*Manuscript
Click here to view linked References

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

# Genome-wide analysis of single nucleotide polymorphisms and copy number variants in fibromyalgia suggest a role for the central nervous system

Elisa Docampo[1,2,3,4], Georgia Escaramís[1,2,3,4], Mònica Gratacòs[1,2,3,4], Sergi Villatoro[1,2,3,4], Anna Puig[1,2,3,4], Manolis Kogevinas[4,5,6,7], Antonio Collado[8,13], Jordi Carbonell[9,13], Javier Rivera[10,13], Javier Vidal[11,13], Jose Alegre[12,13], Xavier Estivill[1,2,3,4*✉] and Raquel Rabionet[1,2,3,4*✉]

[1]Genomics and Disease Group, Centre for Genomic Regulation (CRG), Barcelona, Catalonia, Spain

[2]Universitat Pompeu Fabra (UPF), Barcelona, Catalonia, Spain

[3]Hospital del Mar Medical Research Institute (IMIM), 08003, Barcelona, Catalonia, Spain

[4]Centro de Investigación Biomédica en Red en Epidemiología y Salud Pública (CIBERESP), Barcelona, Catalonia, Spain

[5]Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Catalonia, Spain

[6]IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain

[7]National School of Public Health, Athens, Greece

[8]Fibromyalgia Unit, Rheumatology Service, Hospital Clínic, Barcelona, Catalonia, Spain

[9]Fibromyalgia Unit, Rheumatology Service, Parc de Salut Mar, and Hospital del Mar Research Institute, Barcelona, Catalonia, Spain

[10]Rheumatology Unit, Instituto Provincial de Rehabilitación, Hospital Universitario Gregorio Marañón, Madrid, Spain

[11]Rheumatology Unit, Hospital General de Guadalajara, Guadalajara, Spain

[12]Chronic Fatigue Syndrome Unit, Hospital Vall d'Hebron, Barcelona, Catalonia, Spain

[13]FSGCDB Group, Fibromyalgia and Chronic Fatigue Syndrome, Spanish Genetic and Clinical Data Bank, Foundation FF

[*]RR and XE have equally contributed to this work
✉To whom correspondence should be addressed:
CRG, Dr Aiguader 88, 08003 Barcelona, Spain
Tel: +3493 316 0233 or +3493 316 0233
Fax: +3493 316 0099
Email: kelly.rabionet@crg.eu or xavier.estivill@crg.eu

**Number of pages: 23; Number of Figures: 2; Number of Tables: 4**

**Keywords:** Fibromyalgia, Genome-wide Association Scan, GWAS, Single Nucleotide Polymorphisms, Copy Number Variants, CNV

## INTRODUCTION

Fibromyalgia (FM) is a highly disabling syndrome affecting 2.9% of the European population [3], mainly women in the fourth decade of life, with a female:male ratio of 21:1 [33]. Fibromyalgia is defined by a low pain threshold and a permanent state of pain, accompanied by a constellation of symptoms, such as fatigue, sleep disturbances and cognitive impairment, among others. In the absence of suitable diagnostic tests, FM diagnosis is established by the presence of symptoms for at least three months and the exclusion of somatic diseases [43; 44].

The mechanisms explaining this chronic pain remain unclear. The most established hypothesis underlying FM etiopathogenesis is the existence of a dysfunction in pain processing. FM patients have been shown to present structural differences in the brain [13; 19]. Furthermore, there are several evidences of central sensitization at various levels in the nervous system [11], as well as neurochemical imbalances in the central nervous system leading to a "central amplification" of pain perception [14; 45].

The response to painful stimuli and the FM phenotype have both a genetic component. FM shows family aggregation [4; 5] and higher concordance in monozygotic than dizygotic twins (0.29 vs. 0.16) [18], while the response to painful stimuli has an estimated heritability of 22% to 55% [28]. However, the exploration of the genetic contribution to pain response and chronic pain states is, so far, scarce [27].

