Mutation screening in British 21-hydroxylase deficiency families and development of novel microsatellite based approaches to prenatal diagnosis

Majlinda Lako, Simon Ramsden, R Duncan Campbell, Tom Strachan

Abstract

21-hydroxylase deficiency is a recessively inherited disorder of steroidogenesis, resulting from mutations in the CYP21 gene. This 3.5 kb gene and a highly related CYP21P pseudogene reside on tandemly duplicated 30 kb segments of DNA in the class III HLA region, and the great majority of pathogenic mutations result from sequence exchanges involving the duplicated units. We now describe a comprehensive survey of CYP21 mutations in the British population, encompassing a screen for 17 different mutations in a total of 284 disease chromosomes. The most common mutations were as follows: large scale deletions/ conversions (45% of the affected chromosomes), the intron 2 splice mutation (30.3%), R357W (9.8%), and I172N (7.0%). Mutations were detected in over 92% of the chromosomes examined, suggesting that accurate DNA based diagnosis is possible in most cases using the described strategy. In order to extend highly accurate prenatal diagnosis to all families where samples are available from a previously affected child, we have developed a linkage analysis approach using novel, highly informative microsatellite markers from the class III HLA region.

Keywords: 21-hydroxylase deficiency; CYP21; mutation screening

Steroid 21-hydroxylase deficiency is one of the more common recessive disorders; the severe classical congenital form, occurs with a frequency in white populations of about 1/10 000, while milder non-classical forms, including late onset development of symptoms, are much more common. The disorder manifests as a defect of adrenal steroidogenesis and is characterised by excessive virilisation. The biochemical basis is an abnormal accumulation of 17-hydroxyprogesterone (a major substrate for 21-hydroxylase) which leads to increased metabolic flux in a parallel pathway cumulating in androgen synthesis. Approximately 30% of classical cases show simple virilising features and the remaining 70% or so show additional “salt wasting”, a potentially fatal inability to conserve dietary sodium owing to a parallel defect in aldosterone synthesis. Although both the salt wasting and excessive virilisation components are treatable, difficulties with management of the disease have fuelled a high demand for prenatal diagnosis and early termination of pregnancy.

The different phenotypic variants of the disorder result from mutations in a 3.5 kb gene, CYP21 (formerly called CYP21B), which maps within the HLA complex at 6p21.3. The gene resides on an about 30 kb segment which contains a complement C4 gene and other transcription units, and which has been duplicated in recent evolutionary history. As a result, there is a very high degree of sequence identity between the duplicated 30 kb segments. The duplication has resulted in an additional 3.5 kb long CYP21-like sequence, formerly called CYP21A but more recently termed CYP21P to reflect its presumed status as a pseudogene. The two CYP21 genes show approximately 98% sequence homology, but the capacity of the CYP21P gene for encoding a functional 21-hydroxylase polypeptide has been eroded by the accumulation of a series of deleterious mutations spanning its length.

The very close similarity in sequence of the tandemly duplicated 30 kb segments which contain the CYP21 and CYP21P genes predisposes towards frequent unequal chromosome pairing in this region. Subsequent crossover events can lead to loss or gain of a 30 kb segment and to gene conversion-like sequence exchanges between the duplicated genes. As a result, the vast majority of CYP21 pathogenic mutations involve loss of some essential CYP21 gene sequence: either there is a 30 kb deletion which removes all, or more frequently a portion, of the original CYP21 gene sequence, or there is a gene conversion-like event whereby a small amount of the original CYP21 sequence is replaced by a similar but defective sequence copied from the CYP21P pseudogene. Direct evidence for 30 kb deletions has been obtained by pulsed field gel electrophoresis and such deletions are thought to arise as a result of meiotic unequal crossovers, whereas the gene conversion-like events may arise more frequently in mitotic cells. Where examined, de novo gene conversion events giving rise to 21-hydroxylase deficiency appear to involve conversion of only a few hundred nucleotides. Disease chromosomes bearing a CYP21P pseudogene and a CYP21 gene which bears CYP21P specific sequences over a large component of its length, so called large scale “gene conversion” haplotypes, are also known. Large scale conversion at...
Table 1  Point mutations in CYP21 and primers used for their genotyping

