* These authors contributed equally to this work.

0002-9297/2001/6805-0007$02.00

gesting interactions of different allelic variants in disease etiology. Selection of homogeneous study populations with reduced genetic variability and etiological heterogeneity is therefore important to increasing the possibility of identification of susceptibility genes. Isolated populations have been successfully utilized in mapping Mendelian genes and are likely to be particularly useful for mapping predisposing genes for common, complex diseases (Terwilliger and Weiss 1998; Kruglyak 1999; Wright et al. 1999; Peltonen et al. 2000). A small number of founding individuals and a high rate of past and present consanguineous and endogamous marriages—typical of small communities—reduce the number of susceptibility genes in the population, increasing genetic homogeneity. Furthermore, since individuals are exposed to a common environment and a relatively uniform lifestyle, nongenetic variability is also minimized, and the noise caused by other etiological determinants is therefore reduced. All these characteristics facilitate the search for susceptibility genes involved in complex disease etiology.

We studied an isolated Sardinian village where nephrolithiasis prevalence is higher than in the general Western population. We conducted a multistep genomewide search (GWS) in a deep-rooted pedigree originating from the village and comprising 37 related members affected by uric acid renal stones (UAS). Selecting families with multiple affected individuals originating from populations with restricted genetic variation may result in enrichment of shared genetic determinants and may increase the power of the study. Our study strategy was based on the search for genomic segments shared identical-by-descent (IBD) by affected individuals originating from common ancestors (Houwen et al. 1994; Nikali et al. 1995; Hovatta et al. 1999). GWS was utilized to identify loci underlying this common disorder and to search for major genetic determinants in this population. We identified a chromosomal region on 10q21-q22 that might harbor a susceptibility gene for UAS. Suggestive evidence for linkage to the 20q13.1-q13.3 region was also obtained in our study.

Subjects and Methods

Population and Genealogical Data

We identified an isolated village, Talana, in central Sardinia, with 1,200 inhabitants, that is characterized by slow population growth, high endogamy, and high inbreeding (Wright et al. 1999). Characterization of maternal and paternal lines through mitochondrial DNA and Y-chromosome microsatellite markers allowed us to establish that the majority of the population of present-day Talana descends from <20 founder individuals (authors' unpublished data).

The limited number of founders and isolation increased genetic homogeneity, as compared with that in other larger, more exogamous populations. Geographical and cultural isolation have generated a great deal of homogeneity in lifestyle and eating habits, providing an ideal uniform context for studies of complex diseases. Precise details of births, deaths, marriages, and people's origins have been preserved since 1640 in the parochial *Quinque Libri* and in the municipal archives. All data were transcribed on paper forms and then were transferred into appropriate databases, allowing the creation of genealogical trees for almost any individual.

Subjects

All individuals participating in the study gave informed consent, and all samples were taken according to the Helsinki declaration. Subjects enrolled in this study were examined by the same physicians and were subjected to standardized interviews. Individuals with a history of urinary tract infection or with any secondary condition that might predispose to kidney stones (e.g., inflammatory bowel disease or gout) were not included. All subjects affected by renal stones and their family members underwent renal ultrasound examination to identify asymptomatic cases and to confirm previous diagnoses. Ultrasonography is used to image calculi, such as UAS, that are nonopaque on X-rays (Pollack et al. 1978). Uric acid nephrolithiasis is revealed by the presence of a radiolucent stone and urine acidity ($pH < 5.5$), (Pak et al. 1980; Stojimirovic 1998). Clinical tests included crystallographic analysis of stones; 24-h urinalysis of oxalates, citrates, $NH4^{+}$, $PO4^{--}$, uric acid, calcium, and acidity titration; and serum concentration of calcium, uric acid, urea, creatinine, sodium, and phosphates (Pak et al. 1980; Dawson and Whitfield 1996).

Several dietary factors have been related to stone formation, on the basis of per capita consumption, through epidemiological studies. Dietary calcium, animal proteins, oxalate, and sodium increase the likelihood of stone formation, whereas fiber, magnesium, and potassium seem to reduce the risk (Hassapidou et al. 1999). We collected data on dietary habits for a representative sample of the population of Talana, comprising 40 unaffected unrelated subjects and 60 patients with renal stones. All subjects completed a questionnaire regarding amount, type, and frequency of food intake. The questionnaire was analogous to the analytical tool used by Hassapidou et al. (1999), adapted to include the most common foods of the Mediterranean diet.

Unaffected individuals were checked by ultrasonography to ensure that they did not carry any UAS. The mean age at observation of unaffected controls (~63 years) was sufficiently high to have given an elevated probability of developing stones, since we observed a

low age at onset in affected subjects (mean age at onset ~38 years). Familial relationships among subjects were verified by searches of the genealogical database of our largely inbred study population. No controls had a closer degree of relatedness than second cousins: 2 of 40 controls were second cousins, 13 were third cousins, 4 were third cousins once removed, and 9 were fourth cousins. The remaining controls were more distant than fourth cousins once removed.

Through the archival database common ancestors were identified for 37 of 60 affected individuals. They are linked in an extended 10-generation pedigree comprising 156 subjects (fig. 1), founded by a single ancestral couple born at the beginning of the 18th century. Among the remaining affected subjects, 15 were used in the association study (see Statistical Analysis section), since they constituted an independent sample of unrelated cases and were not connected either to the control subject (i.e., more distant than fourth cousins), or to the familial case subjects.

Genotyping

Genomic DNA was extracted from 10 ml EDTA-treated blood, as described by Miller et al. (1988). PCR reactions were loaded on an ABI Prism 377 DNA Analyzer (PE Biosystems), using denaturating 5.0% polyacrylamide gels. The genotyping data were processed by GENESCAN version 3.1 and GENOTYPER version 2.5 software.

