Am. J. Hum. Genet. 61:1112–1116, 1997

Localization of the Congenital Dyserythropoietic Anemia II Locus to Chromosome 20q11.2 by Genomewide Search

P. Gasparini,[1] E. Miraglia del Giudice,[2] J. Delaunay,[3] A. Totaro,[1] M. Granatiero,[1] S. Melchionda,[1] L. Zelante,[1] and A. Iolascon[4]

[1]Servizio di Genetica Medica, IRCCS-CSS, San Giovanni Rotondo (Fg), Italy; [2]Dipartimento di Pediatria, II Università Napoli, Naples; [3]Service d'Hématologie/U299, Le Kremlin-Bicêtre, France; and [4]Dipartimento di Biomedicina dell'Età Evolutiva, Università di Bari, Bari, Italy

Summary

Congenital dyserythropoietic anemias (CDA) are genetic disorders characterized by anemia and ineffective erythropoiesis. Three main types of CDA have been distinguished: CDA I and CDA III, whose loci have been already mapped, and CDA II (MIM 224100), the most frequent among CDAs, which is transmitted as an autosomal recessive trait and is known also as "HEMPAS" (*h*ereditary *e*rythroblast *m*ultinuclearity with *p*ositive *a*cidified *s*erum). We have recruited a panel of well-characterized CDA II families and have used them to search for the CDA II gene by linkage analysis. After the exclusion of three candidate genes, we obtained conclusive evidence for linkage of CDA II to microsatellite markers on the long arm of chromosome 20 (20q11.2). A maximum two-point LOD score of 5.4 at a recombination fraction of .00 was obtained with marker D20S863. Strong evidence of allelic association with the disease was detected with the same marker. Some recombinational events established a maximum candidate interval of ~5 cM.

Introduction

Congenital dyserythropoietic anemias (CDA) are genetic disorders characterized by anemia and ineffective erythropoiesis. Three main types of CDA have been distinguished: CDA I, recently mapped to 15q15.1-15.3 (Tamary et al. 1996); CDA III, mapped ~20 cM telomeric to the CDA I locus (Lind et al. 1995); and CDA II (MIM 224100), the most frequent among CDAs, which is transmitted as an autosomal recessive trait and is known also as "HEMPAS" (*h*ereditary *e*rythroblast *m*ultinuclearity with *p*ositive *a*cidified *s*erum) (Crookston et al. 1969). It is characterized by normocytic anemia, variable jaundice, and hepatosplenomegaly. Gallbladder disease and hemochromatosis are frequent complications (Iolascon et al. 1996). Bone-marrow histology reveals bi- and multinucleated (10%–40%) erythroblasts with karyorrhexis and additional abnormalities detectable by electron microscopy (Wong et al. 1972). Knowledge about CDA II has undergone two fundamental steps. The first was the discovery of the red blood cell (RBC) membrane-protein abnormality, hypothesized on the basis of increased agglutinability and Ham test positivity. SDS-PAGE of RBC membrane proteins shows a narrower aspect and a faster migration of band 3 (anion-exchange transporter) (Anselstetter et al. 1977; Alloisio et al. 1996). CDA I differs from CDA II because it does not show any band 3 biochemical alteration, whereas CDA III is characterized by different bone-marrow findings and is inherited in a dominant way.

The second step was the discovery of the protein-glycosylation defect. Carbohydrate-structure analysis of both membrane glycoproteins and soluble glycoproteins, as well as biochemical data, indicate that CDA II is associated with defects in the biosynthesis of complex N-linked oligosaccharides (Fukuda et al. 1987, 1990; Fukuda 1990) . The altered distribution of membrane glycoproteins may result in morphologically visible abnormalities and could interfere with functions required for intracellular transport and compartmentalization, cell division, and differentiation. These structural findings could be partially explained by the deficiency of some enzymes, such as N-acetylglucosaminyltransferase II (GnT-II) or alpha-mannosidase II (alpha-Man II) (Fukuda et al. 1987, 1990; Fukuda 1990), a deficiency that already has been demonstrated in vitro in several CDA II patients. Genes encoding for these enzymes have been cloned (Moremen and Robbins 1991; Misago et al. 1995; Tan et al. 1995), but linkage analysis recently has excluded them as candidate genes (Iolascon et al., in press).

