Meta-analysis of genotype–phenotype correlation in X-linked Alport syndrome: impact on clinical counselling

Oliver Gross, Kai-Olaf Netzer, Romy Lambrecht, Stefan Seibold and Manfred Weber

Medical Clinic I, Cologne City Hospital, Cologne-Merheim, Germany

Abstract

Background. Alport syndrome (AS) is a hereditary nephropathy characterized by progressive renal failure, hearing loss and ocular lesions. Numerous mutations of the COL4A5 gene encoding the α5-chain of type IV collagen have been described, establishing the molecular cause of AS. The goal of the present study was to identify the genotype–phenotype correlations that are helpful in clinical counseling. COL4A5-mutations (n = 267) in males were analysed including 23 German Alport families.

Methods. Exons of the COL4A5 gene were PCR-amplified and screened by Southern blot, direct sequencing or denaturing gradient gel electrophoresis. Phenotypes were obtained by questionnaires or extracted from 44 publications in the literature. Data were analysed by Kaplan–Meier statistics, χ² and Kruskal–Wallis tests.

Results. Genotype–phenotype data for 23 German Alport families are reported. Analysis of these data and of mutations published in the literature showed the type of mutation being a significant predictor of end-stage renal failure (ESRF) age. The patients' renal phenotypes could be grouped into three cohorts: (1) large rearrangements, frame shift, nonsense, and splice donor mutations had a mean ESRF age of 19.8 ± 5.7 years; (2) non-glycine- or 3’ glycine-missense mutations, in-frame deletions/insertions and splice acceptor mutations had a mean ESRF age of 25.7 ± 7.2 years and fewer extrarenal symptoms; (3) 5’ glycine substitutions had an even later onset of ESRF at 30.1 ± 7.2 years. Glycine-substitutions occurred less commonly de novo than all other mutations (5.5% vs 13.9%). However, due to the evolutionary advantage of their moderate phenotype, they were the most common mutations. The intrafamilial phenotype of an individual mutation was found to be very consistent with regards to the manifestation of deafness, lenticonus and the time point of onset of ESRF.

Conclusions. Knowledge of the mutation adds significant information about the progress of renal and extrarenal disease in males with X-linked AS. We suggest that the considerable prognostic relevance of a patient’s genotype should be included in the classification of the Alport phenotype.

Keywords: Alport syndrome; genetic counseling; genotype–phenotype correlation; hereditary nephropathy; type IV collagen

Introduction

Alport syndrome (AS) is a hereditary nephropathy characterized by family history of haematuria, progressive renal failure, sensori–neural deafness and typical ocular changes [1,2]. The disease is caused by mutations in type IV collagen genes. Type IV collagen is a major constituent of basement membranes [3,4]. Six genetically distinct α(IV)-chains (α1–α6) have been identified, the corresponding genes of which are located pairwise on chromosomes X, 2, and 13. Each α(IV)-chain contains a C-terminal NC1-domain, a collagenous domain of Gly-X-Y repeats that form the triple-helix structure, and an N-terminal 7S-domain. While the α1- and α2-chains are ubiquitously found in basement membranes, the α3-, α4- and α5(IV)-chains show a restricted distribution and are specifically expressed in the glomerulus, inner ear, and eye [4]. There is evidence that mutations of COL4A5 alter or abolish expression of the α5(IV)-chain. This in turn leads to an abnormal basement membrane with decreased or absent α3- and α4-chains. How this initiates progressive nephritis and scarring observed in AS is not well understood.

To date, more than 300 mutations of all types have been described in the COL4A5-gene [3,5–47]. Valid statistical analysis of the genotype–phenotype relation
Genotype–phenotype in Alport syndrome

has been hampered by a rather variable Alport phenotype and a limited number of mutations reported in individual publications. The aim of this study therefore was to elucidate genotype–phenotype correlations in a review of all COL4A5 mutation studies available in the literature. We evaluated 256 mutations from 44 publications, which provided basic phenotypic information, including 12 mutations from our group. Furthermore, genotype–phenotype data of 11 German Alport patients with previously unpublished mutations were included into the analysis. We found the progression rate of renal disease and hence the age of end-stage renal failure (ESRF) were influenced by the type of the underlying mutation. In the cohort with the most severe mutations, which either abolished protein expression or led to truncated protein chains, patients reached ESRF at the age of 20 years. In contrast, the cohort with least severe mutations reached ESRF at about 30 years of age. The results illustrate the usefulness of molecular genetic testing in Alport syndrome as well as the need for high throughput DNA analysis in the future.