Multistep GWS

Stage I.—The GWS initially was performed on a reduced sample of 17 individuals from a large six-generation pedigree that included 14 members (6 men and 8 women) affected by UAS. The family members analyzed in stage I are highlighted in figure 1. We used the 382 microsatellite-marker mapping set from ABI Prism, version 2. Markers are uniformly distributed over the genome at an average distance of ~10 cM. Genetic maps were derived from Généthon (Dib et al. 1996), and allele frequencies for the population of Talana were estimated from an additional sample of 25 unrelated individuals. The family was analyzed using parametric linkage and nonparametric IBD-sharing approaches. Regions yielding suggestive evidence of linkage (LOD score >1) and/or excess IBD sharing among affected members ($P < .05$) were followed up in the second stage.

Stage II.—Interesting regions identified in stage I were investigated in more-distantly related affected individuals selected from the large pedigree. Eleven regions on 10 chromosomes were followed up, through linkage and IBD-sharing analysis, in the large pedigree shown in figure 1. The pedigree was divided into branches for analysis, so that no meiosis was represented in more than one subfamily: descendants of individuals III:5–III:6, descendants of individuals IV:3–IV:4, and descendants of individuals IV:1–IV:2 (breaking the loops at individuals VII:6 and VII:1 for the two latter pedigrees) (fig. 1). The subdivision was necessary for the computational requirements of linkage analysis, which is not feasible in such a large family. We consider a powerful strategy to use all the information on genealogical data of our study sample, and examination of smaller subfamilies therefore was not attempted at this stage of the analysis.

Stage III.—Linkage and IBD analyses were performed with a denser set of markers in the regions that showed evidence of linkage and/or excess IBD sharing in the extended pedigree. A two-point dominant pseudomarker approach was also performed for markers lying in the identified regions. Six markers spanning the 10q21-q22 region were typed in the entire family sample: D10S1640, D10S1719, and D10S1670 were selected from the Généthon database, and D10S609, D10S1715, and D10S1241 were chosen from the Genome Database. In the 20q13.1-q13.3 region, the following additional markers were chosen: D20S888, D20S902, and D20S102 from the Généthon database and D20S840, D20S52, D20S746, D20S149, D20S158, D20S459, and AFMa202yb9 from the Genome Database. Marker order was derived from the physical map of the Sanger Centre, and genetic distances were derived from the Genetic Location Database or were inferred from physical distances. In order to compare the frequency of specific alleles and haplotypes between case and control subjects, 15 affected and 40 unaffected unrelated subjects, described in the Subjects section, were genotyped at markers lying in the regions that provided significant linkage results.

Statistical Analysis

Mendelian inconsistencies were checked with the PEDMANAGER program, version 0.9. The program was also used to calculate allele frequencies for each marker from the additional sample of 25 unrelated individuals selected from the general population.

Two-point LOD scores were calculated with FASTLINK version 4.1P (Cottingham et al. 1993; Schäffer et al. 1994) in an affected-only strategy, assuming all nonaffected individuals as phenotypically unknown, since unaffected subjects do not provide reliable information on the underlying disease-locus genotype for a complex disease. Model-based linkage analysis was performed under both dominant and recessive models assuming disease-allele frequencies of .01 for the dominant model and .1 for the recessive model. It has been shown that, when the true mode of inheritance is unknown, model-based linkage analysis can still be a robust approach when a small number of different models are assumed.

