Genetic mapping of dinucleotide repeat polymorphisms and von Hippel-Lindau disease on chromosome 3p25–26


Abstract
A genetic map of highly polymorphic microsatellite markers spanning the von Hippel-Lindau region (VHL) of 3p25 was constructed using the CEPH reference pedigrees. A greater than 1000:1 odds map of pter-D3S1038-RAFI1-D3S651-D3S656-D3S110-D3S1255-cen was found. Genotyping of six multigenerational VHL families showed the region surrounding the D3S1038 marker to be the most likely location for the VHL gene with a peak location score of 10-04 with VHL completely linked to D3S1038. These data provide an initial high resolution genetic map of this region; D3S1038 appears to be a highly polymorphic marker that should prove useful in the future for presymptomatic diagnosis.

(J Med Genet 1993;30:487-91)

Von Hippel-Lindau (VHL) disease is an autosomal dominant pleomorphic disorder characterised by haemangioblastomas and cysts in the central nervous system (CNS), visceral, and retina. Renal cysts frequently progress to renal cell carcinoma (RCC) and many subjects have phaeochromocytoma as well. Age of onset of VHL is variable with a broad age of onset of approximately 30 years. Variable expression is also characteristic and non-manifesting obligate carriers over 60 years of age have been observed. At present there is no treatment or cure for VHL. Annual neurological and ophthalmological examinations are recommended. In addition, patients and at risk subjects should undergo at least biannual radiological examinations of the abdomen in order to rule out RCC as well as tumours of the CNS. Since early recognition of symptoms allows preventative intervention, accurate carrier detection before onset of serious complications has both clinical and economic considerations. VHL was originally localised to chromosome 3 by Seizinger et al through linkage to the RAF1 oncogene. Subsequent linkage studies confirmed this localisation. Recent studies have aimed at sublocalisation of VHL to the 3p25 region and VHL was mapped to an 8 cm region between RAF1 and D3S95 (64E2)/D3S719 (479H4), an anonymous DNA probe, was found to be the closest linked marker to VHL in this study with a peak lod score of 8.59 with no recombinants. Comparison of this two point lod score with the peak multipoint score (20-35) indicated that a significant amount of the meiotic information in the VHL families was as yet unrealised. A concurrent report by a separate group also identified flanking markers to the VHL gene. Hosoe et al gave evidence that VHL mapped between the D3S18 (LIB36-42) and the RAF1 loci. With the exception of the RAF1 locus, the markers used were not shared between the two studies, making comparisons difficult. Recently, Liu et al showed that the D3S18 locus was contained in or overlapped the D3S719 locus suggesting the relative order of cen-RAFI1-(VHL/D3S720)-D3S95-(D3S719/D3S18)-tel. Further confirmation of the sublocalisation of VHL to this region was reported by Maher et al who positioned VHL between RAF1 and D3S225 (LIB426), a probe that maps distal to D3S18. All groups agreed on the RAF1 gene as the closest proximal marker to VHL and, although they concurred on the general location of VHL on 3p, fine localisation was not possible. This was undoubtedly in part because of the lack of informativeness of the markers used in the analysis.

The purpose of the present study is threefold: first, to construct a high resolution genetic map of the region by genotyping a series of five newly identified highly polymorphic dinucleotide markers in the Centre d’Etude Polymorphism Humain (CEPH) reference families; second, to orient these markers relative to the consensus proximal flanking marker, RAF1; and third, to examine a set of VHL families for linkage to the RFLP and microsatellite markers in this region. These analyses will provide the framework for the fine mapping of the VHL locus in the 3p25 region with the eventual cloning of the gene as the ultimate goal.

Materials and methods
FAMILY DATA
Patients and families with VHL were ascertained through the Duke Neurogenetics Clinic and through referral from collaborating physicians. All participating family members and key dead subjects in the pedigrees were examined by or had their medical records reviewed by a neurologist (JMV). Criteria for diagnosis were based on Melmon and Rosen and Huson et al, and diagnosis was made in patients with (1) retinal angioma and CNS haemangioblastoma; (2) multiple CNS haemangioblastomas; (3) CNS or retinal lesion with phaeochromocytoma or renal cell carcinoma; and (4) single retinal, CNS, phaeochromocytoma, or renal cell carcinoma in an at risk
family member. At least one member of every family had pathological confirmation of diagnosis. Blood was obtained for lymphoblastoid transformation and for DNA extraction using standard methods on 92 family members (22 affected subjects) from six multiplex VHL pedigrees (table 1). (Pedigrees are available on request.)