Methods

German Alport families

Patients’ data were obtained by standardized questionnaires from cooperating centres in Germany, Austria and Switzerland. Data included family history, haematuria, proteinuria, ESRF, kidney-transplants, ocular changes, deafness, hypertension, and additional symptoms (macrotomatozytina and leiomyomatosis). Ocular changes (lenticus anterior or posterior) were evaluated by consultation of ophthalmologists, sensori–neural deafness was documented by consultation of ENT-specialists. Data from more than 200 male Alport-patients were obtained. The diagnosis of AS was defined as being likely when two of the following four criteria were fulfilled: (i) sensori–neural deafness; (ii) typical ocular changes; (iii) positive family history; and (iv) typical histological changes of the glomerular basement membrane. One-hundred-and-four unrelated families were selected in which at least two of these four criteria were fulfilled, and screened for molecular changes of the COL4A5 gene.

DNA analysis

Genomic DNA was isolated from peripheral blood lymphocytes and Southern blot analysis was performed as described [6]. PCR amplification was performed in a total volume of 50 μl, using DNA 100 ng, 20 pM of each primer, MgCl₂ 2 mM, dNTP 0.2 mM and Taq Polymerase 1 U (Promega, Mannheim, Germany). Positions of primers for each exon are shown in Table 1. Primers marked ‘Ps’ contained a psoralen derivative for photo crosslinking prior denaturing gradient gel electrophoresis (DGGE) analysis [48]. PCR products were purified through filter columns (Mobitech, Göttingen, Germany) and sequenced on an ABI sequencer 373 using the PRISM dye deoxy terminator cycle sequencing kit.

For screening by DGGE, samples of two separate patients were mixed at 95 °C, cooled down to 55 °C, photo-crosslinked and loaded on a 6.8% polyacrylamide gel containing a linear gradient of denaturants (100% being equivalent to 40% formamide and 7 M urea) [48]. The gel was run in 1×TAE buffer (Tris 40 mM, EDTA 1 mM, pH 8.0) at 120 V/h at a temperature of 60 °C. DNA was visualized by silver staining (BioRad). When abnormal band-shifts were found, PCR and DGGE were repeated to exclude artifacts. Corresponding PCR products were purified and sequenced on an ABI sequencer as described above.

Selection of data from the literature

Data on more than 300 mutations were retrieved from 44 publications [3,5–47]. Mutations were excluded when data regarding two or more of the following markers were absent: hearing loss, ocular changes, family history, and changes in the glomerular basement membrane. By this process, 256 mutations could be included in this study.

Statistics

Data were analysed by χ² tests and two-way ANOVA. Data were stratified according to the type of mutation: 5’ glycine substitutions (class 1), 3’ glycine substitutions (class 2), in-frame mutations (class 3), splice donor mutations (class 4), splice acceptor mutations (class 5), frameshift and premature stop mutations (class 6), and large rearrangements (class 7). Data were then analysed by Kaplan–Meier statistics.

Results

Novel COL4A5 mutations

Twenty-three mutations were found in German families (30 exons screened) (Table 2). No particular hot spot was identified, and the mutations were unique to the respective families. Missense mutations comprised the largest fraction in the German study cohort (35%), followed by splice site mutations (26%) and small deletions/insertions (22%). Nonsense mutations and large rearrangements each accounted for 9%. In three out of 21 cases mutations occurred in patients with negative family history. Molecular analyses of family members proved that two of these mutations had occurred de novo.

Phenotypes in the German study cohort

Family history was positive in 18/21 patients (Table 2). The biopsy rate for the whole study cohort was 53%, which is well within the range reported in the literature (44–57%). In families with identified mutations, biopsy rate was not significantly different at 56% (13/21). Average age at onset of ESRF was 25 years, while six patients had not reached ESRF. When stratified according to mutation types (Table 2), mean ESRF ages were 29, 27, 17, 17 and 23 years, respectively, suggesting an impact of the mutation type on the renal outcome. The same stratification did not reveal any differences with regard to hearing loss, which was almost always present. Ocular lesions, however,
Table 1. Primers used for direct sequencing or denaturing gradient gel electrophoresis