Genetic studies performed so far in FM have not been able to establish a clear genetic association. Most of them have been candidate gene studies, focused on genes related to HLA and neurotransmitters [17; 21; 47]. So far, two studies have attempted to explore the genetic contribution to FM in a genome-wide manner. One of them analyzed over 3200 SNPs in 350 genes implicated in pain transmission, inflammatory responses, and in influencing mood and affective states associated with chronic pain conditions, in 496 FM cases and 348 controls. However, the strongest associations did not replicate in independent cohorts [35]. The other one was a linkage scan evaluating 341 markers in 206 affected sibling pairs. They detected a signal in chromosome 17 but no replication analysis was performed [1]. Another recent GWAS study investigating genetic factors involved in chronic widespread pain [30] identified a region of association in chromosome 5, near *CCT5* and *FAM173B*. These two genes also showed a higher RNA-expression in mouse models of inflammatory pain.

The aim of this study was to elucidate genetic susceptibility factors for fibromyalgia. We addressed this objective through two main approaches: a genome-wide association study (GWAS) and the evaluation of copy number variants (CNVs), using genotyping data and array comparative genomic hybridization experiments (aCGH). These analyses were performed on a large and very well characterized cohort of FM patients.

## MATERIALS AND METHODS

### Samples

Fibromyalgia Units of five Spanish Hospitals participated in the collection of samples. An initial set of 313 female samples (FM_discovery), characterized by having low levels of psychiatric comorbidities and being the best fitting the FM diagnosis, was collected at the beginning of the study, and collection continued until an additional set of 1088 female patients (FM_replication) was achieved (total female FM cohort = 1401). All patients fulfilled the 1990's American College of Rheumatology (ACR) criteria for fibromyalgia and were selected by the rheumatologists of the units participating in the study. Patients were then evaluated by another set of physicians trained in the assessment of FM patients. They all passed the same questionnaires and physical examination. For a detailed description of the cohort, see Docampo *et al,*. 2013 [9]. All samples were Spanish of Caucasian origin and had signed informed consent before enrolment. The ethics committees at all recruitment centers approved the project.

We recently performed cluster analysis on the whole cohort of patients, and found that they could be classified into three empirical subgroups, which we then labeled: "FM with low levels of comorbidities and symptomatology" (cluster 1), "FM with high levels of both symptomatology and comorbidities" (cluster 2), and "FM with high symptomatology but low levels of comorbidities" (cluster 3) [9]. A brief summary of the process is provided in the supplementary methods.

Three different control cohorts were used for this study: a cohort of 220 female Spanish samples (ECHRS) from the GABRIEL consortium (http://www.cng.fr/gabriel/index.html) was used in the GWAS analysis (con_ECHRS). In the GWAS replication studies, we genotyped a cohort of 535 female control samples (con_SAL), corresponding to subjects with low levels of pain and fatigue (as assessed by a questionnaire) provided by the National DNA Bank of Salamanca, and a set of 142 female Spanish blood donor samples (con_VH). For the CNV analysis, only the con_SAL set of control samples was used.

A flowchart representing the different cohorts and the analysis in which they were used is provided in the supplementary material (supplementary figure 1).

**Whole genome association study**

*Genotyping*

FM female cases (313 samples, FM_discovery), selected by clinicians for having low levels of psychiatric comorbidities and being the best fitting the FM diagnosis, were genotyped with Illumina 1M-Duo chip. Genotyping was performed in CeGen (Barcelona Node), following the manufacturer's protocol.

Data from 220 general Spanish population samples (GABRIEL consortium http://www.cng.fr/gabriel/index.html) genotyped with Illumina 610-quad chip was used as control dataset (con_ECHRS).

*Quality control*

Quality control (QC) was performed with PLINK [32] (Supplementary methods).

*Allelic association*

Allelic association analysis was performed with PLINK (5% of significance level). QQ plots were performed with the WGAviewer software [12], and Manhattan plot and linkage disequilibrium (LD) evaluation with Haploview software [2].

Power analysis was performed with Quanto (http://hydra.usc.edu/gxe/), showing that for SNPs with MAF≥0.05, and given our sample size, we had over 80% power to detect associations with OR≥2.0 but it showed much lower power to detect associations with smaller OR (1.2) (supplementary methods).

