Homozygosity mapping of disease loci by whole-genome SNP analysis from sporadic consanguineous patients

Christine Kretz
Department of Neurobiology and Genetics, IGBMC, Illkirch, France
Anne Toussaint
Department of Neurobiology and Genetics, IGBMC, Illkirch, France
Jocelyn Laporte
Department of Neurobiology and Genetics, IGBMC, Illkirch, France

Method Article

Keywords: autozygosity, homozygosity mapping, single nucleotide polymorphism, SNP, recessive, consanguinity, amphiphysin, BIN1, disease-causing mutation, centronuclear myopathy

Posted Date: August 8th, 2007

DOI: https://doi.org/10.1038/nprot.2007.343

Abstract

Introduction

Positional cloning is a powerful approach to identify genes mutated in human and animal models monogenic diseases, and is the option of choice when there is no functional candidate genes that appear as good candidates based on their known biochemical functions. The genomic DNA region \(and the embedded polymorphisms) harbouring the disease-causing mutations segregates with the disease in analyzed pedigrees. However, positional cloning relayed until recently on the availability of large pedigrees to reach a significant linkage. This protocol describes the use of whole genome genotyping on sporadic consanguineous patients to identify potential disease loci and subsequent positional candidate genes, by homozygosity mapping \(Autozygosity). It takes advantage of high density single nucleotide polymorphism \(SNP) genotyping arrays, and of the assumption that unrelated patients from several consanguineous families are mutated in the same gene. It was applied to the mapping of a novel gene involved in autosomal recessive centronuclear myopathy, amphiphysin 2/BIN1 \(ref. 1).

Reagents

· Glycogen at 10 mg/ml \(ICN Biomedicals) · Sodium chloride 1M · Ethanol 100% · NE2 digestion buffer \(New England Biolabs) · Bovine Serum Albumine \(BSA) 100X \(New England Biolabs) · XbaI restriction enzyme at 20,000 U/ml \(New England Biolabs) · Sterile filtered water · T4 DNA Ligase \(400 U/µl) and ligase buffer \(New England Biolabs) · Gene chip Mapping 10K Xba Assay Kit N°900441 containing : the adaptators and Xba primers, fragmentation buffer and reagent, labeling reagent, and Terminal deoxynucleotidyl Transferase \(TdT) enzyme and reaction buffer \(Affymetrix) · Ampli Taq Gold DNA polymerase and reaction buffer \(Applied Biosystem) · dNTPs at 10 mM : 2.5 mM of each nucleotide \(Roche) · Agarose electrophoresis grade \(Invitrogen) · TAE Buffer 20X : 40 mM Tris Base + 39 mM glacial Acetic acid + 1 mM EDTA · Elution buffer EB \(Qiagen) · GeneChip Human mapping 10K array Xba 142 2.0 \(Affymetrix)

Equipment

· Centrifuge \(Eppendorf 5804R) · Biosafety laminar flow hood · Gene Amp 9700 thermocycler \(Applied Biosystems) · Vacuum Pump XX5522050 \(Millipore) · 96 well filtration plate Mini Elute 96UF PCR Purification Kit \(Qiagen) · Thermowell Sealing Tape Aluminium 6569 \(Costar) · MS1 minishaker for 96 well plates \(IKA) · 0.2 ml sterile tubes · Material for electrophoresis · Beckman DU640B Spectrophotometer · Gene Chip DNA Analysis Software \(GDAS) Version 3.0 \(Affymetrix)

