# Original Paper

*Christian Clépet*[a]
*Hans J.G. Dauwerse*[b]
*Chantal Desmaze*[a]
*Gert-Jan B. van Ommen*[b]
*Jean Weissenbach*[a]
*Jean Morissette*[c]

[a] Généthon, Evry, France;
[b] Department of Human Genetics, State University, Leiden, The Netherlands;
[c] Centre de Recherche du Centre Hospitalier de l'Université Laval, Québec, Canada

# A 10-cM YAC Contig Spanning GLC1A, the Primary Open-Angle Glaucoma Locus at 1q23-q25

..........................................................................................

**Abstract**
Primary open-angle glaucoma is a complex of ocular disorders characterized by irreversible lesions of the optic nerve, open angle of the anterior chamber of the eye and elevated intraocular pressures. GLC1A, a locus involved in one form of this disease, has been mapped to an approximately 9-cM interval within 1q23-q25, between markers D1S445 and D1S416/D1S480. A 10-cM yeast artificial chromosome (YAC) contig spanning the whole region is described. This contig is based on 67 YACs, and 41 sequence tagged sites comprising 23 genetic markers, 16 YAC ends and 2 expressed sequence tags. The reagents reported in this study should be useful tools for the identification of the GLC1A gene by positional cloning.

......................

## Introduction

Glaucoma is a complex of ocular disease entities and is the second most common cause of blindness in industrialized countries [1]. The most frequently encountered form of the disease is primary open-angle glaucoma (POAG; MIM 137760 [McKusick 1994, pp 559–560]). POAG is characterized by irreversible lesions of the optic nerve, a normal-appearing morphology of the anterior chamber of the eye and is usually associated with increased intraocular pressures. POAG has been subclassified according to age at onset and clinical progression as (1) juvenile open-angle glaucoma (JOAG), which appears between the age

of 3 and early adulthood and manifests an aggressive course, and (2) chronic open-angle glaucoma (COAG), which is distinguished by its mid- to late-age onset and a slow, insidious progression. A dominant JOAG locus was first described on chromosome 1q21-q31 by Sheffield et al. [2]. This locus is now known as GLC1A. More recently, it was shown that both JOAG and COAG were found to be linked to the GLC1A locus within a large French-Canadian family. In this pedigree, ages at onset were distributed over several decades displaying a clinical continuum and all patients, either JOAG or COAG, shared a common 'disease' haplotype encompassing GLC1A [3]. As COAG is a complex non-mendelian disorder, COAG

patients within this pedigree may show variable expression of the JOAG phenotype, and/or expression of the disease may be altered by modifier genes. Furthermore, evidence for genetic heterogeneity for JOAG was observed in a large Swedish family which did not show linkage to 1q [4]. To date, the most refined localization of the GLC1A putative gene is restricted to an approximately 9-cM interval between loci D1S445 and D1S416/D1S480, in 1q23-q25 [3]. The physiological and molecular bases of POAG have not yet been defined. A better understanding of this disease requires identification of the causative gene(s). We describe here a yeast artificial chromosome (YAC) contig, spanning the 8- to 10-cM POAG interval flanked by D1S445 and D1S416/D1S480. After the completion of our work, a YAC contig spanning the same locus was released by the Whitehead Institute/MIT Genome Research Center (WI/MIT), Human Genetic Mapping Project, Release 9, December 1995. However, the contig we report also includes a number of additional polymorphic and nonpolymorphic STSs as well as several YACs not previously described. Most of the YACs are now characterised by fluorescence in situ hybridization (FISH) and/or their insertion ends.

## Materials and Methods

### YAC Library Screening

All YACs were isolated by PCR screening [6] of the CEPH megabase size insert library [7, 8]. All PCR markers are shown in table 1.

PCR was performed in a 25-μl reaction volume, containing 10 m$M$ Tris-HCl, pH 8.8, 50 m$M$ KCl, 1.5 m$M$ MgCl$_2$, 0.1% Triton X-100, 200 μ$M$ of each dNTP (Pharmacia), 0.5 μ$M$ of each primer (synthesized by Genset, France) and 1 unit of $Taq$ DNA polymerase (Cetus). For amplification, 5 ng of YAC DNA were denatured for 5 min at 94°C before adding the $Taq$ polymerase, and the mixture was subjected to 30 cycles as follows: 40 s at 94°C, 30 s annealing at the indicated temperatures (table 1) and 30 s at 72°C. Positive control amplifications were routinely performed on 20 ng human genomic DNA and negative controls on AB1380 yeast DNA and no DNA. Amplification products were analysed on ethidium bromide-stained 2 or 4% agarose gels, depending on their size.

### Characterization of YAC Ends

YAC ends were isolated either by direct sequencing of total reactions of inter-$Alu$-vector PCR [Clépet et al., in preparation], or by a modification of the inverse PCR (IPCR) procedure described by Abderrahim et al. [9]. For IPCR, 200–300 ng of YAC DNA prepared in microtiter dishes [6] were independently digested with 2 units of $Eco$RV, $Hae$III, $Hha$I, $Nla$IV, $Sau$3A, $Msp$I, or $Taq$I, in 40-μl volumes for 7 h. Heat-sensitive enzymes were then denatured for 20 min at 70°C. After ethanol precipitation, DNA pellets were washed, dried at room temperature and carefully resuspended in 40 μl of H$_2$O. 10 μl of digested DNA was denatured 10 min at 65°C and kept on ice before circularization. Ligation was performed in a 40-μl final

