



ARTICLE

Physical and transcriptional map of the hereditary inclusion body myopathy locus on chromosome 9p12-p13

Iris Eisenberg¹, Hagit Hochner¹, Moshe Shemesh¹, Tatjana Levi², Tamara Potikha¹, Menachem Sadeh³, Zohar Argov⁴, Cynthia L Jackson⁵ and Stella Mitrani-Rosenbaum^{*,1}

¹The Unit for Development of Molecular Biology and Genetic Engineering, Hadassah Hospital, The Hebrew University-Hadassah Medical School, Jerusalem 91240, Israel; ²Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA; ³Department of Neurology, Wolfson Hospital, Holon, Israel; ⁴Department of Neurology, Hadassah Hospital, The Hebrew University-Hadassah Medical School, Jerusalem 91240, Israel; ⁵Department of Pathology, Rhode Island Hospital and Brown University, Providence, Rhode Island 02192, USA

Hereditary inclusion body myopathy (HIBM) is a group of neuromuscular disorders characterised by adult-onset, slowly progressive distal and proximal muscle weakness and typical muscle pathology. Previously, we have mapped the gene responsible for a recessive form of HIBM to chromosome 9p1 and narrowed the interval to one single YAC clone of 1 Mb in size. As a further step towards the identification of the HIBM gene, we have constructed a detailed physical and transcriptional map of this region. A high resolution BAC contig that includes the HIBM critical region, flanked by marker 327G74 and D9S1859, was constructed. This contig allowed the precise localisation of 25 genes and ESTs to the proximal region of chromosome 9. The expression pattern of those mapped genes and ESTs was established by Northern blot analysis. In the process of refining the HIBM interval, 13 new polymorphic markers were identified, of which 11 are CA-repeats, and two are single nucleotide polymorphisms. Certainly, this map provides an important integration of physical and transcriptional information corresponding to chromosome 9p12-p13, which is expected to facilitate the cloning and identification not only of the HIBM gene, but also other disease genes which map to this region. *European Journal of Human Genetics* (2001) 9, 501–509.

Keywords: chromosome 9; IBM2; physical and transcript map; BAC contig; polymorphisms; neuromuscular diseases.

Introduction

Hereditary inclusion body myopathy (HIBM) (OMIM 600737) is a unique group of neuromuscular disorders characterised by adult onset, slowly progressive distal and proximal muscle weakness and a typical muscle pathology including rimmed vacuoles and filamentous inclusions.^{1–3} A form of HIBM found in Jews of Persian descent affects most leg muscles except the quadriceps. Our previous studies

localised the gene causing this recessive form of HIBM in Persian Jews within a genomic interval of approximately 1 Mb covered by a single YAC, on chromosome 9p1.^{4–7} Subsequently Christodoulou *et al*⁸ have reported the same locus involved in quadriceps sparing HIBM in additional Jewish patients from Persian origin, and in other ethnic groups, and very recently Mirabella *et al*⁹ have described an Italian family with autosomal recessive quadriceps-sparing inclusion-body myopathy also linked to this locus. Ikeuchi *et al*¹⁰ have also reported that a distal myopathy with rimmed vacuoles observed in Japanese patients maps to the same locus. This disorder is very similar to HIBM in Persian Jews: onset in early adulthood with drop foot; biopsies show rimmed vacuoles and inclusions of the same nature. The

*Correspondence: S Mitrani-Rosenbaum, The Unit for Development of Molecular Biology and Genetic Engineering, Hadassah Hospital-Mount Scopus, Jerusalem 91240, Israel. Tel: 972-2-5844724; fax: 972-2-5819134; E-mail: stella@yam-suff.huji.ac.il
Received 2 January 2001; revised 15 March 2001; accepted 6 April 2001

main difference consists in the progression which eventually involves the quadriceps within 10 years of onset, resulting in early loss of ambulation of the Japanese patients. Today it is agreed that very likely these conditions are allelic forms of the same disorder. Many previous studies have shown that the short arm of chromosome 9 and in particular the region 9p12-p13 plays an important role in a large number of common diseases, such as acute lymphoblastic leukaemia,¹¹ non-Hodgkin's lymphoma,¹² lung cancer,¹³ hepatocellular carcinoma (HCC),¹⁴ as well as in rare diseases, such as arthrogyriposis multiplex congenita, distal, type 1 (AMCD1),¹⁵ acromesomeric dysplasia (AMDM),¹⁶ cartilage hair hypoplasia (CHH),¹⁷ and also in a form of distal hereditary motor neuropathy (HMN) found in Jerash.¹⁸ At present, the availability of a precise physical map of chromosome 9p12-p13 is very restricted, and very few mapped genes have been identified in this region. Therefore, the construction of a contiguous, high-resolution physical and transcriptional map in this defined genomic interval will certainly contribute to the positional cloning of various disease causing genes. In our attempt to identify the HIBM gene, we have established a complete BAC contig of the chromosomal region 9p12-p13, across the β -tropomyosin gene and the 3' end of the PAX5 gene. This contig has allowed the high resolution physical and transcriptional mapping of several genes, expressed sequence tags, and various additional markers in the region. We have precisely positioned 25 transcripts and analysed their expression pattern. Thirteen new polymorphic sites were identified in this study and fine recombination analysis with these markers allowed the restriction of the HIBM critical region between markers 327GT4 and D9S1859, an interval of about 700 kb. Among the genes and ESTs located between these flanking markers, 12 prevail as potential candidate genes for HIBM.