Procedure

DNA purification to obtain purified genomic DNA 1. Take 2-4 µg genomic DNA in a total volume of 20µl, add 1 µl glycogen at 10 mg/ml, 5 µl NaCl 1 M and 40 µl ethanol 100%, and precipitate 30 min \(or

up to overnight) at -20°C. Centrifuge 20 min 13.000 rpm at 4°C. Dry the pellet and resuspend at 250 ng/µl in TE Buffer \(Tris pH 8.0 10 mM, EDTA 0.1 mM). Measure DNA concentration with a spectrophotometer. **_XbaI_ digestion to obtain the digested DNA** 2. Under a sterile laminar flow hood, prepare a master mix on ice with 2 µl NE2 buffer \(10x), 2 µl BSA \(10x), 10.5 µl sterile water and 0.5 µl XbaI restriction enzyme \(20 U/µl). Add 5 µl of the purified genomic DNA \(at 250 ng/µl). Incubate 2 hrs at 37°C, 20 min at 70°C using a thermocycler and keep at 4°C. **Adaptors ligation** 3. On ice, prepare a mix containing 1.25 µl Xba adaptators, 2.5 µl ligase buffer \(10X), and 1.25 µl T4 DNA Ligase. Add the 20 µl Xba digested DNA and incubate 2 hrs at 16°C, 20 min at 70°C using a thermocycler and keep at 4°C. Complete to 100 µl final with sterile water. **PCR reaction to obtain the amplified DNA fragments** 4. On ice, prepare a mix containing 10 µl PCR buffer \(10X), 10 µl dNTP \(2.5 mM of each), 10 µl MgCl2 \(25 mM), 7.5 µl Xba primers \(10 µM), 2 µl Ampli Taq Gold enzyme and 50.5 µl of sterile water. Add 10 µl of the ligated DNA. Launch 4 PCR reactions per DNA sample. 5. PCR program on a GeneAmp 9700 thermocycler : 3 min at 95°C, 35 amplification cycles \(30 sec at 95°C, 30 sec at 59°C, 30 sec at 72°C), 7 min at 72°C, then keep the amplified DNA at 4°C. 6. Check the complexity of the amplified DNA fragments by loading 3 µl of each PCR on a 2% agarose/TAE 1X gel : there should be a DNA ladder centered around 1-2 Kb. 7. Pool the 4 PCR reactions per DNA samples and purify on a filtration plate applying ~800 mbar \(vacuum Pump). Wash three times with 50 µl sterile water, then resuspend the DNA with 45 µl EB elution buffer by shaking moderately for 5 min using a minishaker for 96 well plates. **DNA fragmentation** 8. Measure the DNA concentration of each eluted DNA samples. 40 µg of the amplified DNA fragments is needed for the fragmentation. Add 5 µl of the fragmentation buffer \(10X). Prepare on ice a mix containing 2 µl fragmentation reagent \(3 U), 15 µl fragmentation buffer \(10X) and 133 µl sterile water. Add 5 µl from this mix to the 50 µl of amplified DNA fragments. Incubate 35 min at 37°C, 15 min at 95°C using a thermocycler and keep at 4°C. 9. Load 4 µl of this fragmented DNA on a 4% agarose/TAE 1X gel. Expect a smear of DNA fragments centered around 50 bp. **DNA fragments labeling** 10. On ice, prepare the labeling mix : 14 µl TdT reaction buffer, 2 µl labeling reagent \(7.5 mM) and 3.5 µl Terminal deoxynucleotidyl Transferase \(30 U/µl). Add 50.5 µl of fragmented DNA. Incubate 2 hrs at 37°C, 15 min at 95°C using a thermocycler and keep at 4°C. **Arrays hybridization and treatment** 11. Arrays hybridization, washing, and scanning are performed on a dedicated Affymetrix microarray platform following standardized procedures. **Data analysis** 12. To determine homozygous regions, results from the Affymetrix platform are analyzed with the GDAS 3.0 software. SNP call rate should be greater than 95% and signal detection greater than 99%. Display the loss of heterozygosity \(LOH) parameter, which ranks genomic regions based on the number of contiguous homozygous SNPs, their distance and heterozygosity rate. Higher the LOH is, stronger is the chance that the region is truly homozygous by descent and derived from the same common ancestor. For first degree consanguinity, we consider that the minimal LOH required is 4; on a 10K array, we also use regions with more than 30 SNPs in a row. Count a no call SNP as homozygous. 13. Disease-linked loci are determined by comparing the selected homozygous regions of several consanguineous patients. If patients are from the same family, analysis of unaffected siblings can rule out common homozygous regions by chance. Starting from sporadic cases from different families, we make the assumption that they are mutated in the same gene. The probability that a number of consecutive SNPs would be concordant and homozygous by chance