volume with 0.5 U ligase (Promega) overnight at 4°C, and heat stopped for 15 min at 72°C. To characterize the left end of a YAC insert, $Eco$RV, $Hae$III, $Nla$IV, $Sau$3A, $Msp$I and $Taq$I circularization products were PCR amplified with primers 5U (5'GTTGGTTTAA-GGCGCAAGACTT3') and 5R (5'TGAAGAAAGAGTATACTAC-ATAACA3'). To amplify the right end of a YAC, $Hae$III, $Hha$I and $Nla$IV self-ligated DNAs were used with primers 3U (5'CCCGATC-TCAAGATTACGGAAT3') and 3R (5'CATTCACTTCCCAGA-CTTGCAA3'). PCRs were performed on 10 μl of ligation product, in a 45-μl final volume, with 3 units of $Taq$ DNA polymerase, and the same buffer conditions as described above. In order to improve specificity, hot-start and touch-down PCR were performed: 94°C 5 min denaturation before adding the $Taq$ polymerase; then 3 cycles: 94°C 40 s, 61°C 30 s, 72°C 60 s; followed by 30 cycles 94°C 40 s, 58°C (for 5R/5U) or 59°C (for 3R/3U) 30 s, 72°C 60 s. Final extension was performed at 72°C for 2 min. 10 μl of the PCR reaction was controlled on 2% agarose gel, the rest was filtered through Microcon devices (Amicon Inc., USA) and subjected to dye-dideoxy cycle sequencing with same primers as for PCR, at their respective annealing temperature (using the ABI sequencing system; Foster City, Calif., USA). PCR primers were designed with the Oligo 4.0 program [10]. Sequences were screened for repeated elements using the FASTA program [11] and a local database containing a compilation of known human repeat sequences. Chromosomal assignments of YAC end sequence-tagged sites (STSs) were determined by PCR assay on the somatic cell hybrid panel 2 and/or panel 1 [12], purchased from the National Institute of General Medical Sciences (Camden, Pa., USA).

### Isolation of Other STSs

Isolation of microsatellite markers AFMGLC20, 21, 22, and AFMGLC23 from YAC 792d9 is described in Morissette et al. [in preparation]. Plasmid pHBI40 (D1S66) was purchased from ATCC. This RFLP probe was partially sequenced in order to generate an STS (STS-HBI40) and make its use more convenient.

### Pulsed-Field Gel Electrophoresis and Southern Blotting

Pulsed-field gel electrophoresis (PFGE) was performed in 1% SeaKem agarose gels (Tebu, France), with a CHEF DRIII or a CHEF Mapper apparatus (Bio-Rad, USA). Standard conditions used for analysis of YACs were as follows: 22 h, linear ramp from 60 to 120 s, 6 V/cm, 14°C, TBE 0.5 ×. Gels were depurinated and transfered onto Hybond N+ membranes (Amersham, UK) for 2 h using a vacuum blotter (Appligene, France). Total human DNA was labeled with horseradish peroxidase (using the ECL labeling kit RPN 3001, Amersham, UK) and hybridized at 10 ng/ml final concentration to PFGE filters according to the ECL protocol.

### Fluorescent in situ Hybridization

FISH experiments were performed according to Driesen et al. [13] and Dauwerse et al. [14].