Materials and methods

Construction of a high resolution BAC contig

BAC clones were identified by PCR screening of human BAC libraries (Down to the wellTM—GenomeSystems Inc; CITB library, Release IV—Research Genetics Inc), according to the manufacturers' instructions, or from the released data of the Sanger Center (RPC1-11 library—BACPAC Resources). BAC DNA from single clones was prepared using the QIAGEN Plasmid Maxi kit. The BAC DNA ends were directly sequenced using primers T7 and Sp6, on an ABI Prism 377 automated sequencer. All end sequences were analysed for repetitive elements using the RepeatMasker web server, and subsequently used as STS to screen the BAC libraries again. This process was repeated until a complete BAC contig was established. To confirm the precision of the contig constructed, the content of each clone in the contig was assessed twice by high stringent hybridisation of known markers and by 40 novel STS markers derived from the BACs ends.

Identification of novel polymorphic markers

BAC clones were partially restricted with *Sau3A* (Boehringer Mannheim) and shotgun cloned into *Bam*HI linearised pSPL3 vector (Gibco-BRL). Plated colonies were transferred to nylon membranes (OSMONICS) and hybridised with a Poly(dA-dC)·Poly(dG-dT) probe (Amersham-Pharmacia Biotech). Plasmid inserts from positive colonies were sequenced and PCR amplification primers flanking the repeated region were designed. By searching the Sanger Center recently released BAC sequences, we have been able to identify additional repeat sequences. These sequence data were produced by the Human Chromosome 9 Sequencing Group at the Sanger Center and can be obtained from ftp://ftp.sanger.ac.uk/pub/human/sequences/Chr_9/. Each marker was tested on several unrelated individuals to determine its polymorphic status.

Genotyping with the newly identified markers was performed using amplification with radioactive labelled PCR primers, and denaturing polyacrylamide gel electrophoresis as previously described.⁴

Transcript map

All genes and ESTs previously mapped to the region between D9S163 and D9S50 were selected for further localisation based on the following maps and databases: The human transcript map¹⁹ (<http://www.ncbi.nlm.nih.gov/genemap/>); The Sanger Center—Map 9ctg3 (<http://www.sanger.ac.uk/HGP/Chr9/>); The Stanford RH map (<http://shgc-www.stanford.edu/Mapping/rh/search.html>); The NCBI Unigene database (<http://www.ncbi.nlm.nih.gov/UniGene/index.html>); The Genexpress IMAGE knowledge Base of the Human Muscle Transcriptome: http://idifix-upr420.vjf.cnrs.fr/IMAGE/Page_unique/welcome_muscles.html.

All the selected ESTs were precisely mapped within the HIBM contig by high stringency hybridisation. In addition, all new genomic sequences derived from the BAC clone insert ends and the shotgun sequences were systematically compared with EST databases using the BLASTN 2.0 software.²⁰ Sequences found to match additional ESTs were also tested against the BAC contig by PCR amplification and hybridisation. The resulting integrated map is shown in Figure 1.

Northern blot analysis

To assess tissue distribution and sizes the transcripts mapped inside and in the vicinity of the candidate interval, multiple tissue Northern (MTN, Clontech) blots were used. Hybridisation in ExpressHyb solution (Clontech) and washes were done as recommended by the manufacturer. The blots contain polyA RNA from the following tissues: heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas. Hybridisation probes were prepared by [α -³²P]dCTP labeled PCR generated DNA fragments. Table 1 summarises all primers used for these ESTs amplifications.

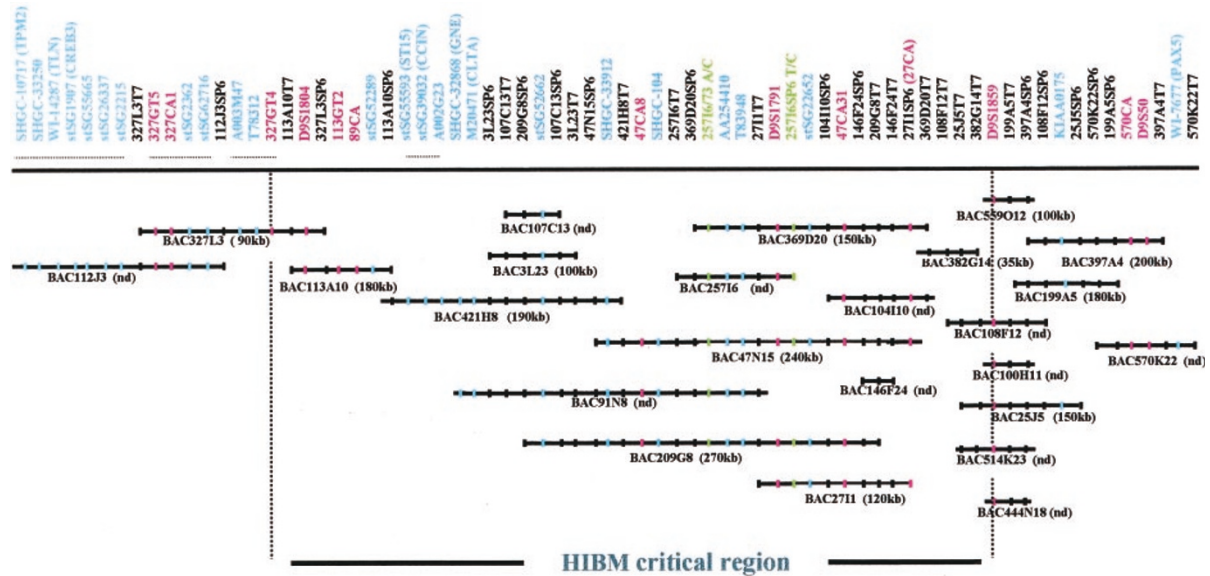


Figure 1 Physical and transcript map of the HIBM region on chromosome 9p12-p13. BAC clones are represented by horizontal lines. The name of each clone and its estimated size is shown below (nd=not done). Filled squares indicate the presence of markers in the clones. Black squares indicate novel STSs generated from BAC ends; blue squares show ESTs and genes; polymorphic markers are in red and the two novel SNPs are indicated in green. 89CA is a polymorphic marker described by Vakkilainen *et al.*³⁴ Non polymorphic microsatellites are not shown. The precise location of BAC59O12, BAC100H11, BAC514K23 and BAC444N18 ends was not tested. The internal order in the underlined clusters is not definitive. Note that the map is not to scale.