<table>
<thead>
<tr>
<th>Exon</th>
<th>Forward primer</th>
<th>Reverse primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>7/8</td>
<td>7A GTTTCTTGTTCTCCATGCTC</td>
<td>8B-Ps CAAACATTGTTGTCCTCCCAAG</td>
</tr>
<tr>
<td>11/12/13</td>
<td>11A AATACATTTTTGATGGGC</td>
<td>R13B2 GCAAGATTTCATTGACTCTCC</td>
</tr>
<tr>
<td>14/15</td>
<td>14A CCCAGGTCTCAACATAGTGTG</td>
<td>R15B CATTATAATGCTACTGATGATAT</td>
</tr>
<tr>
<td>16</td>
<td>16A AGCTGTATTTGGATGGGC</td>
<td>R16B TTTTGTGCACTATGCTCTCCT</td>
</tr>
<tr>
<td>29</td>
<td>29A GGACGGAAAATCTGACCTGGA</td>
<td>30B-Ps GTCGCAAAATAATGCAATTTA</td>
</tr>
<tr>
<td>31</td>
<td>31A CTAGTGTCTGTATCTGACAGG</td>
<td>31B-Ps CTTAATCAACTGAGAAGAAA</td>
</tr>
<tr>
<td>32</td>
<td>32A CCAACCTCTAATGTTTCTCTG</td>
<td>32B-Ps CCTCTCAGATCGTCTGG</td>
</tr>
<tr>
<td>33</td>
<td>33A GCATTAAATTCATTGAGA</td>
<td>33B-Ps CCTCAGATATATCAGTCTGG</td>
</tr>
<tr>
<td>34-Ps</td>
<td>34A-TGATGTAGCTGTCTTTGCCC</td>
<td>34B-Ps TTCAGTGTCAGCTAAAGG</td>
</tr>
<tr>
<td>35-Ps</td>
<td>35A-CCATGGAACGAGAACAACCC</td>
<td>35B-Ps CTTTCCCATTTAATTGGGACT</td>
</tr>
<tr>
<td>36-Ps</td>
<td>36A-CTAACTCGAGTTTGGCAGGAG</td>
<td>36B-Ps ATTTCATATCTGCTCAG</td>
</tr>
<tr>
<td>38</td>
<td>38A GTAAATTTGGAATGGCGTCA</td>
<td>38B-Ps GTTAAATTCAACACAGAG</td>
</tr>
<tr>
<td>39-Ps</td>
<td>39A-CAGTTGTATTATCCACTTGAG</td>
<td>39B-Ps GGTGGAGATGGAAAAATAG</td>
</tr>
<tr>
<td>40-Ps</td>
<td>40A-CAGTGTATTAATCCTACTGAG</td>
<td>40B-Ps GGTGGAATGGAAAAATAG</td>
</tr>
<tr>
<td>41-Ps</td>
<td>41A-Ps GCTTGTACAGTTTGGGAG</td>
<td>41B-Ps GACCTAATCTCCTGCAAG</td>
</tr>
<tr>
<td>42-Ps</td>
<td>42A-Ps GCTTGTACAGTTTGGGAG</td>
<td>42B-Ps Eco CTCTCTTCCTCAGATAT</td>
</tr>
<tr>
<td>43/44</td>
<td>43A GAGTGGATCAGAGCTTACT</td>
<td>43B-Ps CCTTCTTGAGAAACGCTTGG</td>
</tr>
<tr>
<td>45</td>
<td>45A-Ps GGTCTACAGTTTGGGAG</td>
<td>45B-Ps CTGTATCCTTGCAATGATTC</td>
</tr>
<tr>
<td>46-Ps</td>
<td>46A-Ps TTCTACAGTTTGGGAG</td>
<td>46B-Ps CTGTCAGAGTCCTCATGCTA</td>
</tr>
<tr>
<td>47-Ps</td>
<td>47A-Ps GAGTGGATCAGAGCTTACT</td>
<td>47B-Ps GACCTAATCTCCTGCAAG</td>
</tr>
<tr>
<td>48-Ps</td>
<td>48A-CAGTGTACAGTTTGGGAG</td>
<td>48B-Ps GACCTAATCTCCTGCAAG</td>
</tr>
<tr>
<td>49-Ps</td>
<td>49A-GTCTCTTTATTCTCACT</td>
<td>49B-Ps GACCTAATCTCCTGCAAG</td>
</tr>
<tr>
<td>50-Ps</td>
<td>50A-TATGGCAATCGGTTGAGG</td>
<td>50B-Ps2 CATCCTCTGAAGAGCTTTG</td>
</tr>
<tr>
<td>51-Ps</td>
<td>51A-Ps TGTGGATCAGTTTGGGAG</td>
<td>51B-Ps2 CATCCTCTGAAGAGCTTTG</td>
</tr>
</tbody>
</table>

(Ps stands for Psoralen).