**Table 1a.** STSs used for the contig

| marker name | locus symbol | size (bp) | temp. (°C) | oligonucleotide primers (5'->3') | | origin |
|---|---|---|---|---|---|---|
| 650g9L | | 205 | 58 | AACACCGATAAGACAGTGAA | CAGTTGCAGCAGTGGAGAA | |
| 653b3L | | 139 | 50 | TTCATCATATCTNGACACT | TGTTTTAGAAATGCTTCAG | |
| 766e4L | | 150 | 49 | CAGATATGTCTATGTCTNCTAATCT | TATAAGAGGCTTGGTCGTT | |
| 770a2L | | 84 | 53 | CCAAAAAATAACATCATGT | TTCAGAGAGCATATAGACC | |
| 775e10R | | 228 | 58 | CTCCTTCTCACCACCAATAA | TAGGATGCAACCTGAATAGT | |
| 789f12L | | 99 | 55 | GGAGTCTGGCAGCAGCA | ATGACATAATGTGGGAGACATAGT | |
| 809c4R | | 173 | 55 | CACTGAAATCCACTAACAACA | TAACAAATAATCCCCAAACTT | |
| 875b1R | | 220 | 55 | AGTATGGATCTCAGCATTGGCA | TGAGCAAGTGACAGAGTTAGTT | |
| 887a6L | | 136 | 55 | GATGCCTCTTCTCTCCAACTCA | TCTTCTCAGGGTATCCGTGGAA | |
| 895f3R | | 178 | 55 | CCACATTGACTCAATAGTTCTT | TGAGTTTAGACTGTCCTAGAGA | |
| 903f4R | | 197 | 53 | TATGTTTATGTCTCAGGTGTG | TCAGAGAGGAGTAAGCAATCA | |
| 928h2L | | 250 | 55 | CGGCCTTGTTCTTCCAAGTA | TCTAACTCTCATAAGGGTCTCT | |
| 930a12L | | 214 | 55 | ACTGCTGACTACATAAATAGAT | GATATTACCCTNAGCCACCTGAT | |
| 930a12R | | 237 | 55 | CCAGAGTCTGTAGTGAGTCT | GAAGAATGCTGAGTGCCTTT | |
| 933h12L | | 142 | 55 | TGCACAAGATACACTGGGT | CAGAGTGAGACCGTGTCTC | |
| 985e4R | | 296 | 55 | CCATACTCCTTGAAGCACTCCA | ATTTAGTCACAGAGGCATCCTA | |
| STS_HBI40 | | 252 | 55 | GTGTAGCATTCATTCTGCAGG | GAGCTACAAATGTGCCAATG | |
| AFM192ye3 | D1S1569 | 201 | 55 | CGAAAAGATGACACCTCTCAGT | CTCCCACAGGAATACCTGC | § |
| **polymorphic markers** | | | | | | |
| AFMGLC21 | | 114 | 55 | GATCTCTTATCAGTCAGGCA | TTTCTAAGGCTGAATAATATTCG | @ |
| AFMGLC22 | | 152 | 60* | TTAACTCACCACTCCCTGCC | AATTATGGCCTTCGCCC | @ |
| AFM063xg9 | D1S196 | 273 | 55 | GGCTGTGGGTGTTTCTCCTA | AGCTCTCATGNCTTTACATTCT | [18] |
| AFM115ye3 | D1S2887 | 181 | 55 | CATCCAACAAGGGNCTAATA | ATACCTGTGGGCCATTCATA | [15] |
| AFM122xa3 | D1S210 | 119 | 55 | CCTCAGTTCATTCCCCATAA | AGCTGAATCTCACCCAATAACTA | [18] |
| AFM154xc9 | D1S2634 | 133 | 55 | CTTTCTGGATTCANACTCCC | AGCTGAAAAGTGGATCGACT | [3] |
| AFM157xe7 | D1S218 | 276 | 55 | TGTAAAAGCAAACTGTAGACGAT | TTTATGTTATCACCAAGGCTTCT | [18] |
| AFM185xd6 | D1S416 | 148 | 55 | AAGGCGTTGAGTGTTGTGC | AAACCGAACCAATGGGCTA | [18] |
| AFM212xb10 | D1S242 | 219 | 55 | ACCACTCCAGTTTGAGCAAC | AATGCCATGAGGTGTTTCTC | [18] |
| AFM212xf6 | D1S431 | 219 | 55 | CCTAGCACCTAGAGGCAA | GGAGGATAGCATACCAAAAA | [18] |
| AFM217xe11 | D1S433 | 150 | 55 | AGACAAGGTCTGCATTTCAAG | GCCCATCACTACCCATAGG | [18] |
| AFM238xf10 | D1S445 | 156 | 55 | TTATTTCAGACTTCTGACTTCCA | TGAGATCATAGCCAAACCAC | [18] |
| AFM248wg5 | D1S452 | 221 | 55 | TAATGGGTTCAGTGGACCTT | TGCAGTTCCATATTCCAGGT | [18] |
| AFM278ye5 | D1S2814 | 269 | 55 | TGAGCCGAGATTGAGCC | TTGTTACCCAGGGATACCAG | [3] |
| AFM290yf9 | D1S480 | 200 | 55 | CTGATGTATTTTGTCAGGGTTGTC | AGATGGCCTCCTGTGGG | [15] |
| AFM350yh1 | D1S2851 | 189 | 60* | TCTTCCCACCACTGCC | TTTATATTTGTATTCCTACTGCCCA | [3] |
| AFMa152zf5 | D1s2643 | 182 | 55 | GTGTATGATAAATAATTTCAGCCC | CCATTGGTGCATTTTGAA | [15] |
| AFMa197wg1 | D1S2658 | 266 | 55 | AGCATCCATGCACCATGT | CAGGGAATCTCCAGAAGTCA | [15] |
| AFMa203yb9 | D1S2659 | 210 | 55 | ACAACCAAGTCCCAGATTTTAG | GGAAGGTGTCAGAGTGAGTATT | [15] |
| AFMa285zd5 | D1S2691 | 259 | 55 | AGCTGGATATGGCTCAA | CGCTGGTAGCCTTTAATG | [15] |
| AFMb292xb1 | D1S2750 | 110 | 55 | CAGTTTCAAATGCTGATGACAC | GATGATGCCAATTACATCCAAG | [15] |
| AFMb316zb9 | D1S2762 | 244 | 55 | CCTTAATTGTGGTGTTGGT | AAAAATCTGGAAGGCATAAA | [15] |
| AFMb328yg5 | D1S2769 | 227 | 53 | TTGCGAACGACATGACTT | GCCGTGGAACACTGTTG | [15] |
| AFMb352xe9 | D1S2790 | 253 | 55 | AAAATGCTCATTAGTCCAGAAAG | TGGCTATGTTTTACTAGCTCAAG | [15] |
| AFMb359xf5 | D1S2799 | 209 | 55 | AGCAAGACCCTGTCTCAAAA | TGGATAGCTTTCCACCACT | [15] |
| AFMc018yd9 | D1S2815 | 224 | 55 | CTGACATGGAATACCTCTATGATGC | CTCCAAATCTAGTCACACTGGAAG | [15] |
| **ESTs** | | | | | | |
| HSB24C09 | D1S2067E | 123 | 52 | GTTCACAGCCAAGCCATTG | CTGAAGTTTCTGGTTCCTTGG | [20] |
| HSB78G03 | D1S2049E | 247 | 55 | TACTCTCTGCCTTCAGCAATC | CTTCAAGTTCTGCCACTCTTC | [20] |

## Results

From the most recent results [3], the genetic localization of GLC1A is as follows: (D1S196-D1S431, D1S445)-{GLC1A, D1S433-(D1S2851, D1S452, D1S210)-D1S2634-(D1S2814, D1S242, D1S218)}-(D1S416, D1S480). GLC1A is thus located in a region of 8–10 cM which could be subdivided into subintervals with no markers (fig. 1). Subsequently, 10 additional AFM microsatellite markers [15] were added to this map. During the course of this study, some of the data released on public electronic servers by various genome centers could either help, confirm, or conflict with our own data.
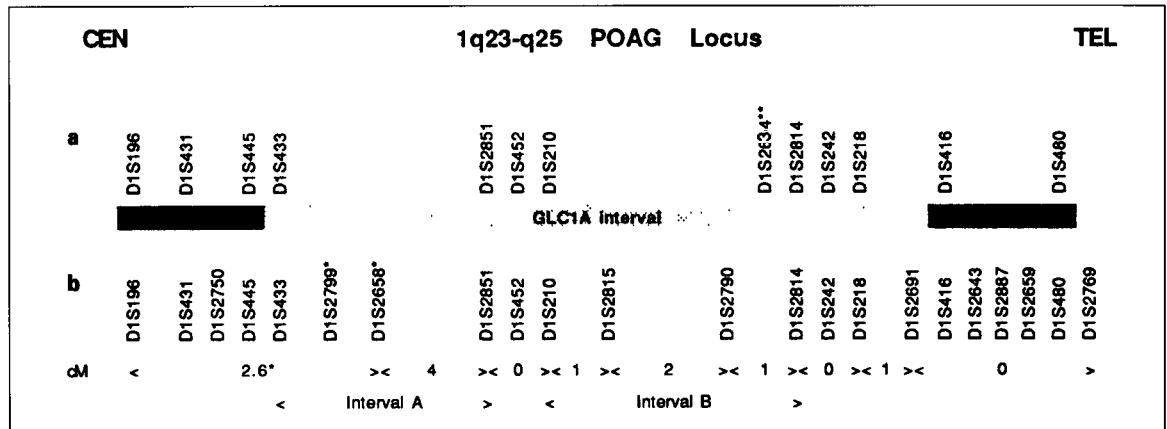
Clépet/Dauwerse/Desmaze/van Ommen/
Weissenbach/Morissette