Electronic database information

Data were taken from the following websites:

- (1) Whitehead Institute for Biomedical Research/MIT Center for Genome Research (WICGR): <http://www.genome.wi.mit.edu>
- (2) The Genome database: <http://www.gdb.org>
- (3) Unigene database chromosome 9: http://www.ncbi.nlm.nih.gov/UniGene/Hs_DATA/ChromLists/Chr9.html
- (4) Blast search: <http://www.ncbi.nlm.nih.gov/BLAST>
- (5) The human transcript map (GeneMap99): <http://www.ncbi.nlm.nih.gov/genemap>
- (6) FASTA searches: <http://www2.igh.cnrs.fr/bin/fasta-guess.cgi>
- (7) The Sanger Center database: <http://www.sanger.ac.uk>
- (8) Summary map of chromosome 9: <http://cedar.genetics.soton.ac.uk/pub/chrom9/map.html>
- (9) The Stanford RH map: <http://shgc-www.stanford.edu/Mapping/rh/search.html>
- (10) The cooperative human linkage center: <http://cgap.nci.nih.gov/CHLC>
- (11) The Genexpress IMAGE knowledge Base of the Human Muscle Transcriptome: http://idifix.upr420.vjf.cnrs.fr/IMAGE/Page_unique/welcome_muscles.html
- (12) The RepeatMasker web server: http://ftp.genome.washington.edu/cgi_bin/RepeatMasker.

Pulsed-field gel electrophoresis

NotI (New England Biolabs, Inc.) digested BAC DNAs were subjected to 1% agarose gels (Seakem, FMC BioProducts), and separated by pulsed-field gel electrophoresis with a CHEF DR-II system (BioRad), in 0.5% TBE at 14°C for 16 h, 6.0 V/cm, with an initial and final switch time of 30 s. Following electrophoresis, the gel was stained with ethidium bromide for visualisation of the digested fragments.

Results

Construction of BAC contig

As a further resource for cloning the HIBM disease gene, a genomic BAC contig flanked by the β -tropomyosin gene and the PAX5 gene, encompassing the HIBM critical region, in the chromosome 9p12-p13 region, was generated. BAC libraries were initially screened using three anchor markers *D9S1804*, *D9S1791* and *D9S50*. Ten BAC clones were identified as positive for these markers (Figure 1). The end sequences of the identified clones were determined by direct sequencing of BAC DNA using primers T7 and Sp6, complementary to the BAC vector sequence. The end sequences were used to develop STS markers for screening BAC libraries again. This process was repeated until all gaps were closed and a complete BAC contig was achieved (Figure 1). Overlap between clones was identified by amplification

Table 1 Primer sequences used in the analysis of the transcripts mapped within chromosome 9p12-p13 BAC contig

EST name	Forward Primer (5', 3')	Reverse Primer (5', 3')	Size (bp)
stSG55665	TGT CCA GAG GCT GAC GTA G	TTA CAC AGA GAC CAG CAG GC	115
stSG26337	GCC GAA TGA TAC AGC CCT TA	AAA GAA GGA GCT GTC CTG AGG	138
SHGC-10717	GGA GCA GAA ATT GCC AAC AT	GGG TGG AAG GGG ATA GGT AA	116
SHGC-33250	TAC AAT TAG GAG CAA GTA AGA AGC G	TAC ACC AGG AAC TGA AAT ATC ACT G	133
WI-14287	ATC ATT TTT ATT GGA GCA TGA CC	GTT CCA GCC AGT CCC CAC	132
stSG1907	TTG TAC CCA TGT GTC TGT CAC	GGG TGA CCT AGG CAG CAT TA	131
stSG2215	GTG ATG AAA ATC TCA GGG AAG G	CTC TGA AGT GGG AGA CGG AG	183
stSG2362	AGG CAC AGA CAT TCC CAA AC	CTG GAA GCT GCA GTC TAC ACC	147
stSG62716	AGT CCC AGC TCC ATC CAT C	CAG CAT CTT TCC AGA GGA GG	159
A003M47	GAA CAG GAA TAA AGA AAT CTG	GAT TCA TCC AAG CTT AAA	124
T78312	CCT TAT GTC CAA ATC TTT GTG C	CCA TGT GGT GAA GTA GGC AG	192
stSG52289	GCA AAG ATG GAG ACG GTT GT	AAG GCA GAT CTG TGC TCT CTG	122
stSG55593	TCC AGT ATA TTG CAT TAA AGA CAT T	TCT GAA ATA CAT AGG GAT ATG CGT A	147
stSG39032	CTG CCC CTT GGA TGT GTC	TCC AGC AAG AAG GCA TTT	132
A002G23	GCA TAG CTT GAG ACT GGT GA	GCA ACT TTA CAA GGT CCA TGT GG	228
M20471	GAA GCT CTT CAC AGT CAT TGG	ACA GGT TGA GGC TAC TCT TGG	189
SHGC-32868	AGA AAT CAT GTG CAT ATA TCT TTC	TTC ACA CTA ACC CAC TAG GAA GG	139
stSG52662	AAC CTC CCC TTA AAA AGG CA	ACC TAA ACT GAG AGC AAG CTG G	159
SHGC-33912	AGA CTC CTA GGT GTG GAC CAT TGG	GAA TTA AGT TTG GGA TGT TTT CCC	182
SHGC-104	TGG TTA CCT CAT TTT GCC GTT TC	TCA AGC CTT CAA ATG ATT TGG TTA C	164
AA254410	CCT AAC AAG GTG ATT TGT GAA AG	CTG GAA GAA AGC TTT GAA GTG AGC	118
T83948	AAT ATC ACG GTT CTG TCT TT	AAA ATC CAA AAT GCT TCT GGT CCC	420
stSG22652	TGA AGA AGA GGG GTC ATG TAC A	CAG GAG GCC AAG AAA GGA G	162
KIAA0175	GGA TGA GTG TGG GTG TGA TAC	TGA CAG ATG GGC TTG ATT TAG	173
WI-7677	CAG AAC AGC CAG GTA GAG CC	ATG ATC CTC TCA CAG CCC C	329

and hybridisation of the same STS markers in multiple clones.