To localize the gene for CDA II, a linkage study using

Received May 29, 1997; accepted for publication September 8, 1997; electronically published October 29, 1997.

Address for correspondence and reprints: Dr. Paolo Gasparini, Servizio di Genetica Medica, IRCCS-"CSS" San Giovanni R. (FG), Italy. E-mail: genetcss@fg.nettuno.it

0002-9297/97/6105-0016$02.00

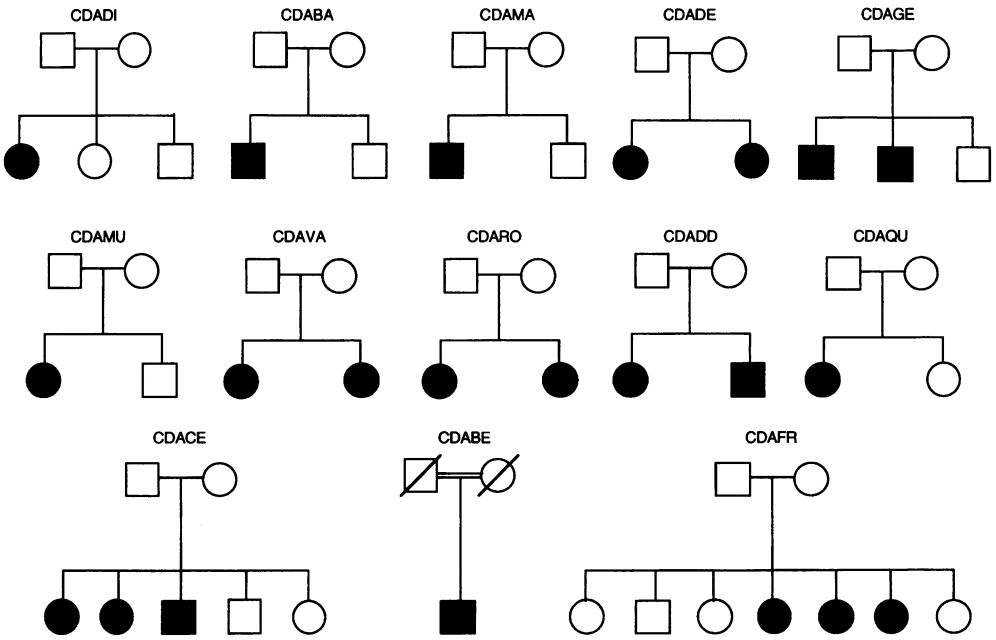

Figure 1 Schematic representation of CDA II pedigrees included in present study

maps of microsatellites was performed. To facilitate the genome search, the screening was initially limited to the affected sibs of a subset of six families. One of these families was the result of a mating between first cousins; thus allele homozygosity was sought in the affected child. In the sibs of the remaining five families, identity by descent was expected. Alleles of these loci were then typed in the corresponding parents and members of the family.

Families and Methods

Families

Twelve Italian families and one French family, all unrelated, were ascertained through pediatric hematology centers and were included in the study. Six of them were each characterized by at least two affected individuals, whereas the remaining six were each characterized by one affected sib and one or more normal sibs. Finally, a consanguineous pedigree, the result of a mating between first cousins, was also selected and included. The pedigrees of the families enrolled in the study are shown in figure 1; clinical details of the patients have been described elsewhere (Alloisio et al. 1996). The age at onset was 1–14 years. All individuals included in this study were accurately investigated at the clinical and laboratory levels (blood-cell count and RBC membrane-protein PAGE). Diagnosis was based on the following criteria: anemia of variable degree, jaundice, and mild hepatosplenomegaly. Bone-marrow observation demonstrated the presence of the typical morphological abnormalities. SDS-PAGE (Wong et al. 1972) showed, in all affected cases, a band 3 with a narrower aspect and a faster migration. The presence of 74-kD and 58-kD proteins was established by means of immunoblotting using polyclonal rabbit antibodies against rat GRP78 (affinity bioreagents; Neshanic Station) and rabbit antibodies against human calreticulin (Alloisio et al. 1996). Peripheral blood was obtained from all subjects, and DNA was isolated from blood leukocytes, according to standard methods.