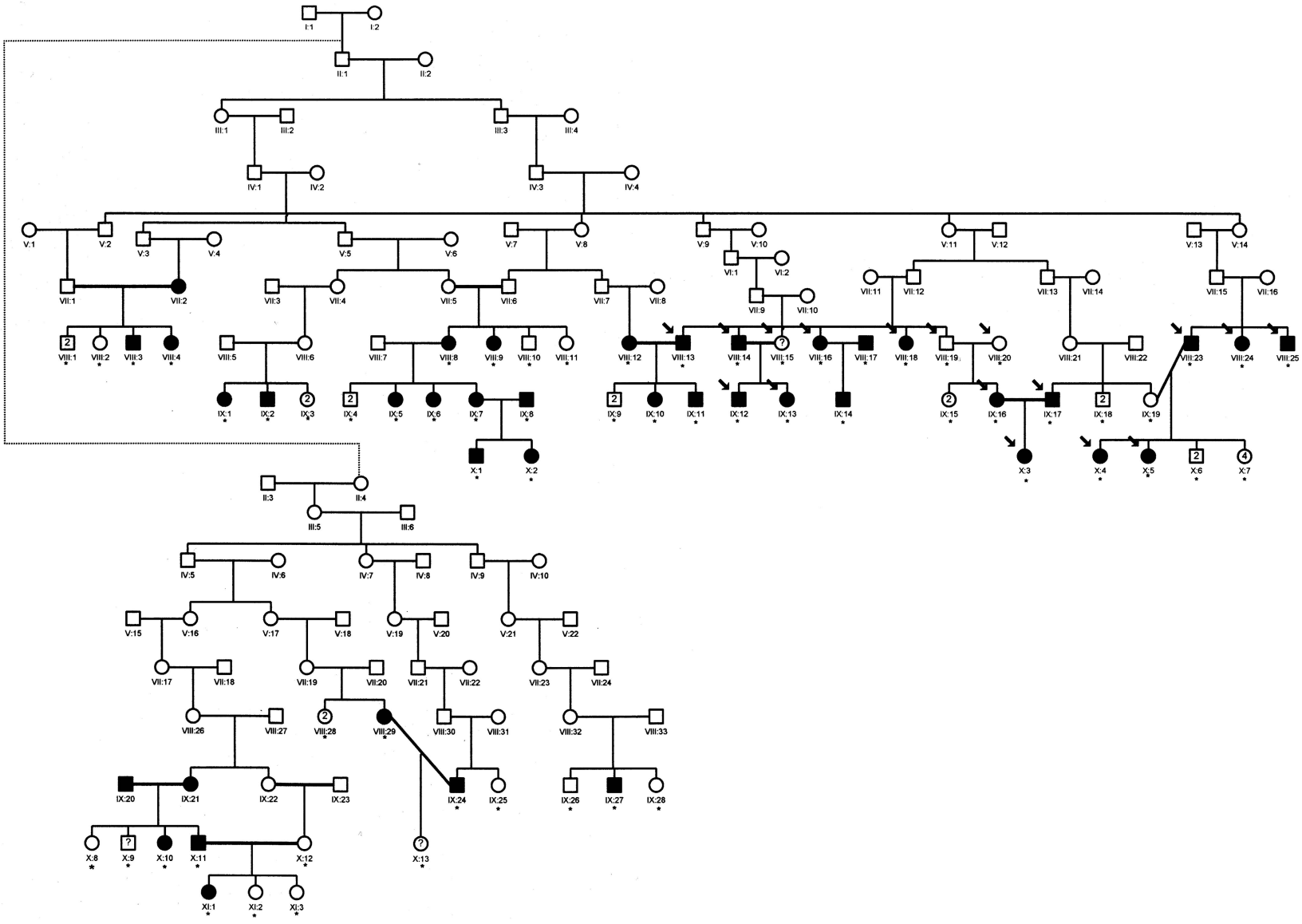

Figure 1 Pedigree used in the linkage analysis. Family members analyzed in stage I are indicated by an arrow. The whole pedigree was studied in stages II and III. Blackened symbols indicate affected individuals; numbers inside symbols indicate the numbers of brothers/sisters present in the sibship; genotyped individuals used in the linkage analysis are indicated by asterisks (*) below their symbols.

Table 1

Anthropometric Characteristics and Food Consumption (Number of Servings per Week of Foods) among Affected and Unaffected Individuals Subjected to the Analysis

Subjects	Mean ± SD Age (years)	Mean ± SD Height (m)	Mean ± SD Weight (kg)	Mean ± SD BMI (kg/m^2)	Mean ± SD Servings/Week of					
					Bread/ Pasta[a]	Meat	Fish	Milk/ Cheese[a]	Fruits/ Vegetables[a]	Sweets
Affected ($n = 40$)	62.75 ± 17.42[b]	1.57 ± .09	65.43 ± 17.36	26.58 ± 4.02	11.25 ± 2.38[b]	4.10 ± 1.48[b]	.70 ± .65	10.50 ± 3.50	11.03 ± 3.45	3.08 ± 2.29
Unaffected ($n = 60$)	54.75 ± 18.70	1.56 ± .09	65.48 ± 12.37	27.02 ± 4.90	12.27 ± 1.87	3.46 ± .84	.81 ± .57	10.80 ± 3.71	10.40 ± 2.80	2.24 ± 1.19

[a] Sum of servings/week.

[b] $.01 < P < .05$ (statistical significance between mean differences was tested by t-test).

Even though the choice of genetic model is arbitrary and, presumably, incorrect, LOD-score models can still outperform nonparametric methods (Hodge et al. 1997; Greenberg et al. 1998; Durner et al. 1999). A pseudomarker analysis that was shown to have properties equivalent to those of an affected-relative IBD-approach (Trembath et al. 1997; Terwilliger 1998; Göring and Terwilliger 2000) was also performed in stage III. Under this approach, every meiosis was made informative for linkage at the disease locus and affected individuals were analyzed as if all possible disease genes were shared IBD. We used a dominant pseudomarker assignment, as described by Göring and Terwilliger (2000).

For multipoint analysis we used a model-free approach based on evaluation of IBD allele sharing among affected individuals. Multipoint IBD-sharing analysis is usually restricted to pedigrees of moderate size and structure. For large pedigrees, like the one we generated, exact methods are not feasible because of the size and complexity of relationships between affected pairs resulting in an enormous number of underlying configurations that are consistent with the observed data. Current methods of analysis for extended and complex pedigrees are based on Markov chain–Monte Carlo (MCMC) sampling. In our study, multipoint nonparametric statistics for allele sharing among affected individuals were evaluated with SIMWALK2 (Sobel and Lange 1996), a statistical-genetics program that uses MCMC and simulated annealing algorithms to analyze pedigree data of any size. The program estimates several IBD-sharing statistics at each position of the map by sampling IBD configurations (consistent with the data) proportionally to their likelihoods. SIMWALK2 reports five different statistics and, in particular, calculates a nonparametric linkage score (NPL) as implemented in GENEHUNTER (Kruglyak et al. 1996). The identified region on 10q21-q22 was also investigated, through exact IBD-sharing statistics, using the computer program ALLEGRO (Gudbjartsson et al. 2000), although, because of computational requirements, the large pedigree was split into five smaller families for the analysis.

Results

Nutritional and Clinical Data

Data on body weight, height, body-mass index (BMI), age, and food consumption have been collected and summarized in table 1. The results of the survey evidenced no major differences of dietary habits among cases and controls in terms of quality of food associated with increased or decreased likelihood of renal-stone formation. Intake of animal proteins from meat, a known risk factor for nephrolithiasis, was actually slightly higher in controls than in affected subjects. It should also be emphasized that most of the population of Talana residents lead a rather uniform and conservative lifestyle, which includes traditional and rural dietary habits. No significant differences in body weight, height, and BMI mean values were observed in the two groups.