<table>
<thead>
<tr>
<th>Exon (Ex) or intron (Int) number</th>
<th>Nucleotide position (in bold; mutant underlined)</th>
<th>Sequence changes</th>
<th>Protein change</th>
<th>Primers used in the first PCR‡</th>
<th>Size of product (bp)</th>
<th>Primers used in the second PCR‡</th>
<th>Size of product (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ex 1</td>
<td></td>
<td></td>
<td></td>
<td>P5+P48</td>
<td>1009</td>
<td>P92/P92+P48+P48 (+P70)</td>
<td>625</td>
</tr>
<tr>
<td>Int 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1385</td>
<td>P1004/P1004A+P55 (+P19)</td>
<td>322</td>
</tr>
<tr>
<td>Ex 4</td>
<td></td>
<td></td>
<td></td>
<td>P47+P4</td>
<td>2064</td>
<td>P2113/P2113+P55 (+P11)</td>
<td>706</td>
</tr>
<tr>
<td>Ex 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2064</td>
<td>P1999/P1999+P55 (+P11)</td>
<td>706</td>
</tr>
<tr>
<td>Ex 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2064</td>
<td>P2113/P2113+P55 (+P11)</td>
<td>706</td>
</tr>
<tr>
<td>Int 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2064</td>
<td>P1784/P1784C+P55 (+P11)</td>
<td>1101</td>
</tr>
<tr>
<td>Ex 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2064</td>
<td>P1990C/P1990+P55 (+P11)</td>
<td>1317</td>
</tr>
<tr>
<td>Ex 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2064</td>
<td>P2113/P2113+P55 (+P11)</td>
<td>1430</td>
</tr>
<tr>
<td>Ex 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2064</td>
<td>P2344A/P2344A+P55 (+P11)</td>
<td>1600</td>
</tr>
<tr>
<td>Int 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2064</td>
<td>P2584A/P2584A+P55 (+P11)</td>
<td>455</td>
</tr>
<tr>
<td>Ex 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2064</td>
<td>P2675G/P2675G+P55 (+P11)</td>
<td>165</td>
</tr>
<tr>
<td>Ex 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2064</td>
<td>P2675G/P2675G+P55 (+P11)</td>
<td>165</td>
</tr>
</tbody>
</table>

Primer names are as described by Wedell and Luthman. Notations for frameshift mutations, splice mutations, and allele-specific primers are indicated in pairs with the normal sequence represented by bold lettering at the end of the primer name while the mutant sequence specific primer ends with bold and underlined lettering. Control primers are indicated in brackets (see Methods).

Table 2  Novel microsatellite markers in the class III HLA region

<table>
<thead>
<tr>
<th>Microsatellite</th>
<th>Range (bp)</th>
<th>Repeat sequence</th>
<th>PCR primers (5’–3’)</th>
<th>Allele number</th>
<th>Heterozygosity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>62</td>
<td>146–197</td>
<td>(TC)n, (CT)n</td>
<td>TCCAATCTCAGCTCTGCTGCCCT TCGAATTCTCAGCCAGCGGG CAGGGGCGAGGGGAGCCACG AGACCCCTGCTCCTGCTCCT</td>
<td>14</td>
<td>82.4</td>
</tr>
<tr>
<td>82–1</td>
<td>86–113</td>
<td>(CA)n, (GAGG)n</td>
<td>CAGGAGGGCAGGGGAGCCACG AGACCCCTGCTCCTGCTCCT</td>
<td>13</td>
<td>81.3</td>
</tr>
<tr>
<td>9N–2</td>
<td>98–120</td>
<td>(CA)n</td>
<td>GCCGTCTCAGGAGGGGAGCCACG AGACCCCTGCTCCTGCTCCT</td>
<td>7</td>
<td>72.3</td>
</tr>
<tr>
<td>D3A</td>
<td>113–144</td>
<td>(CA)18-(GAGG)</td>
<td>AAGCAGGGAGGGGAGCCACG AGACCCCTGCTCCTGCTCCT</td>
<td>6</td>
<td>75.9</td>
</tr>
<tr>
<td>LHI–1</td>
<td>79–103</td>
<td>(CA)n</td>
<td>GCCGTCTCAGGAGGGGAGCCACG AGACCCCTGCTCCTGCTCCT</td>
<td>11</td>
<td>71.0</td>
</tr>
</tbody>
</table>
Table 3  Distribution of point mutations and deletions in 284 disease chromosomes in British 21-hydroxylase deficiency families