Table 2. Mutations identified in the German Alport study cohort

<table>
<thead>
<tr>
<th>Patient number</th>
<th>Mutation</th>
<th>Nucleotide position</th>
<th>Exon</th>
<th>Family history</th>
<th>Biopsy</th>
<th>Hearing loss</th>
<th>Ocular lesions</th>
<th>ESRF type and onset (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DE-014</td>
<td>Gly129→Val*</td>
<td>G588→T</td>
<td>7</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>J (29)</td>
</tr>
<tr>
<td>DE-044</td>
<td>Gly204→Asp</td>
<td>G813→A</td>
<td>11</td>
<td>+</td>
<td>nd</td>
<td>+</td>
<td>+</td>
<td>J (26)</td>
</tr>
<tr>
<td>DE-005</td>
<td>Gly292→Val*</td>
<td>G1077→T</td>
<td>15</td>
<td>+</td>
<td>nd</td>
<td>+</td>
<td>–</td>
<td>J (15)</td>
</tr>
<tr>
<td>DE-034</td>
<td>Gly307→Asp</td>
<td>G1122→A</td>
<td>16</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>A (43)</td>
</tr>
<tr>
<td>DE-232</td>
<td>Arg1563→Gln</td>
<td>G4890→A</td>
<td>48</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>no (26)</td>
</tr>
<tr>
<td>DE-125</td>
<td>Tyr1597→Cys*</td>
<td>A4992→G</td>
<td>49</td>
<td>nd</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>A (31)</td>
</tr>
<tr>
<td>DE-237</td>
<td>Cys1681→Tyr*</td>
<td>G5244→A</td>
<td>51</td>
<td>+</td>
<td>nd</td>
<td>+</td>
<td>+</td>
<td>A (31)</td>
</tr>
<tr>
<td>DE-311</td>
<td>Cys1678→Arg</td>
<td>T5234→C</td>
<td>51</td>
<td>+</td>
<td>nd</td>
<td>+</td>
<td>–</td>
<td>no (12)</td>
</tr>
<tr>
<td>DE-024</td>
<td>Gln287→Stop*</td>
<td>G1061→T</td>
<td>15</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>J (17)</td>
</tr>
<tr>
<td>DE-216</td>
<td>Arg1647→Stop*</td>
<td>C5222→T</td>
<td>51</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>no (20)</td>
<td></td>
</tr>
<tr>
<td>DE-007</td>
<td>AgGGT→ggGGT*</td>
<td>a1094→2→g</td>
<td>16</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>A (39)</td>
</tr>
<tr>
<td>DE-006</td>
<td>CTTGAT−→CTT ... gt*</td>
<td>del gtaa1234+1,2,3,4</td>
<td>18</td>
<td>nd</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>J (28)</td>
</tr>
<tr>
<td>DE-030</td>
<td>TAGgt→TAGct*</td>
<td>g3656+1→c</td>
<td>38</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>J (16)</td>
</tr>
<tr>
<td>DE-116</td>
<td>CTGgt→CTGga</td>
<td>t3896+2→a</td>
<td>40</td>
<td>+</td>
<td>+</td>
<td>nd</td>
<td>–</td>
<td>no (5)</td>
</tr>
<tr>
<td>DE-192</td>
<td>AgGCC→tgGCC</td>
<td>a3807→2→g</td>
<td>41</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>J (28)</td>
</tr>
<tr>
<td>DE-101</td>
<td>AgGTC→aGTC</td>
<td>g4401→1→a</td>
<td>46</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>J (25)</td>
</tr>
<tr>
<td>DE-151</td>
<td>Pro10→Frameshift</td>
<td>del G</td>
<td>7</td>
<td>–</td>
<td>nd</td>
<td>+</td>
<td>+</td>
<td>J (17)</td>
</tr>
<tr>
<td>DE-037</td>
<td>Gly 254→Frameshift*</td>
<td>del AG</td>
<td>13</td>
<td>+</td>
<td>nd</td>
<td>+</td>
<td>+</td>
<td>J (15)</td>
</tr>
<tr>
<td>DE-139</td>
<td>Pro271→Frameshift</td>
<td>del C</td>
<td>14</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>no (8)</td>
<td></td>
</tr>
<tr>
<td>DE-031</td>
<td>Pro865→Pro871→In-frame</td>
<td>del 18 bp A2797→C2814</td>
<td>31</td>
<td>+</td>
<td>+</td>
<td>nd</td>
<td>–</td>
<td>J (18)</td>
</tr>
<tr>
<td>DE-293</td>
<td>Pro1399→Frameshift</td>
<td>ins C</td>
<td>45</td>
<td>de novo</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>no (7)</td>
</tr>
</tbody>
</table>