We identified several subjects in Talana who had renal stones that showed uric acid as the principal component. UAS account generally for 5%–10% of all renal stones, whereas, in our isolate, they were five times more frequent than expected. The risk for UAS formation correlates with the degree of uric acid concentration in urine and low urinary pH (pH <5.5) (Helabe and Sperling 1994; Asplin 1996). The analysis performed in the sample of stone formers evidenced that 85% of the investigated cases had low urinary pH levels (ranging from 5.2 to 5.6), and that acidity titration on 24-h urine was also higher than normal values in 80% of cases (titration >40 mEq/24 h). Excretion parameters on 24-h urine showed that 35% of the cases had also high values of uricosuria (>700 mg/24 h).

Simulation Study

Exact power estimation and assessment of genome-wide significance through simulation is impractical for the whole pedigree, given the complexity of the genealogical structure, the large number of individuals, and the unknown mode of disease inheritance. In any case, a simulation study was performed for the 17-member pedigree used in stage I, to estimate its power and to

evaluate the type I–error rate of specific cutoff values of the LOD scores for following-up regions. The program SLINK (Weeks et al. 1990) was used to generate 200 replicates of the family assuming both linkage (θ = .05) and nonlinkage. Maximum LOD scores of 2.08 and of 3.24 were obtained under the assumption of linkage under the dominant and the recessive models, respectively, with ~50% power. A LOD score ⩾1.0 was observed in 2.6% of the replicates obtained under no linkage, and this LOD score threshold was used for selecting regions to follow up in the second stage.

Stage I.—Results from stage I are shown in table 2. Eleven regions on 10 chromosomes resulted in two-point LOD scores >1 and/or in *P* values <.05 for the multipoint NPL statistic of SIMWALK2. The best results were observed on chromosomes 10 and 20, where both dominant and recessive models provided LOD scores >1 and the NPL score yielded $P < .005$ (peak at marker D10S537, $P = .0037$; peak at marker D20S173, $P = .0016$, respectively). The 11 regions were further evaluated in the second stage. Because our initial GWS was based on a small sample of individuals belonging to a single large family, and given the high level of inbreeding present in the population, we expected to observe a high rate of false-positive regions caused by both chance sharing among individuals and small sample size. Evaluation of the identified regions in additional family members collected from the isolate would allow us to confirm or exclude the involvement of these regions in the second stage of the analysis.

Stage II.—When we extended the analysis to the whole family, we found evidence of linkage and excess IBD sharing for the chromosome 10 and 20 regions identified in stage I. On chromosome 10, LOD scores of 3.07 and 2.17 at recombination fraction (θ) 0 were observed at markers D10S1652 and D10S537 (~5.4 cM apart), respectively, under the affected-only dominant model. Excess IBD sharing (NPL score: $P < .05$) was also observed in the region among affected individuals. For chromosome 20, a suggestive NPL score ($P < .05$) was obtained between D20S196 and D20S173 (~16 cM apart). On chromosome 15, a LOD score of 1.58 was observed at D15S165, but no significant results were obtained with the multipoint allele-sharing approach, and the region was not investigated further. All the remaining regions identified in GWS stage I resulted in two-point LOD score <1 and nonsignificant ($P > .05$) IBD-sharing statistics (data not shown).

Stage III.—Six markers spanning the 10.7-cM region on chromosome 10 and ten markers spanning the 26-cM region on chromosome 20 were genotyped in this stage of the analysis. For the chromosome 20 region, a peak LOD score of 1.09 was obtained at marker D20S902, whereas suggestive NPL scores ($P < .05$) were observed over an ~20-cM region between markers D20S888 and D20S158, with a peak at marker D20S902 ($P = .008$). SIMWALK2 statistic B, which measures the maximum number of alleles among affected individuals descended from any common founder, was the most significant and reached a peak at marker D20S902 ($P = .0025$).

For the chromosome 10 region, results are shown in table 3. No additional marker gave a better result than D10S1652, for which a LOD score of 3.90 was obtained

Table 2

Results of GWS Stage I

Chromosome	Marker Region	Interval (cM)	Peak Marker	Peak NPL (*P*)[a]	Z_{dom}[b]	Z_{rec}[c]
1	D1S2841–D1S207	102.0–110.0	D1S207	.0413	1.40	
2	D2S117–D2S126	198.3–226.0	D2S325	NS	1.76	
3	D3S1263–D3S1300	27.1–70.1	D3S1289	.0043	1.04	
5	D5S2115–D5S408	144.9–202.7	D5S436	.0030		
7	D7S516–D7S502	38.5–81.2	D7S484	.0297		
10	D10S1652–D10S537	81.9–90.3	D10S537	.0037	1.86	1.50
	D10S1651–D10S212	167.7–170.1	D10S212	.0050		1.43
14	D14S261–D14S288	.0–33.2	D14S261	.0292		
15	D15S128–D15S1007	.0–22.0	D15S1002	.0099		
16	D16S423–D16S3068	.0–41.6	D16S3075	.0065		
20	D20S196–D20S171	76.3–101.3	D20S173	.0016	1.80	1.39

[a] Nonparametric linkage score obtained with SIMWALK2 (Statistic E), which measures the degree of clustering of the marker alleles descending from pedigree founders, for affected individuals only (Sobel and Lange 1996). Statistic E is analogous to the NPL score as implemented in the GENEHUNTER program (Kruglyak et al. 1996). Statistic E indicates whether there are a few descent trees that are overly represented among the alleles of affected individuals. NS = not significant ($P > .05$).

[b] Two-point LOD score derived using dominant model.

[c] Two-point LOD score derived using recessive model.