**DNA MARKER GENOTYPING**

The RFLP loci (RAFI, D3S720, D3S18) were genotyped in the VHL family data using standard methodology. In addition, D3S720 was also genotyped in the CEPH family DNA. The RAFI locus was genotyped in the VHL families for both the TaqI (6-8 kb and 6-3 kb) and BglII (4-0 kb and 3-3 kb) polymorphisms; D3S720 was genotyped in both the VHL and CEPH families, for the two StuI polymorphisms (4-8 kb, 4-5 kb, and 2-5 kb, 2-2 kb); and D3S18 was genotyped in the VHL families for the two DraI polymorphisms (3-8, 1-5 kb and 2-8 kb, 2-5 kb). Haplotypes were constructed and the frequencies used in the analyses were as previously reported or were calculated from the CEPH parents. RAFI was genotyped in this laboratory as well as extracted from the CEPH version 4 database. The D3S18 data were extracted from CEPH without error correction.

The dinucleotide repeat markers (D3S1038 (C13-946), D3S651 (C13-312), D3S1110 (C13-1169), D3S656 (C13-326), and D3S1255 (DR7K)) were genotyped in the original 40 CEPH pedigrees as well as the VHL families. The origins of D3S1038, D3S651, D3S1110, and D3S656 have been described elsewhere. D3S1255 was derived from an EMBL library of a chromosome 3 radiation hybrid mapped to this region. The microsatellite data were generated as previously described and gene frequencies were generated as noted above. The incorporation of a multiplex genotyping approach allowed two or three primer pairs to be analysed simultaneously using 30 ng of DNA, accelerating the analysis, and producing 50 to 160 genotypings on a single sequencing gel. All marker, clinical, and family history data for both the CEPH and the VHL families were processed for storage and analysis via the PEDIGENE system.

**LINKAGE ANALYSIS**

**CEPH reference mapping**

The program MAPMAKER was used to construct the initial map of the region incorporating the five dinucleotide markers. All possible orders of the five markers were examined in order to obtain the most likely order of the markers and the associated likelihoods. The computer program CHROMLOOK was used to examine the marker data for possible laboratory data error by detecting all crossover events consistent with the most likely order of markers. All recombination events were re-genotyped in an effort to eliminate sources of error in the genetic map construction and estimation of genetic distances. After several rounds of error correction, map generation was repeated as described above. Once the final order was obtained, differences in recombination between males and females in this region were tested by comparing the likelihood of the data, allowing for sex differences in recombination (H2) to that assuming equal recombination between males and females (H0). Twice the difference in ln likelihoods is asymptotically distributed as a χ² distribution with n degrees of freedom, where n is the difference in the number of recombination fractions estimated under H0 versus H2.

The genetic map was then regeneratated using the RFLP markers together with the dinucleotide repeats. The results were integrated in order to generate as complete a data set as possible. All possible orders of the markers and subsets of the markers were examined using MAPMAKER until a set of loci were found that generated support for order of at least 1000:1 over the next most likely order. The data were again analysed with CHROMLOOK and crossover subjects were retyped for the dinucleotide markers before the calculation of the final map and associated genetic distances. Those markers that could not be absolutely placed with 1000:1 odds were then regionally localised on the genetic map. The MAPMAKER and CHROMLOOK calculations were performed on a SUN 4/370 workstation.

**VHL family linkage analysis**

VHL was analysed as an autosomal dominant trait with age dependent penetrance. A gene frequency of 0.0001 was assumed for the VHL gene. At risk subjects with a normal medical history and examination were included in the analysis by assuming probabilities of affected status using an age of onset correction with eight age dependent classes. The age correction was based on the data of Go et al and our own data. A penetrance of 93% was assumed for all subjects at risk over 50 years of age. The MLINK subprogram of the LINKAGE program package (version 5.1) was used for the two point lod score analysis. The LINKMAP subprogram of LINKAGE was used in the multipoint analysis of the family data. Multipoint location scores were calculated as the log₁₀ likelihood difference of VHL placed at a relative position on a fixed map of loci versus the unlinked state. As the dinucleotide markers were multiallelic systems, those dinucleotide repeats used in the multipoint analysis of the family data were recoded to four allele systems, in order to ease the computational burden of the analysis. The two point

**Table 1 Von Hippel-Lindau pedigree data.**

<table>
<thead>
<tr>
<th>Family members</th>
<th>No of family members</th>
<th>No of family members sampled</th>
<th>No of affected</th>
<th>No of affected sampled</th>
</tr>
</thead>
<tbody>
<tr>
<td>DUK606</td>
<td>21</td>
<td>20</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>DUK771</td>
<td>28</td>
<td>21</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>DUK1073</td>
<td>41</td>
<td>29</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>DUK1340</td>
<td>6</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUK1345</td>
<td>9</td>
<td>9</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>DUK1393</td>
<td>16</td>
<td>7</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>121</td>
<td>92</td>
<td>33</td>
<td>22</td>
</tr>
</tbody>
</table>
results from the recoded data were compared to the two point analysis incorporating all the alleles in order to ensure accuracy of the recording so that the multipoint analysis did not result in loss or alteration of the linkage information.