Four clones from the RPCI-11 library – BAC112J3; BAC327L3; BAC113A10; BAC421H8 – were mapped in this locus by the Sanger Center; we confirmed their location in this contig by hybridization with the relevant markers as illustrated in Figure 1. A total of 23 BAC clones were identified, (four from the RPCI-11 library, as mentioned above, and 19 by library screening: nine from Genome Systems – BAC107C13; BAC3L23; BAC91N8; BAC104I10; BAC27I1; BAC146F24; BAC199A5; BAC25J5; BAC108F12, and the remaining 10 from Research Genetics) and 40 novel sequence tagged sites markers were developed from the end sequences of these BAC clones (Figure 1; primer sequences are available upon request). These STS were used to construct a high resolution physical map for the region from β tropomyosin gene to the 3' end of PAX5 gene (WI-7677) (Figure 1), creating a tiling path of the candidate interval. The sizes of inserts were determined by PFGE after *NotI* digestion, and according to sequence data. The size of the contig was estimated based on the sizes of individual clones and potential overlap between clones.

Identification of novel polymorphic markers

In the process of constructing this detailed physical map we have identified several new polymorphic markers within the established contig. Eleven novel microsatellites and two single nucleotide polymorphisms were detected (Table 2). All repeat markers except 444CA, 327GT6 and 327CA3, were found to be polymorphic in the studied populations. Markers

444CA and 27CA were derived from the Sp6 end sequences of BAC444N18 and 27I1, respectively. 570CA was identified through sequencing a random plasmid subclone derived from BAC570K22, and 47CA31 and 47CA8 were identified in BAC47N15 by hybridisation with a Poly(dA-dC)·Poly(dG-dT) probe, as described in Materials and methods. All the remaining repeats (113GT2, 327GT6, 327GT4, 327CA1, 327GT5) were defined from the published sequences of BACs 113A10 and 327L3 (Accession numbers, AL138834 and AL135841).

The precise location of these markers was determined by STS content mapping in all BAC clones within the contig (Figure 1). Additionally, two single nucleotide polymorphisms were found in BAC257I6. 257Sp6 C/T polymorphism was found in the sequence derived from the Sp6 end of BAC 257I6 and the second polymorphism, a base substitution C to A, was identified while sequencing a subclone of BAC257I6. Both polymorphisms were found to create a restriction site for *MseI* (T'TAA).

Transcript map

In an attempt to obtain a more complete transcription map of the genomic contig generated, we searched several databases for ESTs and genes that potentially map to chromosome 9p12-p13, between the β -tropomyosin gene and marker *D9S50*. More than 50 transcripts were selected for this fine mapping study.

Twenty-five different transcripts were unambiguously positioned and ordered in the transcription map we constructed. Most of those were hybridised on MTN blots,

Table 2 New polymorphic markers identified within chromosome 9p12-p13 BAC contig

Locus name	Parent clone	Forward primer (5', 3')	Reverse primer (5', 3')	Size (bp)
Microsatellites				
327CA1	327L3	CAG ATG AAC TTA TTC CAG G	CTA CCT AGT GCT ATT ACT AG	118
327GT4	327L3	TAT GTC TTC ATT TGG GTG	CTC ATA ACC TAG TTC TCT GC	152
327GT5	327L3	TGA CTT CAC AGC TGC CTC CTC A	GTT ACA CAG TTC CAT GCT GAC	134
113GT2	113A10	GTT CTT GAG AGA AAC AGT AC	CTT GTC AGC CTC CAT AAT CAG	120
47CA8	47N15	CCT GAT GGT TAA TGA TGT CAA GC	GTG GCT TGT ATC TAT AGT CC	220
47CA31	47N15	CCT TAT AGA ACC TGT GTG AG	GGA TGT GAT CAG AAA ATT CC	215
27CA	2711	CAC TCC AGC TGG TGC AAC AG	AGC CAC TTG GTC CTT GAG GG	120
570CA	570K22	GGA CAC ACA GCC ACT CAG	GCA AAC TGA GGA GTA CTC	124
Single nucleotide polymorphisms				
257SP6 T/C	25716	GAT CTC AAA GGG TAG TAA CAT CTC	TGC ACG CTT AGA GTT ATG CAA AAG	270
257/73 A/C	25716	AAC ATA TGA AAG CAC AAT TAC ATA	ATA TTT TTT CAG TTG TGT TTA TAT	270

as described in Materials and methods. Negative results were checked twice. Additional expression data was compiled from various databases. Table 3 shows a summary of all the expression data accumulated for those ESTs both from database information and from our Northern blotting results (all ESTs were assayed by Northern blot except stSG55665, stSG26337, stSG2362, stSG62716, stSG52289, T83948 and WI-7677). Figure 1 indicates the precise location of these transcripts in the genomic contig.

Among those, 17 transcripts are ESTs expressed in various tissues, with low or no homology to known genes, while the remaining eight are ESTs derived from genes with known function.