Table 3

Results of GWS Stage III in Chromosome Region 10q21-q22

			Maximum LOD (θ)[a]			*P*	
cM	Marker	Heterozygosity[b]	Dominant[c]	Recessive[d]	Pseudomarker Assignment	ALLEGRO NPL[e]	SIMWALK2 NPL
0	D10S1640	.57	1.15 (.0)	.43 (.0)	.22 (.23)	.0069	.0361
2.0	D10S1652	.54	3.07 (.0)	.90 (.0)	3.90 (.0)	.0058	.0415
2.5	D10S1719	.75	.43 (.18)	.66 (.04)	.37 (.23)	.0054	.0485
5.9	D10S1670	.67	.83 (.11)	1.67 (.0)	.74 (.14)	.0048	.0495
6.4	D10S1715	.55	.28 (.08)	.19 (.05)	.31 (.14)	.0048	.0463
6.9	D10S1241	.65	.00 (.5)	.13 (.18)	.00 (.5)	.0050	.0545
7.4	D10S537	.76	2.17 (.0)	1.76 (.0)	2.15 (.05)	.0077	.0265
10.7	D10S609	.67	.00 (.5)	.00 (.5)	.00 (.5)	.0057	.0329

[a] No correction for multiple testing due to the different genetic models used was applied.
[b] Marker heterozygosity estimated from genotyped pedigree members.
[c] Gene frequency $q = .01$.
[d] Gene frequency $q = .1$.
[e] Exact *P* values associated with the S-all scoring function.

by the pseudomarker approach. NPL scores estimated with SIMWALK2 yielded *P* values <.05 in the whole region. Exact NPL scores evaluated with ALLEGRO (assuming the S-all scoring function and the exponential model) resulted in *P* values <.01 in the whole region, with peaks at markers D10S1670 and D10S1715 ($P = .0048$).

To investigate the significance of the linkage result, we performed a randomization of the phenotypes in the family members, simulating a disease gene unlinked to the marker for which the best evidence of linkage was obtained (marker D10S1652) and using, in each replicate, true genotypes at marker D10S1652 for available family members. We used SLINK to generate 500 replicates and MSIM to estimate the probability of observing by chance a LOD score exceeding the true LOD score. We observed that, in 3 of 500 replicates, a LOD score >3.07 was obtained in the affected-only approach, yielding an associated *P* value of .006, with an upper endpoint of the 95% confidence interval of 0.0174 (Ott 1999).

Haplotype inspection in familial cases revealed a common ~7.4-cM haplotype (extending from marker D10S1640 to marker D10S537) in many affected members. Portions of this extended haplotype were shared by most affected members belonging to the family, as shown in table 4. For marker D10S1652, 19 affected individuals were homozygous for allele 13 (51.4%) versus 8 unaffected members (24.2%).

The same 10q region, spanning from D10S1640 to D10S537, was investigated in 15 unrelated cases and 40 unaffected controls whose haplotypes were reconstructed from typing of close relatives. A 6.9-cM haplotype between markers D10S1640 and D10S1241 (haplotype 4-13-2-1-3-5), already identified in the family members, was also the most common extended haplotype observed in unrelated cases collected from the general population (table 4) and was present in seven affected haplotypes (23.3%), compared with seven control haplotypes (8.8%). Parts of this haplotype were also more prevalent in affected chromosomes than in control chromosomes: haplotype 4-13, for markers D10S1640–D10S1652, was present in 14 of 30 case haplotypes and in 19 of 80 control haplotypes ($P = .0340$ by two-tailed Fisher's exact test); haplotype 13-2, for markers D10S1652–D10S1719, was present in 11 case haplotypes and in 13 control haplotypes ($P = .0359$ by two-tailed Fisher's exact test). The allele 13 at marker D10S1652 was present in 22 affected haplotypes compared with 30 control haplotypes ($P = .0011$, by two-tailed Fisher's exact test).

Discussion

We studied a large pedigree comprising many individuals affected by UAS, to identify susceptibility genes underlying this complex disease. The genetic contribution to common forms of nephrolithiasis is well established but poorly understood, as the number of genes involved in disease etiology and the magnitude of their effects are still unknown.

Clinical and epidemiological studies have documented that several types of risk factors are involved in disease etiology. Diet, in particular, plays a significant role in disease etiology. Dietary intake was assessed in a representative sample of the population of Talana by a food-frequency questionnaire that was submitted to unaffected and affected subjects. The results of the survey show no major differences between unaffected and affected subjects in terms of quality of food associated with kidney-stone formation.

Kidney stones may form when concentration of uri-

Table 4

Haplotype Analysis in the 10q21-q22 Region in Affected Family Members and in Unrelated Case Subjects