**Results**

**GENETIC MAP**

All possible orders of the five microsatellite orders were examined. The resulting order after extensive error checking, D3S1038-D3S651-D3S656-D3S1110-D3S1255, was favoured over the next most likely order, with a log$_{10}$ likelihood difference of 8·0 (that is, odds >100,000:1). Analysis of the RAF1 locus in conjunction with the dinucleotide markers, examining all possible orders of the six markers, resulted in the order D3S1038-RAFI-D3S651-D3S656-D3S1110-D3S1255 as the most likely order with odds of 3000:1 over the next best order. Evidence for a significant difference in sex specific recombination was found ($\chi^2 = 16·0$, $p < 0·01$). The resulting male and female maps are depicted in fig 1. The D3S720 and D3S18 loci could not be mapped with greater than 1000:1 odds in relation to the other markers in the map. However, they were both localised distal to RAF1 in the region of D3S1038 with odds of 1000:1.

**VHL PEDIGREE DATA**

Two point lod scores for VHL versus the family data are given in table 2. VHL linkage analysis with D3S1038 and D3S656 resulted in the highest lod scores among the markers with $Z(\theta) = 6·04$ and $Z(\theta) = 7·02$, respectively, at a maximum recombination fraction ($\theta = 0·0$). Owing to computational limitations, multipoint analysis was limited to three marker loci in addition to the disease locus. D3S656 and D3S1110 were chosen for inclusion in the map as they are the most informative flanking markers proximal to the RAF1 locus. D3S1038 was used as it mapped distal to RAF1. The distances used in the analysis were based on the male and female genetic maps presented in fig 1. The results for the multipoint analysis (fig 2) indicate the region around the D3S1038 locus as the most likely location for VHL, with a peak location score of 10·04 for VHL completely linked to D3S1038. Owing to the significant sex differences in recombination, the data are depicted relative to the male map in this region. This location was over 10$^3$ times more likely than placing VHL proximal to D3S656, but could not be distinguished statistically from the region just distal to D3S1038 or between D3S1038 and D3S656. A potentially key recombinant subject was identified in DUK771 (fig 3). In subject 9001, the markers D3S720 and D3S1038 and VHL segregate together resulting in a crossover with the marker D3S1255. The order of D3S720 relative to D3S1038 cannot be determined. RAF1, D3S18, D3S656, and D3S110 were not informative in this mating. A second crossover was observed in an unaffected subject, 9006. Subject 9006 is 22 years old and at the time of last examination had no signs or symptoms of VHL. This subject’s probability of being phenotypically normal at 22 years of age while having inherited the VHL genotype is approximately 60%. The loci D3S1038, D3S656, and VHL (normal allele) crossover from the loci D3S1110 and D3S1255. D3S720 and RAF1 were not informative.

**Discussion**

The fine mapping of the gene for VHL is the next step towards the identification of the basic genetic defect. The fine mapping of this disorder has progressed slowly, despite the availability of extensive family material. This was both because of the paucity of available markers in this region and their relative uninformativeness. Although the recently published map by Torp et al. for the short arm of chromosome 3 lists a number of markers mapped distal to the RAF1 gene, the average heterozygosity of the markers included in the manuscript was only 49%. In the present data set, the increase in information gained using the microsatellite markers is realised in both the two point and multipoint analyses. Simulation results for the VHL pedigree data using a four allele marker system approximating a heterozygosity of 0·70 resulted in a summed maximum potential lod score of approximately

---

**Table 2** Summary lod scores for the six von Hippel-Lindau families.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Recombination fraction (θ)</th>
<th>Maximum</th>
<th>Z(θ)</th>
<th>8 males = 9 females.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0·00</td>
<td>0·05</td>
<td>0·10</td>
<td>0·20</td>
</tr>
<tr>
<td>RFLPs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D3S18</td>
<td>3·47</td>
<td>3·01</td>
<td>2·59</td>
<td>2·18</td>
</tr>
<tr>
<td>D3S720</td>
<td>3·55</td>
<td>3·26</td>
<td>2·93</td>
<td>2·56</td>
</tr>
<tr>
<td>RAF1</td>
<td>2·12</td>
<td>2·01</td>
<td>1·86</td>
<td>1·68</td>
</tr>
<tr>
<td>Microsatellites</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D3S1038</td>
<td>6·04</td>
<td>5·31</td>
<td>4·58</td>
<td>3·88</td>
</tr>
<tr>
<td>D3S651</td>
<td>1·19</td>
<td>1·16</td>
<td>1·06</td>
<td>0·93</td>
</tr>
<tr>
<td>D3S656</td>
<td>7·02</td>
<td>6·35</td>
<td>5·63</td>
<td>4·88</td>
</tr>
<tr>
<td>D3S1110</td>
<td>∞</td>
<td>2·92</td>
<td>2·93</td>
<td>2·73</td>
</tr>
<tr>
<td>D3S1255</td>
<td>∞</td>
<td>2·84</td>
<td>2·74</td>
<td>2·48</td>
</tr>
</tbody>
</table>
10-00. In the two point case (table 2), the most informative RFLP (D3S720) realises only a fraction (Z = 3.55) of the potential meiotic information available in the family data versus using the microsatellite markers D3S1038 (Z = 6.04) and D3S656 (Z = 7.02). In the multipoint analysis case using the microsatellite markers, virtually all meiotic information is captured (fig 2).