β -tropomyosin (*TPM2*), an abundant skeletal muscle protein of 284aa,^{21,22} talin (*TLN*), a high molecular weight cytoskeletal protein which plays a key role in the assembly of actin filaments and in the spreading and migration of various cell types,^{23,24} human RNA (*CREB3*), a cAMP response element (CRE)-binding protein/activating transcription factor 1 protein of the basic leucine zipper superfamily,²⁵ *ST15*, which encodes a membrane anchored glycoprotein of about 110 kDa which plays a role in linking oncogenic signals to tumour invasion and metastasis,²⁶ UDP-N-acetylglucosamine-2-epimerase/N-acetylmannosamine kinase (*GNE*), a bifunctional enzyme, which regulates and initiates biosynthesis of N-acetylneuraminic acid (NeuAc),²⁷ calicin (*CCIN*), a polypeptide of 588 amino acids encoded by a 2.2 kb mRNA, expressed in testis as morphogenic cytoskeletal element in spermiogenic differentiation,²⁸ clathrin light polypeptide A (*CLTA*), a regulatory element in clathrin function which is a main structural component of the lattice cover in the cytoplasm face of the coated pits and coated vesicles,^{29,30} paired box homeotic gene 5 (*PAX5*), a 391 amino acids protein that plays an important role in B-cell differentiation as well as in neural development and spermatogenesis.³¹

Refinement of the HIBM candidate region

Previously we restricted the HIBM disease interval between *D9S1791* and *D9S50*, based on recombination events and

linkage disequilibrium.⁷ The distal boundary, *D9S1791*, was defined by a recombination event observed in unaffected individual F34 II-5.⁷ However, very recently, the clinical status of this individual has dramatically changed from unaffected to affected, and therefore his haplotype does not reflect a recombination event, but rather his affected status. Based on these results, *D9S1791* cannot be considered as the distal boundary of the HIBM interval. Moreover, haplotype analysis in other families with additional markers allowed us to relocate the HIBM candidate interval between *327GT4* on the telomeric side, and *D9S1859* on the centromeric side, based on recombination events observed in affected individuals solely (Figure 2): the recombinant chromosome found in affected individual F10 III-4 defined the distal boundary as *327GT4* and the recombination event observed in individual F5 III-4 placed the proximal boundary as marker *D9S1859*.

Based on the newly established BAC contig, we conclude that the HIBM critical interval, between markers *327GT4* and *D9S1859* covers a region of approximately 700 kb of DNA.

Current positional candidates

The BAC physical map allowed us to verify and refine the location of 25 genes and ESTs mapping to chromosome 9p12-p13 (Table 3, Figure 1). As a result we have been able to exclude a large number of genes and ESTs as potential HIBM candidate genes. *PAX5* and β -tropomyosin were already excluded in our previous study,⁷ stSG2215, *CREB3*, SHGC-33250, *TLN*, A003M47, T78312, stSG26337, stSG55665, stSG2362 and stSG62716 were excluded based on the recombination observed with *327GT4*. KIAA0175 transcript was excluded due to recombination event with *D9S1859*. The 12 remaining transcripts (Table 3) located between these two flanking markers can be considered as candidate disease causing genes in HIBM.

Among those, eight are ESTs with no similarity to any known gene, and four (*ST15*, *CCIN*, *CLTA* and *GNE*) are genes with a known function. *CCIN* is expressed only in testis. The other three genes are expressed in a variety of tissues, including skeletal muscle, the affected tissue in HIBM. Clathrin light chain A (*CLTA*) is a regulatory element

Table 3 Transcripts assigned to chromosome 9p12-p13 BAC contig

<i>EST name</i>	<i>GenBank Acc. No.</i>	<i>Gene/similarity</i>	<i>Known function</i>	<i>Expression data</i>
stSG55665	R45884	Similarity to KIAA0258 protein		Blood, brain, CNS, colon, kidney, lung, ovary, parathyroid, prostate, uterus
stSG26337	H65042	No similarity		Blood, lung
SHGC-10717	X06825	Tropomyosin beta chain (TPM2)	Cytoskeletal protein	Blood, bone, brain, colon, skeletal muscle, heart, lung, ovary, pancreas, prostate, kidney, spleen, uterus, stomach, thymus
SHGC-33250	G27778	Low similarity to protein phosphatase PP2A, regulatory subunit (PR130)		Lung
WI-14287	AF177198	Talin (TLN)	Cytoskeletal protein and putative tumor suppressor	Bone, brain, lung, foreskin, ovary, stomach, uterus, blood, colon, skeletal muscle, heart, placenta, liver, kidney, pancreas
stSG1907	AF009368	Luman (CREB3)	Transcription factor LZIP	Blood, bone, brain, breast, CNS, colon, heart, kidney, lung, lymph, skeletal muscle, ovary, pancreas, parathyroid, placenta, prostate, testis, uterus, liver
stSG2215	R37397	No similarity		Brain, heart, kidney, ovary, prostate, uterus, parathyroid, stomach
stSG2362	R38763	Similarity to NAG-5 protein		Brain, kidney, ovary, placenta
stSG62716	W42412	No similarity		Colon, heart, lung, ovary
A003M47	AF052116	No similarity		Brain
T78312	T78312	No similarity		Brain
stSG52289	AA634361	No similarity		Testis
stSG55593	D50406	RECK (ST15)	Membrane-anchored glycoprotein (metastasis and invasion)	Brain, kidney, bone, heart, ovary, pancreas, parathyroid, placenta, prostate, testis, uterus
stSG39032	Z46967	Calicin (CCIN)	Morphogenic cytoskeletal element in spermiogenic differentiation	Testis
A002G23	G26403	No similarity		Blood, breast, brain, colon, heart, kidney, lung, ovary, prostate, thymus, uterus, placenta, lung, liver, skeletal muscle, pancreas
M20471	M20471	Clathrin light chain A (CLTA)	Regulatory element in clathrin function	Brain, breast, colon, heart, bone, kidney, lung, lymph, skeletal muscle, ovary, pancreas, parathyroid, placenta, prostate, spleen, testis, stomach, liver, uterus
SHGC-32868	G27648	UDP-N-acetylglucosamine-2-epimerase (GNE)	Bi-functional enzyme in the biosynthesis of N-acetylneuraminic acid (NeuAc)	Breast, colon, liver, lung, parathyroid, placenta, prostate, uterus, brain, kidney, skeletal muscle, heart, pancreas
stSG52662	N71485	No similarity		Foetal liver/spleen
SHGC-33912	G29562	No similarity		Foetal liver/spleen
SHGC-104	D11961	No similarity		Brain, heart, kidney, lung, skeletal muscle, parathyroid, placenta, prostate, thymus, uterus, liver, pancreas
AA254410	AA254410	No similarity		Foetal liver, spleen
T83948	T83948	No similarity		Mouse spleen
stSG22652	R95130	No similarity		Foetal liver/spleen
KIAA0175	D79997	Similarity to <i>X.laevis</i> protein kinase, a putative transmembrane domain		Bone, colon, heart, skeletal muscle, stomach, uterus, tonsil
WI-7677	M96944	PAX5	B-cell specific transcription factor	Blood, lung, testis, tonsil