Search Strategy and Microsatellite Analysis

For the genomewide search, the ABI PRISM linkage mapping set (Perkin-Elmer) was used. It is characterized by >375 markers that define a 10-cM-resolution human index map (Gyapay et al. 1994). PCR reactions using fluorescently labeled primers were run under the conditions suggested by the supplier. An aliquot of PCR reaction was run in an ABI PRISM 373 or an 377 DNA sequencer, and results were processed by GENESCAN software. Allele assignment was performed by use of Genotyper software. To saturate the chromosome 20–positive region, additional pairs of fluorescently labeled primers were specifically synthesized for the following markers: D20S912, D20S200, D20S890, D20S884, D20S863, and D20S908 (Dib et al. 1996). All living individuals of each family were genotyped and thus contributed to the linkage calculations.

Table 1

Pairwise LOD Scores between CDAII and Chromosome 20 Markers

	LOD Score at Recombination Fraction of								
Marker	.0	.01	.05	.10	.20	.30	.40	Z_{max}	θ_{max}
D20S912	∞	2.01	2.24	1.97	1.22	.57	.16	2.27	.037
D20S200	∞	.47	1.92	2.10	1.62	.89	.26	2.11	.090
D20S890	∞	−1.72	.58	1.17	1.15	.69	.21	1.26	.140
D20S195	4.50	4.39	3.92	3.34	2.18	1.14	.34	4.50	.000
D20S863	5.41	5.25	4.65	3.91	2.50	1.28	.37	5.41	.000
D20S884	∞	4.74	4.76	4.22	2.86	1.53	.45	4.89	.026
D20S908	∞	.36	1.86	2.04	1.58	.88	.26	2.05	.090
D20S107	∞	−1.06	.69	1.14	1.08	.65	.20	1.20	.130

Linkage Analysis

Statistical analysis was performed on the basis of an autosomal disease with complete penetrance. The disease-gene frequency was set to .012, and all marker alleles were considered to be equally frequent. Two-point linkage analysis was performed by use of the MLINK program, version 5.1, from the LINKAGE computer package (Ott 1992). Loops of consanguinity were accommodated as suggested by Ott (1992). Values for maximum LOD score (Z_{max}) were calculated with the ILINK program, from the same computer package. The ~95% confidence limits for the maximum recombination fraction (θ_{max}) at the Z_{max} were calculated by the 1 − LOD–down method (Ott 1992). Alleles were down-coded, without loss of informativity, to reduce computing time. The HOMOG program was used to test for nonallelic heterogeneity, by use of pairwise LOD scores between D20S863 and the disease locus. The statistical significance of linkage disequilibrium with the disease was tested by utilization of the χ^2 test of homogeneity for each marker and then was standardized to obtain the delta coefficient, whose range is 0–1, to measure the association. The delta coefficient is defined as $(ad - bc)/\sqrt{[(a + b)(c + d)(a + c)(b + d)]}$, where a and b are the frequencies of allele 1 in affected and unaffected chromosomes, respectively, and c and d are the frequencies of allele 2 in the same sets of chromosomes.