**Fig. 1.** Genetic map of the GLC1A locus. **a** From Morissette et al. [3]; the gray box shows the boundaries of the GLC1A interval. **b** Genetic map as in Dib et al. [15]. (*) The position of these markers conflict with physical data (table 3) but their order was determined with odds < 1,000:1 on this genetic map. ** This marker was not retained in the final map [15]. CEN = Centromere; TEL = telomere.

**Table 1b.** STSs generated during this study and not assigned to the GLC1A locus

| marker | size (bp) | temp. (°C) | oligonucleotide primers (5'->3') | | assignment Chrom. |
|---|---|---|---|---|---|
| 747b1L | 205 | 60 | GCAGAACTGAGTTAGATTACA | AAGCCTTAGAGCCCTCTGT | X |
| 756c3R | 162 | 60 | TGGATATATGAGAGACCACGTT | AGGAGTTTTCCAGGCAGTAG | 5 |
| 762g10L | 170 | 55 | ATAAAAACACAAGAGGATCT | TATTATCCCTGAAAATGATGAG | 4 |
| 765g6L | 294 | 62 | TGATTCCTTCTGTGTTCCCA | TCTTGTCTCCGTGATGCTGA | 4 |
| 765h10L | 300 | 62 | ATTCAGCAAAGGAGGCACAG | AAACACCCAGTGATGCTCCA | 1 |
| 775e10L | 150 | 55 | AGTCAGAATCAGAGAGTTTT | CTGAGCTGTTTTGCTTAATT | 4 |
| 778f8L | 132 | 55 | TCGATAGAGTATAATCACATGT | GAATTCCCCTTGAAGACAC | X |
| 809c4L | 330 | 55 | ATAGCAATGGCTCTAAGTCTTC | CAGTGAAAGTTCTCCATGTAGT | 10 |
| 883h6L | 228 | 58 | GAGCATATATGGCTTATGCA | AGGTCATTGCAGTTCAGTTC | 2 |
| 905a3R | 175 | 49# | TATTCACTTTTTCTCCGACAT | TTAACATTAGCTTTAGGCTCCT | 5 or 6 |
| 955c7R | 258 | 60 | CCATGAGGTGATGTCATTTT | CATCTCTGCTGTGCTGGTT | 4 |
| 974a12R | 224 | 55 | TTCAGCGTGTGGTGTCAAA | TTTCCACGTTCTCCTCCTT | 4 |
| AFMGLC20 | 94 | 55 | GATCACTTCCCACATAAAT | TCATCTGCTTCCAAACCGTG | 3 |
| AFMGLC23 | 185 | 55 | ATCCTGACTAATAAAGCCCAT | GGAAGTACAATGATAAAGATGCT | 5 or 6 |

YAC end STSs retain the name of the YAC; L and R respectively stand for insert end adjacent to the Trp arm or Ura arm of pYAC4. AFMGLC20 and AFMGLC23 are microsatellite markers derived from YAC 792d9. Numbers in square brackets indicate previously published markers. New STSs have been submitted to GDB.

§ Marker developed at Généthon and assigned to some of the YACs of the contig at the WI/MIT (Release 3, July 1994).
@ Morissette et al. [in preparation].
* PCR at 60°C for STS content analysis and at 55°C for genotyping.
# 35 PCR cycles rather than 30 are required.

## Isolation of YAC End STSs

A total of 67 YAC clones containing at least one STS from the GLC1A interval were identified by PCR screening. Eighteen of these YACs have not been previously reported to be linked to an STS of the GLC1A interval. Most of the 67 YACs were used as templates for isolation of sequences corresponding to the insert ends using inter

*Alu* vector PCR and inverse PCR methods (see Methods). Table 1 summarizes the characterization of YAC ends. In total, 38 YAC insert end sequences were characterized; 22 of them were obtained by inverse PCR and 16 by direct sequencing of *Alu* vector PCR reactions [Clépet et al., in preparation]. Out of the 38 YAC end sequences, 8 (21%) corresponded to part of a repeated element and could not