*Imputation*

For GWAS regions showing positive signals, we performed imputation in a window span of 100kb with Impute v2 (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html) (Supplementary methods).

*SNPs annotation and Pathway analysis*

SNPs showing strongest association were annotated with WGAViewer [12]. The relation to disease of the SNPs and their genomic regions was evaluated with the Decipher database (http://decipher.sanger.ac.uk/). These SNPs were also analyzed with Ingenuity Systems

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Pathway analysis (IPA) software (http://www.ingenuity.com/) and GeneSet analysis Toolkit v2 (http://bioinfo.vanderbilt.edu/webgestalt/). The function of associated SNPs was evaluated with Genevar (http://www.sanger.ac.uk/resources/software/genevar/) (1 MB window centred in the transcription start site of the gene, Spearman Rank Correlation (p <0.001)), Puppasuite (http://pupasuite.bioinfo.cipf.es/), and the Regulome database (http://www.regulomedb.org/).

*Replication*

Twenty-one SNPs showing strongest association values were genotyped in a replication set from our cohort, consisting in 952 female FM cases (FM_replication) and 644 female controls (529 from con_SAL + 115 from con_VH). These SNPs were prioritized based on their p-values, the presence of association peaks (clustering of SNPs with strong associations in a genomic region), their function (only including SNPs located in genes or gene regions recognizedby UCSC genome browser and WGA viewer software), and pathway analysis results, focusing on genes from the highlighted neurological disease pathway.

Genotyping was performed by Kbiosciences (KASPar®). As quality control, duplicates (two HapMap samples in each plate) and negative controls where included. SNPs not fulfilling HWE, with a MAF <5% or a low genotyping rate (<95%), as well as samples with low genotyping rate (<95%), were excluded from the analysis. This resulted in the exclusion of one SNP and 55 samples (12 cases and 43 controls). Association analysis was performed with SNPassoc R package.

**CNV analysis**

CNVs were assessed by aCGH. Experiments were performed with the Agilent® 400K using a pooling strategy to dilute common CNV polymorphisms due to inter-individual variability and highlight those common variants with a different frequency between cases and controls [8]. From the FM_discovery set, three pools of FM samples were designed, two with 20 FM cases each (FM_1 and FM_2) and one with 30 cases (FM_3). All the samples included in the pools had a family history of FM, and low levels of psychiatric comorbidities, and were grouped based on characteristics such as absence (FM1) or presence (FM2) of Chronic Fatigue Syndrome, or early onset of symptoms (FM3). FM pools were hybridized against one pool of controls from the con_SAL set (50 samples).

The resulting data was analyzed with Agilent's Genomic Workbench software, using the ADM-2 algorithm. We selected regions with at least three aberrant probes and a smoothed log2ratio >0.3 (2 standard deviations).

*Validation of array results*

Breakpoints were detected by PCR amplification of the deleted allele based on data from Conrad *et al.* [6], followed by sequencing of the amplified products and BLAT alignment. Multiplex reactions for each CNV, including primers for both the deleted and non-deleted forms, were designed for genotyping. PCR conditions for the different experiments are described in supplementary methods.

CNV genotyping reactions included both samples from FM cases and controls in order to avoid possible bias, non-template controls and, when available, positive controls (HapMap samples with validated genotypes for each CNV [6]). NRXN3_del was also assessed by a Veracode assay as described [10].

Allelic association analysis was performed with Fisher exact test. Genotypic association analyses were performed using binomial and multinomial regression models for case-control and among fibromyalgia clusters studies respectively. Quality control and association analyses of the Veracode assay were performed with PLINK software.

## RESULTS

### Genome-wide association study

*SNP genotyping analysis highlights the involvement of genes in neurological pathways*

After QC procedures, 505,454 SNPs were considered for allelic association analyses in 300 FM and 203 controls. These samples did not show evidence of population stratification, as illustrated by the QQ plot (Figure 1) and the genomic inflation value ($\lambda = 1.013$). Association analysis did not yield any SNP association over the GWAS significance threshold, as shown in the Manhattan plot (Figure 2). Nevertheless, some possible association peaks emerged at chromosomes 3 and X, and eight SNPs showed p-values $<1x10^{-5}$, while another 69 had p-values $<1x10^{-4}$.