<table>
<thead>
<tr>
<th>Type</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large scale deletions/S' conversions</td>
<td>45.0 %</td>
</tr>
<tr>
<td>Intron 2 splice mutation</td>
<td>30.3 %</td>
</tr>
<tr>
<td>R357W</td>
<td>9.8 %</td>
</tr>
<tr>
<td>H172N</td>
<td>7.0 %</td>
</tr>
<tr>
<td>Unidentified mutations</td>
<td>7.8 %</td>
</tr>
</tbody>
</table>

neonates and confirmed by standard hormone assays, notably of serum 17-hydroxyprogesterone following intravenous administration of ACTH, as described previously. Blood samples obtained with informed consent were used to prepare standard or high molecular genomic DNA, using conventional methods. In some cases genomic DNA samples were referred from UK laboratories who had transmitted requests for prenatal diagnosis for the condition.

Figure 1 Genotyping of a 21-hydroxylase deficiency family for the intron 2 splice site mutation and the R357W mutation. (A) ARMS based typing of the intron 2 splice mutation. Normal alleles carry either an A or a C at the mutation site; mutant alleles have a G. For each individual the types of allele which have been detected are shown at the top, with the mutant allele underlined. 1=father, 2=mother, 3=child. The 423 bp product is a G. For each individual the types of allele which have been detected are shown at the top. R357W mutation. Normal alleles carry either an A or C at the mutation site; mutant alleles have a T. For each individual the allele types which have been detected are shown at the top with the mutation site. (B) ARMS based genotyping of the intron 2 splice site mutation. The presence of a 1430 bp product after the second amplification with P55 and P2114T indicates the presence of at least one normal allele, while the presence of a 86 bp product was obtained with the P48 and P659 primers (table 1). The presence of an 86 bp product indicates the presence of at least one mutated allele. In this case, the father carries two normal alleles (genotype A, G), the mother carries one normal and one mutated allele (genotype A, C), while the child carries one normal and one mutated allele (genotype C, T).
Normal alleles carry a T at the mutation site, mutant alleles carry an A. For each person, the allele types which have been detected are shown at the top.

Figure 2. Microsatellite typing and mutation screening for a 21-hydroxylase deficiency family. (A) ARMS based genotyping of the I172N mutation. Lako, Ramsden, Campbell, et al.

MICROSATELLITE TYPING

Genomic DNA was amplified with the oligonucleotide pairs shown in table 2.35 The forward primer was end labelled with 32P-ta DNA ligase (Amersham International plc) and T4 polynucleotide kinase (New England Biolabs). PCR was carried out in a thermocycler in a 10 µl volume containing 50-100 ng of genomic DNA template, 0.5 µmol/l of each primer, 50 µmol/l of each dATP, dCTP, dGTP, dTTP, 1 unit of Taq polymerase, and 1 x Taq polymerase buffer (Promega). Samples were overlaid with 20 µl mineral oil. The (CA) repeat microsatellites were amplified through 30 cycles, each consisting of 30 seconds at 96°C, 15 seconds at 62°C, and 30 seconds at 72°C. The last elongation step was carried out for a further three minutes at 72°C. Three µl of the amplified products were mixed with 3 µl of dye and electrophoresed in denaturing polyacrylamide gel containing 6% acrylamide and 8 mol/l urea at room temperature at 50 W for two to three hours. Gels were fixed, dried, and exposed at −70°C for 10 hours.