(30 out of 51 COL4A5 exons were screened. *Previously published mutations [6,9,46–48]; nd, no data; –, negative; +, positive.)
seemed to be more common in patients with large rearrangements and frameshift-causing small deletions/insertions, as compared with all other mutations (57 vs 25%).

**Analysis of published mutations**

In order to perform more valid statistical analyses, 44 publications on AS were screened [3,5–47]. Family history was positive in 84%. All de novo mutations in German families were proven to be de novo on a molecular basis. In the literature, no information was found concerning how the diagnosis of de novo mutations was made. Data regarding onset of ESRF were available in 169 patients, 82% of which suffered from juvenile onset (<30 years of age). Sensori–neural deafness was present in 73% of patients (n = 235). Data regarding ocular lesions were available in 49 patients, 57% of which had lenticular anterior.

For genotype–phenotype analyses, mutations were divided into five groups, depending on expected genetic damage: (1) glycine substitution, (2) in-frame and missense (other than glycine), (3) splice site, (4) frameshift and nonsense mutations (including small deletions/insertions) and (5) large rearrangements (Table 3).

**Large rearrangements**

Forty-four large rearrangements were described in the literature. Rearrangements lead to the juvenile type of AS in more than 95% (n = 27, P = 0.05). Seventy-seven per cent of patients had sensori–neural deafness. All five patients with sufficient data about ocular lesions had a lenticular.

**Missense mutations**

Eight single base exchanges were found in German patients (Table 2). Four of these were non-glycine missense mutations resulting in minor changes in the α5(IV)-chain within the NC1-domain, and in an adult onset of ESRF (Table 2). Four missense mutations resulted in changes of a glycine residue in the Gly-XY repeat sequences of the collagenous domain.

Glycine-XY mutations led to a highly significant later onset of ESRF (66 vs 90% in the juvenile type of AS, n = 56, P = 0.001) (Table 3). The zipper-like folding mechanism of the triple helix of type IV collagen is believed to start from the C-terminal end. In order to evaluate if the distance of the mutation from the C-terminal influences the phenotype, glycine-XY missense mutations (n = 98) were divided into two groups: (i) location within exon 1–20 of the α5(IV)-chain and (ii) location within exon 21–47. Mutations located in exons 1–20 influenced the phenotype in a significantly less severe manner than mutations located in exons 21–47 (55 vs 72% juvenile type of AS, n = 56, P = 0.05). Glycine-XY mutations also resulted in lower numbers of ocular changes (25 vs 73%, n = 16) and hearing loss (69 vs 75%, n = 90).

De novo mutations significantly involved the glycine-XY domain less often than other mutations (5.5 vs 13.9%, P = 0.05).

**Nonsense mutations**

Two nonsense mutations in German families led to a premature stop codon in exon 15 (DE-024) and 51 (DE-216). All 18 patients with nonsense mutations described in the literature suffered from a juvenile onset of ESRF and a high frequency of hearing loss and ocular lesions (data not significant).

**Donor and acceptor splice site mutations**

Six splice site mutations were found in the German families and 33 in the literature. Donor splice site mutations (n = 21) led to a high number of juvenile ESRF cases (94%), hearing loss (86%) and ocular lesions (80%). Acceptor splice site mutations (n = 18) resulted in a significantly lower number of juvenile cases of AS (63 vs 94%, n = 34, P = 0.05) and hearing impairment (61 vs 86%, n = 39, P = 0.05).

**Small insertions and deletions**

An insertion of one C was found in patient DE-293 resulting in a shift of the reading frame. His mother and other family members had a normal DGGE- and
direct sequencing-pattern indicating a de novo mutation. Three frameshift mutations and one in-frame mutation (deletion of 18 basepairs in family DE-031) were found. All patients suffered from hearing loss and juvenile onset of ESRF. Four out of five patients with in-frame mutations described in the literature had juvenile onset of ESRF (data not significant). Fifty-three frameshift mutations are described in the literature. As glycine-XY mutations influence the phenotype according to their distance from the C-terminal, frameshift mutations were also divided into two groups: (i) location within exon 1–30 and (ii) location within exon 31–51. No significant differences were found in the frequency of hearing loss (n = 53), ocular lesions (n = 12) or in onset of ESRF (n = 33).