These top 77 SNPs (p-value $<1x10^{-4}$) (Supplementary Table 4) were selected to perform pathway analysis. Out of the 77 gene IDs introduced in IPA, 53 were mapped. IPA pathway

analysis identified two top networks that were overrepresented in our geneset: reproductive system disease (16 genes) and neurological disease (12 genes). Thirty-one of the 77 gene IDs were mapped in Geneset analysis Toolkit to perform GO analysis, identifying two main molecular functions: protein binding and ion binding, with 16 and 14 genes, respectively; in these categories, the analysis also identified metal ion binding as the only statistically significant overrepresented molecular function (p-adjusted <0.05, Benjamini correction).

*Replication*

Twenty-one of the top 77 SNPs were selected for replication in a larger cohort, after a prioritization step based on imputation data and gene function. Replication was performed only in females. After QC procedures, 20 SNPs in 940 cases and 601 controls were considered for the replication analysis. We performed allelic association tests. In the replication data we observed significant evidence of SNP-effect sizes in the same direction as the discovery set (9/17). Only one of these SNPs, rs11127292, located in the *MYT1L* (myelin transcription factor 1-like) gene, showed association (p <0.05) (Table 1). Based on 17 (21, 1 no HWE, 3 not independent) independent SNPs taken forward, we would expect 0.05x17 = 0.85 to reach p = 0.05 by chance.

We then applied the subphenotyping of the samples [9], and three subphenotype cohorts were tested for association for the four SNPs showing the strongest joint p-values. Again, rs11127292 showed association in female FM cases with low levels of comorbidities and high levels of symptomatology (joint p-value = $4.28 \times 10^{-5}$) (Table 2). Potential functionality of this SNP was tested by Genvar cis-eQTL-SNP analysis, which showed a correlation trend (p=0.0091, while the empirical threshold for significance is 0.0086) between this SNP and expression levels of *STNG2* (one expression probe) (supplementary figure 3). Assessing functional significance of this SNP with Puppasuite or Regulomedb did not highlight any other potential functionality.

**CNV analysis identifies *NRXN3* as a novel FM gene**

We detected seven regions showing differential hybridization between FM and control pools both in direct and dye swap hybridizations (Supplementary table 5). Three of these regions (*WDR60*, *DOCK5* and *SIRPB1),* were not considered for replication since they repeatedly appeared in the aCGH experiments performed in the laboratory for different

disorders (data not shown). An additional region detected by the less restrictive ADM-1 was also considered for validation as a functional candidate.

We attempted to validate the remaining four regions coming from ADM-2 analysis, and the ADM-1 detected region, *MYO5B*. We mapped or confirmed CNV breakpoints (*GALNTL6*: rs67651552; *PTPRD*: rs71315285; *MYO5B*: rs72192652; WWOX: hg18_chr16: 76929139-76942400; and *NRXN3*:hg18_chr14:79175982-79184862) and designed multiplex PCR experiments for genotyping. Genotyping was performed first in a subset of 345 FM samples (selected randomly from FM_discovery and FM_replication) and 133 controls from con_SAL, and, if aCGH findings were supported, in our entire cohort (FM_discovery + FM_replication and con_SAL). Only one CNV (NRXN3) was considered for genotyping in the entire cohort.

The CNV located in the intronic region of *NRXN3* showed support for association in the initial genotyping step. Genotyping of the initial subset of 345 FM cases and 133 controls confirmed the results observed in the pools, showing an association of the deleted allele with FM (genotypic association, recessive model $p = 9.215 \times 10^{-5}$, OR (95%CI) = 3.29 (1.69-6.41); allelic association p = 0.16 Fisher Test). We then completed the genotyping of our entire cohort (genotypic association, recessive model p = 0.021, OR (95%CI) = 1.46 (1.05-2.04); allelic association Fisher Test p = 0.015, OR (95%CI) = 1.22 (1.03-1.43)), which validated the association.