Results

MOLECULAR CHARACTERISATION OF PATHOGENIC MUTATIONS IN BRITISH 21-HYDROXYLASE DEFICIENCY PATIENTS

Southern blot analysis was used to detect CYP21 gene deletions or large 5' CYP21 "gene conversions" and ARMS based allele specific amplification was used to screen for 16 known point mutations in 284 disease chromosomes in British 21-hydroxylase deficiency families (see Methods). Based on these two methods, the disease mutation was established in 92.3% of the disease chromosomes (table 3). As established by a combination of pulsed field gel electrophoresis and conventional Southern analyses in a previous smaller study,33 large scale deletions and large scale "gene conversions" are very common in the UK population, and the present study endorsed this finding. The diagnostic 3.7 kb Taq restriction fragment associated with the CYP21 gene was absent in 45% of the affected chromosomes (table 3). Of the various point mutation classes observed, there was a very frequent representation of the established intron 2 splice site mutation, being found in nearly one third of the affected chromosomes (table 3). All other mutations tested did not appear to be evident in the chromosomes sampled, leaving a total of 7.8% of disease chromosomes where the pathogenic mutation could not be established by the tests used.

USE OF NOVEL MICROSATELLITE MARKERS IN THE CLASS III HLA REGION IN PRENATAL DIAGNOSIS OF 21-HYDROXYLASE DEFICIENCY PATIENTS

The above mutation screening protocol identified pathogenic mutations in over 92% of the disease chromosomes sampled. In order to extend accurate prenatal diagnosis to all 21-hydroxylase deficiency families where samples are available from a previously affected child, indirect linkage analyses using extremely closely linked markers are required. Ten highly polymorphic (CA)n/(TG)n microsatellite markers have very recently been identified in the class III HLA region36 and in the current study we have tested five of the new microsatellite markers, LH-1, D3A, 9N-2, 62, and 82-1 (table 2), by typing 50 unrelated subjects and 26 subjects belonging to six families with a previously affected child. Three of the markers are located distal to the CYP21 gene and two are proximal, with the distance between the marker and CYP21 varying from 100 to 400 kb (fig 3). The distribution and frequencies of different alleles calculated from the study of 50 unrelated subjects are presented in table 2. The segregation of alleles defined by microsatellite typing was always found to correspond with segregation of mutations defined by direct
analysis of the CYP21B gene. Fig 2 shows one example. In this case ARMS screening identified both parents to be heterozygous carriers of the I172N mutation, while the CVS sample and a sample from the previously affected child were shown to be homozygous for this mutation, confirming the result obtained by microsatellite typing. The use of such highly polymorphic very closely linked markers should permit accurate prenatal diagnosis in essentially all 21-hydroxylase deficiency families where a sample is available from a previously affected child.

Discussion
The molecular pathology of 21-OH deficiency is, in very large measure, the result of sequence exchanges between the duplicated CYP21 and CYP21P loci. As a result of this, and the very high extent of sequence homology between these genes, there are comparatively few types of pathogenic mutation,11 12 prompting the diagnostic use of direct mutation typing. In addition to allele specific oligonucleotide hybridisation approaches, rapid PCR based methods have been used to identify point mutations in the CYP21 gene by selective amplification of CYP21 sequences using locus specific primers, followed by ARMS based allele specific amplification15 17 or amplification created restriction site approaches.18 19 The current study, in addition to testing for deletions, was designed to type for a total of 16 pathogenic point mutations. Some of these were known to be associated with the non-classical form of 21-hydroxylase deficiency11 12 (table 1) and would not have been expected in classical disease chromosomes, and some represented rare point mutations which have not been introduced from the CYP21P.20 21 The single largest mutation class was the large scale deletion and large scale 5' CYP21 gene conversion category which accounted for over 40% of the disease chromosomes sampled. Although Southern blotting of conventional TaqI digests cannot readily distinguish large scale deletions from large scale conversions, direct quantitation of deletions is possible using PFGE based BsuHII restriction mapping and has suggested a ratio of 3:1 for large scale gene deletion to large scale gene conversion events in the British population.22 European populations often show broadly similar frequencies of gene deletion/large scale 5' CYP21 gene conversion, but in other populations the reported figure may be quite low.23 The high frequency of the intron 2 splice site mutation described in the present report also parallels that observed in several other population surveys, but the R357W mutation appears particularly frequent in the British population when referenced against other reports.