The ability to generate highly polymorphic marker systems with heterozygosities greater than 50% is now a reality.22 The markers used in the present analysis were chosen both for their physical localisation in the region of the VHL gene as well as for their high informativeness. With the exception of the D3S651 locus (Het = 0.34), the heterozygosities of the markers tested were well over 50%. Of particular note is D3S1038, the marker which localises distal to RAFI in the VHL region, whose heterozygosity is 0.80.9 The sublocalisation of the D3S1038 locus in this region together with its high informativeness make it a critical marker for genotyping in all available VHL family data, but in particular those crossover subjects that have been reported for the D3S18 and RAFI loci.157 Crossover subjects such as those depicted in fig 3 need to be identified and fully characterised by all available markers in this region in an effort to localise the VHL gene further.

The use of linkage analysis for carrier detection studies in VHL is of clinical significance. The extent of clinical surveillance necessary on a routine basis in order to assure a person of the absence of life threatening symptoms is very time consuming and expensive. Many patients for just these reasons fail to undergo the necessary battery of tests. The ability to use linkage analysis in carrier detection is a well established procedure.24 Carrier detection in VHL is at present possible but limited owing to the relative un informativeness of the RFLP markers. The availability of highly informative microsatellite markers will make the application of such studies both more informative as well as less tedious to perform. The identification of asymptomatic VHL gene carriers will allow for early identification and hence early prevention. It will also indicate those subjects who have a high probability of not carrying the VHL gene (> 99%), and hence on whom surveillance can be relaxed. Care must be exercised though in the application of linkage to genetic counselling owing to the confounding factor of heterogeneity. Al-

Figure 3 VHL family DUK771 with crossover shown in subject 9001 as (D3S720-D3S1038-VHL) XO (D3S1255), and subject 9006 as (D3S1038-VHL (normal alleles)-D3S656) XO (D3S1110-D3S1255). The order of D3S720 with respect to D3S1038 is not known. The sexes are masked to provide confidentiality.
though to date there has been no documented evidence for heterogeneity in VHL, Seizinger et al.\(^7\) did report a single pedigree that excluded the region between *RAFI* and *D3S95* but showed positive lod scores for markers on both the distal and proximal of these loci. The exact interpretation of these data were inconclusive. In general, all other families studied to date show linkage to 3p25–26.\(^7\)\(^8\)

Our results indicate a significant sex difference in recombination in the genetic map. These data were incorporated into the multi-point analysis. Previous maps of this region using the RFLP data\(^9\) failed to give evidence for significant differences in male and female recombination, although the ratio of the male to female maps previously reported is similar to that found in the present study.\(^9\) The significance of our findings could be the result of the increased heterozygositites of the markers used in the analysis. Multipoint analysis of the VHL family data assuming equal recombination between males and females (data not shown) did not change the results as to the sublocalisation of the VHL gene.

In summary, we present an initial microsatellite map of the 3p25 region in the area of the VHL gene. This map will be useful for studies of both VHL and in the general mapping of chromosome 3p. The highly polymorphic marker *D3S1038* maps distal to *RAFI* in the CEPH reference pedigrees and thus represents a critical marker for future studies in an effort to identify the basic defect in VHL. Concurrent studies are under way to identify VHL families with additional informative crossovers, as well as the development of further dinucleotide repeats mapping to the *RAFI*-*D3S1038*-*D3S719* region. This will allow further meiotic mapping to narrow the region flanking the VHL gene.

The authors wish to thank Tina Brantley for her expertise in preparing the manuscript and Carol Haynes and Marc Speer for computational support, Peggy Pate and Helen Harbett for data entry, and the DNA Bank personnel. This work was supported by NINDS grants NS26330 (MAP-V), RO1 NS29416 (JMV), KO8 NS01289 (JMV), RO1 HG00358 (EW, HD), and March of Dimes grant 1-1-227 (DS).


doi: 10.1136/jmg.30.6.487

Updated information and services can be found at:
http://jmg.bmj.com/content/30/6/487

**Email alerting service**

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

**Notes**

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

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

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