We then applied the cluster classification described previously [9] and performed association analyses for each of the FM clusters against the controls. The significance of the association of the deleted allele with FM was maintained when selecting samples with low levels of comorbidities (clusters 1 and 3) (genotypic association, recessive model p = 0.019, OR (95%CI) = 1.49 (1.06-2.11); allelic association Fisher Test p = 0.004, OR (95%CI) = 1.27 (1.07-1.50)). The genotypic distribution in the different subsets is summarized in Tables 3 and 4.

**DISCUSSION**

We have explored the genetic susceptibility to FM through GWAS and aCGH assessment of CNVs. Our findings identify two associated variants, a SNP in *MYT1L,* and an intronic CNV in *NRXN3.* Both results suggest a possible role for the central nervous system in FM genetic susceptibility.

None of the SNPs reached genome-wide significance, but we observed signal peaks at chromosomes 3 and X, and identified 77 SNPs with p-values $<1x10^{-4}$. Since recent studies detected linkage to chr17 [1], and association in chr5 [30], we specifically checked our GWAS signals in these regions. Among the top 77 SNPs, four were located in chr17 and one in chr5, although outside the reported linkage or association regions. Twenty-one SNPs, including one in chr17, were selected for replication in a large independent FM cohort. Nine of the SNPs showed an effect in the same direction than those identified in the discovery cohort, and one, rs11127292, had a nominally significant p-value.

rs11127292 is located in the third intron of *MYT1L*, and may be correlated to expression levels of *SNTG2* in lymphocytes. Both *MYT1L* and *SNTG2* are good candidates for FM: they are involved in neuronal pathways, and FM could be characterized by a central nervous system (CNS) dysfunction [36]. *MYT1L is* involved in neuronal differentiation [31; 40], and variants in this gene have been associated with neuropsychiatric disorders [25; 41; 42], while variants in *SNTG2* have been associated to autism [46] and to suicide attempts in major depression [29].

The potential role for the CNS dysfunction in FM was not only highlighted by the association of rs11127292. Neurological disease was one of the two top networks identified by IPA analysis of the 77 GWAS top SNPs. Furthermore, Geneset GO analysis showed ion binding (and calcium in particular) as one of the two main molecular functions, calcium channel complex as the top molecular component, and regulation of calcium mediated signalling and calcium ion transport and neurogenesis among the top ten biological functions. These results suggest an involvement of the CNS in fibromyalgia, supporting the hypothesis that FM could be due to a dysfunction in pain processing.

Nevertheless, we have to take these results with care since this GWAS presents as its main limitation the reduced number of samples considered. Although the samples were carefully selected to attain a more homogeneous phenotype, the high genetic and phenotypic heterogeneity of FM, coupled with the small discovery sample size, could explain why we were not able to detect an association reaching GWAS significance. The small sample size could also result in the detection of false positives and a consequent lack of replication of associated variants. Since we used a different genotyping platform for the controls, despite performing careful QC, some technical bias could remain leading to false positives. Nevertheless, the association of rs11127292 was confirmed in the replication phase, and stratified analysis improved the replication results, highlighting the importance of a precise

phenotyping of the samples in dissecting the genetic components of extremely complex disorders like FM.

We also explored the contribution of CNVs to FM genetic susceptibility using aCGH with a pooling strategy. Seven regions showed differential hybridization in aCGH experiments, one of which, an intronic 8kb deletion within the *NRXN3* gene, was validated on follow-up. Further subsetting of the samples also increased the significance of the association in those samples with low levels of comorbidities. We could not confirm an association for the CNVs within *GALNTL6, WWOX*, *PTPRD* and *MYO5B,* but their participation in FM cannot be completely excluded, as they could be low effect variants, requiring larger cohorts for validation. These genes have been shown to play a role in the CNS [22; 23; 34; 38], and considering our GWAS and aCGH results, this would make them good candidates for FM susceptibility.

The use of pooled samples presents clear economic advantages, but it can also lead to a higher number of false positives. Causes of false positives can be due to real but random differences in the frequencies between cases and controls, or to unequal sample contributions to the pools, leading to biased overrepresentation of rare variants. In order to overcome these limitations, we used three independent pools of FM as biological replicates. All three FM pools were hybridized against the same control pool, which had been frequently tested in the laboratory. Variants recurrently appearing in other hybridizations with this control pool were filtered out. In addition, direct and 'dye-swap' experiments were performed. Nevertheless, only one of the five regions selected for follow-up showed association with FM in the larger cohort.