Although in the present study deletions and large scale conversions were sampled by Southern blot hybridisation typing, rapid PCR based alternatives for screening for such mutations could also be used. In this case primers can be designed to amplify simultaneously equivalent segments of the CYP21 and CYP21P genes at a position immediately upstream of the coding sequence where the presence (in CYP21P) or absence (in CYP21) of a TaqI site provides the basis for the diagnostic 3.2 kb and 3.7 kb TaqI RFLPs. Subsequent TaqI digestion of the CYP21/CYP21P amplification product will yield additional small digestion products in the presence of the CYP21 gene (false negative results can be discounted by adding an internal control of a DNA fragment known to possess a TaqI site resulting in additional diagnostic band sizes after TaqI digestion). Because of the ability to PCR type for deletions and large scale conversions and because these two mutation classes plus only three other point mutations (the intron 2 splice mutation, R357W, and I172N) account for the pathogenesis in over 92% of disease chromosomes in British 21-hydroxylase deficiency families, accurate DNA based diagnosis can be carried out simply and rapidly in the great majority of such cases. The pathogenic mutations in a significant proportion (7.8%) of disease chromosomes in the present study could not be identified, however, and this may suggest that there are several other classes of rare mutation, in addition to
the known ones that were typed in the present study. Therefore, while direct mutation typing may be useful in prenatal diagnosis in families where the pathogenesis has previously been determined, the prospect of prenatal diagnosis by rapid multiplex mutation screening without previous knowledge of the pathogenesis is not an attractive one. In addition, such multiplex typing may not always be straightforward and some PCR artefacts can lead to difficulty in interpretation, as in the case of the intron 2 splice mutation.34

The usefulness of alternative linkage based approaches for prenatal diagnosis is determined by the number of markers available for testing and their heterozygosity indices, and the accuracy of the tests depend on how closely linked the markers are, and on whether flanking markers are available. Traditionally, linkage analyses have relied heavily on class I and class II HLA gene markers,35 which are often disadvantaged by not being straightforward to use, and although located close to the CYP21 locus, a significant frequency of recombination with the disease locus is expected. During this study, however, 10 new highly polymorphic microsatellite markers were identified within 100–400 kb from the CYP21 gene36 (fig 3, table 1) and we typed five of these in 50 unrelated subjects and 26 subjects from six families with at least one previously affected child with CAH. These highly informative markers have proved very simple and easy to apply, resulting in a novel, rapid approach to prenatal diagnosis of 21-hydroxylase deficiency which requires minimal amounts of DNA from the parents and a previously affected child.

We are grateful to the Wellcome Trust for financial support of this work and to A Wedell for providing positive control samples for screening of some point mutations.

3 Carroll MC, Campbell RD, Porter RR. Mapping of steroid 21-hydroxylase genes adjacent to complement component C4 genes in HLA, the major histocompatibility complex in man. Proc Natl Acad Sci USA 1985;82:1089-93.
4 White PC, Grossberger D, Osnufer BJ, et al. Two genes encoding steroid 21-hydroxylase are located near the genes encoding the fourth component of complement in man. Proc Natl Acad Sci USA 1985;82:1089-93.
7 Speiser PW, Barry F, Miller WL. Alternate promoters and 3' ends buried in other genes.
14 Tusie-Luna M, White PC. Gene conversions and unequal crossovers between CYP21 (steroid 21-hydroxylase) and CYP21P involve different mechanisms. Proc Natl Acad Sci USA 1995;92:10796-800.
33 Tusie-Luna M, White PC. Gene conversions and unequal crossovers between CYP21 (steroid 21-hydroxylase) and CYP21P involve different mechanisms. Proc Natl Acad Sci USA 1993;90:1272-3.
39 Tusie-Luna M, White PC. Gene conversions and unequal crossovers between CYP21 (steroid 21-hydroxylase) and CYP21P involve different mechanisms. Proc Natl Acad Sci USA 1991;88:10796-800.