Kaplan–Meier statistics (Figure 2) showed the type of mutation as being a significant predictor of the onset of ESRF (P < 0.0001). Splice acceptor and glycine-XY mutations (exons 1–20 and exons 21–47) on their own are significant predictors of the estimated time point of onset of ESRF.

Mutations were separated into three groups, according to their likely effects on protein structure as follows.

(i) Juvenile type of AS in >90% of patients: large rearrangements, premature stop, frameshift mutations and mutations involving the donor splice site or NC1-domain (n = 92) (‘truncated protein group’).

(ii) Juvenile type of AS in ~75% of patients: glycine-XY missense mutations of exons 21–47, in-frame mutations and mutations involving the acceptor splice site (n = 57) (‘altered protein structure group’).

(iii) Juvenile type of AS in ~50% of patients: glycine-XY mutations involving exons 1–20 (n = 20).

In all three groups, numbers of patients reaching ESRF differed significantly between 17 and 37 years of age (P = 0.01) as well as the numbers of patients with adult type of AS (truncated protein group: 7.5%; altered protein structure group: 24.6%; glycine-XY mutations of exons 1–20: 45%; P = 0.01) (Figure 3). Mean age when reaching ESRF differed significantly (19.8 vs 25.7 vs 30.1 years, respectively, P = 0.01). Seventy-five per cent of patients with a truncated z-chain reached ESRF with the age of 23, whereas only 40% of patients with an altered z-chain structure and 20% of patients with glycine-XY mutations of exons 1–20 have proceeded to ESRF at the same age (P = 0.001). The same tendency (not significant) exists regarding the onset and total number of hearing losses (data not shown).

In order to validate our previous findings, the phenotypic consistency of all affected members of German Alport families with known mutations was evaluated out of a total number of 45 Alport patients in 23 families. All 14 affected patients of eight different families within two different generations had lenticulo. Twenty-one families reported hearing loss; all 38 Alport patients in three different generations were affected. In 13 families two or more affected patients could be analysed for the consistency of the

### Table 3. Correlation between genotype and phenotype according to type and location of mutation

<table>
<thead>
<tr>
<th>Patient number</th>
<th>Mutation</th>
<th>Family history</th>
<th>Biopsy done</th>
<th>Hearing loss</th>
<th>Ocular lesions</th>
<th>ESRF juvenile onset</th>
</tr>
</thead>
</table>
| 30             | Gly-XY mutation Exon 1–20 | 93% | 44% | 70% | 2/8 (25%) | 11/20 (55%)*
| 68             | Exon 21–47 | 92% | 67% | 68% | 2/8 (25%) | 26/36 (72%)*
|                | Σ de novo | 3/55 (5.5%)* | Σ 62/90 (68.9%) | Σ 4/16 (25%)* | Σ 37/56 (66.1%)* |
| 20             | Nonsense | – | 90% | 47% | 85% | 4/6 (67%) | 12/12 (100%)
| 21             | Splice site Donor splice site | 81% | 52% | 61%* | 4/5 (80%) | 17/18 (94%)
| 18             | Acceptor splice site In-frame | 83% | 60% | 79% | 4/5 (80%) | 10/16 (63%)*
| 9              | Frameshift Exon 1–30 | 67% | 44% | 50% | nd | 4/5 (80%)
| 32             | Exon 31–51 | 77% | 48% | 73% | 1/2 (50%) | 15/17 (88%)
| 44             | Large rearrangement | – | 77% | 52% | 77% | 5/5 (100%) | 26/27 (96%)
| Σ 267          | Total | – | 84%* | 53% | 171/235 (73%) | 28/49 (57%) | 139/169 (82%)

244 patients described in the literature were included, plus 23 patients in this study. Young patients below 20 years of age who had not yet reached ESRF were excluded from the last column. nd, no data; *P < 0.05; **P < 0.01; ***P < 0.001 in x² tests; *sum of de novo mutations (without Gly-XY) 17/122 (13.9%)*.
time point of onset of ESRF ($n = 35$). All patients had developed ESRF within 4 years of the mean time point of ESRF of other affected family members (standard deviation, $< 3$ years). In two families this was also true for all male Alport patients in three different generations.