in clathrin gene function known to be involved in several pathways of lysosomal proteolysis, in various cell types. Clathrin has also been implicated in the pathology of Alzheimer's disease and Pick's disease. Immunohistochemical stainings of muscle fibre from Japanese HIBM patients did not lead to definitive conclusions about differential expression of this protein between those and control muscle tissue.³² According to these data, the involvement

of CLTA in HIBM cannot be ruled out at this point. The involvement of the three known additional genes located in the interval in the pathogenesis of HIBM seems less likely: mutations in the bifunctional enzyme GNE have been related to sialuria, a rare inborn error of metabolism characterised by cytoplasmic accumulation and increased urinary excretion of free neuraminic acid; ST15 has been recently recognized as a matrix protein regulated tumour

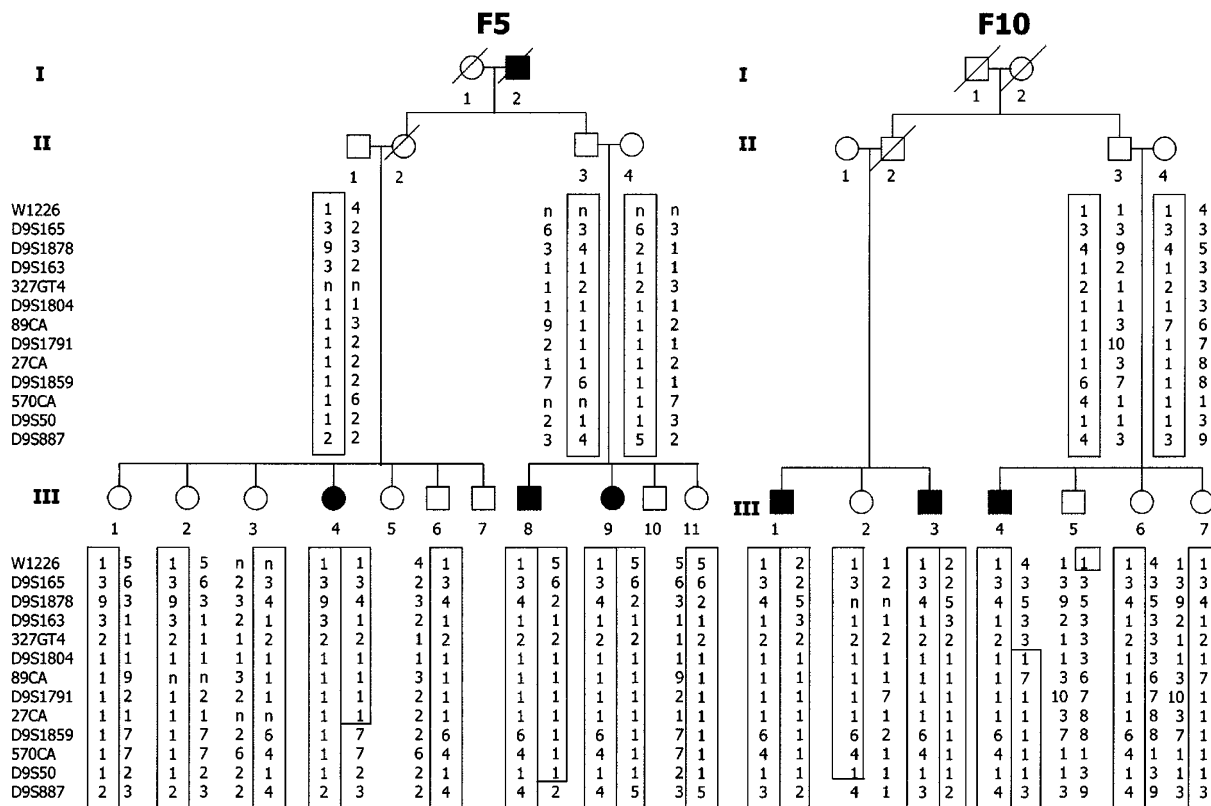


Figure 2 Haplotype analysis in two HIBM families showing recombination events in the vicinity of the HIBM locus. Haplotypes comprise alleles at the listed loci ordered telomeric to centromeric *W1226*, *D9S165*, *D9S1878*, *D9S163*, *327GT4*, *D9S1804*, *89CA*, *D9S1791*, *27CA*, *D9S1859*, *570CA*, *D9S50* and *D9S887*, (n=not tested). Individuals are designated with generation and pedigree numbers. Disease status is indicated: affected, filled symbols; unaffected, clear symbols. Deceased individuals are denoted by slashes.

suppressor gene which can mediate a variety of signals to tumour invasion and metastasis. CCIN, a morphogenic cytoskeletal protein, is expressed solely in testis and related to differentiation.

Although none of those four genes has been directly involved in any known muscle function, their potential relevance or irrelevance to HIBM has still to be established.

Discussion

The positional cloning approach still represents a very potent tool for the identification of a disease gene of unknown biological or biochemical function. The first step towards the identification of genes located in a specific chromosomal fragment is the construction of a high resolution genomic map within this region.