**Table 2.** Characterization of YAC clones from the GLC1A region

| | Size | Left-End | Right-End | FISH |
|---|---|---|---|---|
| 650g9 | 490* | + | KpnA | 1q, 1p21-22 |
| 653b3 | 330* | 1,+ | | |
| 670f3 | 1160* | | | 1q |
| 672a9 | 300* | | | |
| 691b5 | 950* | | | 1q |
| 747b1 | 1760* | X | | 1q, 1p21-22, 1p43-44, Xq, 2p(weak) |
| 749g9 | | KpnA | | |
| 750a1 | 1720* | | | |
| 752h3 | 1040 | | Kpnl | 1q |
| 754a11 | 1120*&1740* | | | 1q |
| 756c3 | 700* | | 5 | 1q & 2 signals on 13q,14q or 15q |
| 757a7 | 1270* | | | 1q |
| 762g10 | 1720* | 14 | | |
| 765c8 | 1440* | Kpnl | | 1q |
| 765g6 | 1430* | 4 | | 4 or 5 |
| 765h10 | 1790* | 1,(!!) | | |
| 766 e4 | 690* | + | | 1q |
| 770a11 | 1460* | | | |
| 770a2 | 1730* | 1,+ | | 1q |
| 775d11 | 80*&1620* | | | |
| 775 e10 | 1400* | 4 | + | |
| 778f8 | 1120* | X | | |
| 783d2 | 730* | | | |
| 789f12 | 820*&1300* | 1,+ | | 1q |
| 791b2 | 1230* | | | |
| 792d9 | 1640 | Kpnl | | |
| 793 e7 | 1380* | | | 1q |
| 809c4 | 1740 | 10 | + | 1q, (4q or 5q), (8q or 9q) |
| 851d11 | 1660* | | | 2* |
| 854b9 | 1060* | | | 1, (4 or 5) |
| 857c8 | 420* | | | 9 |
| 859g4 | 1390* | | | 1q, 11q |
| 870 e1 | 550 | | | 1q, 2pter |
| 871h3 | 1680* | | | |
| 874b11 | 810* | | | |
| 875b1 | 1790 | Kpnl | + | 1q |
| 883h6 | 850 | 2 | | 1q |
| 887a6 | 850 | + | | 1q |
| 893c6 | 700* | | | |
| 894a9 | 900* | | | 1q |
| 895f3 | 1730 | | + | 1q |
| 902c7 | 1790* | | | |
| 903f4 | 850 | | + | 1q |
| 905a3 | 730*&1160* | | 5 or 6 | |
| 911 e5 | 1090* | | | 1q, (4q or 5q) |
| 912h6 | 1090* | | | |
| 913 e6 | 1080*&1230* | | | |
| 916 e2 | 1370* | | | 1q |
| 926g9 | 1660* | | | 1q, 3qter |
| 928a5 | 390* | | | |
| 928c8 | 1510* | | | 1q |
| 928g11 | 1600* | | | 1q |
| 928h2 | 1490 | + | | 1q, (4q or 5q) & all acrocentric p-arms |
| 930a12 | 1580 | + | + | 1q |
| 933h12 | 1180* | 1,+ | | 1q |
| 941f11 | 1270* | | | 1q |
| 945g6 | 1230* | | | 1q |
| 948g10 | 1770* | | | |
| 950a5 | 860* | | | |
| 953 e8 | 1610* | | | |
| 955c7 | 1170* | | 4 | |
| 958c8 | 830* | | | 1q, (13, 14 or 15 -qter) |
| 974a12 | 1570* | | 4 | 1q, (4 or 5 -qter) |
| 981c4 | | | | 1q, 11q |
| 982h7 | 1200* | | | 1q |
| 984f7 | 1550* | | | 1q |
| 985 e4 | 1750 | Kpnl | 1,+ | 1q |

YAC sizes are in kilobases. An asterisk (*) indicates when size or FISH data are from the CEPH-Généthon database [16]. Left end and right end characterizations: (+) insert-end STS mapping in the GLC1A contig; (1,+) assigned to chromosome 1 and to the contig; (!!) chromosome-1 chimera; (KpnA/I) sequences containing part of a repeat element; other indications correspond to chromosomal assignments.

be used for designing primers (table 2). Overall, 28 YAC end STSs have been generated. Most of these were assigned to a chromosome by PCR assays on a somatic cell hybrid panel [12]. When a YAC end STS was shared by at least 3 different YACs on the contig, it was considered as part of the interval. Of the 28 YAC end STSs, only 16 (55%) could be used to construct the contig; the other 12 eventually proved to originate from chimeric parts of the YACs. As an example, 2 cases of 'chimeric STSs' 765h10L and 955c7R are described in more detail: 765h10L has been assigned to chromosome 1 but is absent from all other YACs from the contig except 765h10. Similarly, AFM114xa5 is a chromosome-1-specific marker present on 765h10 (WI/MIT, Release 7) and absent from all other YACs from the GLC1A contig. Therefore YAC 765h10 appears as an intrachromosomal chimera.

955c7 is D1S2814 positive and has its right end (955c7R) on chromosome 4 (tables 1, 3). 928a5 shows the same STS content as 955c7, and contains 955c7R in particular (data not shown). It seems very likely that both YACs derive from the same initial clone; however, their sizes are different (1,170 kb for 955c7 versus 390 kb for 928a5). One possibility is that both YACs derive from the same initial clone 955c7, and that a deletion occurred subsequently in a derived cell, giving rise to 928a5.

## Characterization of the YACs

Most of the YACs (a total of 41) were subjected to FISH analysis on metaphase chromosomes to ascertain their DNA content. Eighteen YACs gave hybridization signals outside chromosome 1q, suggesting chimerism (table 2). Two YACs in particular (857c8 and 765g6) did not show any hybridization signal in 1q, despite their STS content (D1S2815 and D1S2814, respectively), indicating that the chromosome 1 portion of both of these YACs only represents a minor part of the insert. In addition to FISH data, left-end and right-end STSs were character-

Clépet/Dauwerse/Desmaze/van Ommen/
Weissenbach/Morissette

ized (table 2). Overall, 49 YACs were analysed either by FISH or insert end chromosomal assignment or both; 25 of these YACs appear to be chimeric.

### Construction of a YAC Contig

A contig of the proximal part of the GLC1A region (interval A) was derived from published information and data obtained in this study. The CEPH Généthon map release (version 1.1 [16]) on which we initially relied for contig construction proposed one level-6 contig between microsatellites D1S431 and D1S210. This path, based on fingerprint and Alu-PCR hybridization data, comprised the following 6 YACs: 903f4, 875b1, 930a12, 895f3, 985e4, 809c4. To confirm this contig, a series of YAC end STSs were generated from several YACs including those listed in Quickmap. Ten YAC end STSs were characterized in that region and a contig containing 3 YACs (table 3) was sufficient to span interval A, thus reducing the initial Quickmap contig. In the meantime, this contig has been reduced to level 4 (CEPH-Généthon map, Version 2; [5]).

The contig was extended to the distal part of the GLC1A region (interval B) using a YAC walk initiated from four starting points: the flanking markers D1S210 and D1S2814 and internal markers D1S2815 and D1S2790. On average, 10 mega YACs were isolated from each landmark. Overlaps between YACs were established with YAC end STSs except for junction D1S2790-D1S2814, for which all 11 isolated YAC ends were irrelevant (8 being from other parts of the genome) (see tables 1, 3). The link between D1S2790 and D1S2814 was shown by hybridization of inter Alu-PCR products (data not shown), and confirmed with an expressed sequence tag (EST) marker mapped to this region using radiation hybrids [17].