	Marker (Distance in cM)						
Subject	D10S1640 (0)	D10S1652 (2.0)	D10S1719 (2.5)	D10S1670 (5.9)	D10S1715 (6.4)	D10S1241 (6.9)	D10S537 (7.4)
Affected family members:							
IX:27[a]	2	13	2	1	2	1	7
IX:24	3	13	2	1	2	3	4
VIII:29	3	13	2	1	2	3	8
VIII:18	4	13	2	1	3	5	10
IX:12[a]	4	13	2	1	3	5	10
VIII:14	4	13	2	1	3	5	10
VIII:16	4	13	2	1	3	5	10
X:4[a]	4	13	2	1	3	5	10
X:4[a]	4	13	2	1	3	5	10
VIII:24	4	13	2	1	3	5	10
VIII:23	4	13	2	1	3	5	10
X:5[a]	4	13	2	1	3	5	10
VIII:25	4	13	2	1	3	5	10
IX:17	4	13	2	1	3	5	10
X:3	4	13	2	1	3	5	10
IX:14[a]	4	13	2	1	2	5	10
X:5[a]	4	13	2	3	3	9	8
IX:1[a]	4	13	3	4	3	1	8
IX:2[a]	4	13	3	4	3	1	8
VIII:4	4	13	3	1	5	1	4
IX:6[a]	4	13	4	1	3	5	8
IX:7[a]	4	13	4	1	3	5	8
X:2	4	13	4	1	3	5	8
VIII:3[a]	4	13	4	3	3	5	4
IX:12[a]	4	13	7	3	3	9	4
IX:13	4	13	7	3	3	9	4
IX:27[a]	4	13	8	1	3	5	7
IX:5	4	13	8	4	3	1	4
IX:6[a]	4	13	8	4	3	1	4
IX:7[a]	4	13	8	4	3	1	4
IX:1[a]	4	13	8	9	3	5	5
IX:2[a]	4	13	8	9	3	5	5
VIII:9	4	13	8	4	3	1	4
X:1	4	13	8	4	3	1	4
VIII:8	4	13	8	4	3	1	4
VIII:3[a]	4	13	8	3	3	5	8
IX:14[a]	4	13	8	3	3	9	8
Unrelated case subjects:							
PT1[a]	2	13	2	1	3	1	8
PT2[a]	2	13	2	1	2	1	4
PT4	2	13	2	1	2	1	9
PT5	4	13	2	1	3	1	8
PT6	4	13	2	1	3	5	9
PT1[a]	4	13	2	1	3	5	8
PT7	4	13	2	1	3	5	8
PT8	4	13	2	1	3	5	4
PT3[a]	4	13	2	1	3	5	4
PT9	4	13	2	1	3	5	10
PT2[a]	4	13	2	1	3	5	4
PT10	4	13	3	1	3	1	10
PT11	4	13	3	3	3	1	4
PT12	4	13	4	1	3	1	4
PT3[a]	4	13	4	1	3	5	8
PT13	4	13	8	3	3	5	8
PT14	4	13	8	4	3	5	8

Note.—Boxed haplotypes indicate the haplotype shared by affected family members (see text). Among unrelated cases, 7 patients of 15 shared the 4-13-2-1-3-5 haplotype found in the family. The same haplotype was found only in seven control haplotypes (8.8%) (not shown).

[a] For these subjects, both haplotypes are shown.

nary constituents exceeds their solubility. Although hyperuricosuria is a risk factor in only 30% of cases, a common denominator in normo- and hyperuricosuric patients is a low urinary pH (Smith 1998). Uric acid solubility decreases abruptly at urinary pH <5.5. Thus, a long period of urine acidity leads to UAS formation. It must be considered that a proportion of healthy people generally exhibit low urinary pH values and the same uricuria as some UAS formers. Consequently, even if low PH is likely to be a necessary determinant for UAS, it is not a sufficient factor (Koka et al. 2000). Very little is known about uric acid crystallization inhibitors. Urinary glycoproteins similar to mucine inhibit uric acid crystallization (Grases et al. 1999), although genes encoding these proteins have not yet been identified.

In a exploratory study performed on our case subjects, we observed low urinary pH levels (mean pH 5.55), and these were the only abnormal values identified, suggesting that reduced urinary pH could be an important risk factor for UAS formation in our patients, although further evaluation is needed to confirm this hypothesis in a more exhaustive population sample.

For acidification of urine, several processes are responsible. One mechanism is NH_3 production from glutamine, which takes place in the proximal tubule (Hamm 1987; Hayashi 1998). A reduced NH_3 production resulting in a decreased buffer capacity for H^+ ions (caused by lack of glutamine or by an enzymatic defect) is considered a potential mechanism for permanently low urinary pH in UAS formers. Nevertheless, we did not observe a significantly reduced NH_3 level among our cases. Uric acid is the end product of purine degradation in humans, and the kidney plays a predominant role in its elimination. The molecular basis for renal urate transport, however, has not been completely defined.

We performed a GWS, using both parametric and nonparametric linkage approaches to identify susceptibility genes for UAS, on an extended family comprising many affected individuals. The family was collected from a small, isolated Sardinian village, characterized by genetic and environmental homogeneity. Although parametric linkage analysis can be problematic for complex traits, the approach can also be especially powerful when affected individuals arise from genetically isolated populations and when common ancestors can be identified (Houwen et al. 1994; Nikali et al. 1995; Hovatta et al.1999). A two-point LOD score analysis was also performed between a pseudomarker locus that represents the maximum possible IBD sharing at the susceptibility locus among affected members and marker loci. The method was shown to approximate nonparametric affected-relative approaches (Göring and Terwilliger 2000). Nonparametric methods, based on IBD sharing among affected relatives, are more commonly used in the genetic analysis of complex traits, using sib pairs or general pedigrees of moderate size. In our study, we used an approximation method for multipoint allele-sharing analysis based on MCMC that was implemented in the computer program SIMWALK2 (Sobel and Lange 1996) and that allowed us to make full use of our family material. Large pedigrees that include distantly affected relatives provide greater power for linkage detection, since the prior probability of sharing any genome segment is small for distant relatives and evidence of IBD sharing indicates the presence of a genetic factor contributing to the trait. Furthermore, large pedigrees with many affected relatives may insure an increased genetic homogeneity of the trait and, therefore, a more powerful analysis. The study of large families collected from a genetic isolate can be powerful even for common allelic variants that are likely to have the same ancestral origin.

A multiple-stage strategy was used in our study. In the first stage, a subpedigree was genotyped and analyzed using a relatively sparse map of markers. Regions with evidence of linkage were followed up in the entire pedigree increasing marker density. We identify a chromosomal region on 10q21-q22, where LOD scores of 3.07 and 3.90 at $\theta = 0$ were obtained for marker D10S1652 under dominant-model and pseudomarker approaches, respectively. Multipoint IBD-sharing analysis with SIMWALK2 yielded suggestive evidence ($P <$ .05) with the NPL statistic, and exact methods resulted in P values <.01 for all markers spanning the 10.7-cM region. The localization of a new candidate locus on chromosome 10 is supported by haplotype sharing observed in affected pedigree members, since it is extremely unlikely that distantly related affected subjects would share long chromosomal intervals just by chance. Therefore, this result would indicate IBD sharing of marker haplotypes surrounding a disease gene, although it is difficult to evaluate the exact significance of haplotype sharing observed in the family, given the background inbreeding present in the population.