Results

After analysis of data from 145 markers, representing ~50% of the human genome, evidence for linkage was obtained with markers D20S195 and D20S107, whereas negative results were obtained for the remaining markers, including those of chromosome 15, where both the CDA I locus and the CDA III locus have been mapped. Six additional markers from the region—D20S912, D20S200, D20S890, D20S884, D20S863, and D20S908—were then typed, both in all individuals of the six families used for the genomewide search and in an additional six families each containing a single affected individual and one or two unaffected siblings. Finally, a family characterized by three affected and three healthy children was included too. The order and the distances of the aforementioned markers, as deduced from published maps (Dib et al. 1996), are cen–D20S912–5 cM–D20S200–3 cM–D20S890–D20S195–D20S863–D20S884–3 cM–D20S908–1 cM–D20S107–tel. Pairwise LOD scores for the eight markers are shown in table 1, and the linkage map for these markers is shown in figure 2.

The highest LOD score (Z_{max} = 5.41 at θ = .00) was obtained with D20S863 , but no recombinants were detected also with D20S195 (Z_{max} = 4.5 at θ = .00). Positive results were also obtained with the remaining six markers, at $\theta > 0$. The HOMOG program detected no

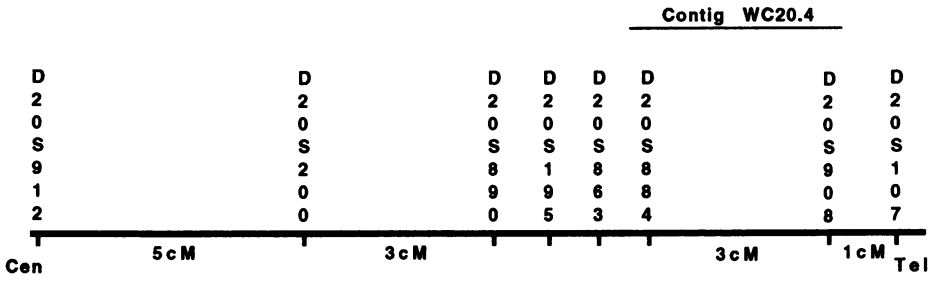

Figure 2 Linkage map of microsatellite markers from chromosome 20q11.2. Genetic distances are indicated (in cM). The position of contig WC20.4 (Whitehead Institute/MIT Center for Genome Research) is also reported.

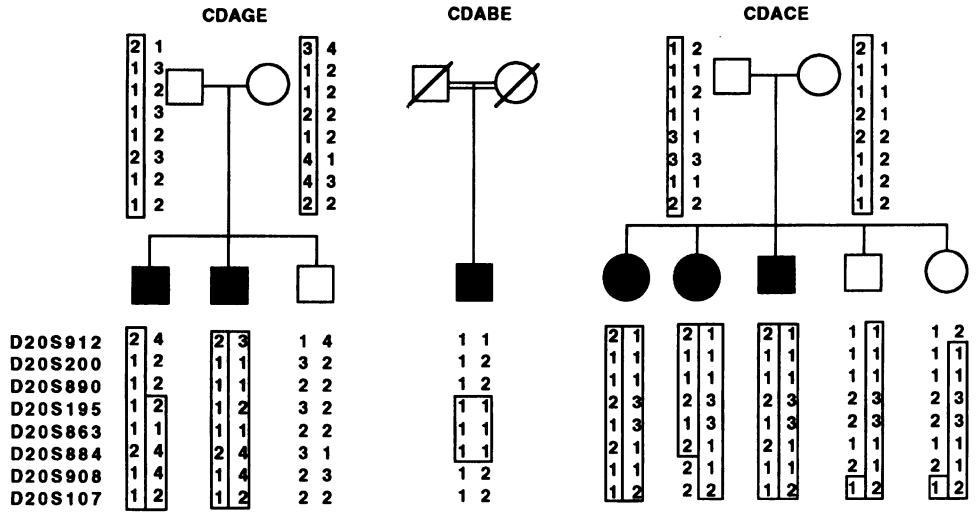

Figure 3 Examples of some recombinational events that help in refining the CDA II gene location (families CDAGE and CDACE). A consanguineous family (with mating between two first cousins), CDABE, is also reported to confirm, by homozygosity, the region defined by the recombinants. The order of markers is D20S912-D20S200-D20S890-D20S195-D20S863-D20S884-D20S908-D29S107.