depends on the number of patients and families and is given in ref 2. Positional candidate regions are selected if homozygous regions are common to at least three unrelated consanguineous patients. 14. Positional candidate regions are defined by the nucleotide position of the first heterozygous SNPs that frame the smallest common homozygous region. Copy-paste these coordinates in a human genome web browser \(for example "http://genome.ucsc.edu/cgi-bin/hgGateway":http://genome.ucsc.edu/cgi-bin/hgGateway) to display the list of positional candidate genes that remain to be analysed in search of sequence variants.

Timing

- First day \(9 hrs) : preparation of genomic DNA, 1 hr; DNA digestion, 2.5 hrs; ligation, 2.5 hrs; and PCR, 2.5 hrs - Second day \(7 hrs) : electrophoresis, 1 hr; DNA purification, 2 hrs; DNA fragmentation, 0.75 hr; electrophoresis, 1 hr; DNA labeling, 2.25 hrs - Third day \(16 hrs) : DNA microarray hybridisation, washing and scanning on a dedicated Affymetrix platform - Forth day \(2 hrs) : data analysis

Critical Steps

2 to 7. Minimize the risk of cross contamination. Mix can be prepared under a sterile laminar flow hood, while DNA is added on another bench. Separate pre-PCR reaction and post-PCR handling of the products in two different rooms and use dedicated pipettes for each step. 2 and 4. Perform negative controls without DNA for both the Xba digestion and PCR amplification. 6 and 9. Verify on agarose gel that the DNA amplification and fragmentation have performed as expected. If there is no DNA ladder at step 6, or if DNA fragments of the expected sizes are not obtained, the complexity of the DNA probes will not be sufficient for whole genome analysis. 8. Be sure to have enough DNA for fragmentation: 40 µg are needed. 12. Patients from consanguineous families should be used for homozygosity mapping. 13. When comparing unrelated consanguineous patients, phenotypic assessment is crucial to choose the patients that will be compared. However, keep in mind that genetic heterogeneity exists in cases with similar phenotypes.

Troubleshooting

In case of low quality of the starting genomic DNA, or if there is a contamination between 2 different DNAs, the percentage of SNP call might be reduced considerably. Do not analyse results below 95% SNP call.

Anticipated Results

We identified mutations in the amphiphysin 2 \(BIN1) gene in patients with autosomal recessive centronuclear myopathy through two parallel approaches : functional candidate and positional candidate genes \(ref. 1). For the second approach, we followed the described procedure to define a common homozygous region on chromosome 2 in three sporadic consanguineous patients from unrelated

families : at 2q14, patient AAT68 had 66 homozygous SNPs for a LOH of 12.3; patient ADR71 94 SNPs, LOH=17.4; patient ABJ25 60 SNPs, LOH=11.3. The common homozygous region encompassed 66 SNPs for a length of 21 Mb. Direct sequencing of the BIN1 gene uncovered mutations in patients AAT68 and ADR71. Note that we have not found a disease-causing sequence variant in patient ABJ25 in all coding exons of the gene and that, through the functional candidate approach, we found a homozygous mutation in a forth unrelated patient that was not obviously linked to this region \(9 homozygous SNPs around the BIN1 gene, LOH=1.2). All analyzed patients were from first degree consanguinity and displayed several large homozygous regions.