**Discussion**

In agreement with previous studies, the mode of inheritance of AS in German families was X-chromosomal in $85\%$, and autosomal recessive in about $15\%$ of patients. A total of $9.5\%$ of our patients had *de novo* mutations ($11.3\%$ in the literature). Mutations varied from single base exchanges to large rearrangements and were spread over the entire gene; no hot spot could be identified. The percentage of kidney biopsies done in all our patients, only our patients with known mutations or patients with mutations described in the literature was equal, being between $52$ and $57\%$. This indicates that knowing the positive result of kidney biopsy did not lead to an inadequate pre-selection of patients for this analysis or previous studies.

The European study by Jais *et al*. [49] summarizes the data of 312 mutations based on research done between 1994 and 1997. In contrast, the present study also includes publications after 1997 leaving only a minority of identical mutations being analysed in both studies. Additionally, more than 100 mutations from the US and Japan were included [5,12,23,26,31,33,34,36,37,40–44]. The previous study by Jais *et al*. [49] reports the genotype-phenotype correlation with regard to large rearrangements, missense, splicing and 'small' mutations, focusing on the differences between major rearrangements and small mutations. It does not distinguish between non-glycine and glycine-missense mutations as the most common form of mutation in AS. Furthermore, in contrast to all previous studies, the present one focuses on the effect of the location of small mutations within the $\alpha_5$(IV)-chain on the phenotype in detail. As a new finding, different types of mutations and their locations on the COL4A5-gene, and therefore their predicted different effect on protein structure, were shown to be a significant predictor of the severity of disease.

The detection rate for mutations was $74\%$ in the 41 patients, fulfilling the diagnostic criteria for AS [5,50], and $39\%$ in all 104 German families. Therefore, our detection rate is in agreement with most previous studies, showing a rate above $50\%$ in patients fulfilling three or all clinical criteria for AS [16]. Starting in 1991, numerous small children without renal failure
or renal biopsy were included in our study, explaining the high number of patients with little clinical data. However, it still remains an unsolved mystery as to why one is not able to find all the causative mutations. The inter-genic region between the COL4A5- and COL4A6-gene was not affected in 30 patients [51].

Around 8.7% of mutations in the German patients were large rearrangements. This percentage is smaller than in previous studies (up to 20%, Table 3). However, previous studies may have been biased toward large rearrangements, since small mutations are more time consuming and difficult to detect. Two different methods were used to screen for small mutations: direct sequencing and DGGE [48]. Sensitivity and specificity of DGGE used in exons 29–51 was evaluated by additional direct sequencing of exons 43 and 44 and 50 [48]. However, direct sequencing remains the gold standard in screening for small mutations, as shown previously [16].

Throughout recent history, including the family history and clinical screenings of all family members is still the gold standard for making the diagnosis of AS. Hearing loss and ocular changes might be more common in AS than previously thought [50]. A total of 39% of the German patients with mutations, and more than 40% of patients in the literature, had typical ocular changes. Therefore, consultation of specialists adds important information about extra-renal manifestations for making the exact diagnosis of AS. Screening for mutations in young patients with uncertain clinical data is a helpful and sensitive tool. However, as a very large gene is to be screened for point mutations, newer techniques such as microarrays need to be developed for rapid molecular diagnosis.

The effect on protein expression was biochemically evaluated for very few mutations, and therefore only predictive data can be used to further analyse the effect of a specific mutation. Further studies based on RNA-analyses or immunohistochemical data of the skin with α5(IV)-specific antibodies are to be used to discriminate between incorporation and non-incorporation. Immunohistochemistry has previously been shown to be a suitable diagnostic and prognostic tool in AS.

Premature stop-codons, frameshift mutations and large rearrangements are likely to result in a truncated (or absent) protein. They cause juvenile onset of ESRF, hearing loss and ocular changes in most patients (>92%, Table 3). Glycine-XY missense and in-frame mutations, however, less frequently result in the juvenile type of AS (73%). This difference is significant between 17 and 37 years of age. The mean age of onset of ESRF is 19.8 in the first group, 25.7 in the second and 30.1 in the third (Figure 3). Similar differences in hearing loss and ocular lesions make an artifact unlikely. Therefore, the impact of individual groups of mutations on the gene product must be different and result in a distinct phenotype.