evidence in support of genetic heterogeneity (χ^2 = 24.928; α = 1). Several recombinations were detected that help in definition of the limits of the CDA II gene location. The centromeric limit is clearly defined by marker D20S890, which shows a recombinational event in family CDAGE (fig. 3), whereas additional recombinants have been detected in the sample with markers located more centromeric (D20S200 and D20S912). A recombinant in family CDACE indicates that the CDAII gene is located proximal to D20S908 (fig. 3). These limits are confirmed by the consanguineous family, in which allele homozygosity was expected in the affected child. As is shown in figure 3, he is homozygous for the alleles of D20S863, D20S195, and of D20S884, which are located within the region previously defined. Haplotype analysis based on the use of these three markers revealed in this subject the presence of a rare haplotype (141-227-137), further confirming the role of homozygosity in the definition of the limits for the CDA II locus in this family. The region spanning D20S890–D20S908 is ~5 cM. Since these two markers are located on chromosome 20q11.2, we propose this region as the location for the CDA II gene.

Further evidence for the location of the CDA II gene in this interval was obtained by linkage-disequilibrium analysis performed by use of one randomly affected individual for each unrelated family. As already mentioned, no recombinants for loci D20S195 and D20S863 were detected in the families that we studied. This last locus also exhibited strong allelic association with the disease (χ^2 = 12.74; $P < .0001$). Allele 2 was present in 20/26 (.769) of the inferred disease chromosomes, compared with 5/23 (.217) of the nondisease chromosomes, giving a disequilibrium coefficient of δ = .552. A different, but not statistically significant, allele distribution was also observed with D20S195 (χ^2 = 2.07; P = .15), with allele 6 being present in 4/26 (.153) disease chromosomes and never in normal chromosomes (δ = .288) and with D20S908 (χ^2 = 1.05; P = .22), with allele 5 being present in 11/26 (.423) affected chromosomes and in 5/23 (.217) nondisease chromosomes (δ = .22). The strong association of the disease with D20S863 allele 2 suggests the presence of a major mutation and, most likely, a founder effect. This finding is in agreement with the origin of our sample, which comes mainly from southern Italy (i.e., Campania and the areas surrounding it).

Discussion

Very recently we have excluded, by linkage analysis, GnT-II, alpha-Man II, and alpha-mannosidase X as candidates for the disease (Iolascon et al., in press). In the present paper, we report the identification of the locus for CDA II, the most common form of CDA, after the exclusion of ~50% of the human genome by a genome-wide search performed by use of panels of fluorescently labeled microsatellites. This finding is an important step toward the cloning of the gene itself. The candidate region here defined is quite small, ~5 cM, but additional analysis with new markers in an expanded set of families should lead to a further refinement of the CDA II location. Of the several genes mapped within this region, none has any apparent relationship with the disease. In any case, the region marked by D20S884, D20S195, and D20S908 is already contained in an available YAC con-

tig (WC20.4; Whitehead Institute/MIT Center for Genome Research). As soon as it is extended to markers D20S863 and D20S890, it will be useful for cloning the CDA II gene. cDNA selection and exon-trapping techniques will be useful for isolation of expressed sequences and, together with already cloned ESTs, for construction of a transcriptional map of the region. The identification of the CDA II gene will lead to the understanding of the molecular basis of one of the few inborn errors of metabolism that involve the biosynthesis of a glycoprotein. Finally, our results clearly demonstrate that the clinical and laboratory differences detectable between CDA II and the two other forms of CDA (CDA I and CDA III) are certainly due to a different genetic background.

Acknowledgments

We thank the families and the physicians associated with Associazione Italiana Ematologia Oncologia Pediatrica who collaborated in this work. This work was supported in part by the Ministero Italiano della Sanitá (support to P.G.) and the Associazione Italiana Ricerca Cancro (support to A.I.).