References

1. Nicot, A.S., Toussaint, A., Tosch, V., Kretz, C., Wallgren-Pettersson, C., Iwarsson, E., Kingston, H., Garnier, J.M., Biancalana, V., Oldfors, A., Mandel, J.L., Laporte, J. Mutations in amphiphysin 2 \(BIN1) disrupt its membrane tubulation properties and its interaction with dynamin 2, and cause autosomal recessive centronuclear myopathy. _Nature Genetics_ **39**, 1134-1139 \(2007) 2. Woods, C.G., Valente, E.M., Bond, J., Roberts E. A new method for autozygosity mapping using single nucleotide polymorphisms \(SNPs) and EXCLUDEAR. _J. Med. Genet_. **41**, 101-104 \(2004).

Acknowledgements

We thank Christelle Thibault, who is running the Affymetrix common platform at IGBMC, for technical assistance and advices.

Figures

SNPs	Chromosomal regions	AAT68	ADR71	LF41	ABJ25
rs725139	2 q13		AB		
rs724496	2 q13	AB	AA		
rs1374161	2 q13	BB	…		
…	… …	…	…		
rs951431	2 q14.1	AA	AA	AA	AB
rs959031	2 q14.1	BB	BB	AB	BB
rs332704	2 q14.2	BB	BB	AA	BB
rs1534100	2 q14.2	BB	AA	AB	BB
rs874816	2 q14.2	AA	BB	AB	AA
rs959111	2 q14.2	BB	BB	AA	BB
rs2311398	2 q14.2	AA	AA	AA	AA
rs724691	2 q14.3	BB	AA	BB	AA
rs724692	2 q14.3	AA	AA	AA	BB
rs1368107	2 q14.3	AA	AA	AA	BB
rs723839	2 q14.3	BB	BB	BB	AA
rs1371494	2 q14.3	BB	BB	AB	BB
rs1607327	2 q14.3	BB	AA	AB	BB
rs294665	2 q14.3	AA	AA	AA	AA
rs167164	2 q14.3	AA	AA	AA	AA
rs951528	2 q14.3	BB	BB	BB	BB
rs2203772	2 q14.3	AA	AA	AA	AA
rs2419559	2 q14.3	AA	BB	AA	BB
rs1980356	2 q14.3	AA	AA	AB	AA
rs1405090	2 q14.3	AA	AA	AB	AA
rs1589908	2 q14.3	BB	BB	AB	BB
rs1343822	2 q14.3	BB	BB	AA	BB
rs2090030	2 q14.3	AA	BB	BB	AA
rs779991	2 q14.3	AA	AA	AA	BB
rs1186396	2 q14.3	AA	AA	AA	BB
rs1170578	2 q14.3	BB	BB	BB	AA
rs1367244	2 q14.3	AA	AA	BB	AA
rs2421084	2 q14.3	AA	BB	AA	AA
rs717601	2 q14.3	BB	BB	AB	BB
rs1583471	2 q14.3	BB	BB	BB	BB
rs763829	2 q14.3	BB	BB	BB	BB
rs1820556	2 q14.3	AA	BB	AA	BB
rs2404175	2 q14.3	AA	AA	AA	AA
rs1370229	2 q14.3	BB	AA	BB	BB
rs1316775	2 q14.3	BB	BB	BB	AA
rs1404073	2 q21.1	AA	AA	AA	AA
rs1097703	2 q21.1	BB	BB	BB	BB
rs1349734	2 q21.1	AA	AA	BB	AA
rs1946798	2 q21.2	AA	AA	AB	AA
rs2320399	2 q21.2	BB	BB	AB	BB
rs2321201	2 q21.2	AA	AA	AB	AA
rs3886664	2 q21.2	BB	BB	BB	BB
rs1867898	2 q21.2	BB	BB	AA	BB
rs1564935	2 q21.2	AB	AA	BB	AA
…	… …		…		…
rs1357157	2 q22.1		…		BB
rs1395010	2 q22.1		…		AB
…	… …		…		
rs1384662	2 q22.1		AA		
rs2217964	2 q22.1		AB		

← *BIN1*

Figure 1

SNP haplotype analysis of 4 unrelated consanguineous centronuclear myopathy patients for the 2q14 region containing the BIN1 gene. Patients AAT68, ADR71 and LF41 have mutations in BIN1, while we did not find nucleotide variant in ABJ25 by sequencing all exons and intron-exon boundaries. A and B denote the 2 alleles of each SNP: AA and BB, homozygous alleles, AB, heterozygous. Homozygous regions were

considered significant from 25 consecutive homozygous SNPs on a 10K array. Data first published in ref. 1 and reproduced with permission from Nature Publishing Group.