Acceptor splice site mutations result in a significantly lower number of juvenile type cases of AS (63 vs 94%) and hearing loss than donor splice site mutations. The reason for this different phenotype remains
Genotype–phenotype in Alport syndrome

unclear without information on protein-expression and structure. Acceptor splice site mutations might cause a relatively ‘benign’ skipping of one exon more often. In contrast, donor splice site mutations might cause a premature stop more often, leading to a more severe change in the protein structure. Further studies need to be done, investigating the distinct gene-products, to elucidate this puzzling phenomenon.

Glycine-substitutions are likely to alter the folding of the triple helix of type IV collagen [52]. They cause juvenile type of AS less frequently. Interestingly, the distance of the same kind of mutation from the NC1-domain influences the time-point of onset of ESRF. This may be due to triple-helix formation starting at the NC1-domain and proceeding in a zipper-like feature to the N-terminus. Mutations in exon 1–20 may lead to a less critical disruption of this process. Similar observations have been made in osteogenesis imperfecta in type I collagen folding [52].

De novo glycine-XY mutations are less frequent than other de novo mutations (5.5 vs 13.9%). This is in contrast to the fact that glycine-XY substitutions are the most common mutation in AS (40%) [53]. The later onset of renal failure may increase the fitness of reproduction in these patients resulting in an evolutionary advantage. Therefore, glycine-XY mutations are transmitted to the next generation more often. This theoretical evolutionary advantage may no longer exist nowadays, because of the possibility of transplantation.

In summary, predicted major changes of protein structure nearly always cause early onset of ESRF. With some limitations, these data also correlate with severity and time of onset of extrarenal symptoms. These results might be of special interest in families without a known family history since de novo mutations and small numbers of male family members are common in X-linked hereditary diseases. For example, 50% of patients with a predicted truncated protein reach ESRF by the age of 19 and 75% do by 24 years of age. In contrast, only 26% of patients with an altered protein structure require dialysis at the same age and only 38% do by 24 years of age. The effect of glycine-substitutions on the phenotype depends on the distance of the mutation from the NC1-domain. De novo glycine mutations are less frequent than other de novo mutations. However, due to previous evolutionary advantage, glycine substitutions are the most common mutations.

According to our results, different types of mutations result in distinguishable Alport phenotypes. The consistency of the clinical phenotype was evaluated further in all German Alport families. Eight families had ocular changes and all 14 affected patients in two different generations had lenticous. Twenty-one families reported hearing loss and all 38 patients in three different generations were affected. In 13 families with two or more Alport patients with ESRF (n = 35), all patients developed ESRF within 4 years of the mean time-point of ESRF of other affected family members (standard deviation, <3 years). In two families this was also true for all patients in three different generations. Therefore, as can be expected, the intrafamilial standard deviation of the effect of the individual mutation is lower (<3 years) than the interfamilial standard deviation (6–7 years). It seems remarkable that, despite the improvement of overall health in humans during the last century, the phenotype of AS has not improved significantly in younger generations.

The distinction between the juvenile and adult types of AS from 1988 by Atkin et al. [54] does not reflect the new possibility for distinguishing phenotypes of X-chromosomal AS by knowing their genotypes. We suggest that the considerable prognostic relevance of the patients’ genotypes should be included in classification of the phenotype as follows.

- **Type S (Severe)**
  Genotype: large rearrangements, premature stop, frameshift, donor splice site and mutations involving the NC1-domain, 15% de novo mutations.
  Phenotype: ESRF ~ 20 years of age, 80% hearing loss, 40% ocular lesions.

- **Type MS (Moderate–Severe)**
  Genotype: non-glycine XY-missense, glycine-XY involving exon 21–47, in-frame and acceptor splice site mutations, 15% de novo mutations (5% de novo glycine-XY mutations).
  Phenotype: ESRF ~ 26 years of age, 65% hearing loss, 30% ocular lesions.

- **Type M (Moderate)**
  Genotype: glycine-XY mutations involving exon 1–20, 5% de novo mutations.
  Phenotype: ESRF ~ 30 years of age, 70% hearing loss, 30% ocular lesions.

**Acknowledgements.** This work was supported by a DFG grant to K.-O. Netzer (Ne 350/7-1), and by the SFB 263, C6 to M. Weber. Parts of this work were presented at the Annual Meetings of the American Society of Nephrology in 1995, 1996 and 1997, and of the European Renal Association in 2000. The authors wish to thank C. Merkl for her technical assistance and Dr Müller-Myhsk for his excellent help with the statistical analysis.

**References**

Genotype–phenotype in Alport syndrome


Received for publication: 15.9.01
Accepted in revised form: 07.2.02