Observation of the presence or absence of each marker in each YAC enabled marker ordering as shown in table 3. This order assumes minimal breakage and rearrangement of the YACs. However, a few deletions still had to be assumed for some YACs: 928h2 and 928g11 show internal deletions encompassing several adjacent STSs (5 and 4, respectively); 3 other YACs show single-STS deletions. Moreover, this physical map is not in total agreement with the genetic map [18] where D1S433 was given as 1 cM centromeric to D1S445; such an order would require gap introduction into two different YACs. In contrast, the reverse order (centromere-D1S445-D1S433) is consistent with the STS content of all 5 YACs spanning this region and with genetic data obtained more recently [3]. However this more recent map also shows a discrepancy with

the present physical map for marker D1S2634 which we now place distal to D1S2814. (D1S2634 was not retained on the new Généthon genetic map [15].)

The physical map also resolves markers that could not be separated by linkage analysis or that were ordered with odds of less than 1,000:1 as in the case of the proximal group of markers, containing D1S2799 and D1S2658 (fig. 1, table 3). In addition some other markers were also integrated into the contig: D1S66 located in 1q21-q31 [19] is present in the most proximal YACs (see STS-HBI40, table 3). Indication that D1S1569 (a marker with low polymorphism) mapped in the contig was at first obtained from the WI/MIT (Release 3, July 1994). This marker was then localized inside the centromeric interval A (table 3).

Two ESTs (D1S2067E and D1S2049E) previously assigned to the GLC1A locus using radiation hybrids [17] were mapped to the contig. Both of them belong to the same 'cDNA families' as 2 ESTs (respectively D1S2359 and D1S14662) that were reported independently on contigs (WC-1.33 and WC-215) of the GLC1A region by the WI/MIT (Release 6, March 1995).

A tentative estimate of the physical size of the contig was assessed from nonchimeric YACs. The upper size limit was estimated from the smallest path of overlapping YACs. Conversely, the lower limit was calculated by adding the size of the largest nonchimeric YACs showing no STS overlaps. As shown in figure 2 (see also tables 2, 3), the length of the 10-cM GLC1A interval should be comprised between 7.5 Mb (D1S2658-D1S2691) and 9.4 Mb (D1S196-D1S416). This size estimate is consistent with the average of 0.85 Mb per centimorgan observed in the human genome [15].

### Discussion

We have established a YAC contig spanning the 8–10 cM 1q23-q25 region containing GLC1A. This continuum of YACs was achieved after a series of bidirectional walks initiated from AFM genetic markers which mapped in this region. All YAC clones were derived from the CEPH-mega-YAC library whose large-insert size enabled contig completion with a minimal number of steps. D1S196 and D1S416 are connected by a minimal set of 6 YACs (table 3). This YAC contig is entirely STS based. In total, 41 STSs are physically linked to this region including 23 genetic markers, 16 YAC ends and 2 ESTs. Sixty-seven YACs were used to construct this contig. The chimerism of most of the YACs was assessed by FISH analysis on

**Table 3.** STS content of YAC clones from the GLC1A region

Clépet/Dauwerse/Desmaze/van Ommen/
Weissenbach/Morissette

YAC clones (columns, listed from top): 887a6, 903f4, 883h6, 875b1, 930a12, 928h2, 895f3, 793e7, 926g9*, 985e4*, 974a12*, 789f12, 984f7*, 752h3, 754a11, 928c8, 870e1, 809o4, 928c8, 757a7, 792d9, 851d11, 953e8, 981c4*, 928g11, 749g9, 941f1, 766e4, 902c7, 670f3*, 874b11, 650g9*, 778f8, 653b3*, 857c8, 911e5, 912n6, 933h12, 770a2, 750a1, 770a11*, 775d11, 791b2*, 893c6*, 691b5*, 765c8, 854b9, 894a9, 982h7*, 747b1, 756c3, 958c8, 672a9, 762g10*, 765g6*, 928e5*, 955c7*, 905e3, 775e10, 765h10, 916e2, 945g6, 948g10, 913e6, 871h3*, 783d2, 950a5, 859g4

STS markers (rows, listed from top): D1S2762, D1S196, D1S431, D1S2750*, D1S445, D1S2067E, D1S2799*, D1S2658, STS_HBI40*, 930a12R*, D1S433, 887a6L*, 928h2L*, 903f4R*, 895f3R*, D1S1569, 985e4R*, 789f12L*, 875b1R*, 930a12L*, 809o4R*, D1S2851, D1S452, D1S210, AFMGLC22*, AFMGLC21*, 653b3L*, D1S2815, 650g9L*, 933h12L*, 770a2L*, 766e4L*, D1S2790*, D1S2049E, D1S2814, D1S2634*, D1S242, D1S218, 775e10R*, D1S2691, D1S416, D1S2643*, D1S2887, D1S2659, D1S480, D1S2769

metaphase spreads, as well as by STS content. Overall, 50% of the YACs that were characterized appear to be chimeric. To reduce the risk of integrating false data from chimeric YACs, the depth of the contig was kept as large as possible. There are on average 10 YACs per STS with a minimal depth of 3 clones. Assessed from the size and STS content of nonchimeric YACs, the physical length of the 10-cM GLC1A interval was tentatively estimated to be between 7.5 and 9.4 Mb. Initially and during the course of this study, data released from Human Genome Map-