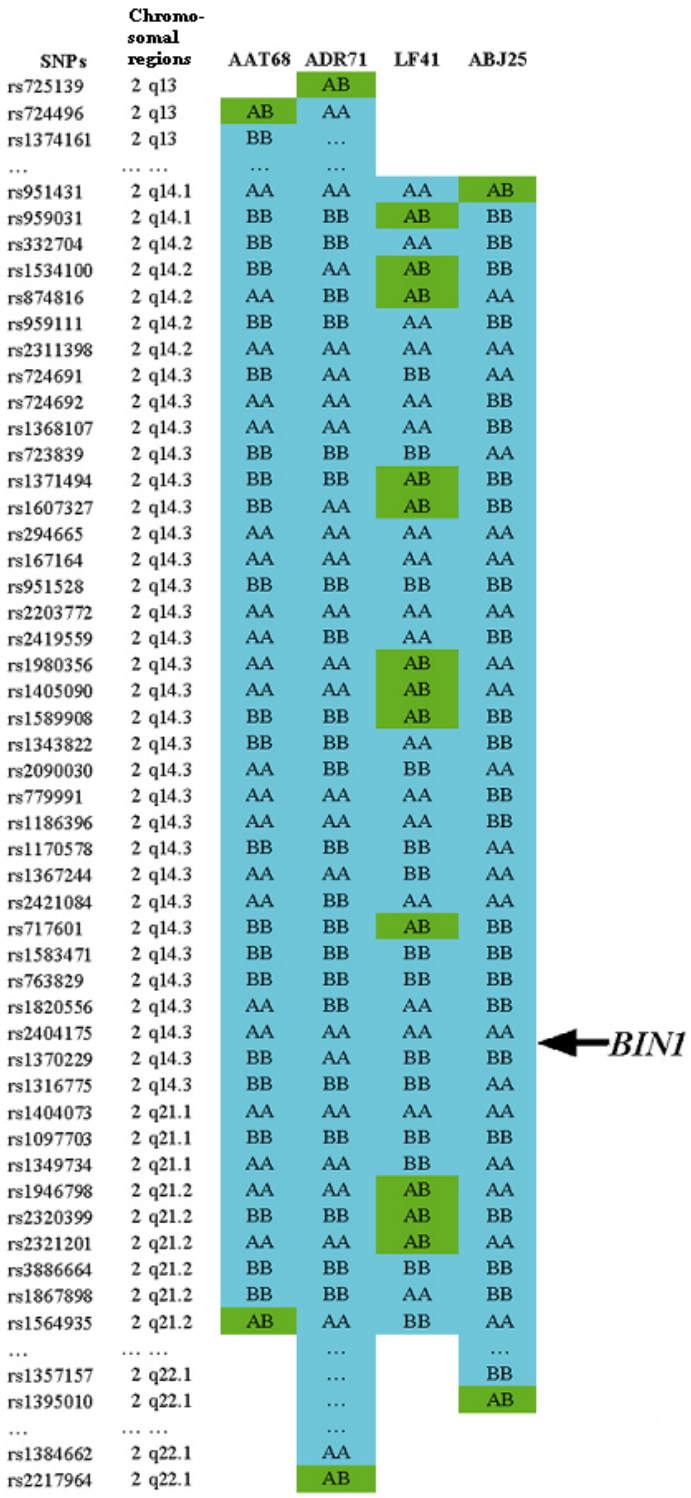

Figure 2