Our previous studies localised the HIBM gene into a region of chromosome 9p12-p13 flanked by markers *D9S1791* and *D9S50* in a 1 Mb interval laying on a single YAC.^{7,33} BAC end walking was used for the construction of a contig covering the relevant region. Analysis of single BAC libraries revealed

the presence of gaps at different locations in the region. Only by screening three different BAC genomic libraries, we were able to establish a complete contig spanning the genes β tropomyosin and the 3' end of *PAX5*, which includes the HIBM candidate region. This continuous contig was used as the basis for the generation of a highly detailed physical and transcriptional map of the region. We have precisely ordered 40 different loci, and identified 13 new polymorphic markers which allowed the refinement of the haplotype analysis in HIBM families. Although four of the clusters could not be definitely ordered internally by BAC hybridisation, we have established their location from combined hybridisation results and data of all available databases relating to them (GeneMap99, The Sanger Center, GDB). Our previous genotyping studies with eight polymorphic markers located the HIBM disease in the interval flanked by *D9S1791* and *D9S50*, based on both recombination and linkage disequilibrium data. The disease status in each subject was determined by personal clinical examination of one of us (Z Argov and M Sadeh). Unfortunately, there are no known pre-symptomatic tests in HIBM to predict the status of a 'currently healthy'

individual (like the markedly elevated serum creatine kinase before onset of muscle weakness in various muscular dystrophies). Since all of the patients in the HIBM families analysed were diagnosed before age 35 years, unaffected family members below the age of 35 years were excluded from analysis in our studies. The recombination event which located the distal boundary at *D9S1791* was recognised in an unaffected individual which was examined at age 43 years and showed no sign of muscle disease then. Since in this family the age of onset was 33 years, that is 10 years before, the data observed pointed to a recombination event. However, a recent reexamination of the patient revealed the clinical criteria for HIBM diagnosis, so that we are now aware that some of the presently healthy members of the families studied may become symptomatic at a later time. Therefore only recombination events observed in affected individuals are taken into consideration. At present, we cannot consider *D9S1791* as the distal boundary defining the HIBM interval; the closest unambiguous recombination event has been recognised at marker *327GT4* in an affected individual. Twenty-five genes and ESTs derived from all screened databases have been analysed for their precise location in the 9p12-p13 region, and for their expression pattern in a variety of tissues. Our results show that most of those transcripts are present in more than one tissue type, some are more specific to certain types, such as foetal liver or spleen, brain or testis. Furthermore no transcript was found to be expressed solely in skeletal muscle tissue. Among the 25 transcripts identified in the region, two cytoskeletal proteins, β -tropomyosin and talin, are the only genes with a recognised function in muscle, and therefore they were the most attractive genes to be potentially involved in HIBM. Both genes were subsequently excluded by position, by linkage disequilibrium data⁷ and, for talin, because a recombination event has been recognised with a polymorphic site within the gene²⁴ in one affected individual. However, the function of many of the genes and ESTs mapped to 9p12-p13 has not yet been established, and the continuous flow of sequence and transcriptional data demands a constant update of our analysis, till the systematic sequencing of the entire interval is completed.

Our goal is the identification of all the genes in human chromosome 9p12-p13, including the HIBM interval laying at present between *327TG4* and *D9S1859*. The high resolution physical and transcriptional map established in these studies will contribute to the positional cloning of various disease causing genes located in this region, among them the gene responsible for HIBM. Certainly the identification of the gene will have immediate implications such as efficient genetic counseling for HIBM families, and precise knowledge of all HIBM allelic disorders. Moreover, further characterisation of the HIBM causing gene will provide important insights into the mechanisms of inclusion body myopathies and increase our understanding of the molecular basis for impaired function of the muscle tissue.

Acknowledgments

We would like to thank Mira Korner and all the staff from The Laboratory of DNA Analysis at The Institute of Life Sciences, The Hebrew University of Jerusalem, for their skilful assistance in sequencing, and Anka Sanilevitch for her invaluable help in oligonucleotide synthesis. This study was supported by Hadasit (Medical Research Services Development Co. Ltd., a subsidiary for R&D of Hadassah Medical Organization), by Grant 96-00268 from the United States-Israel Binational Science Foundation (BSF), by a grant from the Association Francaise contre les Myopathies (AFM), and by a special donation from Hadassah Southern California, the Persian Group Council, Vanguard II, Healing Spirit and Haifa Group, and from the ARM organisation.