In small, founder populations, like the one we studied, linkage disequilibrium may extend over relatively large intervals (Terwilliger et al. 1998) and therefore might be detectable in the proximity of a disease gene. A large proportion of affected individuals are likely to have inherited the same predisposing genes from common ancestors and to share IBD portions of chromosomal regions around susceptibility genes. The same ~7-cM haplotype observed in several cases of the large family reported here was also present at higher frequency among unrelated affected subjects than in unaffected controls collected from the population. Significant association was observed with haplotype 4-13 at markers D10S1640–D10S1652 ($P = .0340$) and haplotype 13-2 at markers D10S1652–D10S1719 ($P =$

.0359). Stronger association ($P = .0011$) was observed with allele 13 at marker D10S1652, which provided the highest LOD score in the linkage analysis. Further evaluations, using a more exhaustive sample and combining linkage and association approaches, are needed to confirm these preliminary results.

In this study, we also obtained suggestive evidence of linkage in the 20q13.1-q13.3 region, where excess IBD sharing among affected subjects was observed over a large 20-cM region and a peak LOD score of 1.09 was obtained at marker D20S902.

Analysis of a large pedigree allowed us to include all known connections among individuals and thus to control, as much as possible, "hidden" inbreeding that might inflate the LOD score, creating false-positive regions (Miano et al. 2000). Deviation from Hardy-Weinberg equilibrium was also tested in the control sample and provided no significant results (data not shown).

Because the first stage of the study was performed in a reduced sample of affected individuals belonging to a single large family, our approach was only suitable for detection of major genes that confer a relatively high disease risk in the population. However, it might not have had enough power to detect weak-effect disease genes, which are likely to be missed when a small sample of individuals is investigated through GWS. We are currently extending our study to include more affected subjects and unaffected control subjects, to confirm the involvement of the identified region in genetic predisposition to UAS formation through linkage disequilibrium analysis using a denser set of markers.

Analysis of the genes located in this region, based on the most recent chromosome 10 map, did not reveal obvious candidates. The lack of a known gene at these chromosomal regions leads to the possibility that the identified candidate region might harbor a new susceptibility gene involved in kidney stone formation.

Acknowledgments

We are grateful to the population of Talana and to all the individuals who participated in this study. We thank Anna Rosa Tegas, for collection of samples, and Marco Paulis, Marco Cossu, and Paolo Cirillo, for ultrasonography investigations.

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

Généthon, http://www.genethon.fr/genethon_en.html
Genetic Location Database, http://cedar.genetics.soton.ac.uk/public_html/ldb.html
Genome Database, http://gdbwww.gdb.org/
Helsinki declaration, http://www.wma.net/e/approvedhelsinki.html
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for CSNU [MIM 220100]; for hypercalciuria [MIM 143870]; for oxalate [MIM 167030])
Sanger Centre, http://www.sanger.ac.uk/HGP/

References

Asplin JR (1996) Uric acid stones. Semin Nephrol 16:412–424

Baggio B (1999) Genetic and dietary factors in idiopathic calcium nephrolithiasis: what do we have, what do we need? J Nephrol 12:371–374

Bisceglia L, Calonge MJ, Totaro A, Feliubadalo L, Melchionda S, Garcia J, Testar X, Gallucci M, Ponzone A, Zelante L, Zorzano A, Estivill X, Gasparini P, Nunes V, Palacin M (1997) Localization, by linkage analysis, of the cystinuria type III gene to chromosome 19q13.1. Am J Hum Genet 60: 611–616

Cottingham RW Jr, Idury RM, Schäffer AA (1993) Faster sequential genetic linkage computations. Am J Hum Genet 53: 252–263

Curhan GC, Willett WC, Rimm EB, Stampfer MJ (1997) Family history and risk of kidney stones. J Am Soc Nephrol 8: 1568–1573

Dawson C, Whitfield H (1996) ABC of urology: urinary stone disease. BMJ 312:1219–1221

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature 380:152–154

Durner M, Vieland VJ, Greenberg DA (1999) Further evidence for the increased power of LOD scores compared with nonparametric methods. Am J Hum Genet 64:281–289

Dussol B, Berland Y (1998) Urinary kidney stone inhibitors: what is the news? Urol Int 60:69–73

Göring HHH, Terwilliger JD (2000) Linkage analysis in the presence of errors. IV: Joint pseudomarker analysis of linkage and/or linkage disequilibrium on a mixture of pedigrees and singletons when the mode of inheritance cannot be accurately specified. Am J Hum Genet 66:1310–1327

Grases F, Ramis M, Villacampa AI, Costa-Bauzà A (1999) Uric acid urolithiasis and crystallization inhibitors. Urol Int 62: 201–204

Greenberg DA, Abreu P, Hodge SE (1998) The power to detect linkage in complex disease by means of simple LOD-score analyses. Am J Hum Genet 63:870–879

Gudbjartsson DF, Jonasson K, Frigge ML, Kong A (2000) Allegro, a new computer program for multipoint linkage analysis. Nat Genet 25:12–13

Hamm LL, Simon EE (1987) Roles and mechanisms of urinary buffer excretion. Am J Physiol 253:595–605

Hassapidou MN, Paraskevopoulos ST, Karakoltsidis PA, Petridis D, Fotiadou E (1999) Dietary habits of patients with renal stone disease in Greece. J Hum Nutr Dietet 12:47–51

Hayashi M (1998) Physiology and pathophysiology of acid-base homeostasis in the kidney. Intern Med 37:221–225