References

Alloisio N, Texier P, Denoroy L, Berger C, Miraglia del Giudice E, Perrotta S, Iolascon A, et al (1996) The cisternae decorating the red blood cell membrane in congenital dyserythropoietic anemia (type II) originate from the endoplasmic reticulum. Blood 87:4433–4438

Anselstetter V, Horstmann HJ, Heimpel H (1977) Congenital dyserythropoietic anemia type I and II: aberrant pattern of erythrocyte membrane proteins in CDA II, as revealed by two-dimensional polyacrylamide gel electrophoresis. Br J Haematol 35:209–217

Crookston JH, Crookston MC, Burnie KL, Fracombe WH, Dacie JV, Davis H, Lewis SM (1969) Hereditary erythroblastic multinuclearity associated with a positive acidified-serum test: a typical congenital dyserythropoietic anemia. Br J Haematol 17:11–23

Dib C, Fauré S, Fizames C, Samson D, Drout N, Vignal P, Millasseau P, et al (1996) A comprehensive genetic map of the human genome based on 5264 microsatellites. Nature 380:152–154

Fukuda MN (1990) HEMPAS disease: genetic defect of glycosilation. Glycobiology 1:9–15

Fukuda MN, Dell A, Scartezzini P (1987) Primary defect congenital dyserythropoietic anemia type II: failure in glycosylation of erythrocyte lactosaminoglycan proteins caused by lowered N acteylglucosaminyltransferase II. J Biol Chem 262:7195–7199

Fukuda MN, Masri KA, Dell A, Luzzatto L, Moremen KW (1990) Incomplete synthesis of N-glycans in congenital dyserythropoietic anemia type II (HEMPAS) caused by a gene defect encoding a mannosidase II. Proc Natl Acad Sci USA 87:7443-7448

Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, Samson D, et al (1994) The 1993–1994 Généthon human genetic linkage map. Nat Genet 7:246–339

Iolascon A, D'Agostaro G, Perrotta S, Izzo P, Tavano R, Miraglia del Giudice E (1996) Congenital dyserythropoietic anemia type III: molecular basis and clinical aspects. Haematologica 81:543–559

Iolascon A, Miraglia del Giudice E, Perrotta S, Granatiero M, Zelante L, Gasparini P. Exclusion of three candidate genes as determinants of congenital dyserythropoietic anemia type II. Blood (in press)

Lind L, Sandstrom H, Wahlin A, Eriksson M, Nilsson-Sojka B, Sikstrom C, Holmgren G, et al (1995) Localization of the gene for congenital dyserythropoietic anemia type III, CDAN3, to chromosome 15q21-25. Hum Mol Genet 4: 109–112

Misago M, Liao YF, Kudo S, Eto S, Mattei MG, Moremen KW, Fukuda MN (1995) Molecular cloning and expression of cDNAs encoding human a-mannosidase II and a previously unrecognized a-mannosidase IIx isozyme. Proc Natl Acad Sci USA 92:11766–11770

Moremen KW, Robbins PW (1991) Isolation, characterization and expression of cDNAs encoding murine a-mannosidase II, a Golgi enzyme that controls conversion of high mannose to complex N-glycans. J Cell Biol 115:1521–1534

Ott J (1992) Analysis of human linkage, 2d ed. John Hopkins University Press, Baltimore

Tamary H, Shalmon L, Shalev H, Halil A, Shaft D, Zoldan M, Resnitzky P, et al (1996) Localization of the gene for congenital dyserythropoietic anemia type I to chromosome 15q15.1-15.3. Blood 88 Suppl 1:144a

Tan J, D'Agostaro G, Bendiak B, Reck F, Sarkar M, Squire JA, Leong P, et al (1995) The human UDP-N-acetylglucosamine: a-6-D-mannoside-b-1,2-N-acetylglucosaminyltransferase gene (MGAT2): cloning of genomic DNA, localization to chromosome 14q21, expression in insect cells and purification of the recombinant protein. Eur J Biochem 231: 317–328

Wong KY, Hug G, Lampkin BC (1972) Congenital dyserythropoietic anemia type II: ultrastructural and radioautographic studies of blood and bone marrow. Blood 39: 23–30