References

- Argov Z, Yarom R. "Rimmed vacuole myopathy" sparing the quadriceps: a unique disorder in Iranian Jews. *J Neurol Sci* 1984; **64**: 33–43.
- Sadeh M, Gadoth M, Hadar H, Ben David E. Vacuolar myopathy sparing the quadriceps. *Brain* 1993; **16**: 217–232.
- Askanas V, Engel WK. New advances in the understanding of sporadic inclusion-body myositis and hereditary inclusion-body myopathies. *Curr Opin Rheumatology* 1995; **7**: 486–496.
- Mitrani-Rosenbaum S, Argov Z, Blumenfeld A, Seidman CE, Seidman JG. Hereditary Inclusion Body Myopathy maps to chromosome 9p1-q1. *Hum Mol Genet* 1996; **5**: 159–163.
- Argov Z, Tiram E, Eisenberg I *et al*. Various types of hereditary inclusion body myopathies map to chromosome 9p1-q1. *Ann Neurol* 1997; **41**: 548–551.
- Argov Z, Sadeh M, Eisenberg I, Karpati G, Mitrani-Rosenbaum S. Facial weakness in hereditary inclusion body myopathies. *Neurology* 1998; **50**: 1925–1926.
- Eisenberg I, Thiel C, Levi T *et al*. Fine Structure Mapping of the Hereditary Inclusion Body Myopathy Locus. *Genomics* 1999; **55**: 43–48.
- Christodoulou K, Papadopoulou E, Tsingis M *et al*. Narrowing of the gene locus for autosomal-recessive quadriceps sparing inclusion body myopathy (ARQS-IBM) to chromosome 9p1. *Acta Myol* 1998; **2**: 7–9.
- Mirabella R, Christodoulou K, Di Giovanni S, Ricci E, Tonali P, Servidei S. An Italian family with autosomal recessive quadriceps-sparing inclusion-body myopathy (ARQS-IBM) linked to chromosome 9p1. *Neurol Sci* 2000; **21**: 99–102.
- Ikeuchi T, Asaka T, Saito M *et al*. Gene locus for autosomal recessive distal myopathy with rimmed vacuoles maps to chromosome 9. *Ann Neurol* 1997; **41**: 432–437.
- Sessarego M, Defferrari R, Fugazza G, Comelli A, Salvadio E, Ajmar F. Involvement of the short arm of the derivative chromosome 9 in Philadelphia-positive acute lymphoblastic leukemia. *Cancer Genet Cytogenet* 1991; **52**: 43–49.
- Gaidano G, Dalla-Favera R. Biologic and molecular characterization of non-Hodgkin's lymphoma. *Curr Opin Oncology* 1993; **5**: 776–784.
- Kim SK, Ro JY, Kemp BL *et al*. Identification of three distinct tumor suppressor loci on the short arm of chromosome 9 in small cell lung cancer. *Cancer Res* 1997; **57**: 400–403.
- Liew CT, Li HM, Lo KW *et al*. Frequent allelic loss on chromosome 9 in hepatocellular carcinoma. *Int J Cancer* 1999; **81**: 319–324.
- Bamshad M, Watkins WS, Zenger RK *et al*. A gene for distal arthrogryposis type I maps to the pericentromeric region of chromosome 9. *Am J Hum Genet* 1994; **55**: 1153–1158.
- Kant S, Polinkovsky A, Mundlos S *et al*. Acromesomelic dysplasia Maroteaux type maps to human chromosome 9. *Am J Hum Genet* 1998; **63**: 155–162.

- 17 Sulisalo T, Sistonen P, Hastbacka J *et al*. Cartilage-hair hypoplasia gene assigned to chromosome 9 by linkage analysis. *Nature Genetics* 1993; **3**: 338–341.
- 18 Middleton LT, Christodoulou K, Mubaidin A *et al*. Distal hereditary motor neuropathy of the Jerash type. *Ann NY Acad Sci* 1999; **883**: 439–442.
- 19 Deloukas P, Schuler GD, Gyapay G *et al*. A physical map of 30,000 human genes. *Science* 1998; **282**: 744–746.
- 20 Altschul SF, Madden TL, Schaffer AA *et al*. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997; **25**: 3389–3402.
- 21 Widada JS, Ferraz C, Capony JP, Liautard JP. Complete nucleotide-sequence of the adult skeletal isoform of skeletal muscle beta-tropomyosin. *Nucleic Acids Res* 1988; **16**: 3109.
- 22 Tiso N, Rampoldi L, Pallavicini A *et al*. Fine mapping of five human skeletal muscle genes: alpha-tropomyosin, beta-tropomyosin, troponin-I slow twitch, troponin-I fast twitch, and troponin-C fast. *Biochem Biophys Res Commun* 1997; **230**: 347–350.
- 23 Rees DJ, Ades SE, Singer SJ, Hynes RO. Sequence and domain structure of talin. *Nature* 1990; **347**: 685–689.
- 24 Ben-Yosef T, Francomano CA. Characterization of the human talin (TLN) gene: genomic structure, chromosomal localization, and expression pattern. *Genomics* 1999; **62**: 316–319.
- 25 Lu R, Yang P, O'Hare P, Misra V, Luman. A new member of the CREB/ATF family, binds to herpes simplex virus VP16-associated host cellular factor. *Mol Cell Biol* 1997; **17**: 5117–5126.
- 26 Takahashi C, Sheng Z, Horan, TP *et al*. Inhibition of tumor invasion by the membrane-anchored glycoprotein RECK. *Proc Natl Acad Sci USA* 1998; **95**: 13221–13226.
- 27 Hinderlich S, Stasche R, Zeitler R, Reutter W. A bifunctional enzyme catalyzes the first two steps in N-acetylneuraminic acid biosynthesis of rat liver. Purification and characterization of UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase. *J Biol Chem* 1997; **272**: 24313–24318.
- 28 von Bulow M, Heid H, Hess H, Franke WW. Molecular nature of calicin, a major basic protein of the mammalian sperm head cytoskeleton. *Exp Cell Res* 1995; **219**: 407–413.
- 29 Jackson AP, Parham P. Structure of human clathrin light chains. Conservation of light chain polymorphisms in three mammalian species. *J Biol Chem* 1988; **263**: 16688–16695.
- 30 Ponnambalam S, Jackson AP, LeBeau MM *et al*. Chromosomal location and some structural features of human clathrin light-chain genes (CLTA and CLTB). *Genomics* 1994; **24**: 440–444.
- 31 Adams B, Dorfler P, Aguzzi A *et al*. Pax-5 encodes the transcription factor BSAP and is expressed in B lymphocytes, the developing CNS, and adult testis. *Genes Dev* 1992; **6**: 1589–1607.
- 32 Kumamoto T, Abe T, Nagao S, Ueyama H, Tsuda T. Immunohistochemical study of clathrin in distal myopathy with rimmed vacuoles. *Acta Neuropathol* 1998; **95**: 571–575.
- 33 Argov Z, Eisenberg I, Mitrani-Rosenbaum S. Genetics of inclusion body myopathies. *Curr Opin Rheumatology* 1998; **10**: 543–547.
- 34 Vakkilainen T, Kivipensas P, Kaitila I, de la Chapelle A, Ridanpaa M. Integrated high-resolution BAC, P1, and transcript map of the CHH region in chromosome 9p13. *Genomics* 1999; **59**: 319–325.