Helabe A, Sperling O (1994) Uric acid nephrolithiasis. Miner Electrolyte Metab 20:424–431

Hodge SE, Abreu PC, Greenberg DA (1997) Magnitude of type I error when single-locus linkage analysis is maximized over models: a simulation study. Am J Hum Genet 60: 217–227

Houwen RHJ, Baharloo S, Blankenship K, Raeymaekers P, Juyn J, Sandkuijl LA, Freimer NB (1994) Genome screening by searching for shared segments: mapping a gene for benign recurrent intrahepatic cholestasis. Nat Genet 8:380–386

Hovatta I, Varilo T, Suvisaari J, Terwilliger JD, Ollikainen V, Arajärvi R, Juvonen H, Kokko-Sahin ML, Väisänen L, Mannila H, Lönnqvist J, Peltonen L (1999) A genomewide screen for schizophrenia genes in an isolated Finnish subpopulation, suggesting multiple susceptibility loci. Am J Hum Genet 65:1114–1124

Jaeger P (1996) Genetic versus environmental factors in renal stone disease. Curr Opin Nephrol Hypertens 5:342–346

Koka RM, Huang E, Lieske JC (2000) Adhesion of uric acid crystals to the surface of renal epithelial cells. Am J Physiol Renal Physiol 278:989–998

Kruglyak L (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. Nat Genet 22:139–144

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 58:1347–1363

Miano MG, Jacobson SG, Carothers A, Hanson I, Teague P, Lovell J, et al (2000) Pitfalls in homozygosity mapping. Am J Hum Genet 67:1348–1351

Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16:1215

Nikali K, Soumalainen A, Terwilliger JD, Koskinen T, Weissenbach J, Peltonen L (1995) Random search for shared chromosomal regions in four affected individuals: the assignment of a new hereditary ataxia locus. Am J Hum Genet 56:1088–1095

Ott J (1999) Analysis of human genetic linkage, 3d ed. Johns Hopkins University Press, Baltimore

Pak CY, Britton F, Peterson R, Ward D, Northcutt C, Breslau NA, McGuire J, Sakhaee K, Bush S, Nicar M, Norman DA, Peters P (1980) Ambulatory evaluation of nephrolithiasis: classification, clinical presentation and diagnostic criteria. Am J Med 69:19–30

Peltonen L, Palotie A, Lange K (2000) Use of population isolates for mapping complex traits. Nat Rev Genet 1:182–190

Pollack HM, Arger PH, Goldberg BB, Mulholland SG (1978) Ultrasonic detection of nonopaque renal calculi. Radiology 127:233–237

Pras E, Arber N, Aksentijevich I, Katz G, Schapiro JM, Prosen L, Gruberg L, Harel D, Liberman U, Weissenbach J, Pras M, Kastner DL (1994) Localizaton of a gene causing cystinuria to chromosome 2p. Nat Genet 6:415–419

Reed BY, Heller HJ, Gitomer WL, Pak CY (1999) Mapping a gene defect in absorptive hypercalciuria to chromosome 1q23.3-q24. J Clin Endocrinol Metab 84:3907–3913

Rivers K, Shetty S, Menon M (2000) When and how to evaluate a patient with nephrolithiasis. Urol Clin North Am 27: 203–213

Schäffer AA, Gupta SK, Shriram K, Cottingham RW Jr (1994) Avoiding recomputation in linkage analysis. Hum Hered 44: 225–237

Scheinman SJ (1999) Nephrolithiasis. Semin Nephrol 19: 381–388

Scott P, Ouimet D, Valiquette L, Guay G, Proulx Y, Trouvê ML, Gagnon B, Bonnardeaux A (1999) Suggestive evidence for a susceptibility gene near the vitamin D receptor locus in idiopathic calcium stone formation. J Am Soc Nephrol 10:1007–1013

Serio A, Fraioli A (1999) Epidemiology of nephrolithiasis. Nephron Suppl 81:26–30

Siddiqui AA, Sultana T, Bucholz NP, Waqar MA, Talati J (1998) Proteins in renal stones and urine of stone formers. Urol Res 26:383–388

Smith CL (1998) Renal stone analysis: is there any clinical value? Curr Opin Nephrol Hypertens 7:703–709

Sobel E, Lange K (1996) Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker sharing statistics. Am J Hum Genet 58:1323–1337

Soucie JM, Coates RJ, McClellan W, Austin H, Thun M (1996) Relation between geographic variability in kidney stones prevalence and risk factors for stones. Am J Epidemiol 143: 487–495

Stojimirovic B (1998) Functional evaluation in patients with kidney calculi. Srp Arh Celok Lek 126:394–398

Terwilliger JD (1998) Linkage analysis, model-based. In: Armitage P, Colton T (eds) Encyclopedia of biostatistics. John Wiley & Sons, Chichester

Terwilliger JD, Weiss KM (1998) Linkage disequilibrium mapping of complex diseases: fantasy or reality? Curr Opin Biotechnol 9:578–594

Terwilliger JD, Zöllner S, Laan M, Pääbo S (1998) Mapping in small populations with no demographic expansion. Hum Hered 48:138–154

Trembath RC, Clough RL, Rosbotham JL, Jones AB, Camp RDR, Frodsham A, Browne J, et al (1997) Identification of a major susceptibility locus on chromosome 6p and evidence for further disease loci revealed by two stage genome-wide search in psoriasis. Hum Mol Genet 6:813–820

Weeks DE, Ott J, Lathrop GM (1990) SLINK: a general simulation program for linkage analysis. Am J Hum Genet Suppl 47:A204

Wright AF, Carothers AD, Pirastu M (1999) Population choice in mapping genes for complex diseases. Nat Genet 23: 397–404