ping Centers were used to some extent. The following groups of markers are integrated at level 1 on YACs in the CEPH-Généthon map (Quickmap, Version 2; [5]), {D1S196, D1S431, D1S445} {D1S433} {D1S210, D1S452, D1S2851} {D1S2814} {D1S218, D1S242} {D1S416}. Based on fingerprint and *Alu*-PCR hybridization data, a few contigs spanning the proximal part of the GLC1A interval flanked by D1S196 and D1S210 (interval A, fig. 1) are available on Quickmap. The shortest Quickmap contig shown is of the 4th order, and is
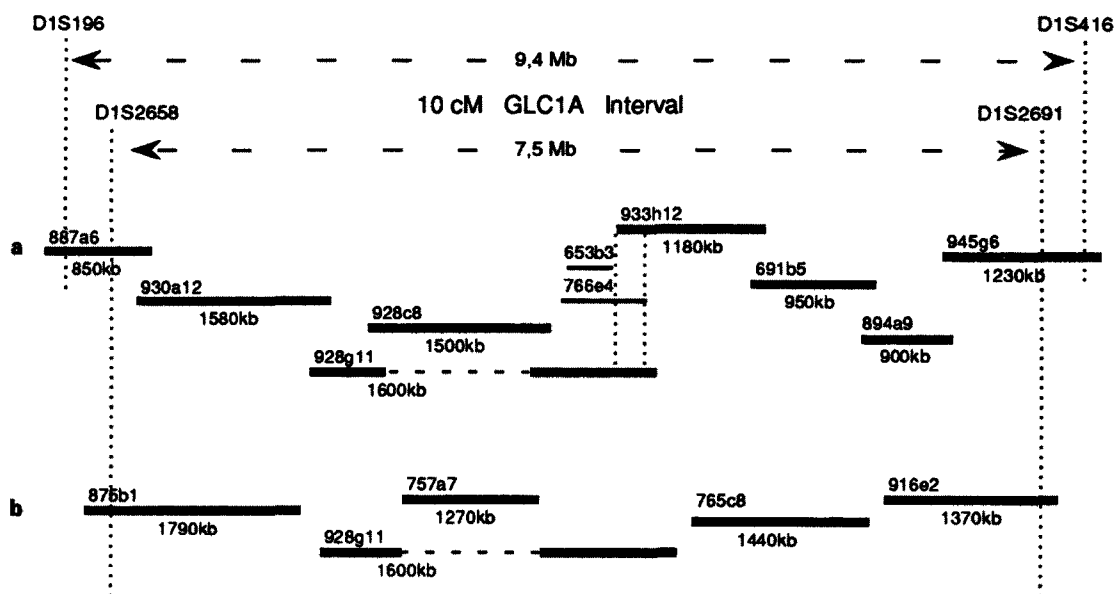


**Fig. 2.** Tentative estimate of the physical size of the GLC1A YAC contig. **a** Upper size limit: the smallest path of overlapping YACs found between markers D1S196 and D1S416 is shown. Nonchimeric YACs were chosen according to the present results (FISH and STS content). The deletion (including 4 STSs) of YAC 928 g11 is compensated by 928c8. The 3-STS overlap between 933h12 and 928g11 was subtracted from the overall length as it can be estimated as being roughly equivalent to the (360 kb) size difference between 766e4 (690 kb) and 653b3 (330 kb). The overall length of this path is 9,430 kb (850 kb + 1,580 kb + 1,600 kb + 1,500 kb + 1,180 kb – 360 kb + 950 kb + 900 kb + 1,230 kb). **b** The lower size limit is estimated between markers D1S2658 and D1S2691. A minimal length of 7470 kb for this interval was calculated using the largest nonchimeric YACs showing no STS overlaps (1,790 kb + 1,600 kb + 1,270 kb + 1,440 kb + 1,370 kb).

The PCR markers are listed along the top and the YACs down the left-hand side. A plus or a minus indicates that a marker is respectively present or absent from a YAC. A blank indicates that a marker was not tested on a YAC. L (left end) and R (right end) denote a positive assay and indicate the YAC used to design the corresponding insert end STS. Asterisks (*) denote YACs and STSs that were not previously assigned to the GLC1A locus. The gray top bar shows published limits of the GLC1A locus. In cases where markers were not resolved (D1S431–D1S2750; D1S445–D1S2799; D1S2658-STS-HBI40, and D1S242–D1S218), the relative order shown is arbitrary.

anchored on the marker in between, D1S433. Although this 4-YAC contig is close to the real situation, this path can actually be shortened to 3 YACs. This part of our contig was mostly assembled from YACs listed in Quickmap. In the more telomeric interval, running from D1S210 to D1S218/D1S242 (interval B, fig. 1), no contigs are shown on the Quickmap server. After the completion of this work, a YAC contig (WC1.23) spanning the whole GLC1A region was released by the WI/MIT (Release 9, December 1995). The GLC1A portion of WC1.23, flanked by D1S196 and D1S416, is based on 30 STSs and more than 69 YACs. Contig WC1.23 shares 14 STSs and 38 YACs with our contig, two ESTs from WC1.23 (D1S2359 and D1S14662) belong to the same cDNA families as D1S2067E and D1S2049E which we mapped independently, using radiation hybrids. The order of the shared markers is consistent except for the doublet (D1S242, D1S2814) which is reversed. The order proposed on WC1.23 for these 2 STSs appears very unlikely as it requires introducing a gap within 7 YACs in our contig. Moreover, in the MIT data, YACs 756c3 and 775e10 are given as negative for D1S2814, and YACs 945g6 and 948g10 as negative for D1S242 whereas we verified that all four YACs were positive for both markers. Therefore this small discrepancy regarding the order of the markers most probably results from false-negative results in the MIT data set, and we believe the order we have deter-

mined is more likely. About 14 STSs and 31 YACs from WC1.23 were not used in the present study. Conversely, our YAC continuum was generated with 24 STSs and 18 YACs not previously described (see table 3). In addition, each YAC/STS hit was confirmed by individual PCR assay and most of the YACs (49 in total) are now also characterized either by FISH or insert-end chromosomal assignment or both.

At the present time, the disease locus interval is still very large and this might hamper the gene search. Refinement of the GLC1A localization has been attempted by assaying all microsatellite markers mapped in this contig on our families of patients (data not shown). Although some recombination events were found among unaffected members (confirming the order of some markers), it has not yet been possible to narrow the GLC1A interval. Recruitment of new patients and linkage analysis remain a priority for identifying the GLC1A gene.

## References

1 Leske MC: The epidemiology of open-angle glaucoma: A review. Am J Epidemiol 1983; 118:166–191.

2 Sheffield VC, Stone EM, Alward WL, Drack AV, Johnson AT, Streb LM, Nichols BE: Genetic linkage of familial open-angle glaucoma to chromosome 1q21-q31. Nat Genet 1993;4: 47–50.

3 Morissette J, Cote G, Anctil JL, Plante M, Amyot M, Heon E, Trope GE, Weissenbach J, Raymond V: A common gene for juvenile and adult-onset primary open-angle glaucomas confined on chromosome 1q. Am J Hum Genet 1995;56:1431–1442.

4 Graff C, Urbak SF, Jerndal T, Wadelius C: Confirmation of linkage to 1q21-q31 in a Danish autosomal dominant juvenile-onset glaucoma family and evidence of genetic heterogeneity. Hum Genet 1995;96:285–289.

5 Chumakov IM, Rigault P, Le Gall I, Bellanné-Chantelot C, Billault A, Guillou S, Soularue P, Guasconi G, Poullier E, Gros I, Belova M, Sambucy JL, Susini L, Gervy P, Glibert F, Beaufils S, Bui H, Massart C, Detand MF,

Dukasz F, Lecoulant S, Ougen P, Perrot V, Saumier M, Soravito C, Bahouayila R, Cohen-Akenine A, Barillot E, Bertrand S, Codani JJ, Caterina D, Georges I, Lacroix B, Lucotte G, Sahbatou M, Schmit C, Sangouard M, Tubacher E, Dib C, Faure S, Fizames C, Gyapay G, Millasseau P, NGuyen S, Muselet D, Vignal A, Morissette J, Menninger J, Lieman J, Menninger H, Desai T, Banks A, Bray-Ward P, Ward D, Hudson T, Gerety S, Foote S, Stein L, Page DC, Weissenbach J, Le Paslier D, Cohen D: A YAC contig map of the human genome. Nature Genome Directory 1995;377:175–297.

6 Chumakov IM, Le Gall I, Billault A, Ougen P, Soularue P, Guillou S, Rigault P, Bui H, De Tand MF, Barillot E, Abderrahim H, Cherif D, Berger R, Le Paslier D, Cohen D: Isolation of chromosome 21-specific yeast artificial chromosomes from a total human genome library. Nat Genet 1992;1:222–225.

7 Bellanne-Chantelot C, Lacroix B, Ougen P, Billault A, Beaufils S, Bertrand S, Georges I, Glibert F, Gros I, Lucotte G, Susini L, Codani JJ, Gesnouin P, Pook S, Vayssex G, Lu-Kuo J, Ried T, Ward D, Chumakov I, Le Paslier D, Barillot E, Cohen D: Mapping the whole human genome by fingerprinting yeast artificial chromosomes. Cell 1992;70:1059–1068.

8 Dausset J, Ougen P, Abderrahim H, Billault A, Sambucy JL, Cohen D, Le Paslier D: The CEPH YAC library. Behring Inst Mitt 1992;91: 13–20.

9 Abderrahim H, Sambucy JL, Iris F, Ougen P, Billault P, Chumakov IM, Dausset J, Cohen D, Le Paslier D: Cloning the human major histocompatibility complex in YACs. Genomics 1994;23:520–527.

10 Rychlik W, Rhoads RE: A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and in vitro amplification of DNA. Nucleic Acids Res 1989;17: 8543–8551.

11 Pearson WR, Lipman DJ: Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 1988;85:2444–2448.

12 Drwinga HL, Toji LH, Kim CH, Greene AE, Mulivor RA: NIGMS human/rodent somatic cell hybrid mapping panels 1 and 2. Genomics 1993;16:311–314.

13 Driesen MS, Dauwerse JG, Wapenaar MC, Meershoek EJ, Mollevanger P, Chen KL, Fischbeck KH, van Ommen GJB: Generation and fluorescent in situ hybridization mapping of yeast artificial chromosomes of 1p, 17p, 17q and 19q from a hybrid cell line by high-density screening of an amplified library. Genomics 1991;11:1079–1087.

14 Dauwerse JG, Jumelet EA, Wessels JW, Saris JJ, Hagemeijeir A, Beverstock GC, van Ommen GJB, Breuning MH: Extensive crosshomology between the long arm and short arm of chromosome 16 may explain leukemic inversions and translocations. Blood 1992;79:1299–1304.

15 Dib C, Fauré 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: A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature 1996;380:152–154.

16 Cohen D, Chumakov I, Weissenbach J: A first-generation physical map of the human genome. Nature 1993;366:698–701.

17 Gyapay G, Schmitt K, Fizames C, Hywel J, Vega-Czarny N, Spillett D, Muselet D, Prud'Homme J-F, Dib C, Auffray C, Morissette J, Weissenbach J, Goodfellow PN: A radiation hybrid map of the human genome. Hum Mol Genet 1996;5:339–346.

18 Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, Marc S, Bernardi G, Lathrop M, Weissenbach J: The 1993–94 Généthon human genetic linkage map. Nat Genet 1994;7:246–339.

19 Dracopoli NC, O'Connell P, Elsner TI, Lalouel JM, White RL, Buetow KH, Nishimura DY, Murray JC, Helms C, Mishra SK, Donis-Keller H, Hall JM, Lee MK, King MC, Attwood J, Morton N, Robson E, Mahtani M, Willard HF, Royle NJ, Patel I, Jeffreys AJ, Verga V, Jenkins T, Weber J, Mitchell A, Bale AE: The CEPH consortium linkage map of human chromosome 1. Genomics 1991;9:686–700.

20 Auffray C, Behar G, Bois F, Bouchier C, Dasilva C, Devignes MD, Duprat S, Houlgatte R, Jumeau MN, Lamy B, Lorenzo F, Mitchell H, Mariage SR, Pietu G, Pouliot Y, Sebastiani KC, Tessier A: IMAGE: integrated molecular analysis of the human genome and its expression. C R Acad Sci III 1995;318:263–272.