An intronic CNV in *NRXN3* was associated with FM with low levels of comorbidities. NRXN3 is a good candidate for FM as it is essential for neuronal development and for signal transmission. Neurexins are among the most widely studied adhesion and scaffolding molecules involved in synapse stability and function. They are transmembrane proteins located in the presynaptic neuron. They have three extracellular binding partners: neuroligins, dystroglycan and neurexophilins [7]. In particular, the binding with neuroligins is essential for the development and function of GABAergic and glutamatergic synapses [37]. Several variants in neurexin genes (rare CNVs and SNPs) have been associated with different phenotypes, mainly neuropsychiatric disorders, including addictive behavior [16; 20] and obesity [15]. Changes in *NRXN3* and therefore in signal transmission could explain the central nervous pain dysfunction characteristic of FM.

The detected *NRXN3* variant is an intronic CNV. We tried to identify a possible link to other functional variants by genotyping additional *NRXN3* variants in the FM samples, unsuccesfully (data not shown). We then considered possible functional consequences of the CNV at the mRNA level. NRXN3 has several different isoforms generated by alternative splicing. One of the alternative splicing sites is located relatively close to the CNV. Preliminary data from two neuronal cell lines indicate that the CNV could affect differential skipping of exon 20, which encodes a protein domain involved in binding to neuroligins [44], although data from lymphoblastoid cell lines do not support this hypothesis. In order to extract a definite conclusion regarding the correlation between exon-skipping and the CNV, it would be advisable to test additional neuronal cells with different genotypes for the deletion, as it is possible that the splicing behavior changes between lymphoblastoid and neuronal lineages.

Although our results correspond to the analysis of a small initial dataset, we have found further support in a larger, well-characterized cohort, especially when selecting for a specific phenotype. It would be particularly interesting to follow up on our results by validation of the specific findings in additional large, well characterized, cohorts. This is challenging in FM since, to our knowledge, this cohort is the largest existing FM DNA collection.

The two identified FM associations, the SNP in *MYTL1* and the CNV in *NRXN3*, were slightly improved after subphenotyping. This indicates that identifying homogeneous phenotypes constitutes a key step for the identification of FM genetic susceptibility factors. The slight improvement in the association scores further validates the cluster-based classification used, indicating that cases with low levels of comorbidities form a more genetically homogeneous subset of FM cases.

In summary, our GWAS and aCGH results point at a role for the CNS in FM genetic susceptibility. In fact, variants detected by both studies are linked: calcium transport appears as one of the main GWAS molecular functions, and the neurexins-neuroligins complex formation is dependent on calcium [26]; also, the SNP showing the strongest association has been correlated with expression levels of *SNTG2* (in blood) and SNTG2 interacts with neuroligins 3 and 4, which are neurexin partners [46]. Furthermore, another study evaluating gene expression of FM individuals detected changes in genes implicated in pain transmission, which would support our findings [24]. This CNS implication in FM would be further supported by recently published findings on pain, showing evidence for specific

neurophysiologic alterations in FM patients [39]. Of particular interest, functional and morphological impairment of small fibers have been reported in FM cases. Individuals presenting those electrophysiological changes would be the ideal models to explore the association and functionality of our detected genetic variants in order to try to establish a correlation with clinical severity, outcome and response to treatment.

In spite of the difficulties encountered in the study of genetic factors of FM (clinical heterogeneity, reduced availability of replication cohorts and non-availability of target tissue) we have been able to detect variants that can shed a light on genetic factors determining FM susceptibility. To our knowledge, only neurotransmitter related genes (including receptors, transporters and enzymes implicated in neurotransmitters metabolism) had been tested as FM susceptibility candidates. The possible role of synaptic structural molecules such as NRXN3 and molecules implicated in CNS development and functioning, such as MYTL1, open a new wide field of research on aetiology and drug targets. One consideration that we have to take into account is that all of these molecules have been previously associated with neuropsychiatric disorders. If these synapse genes associations are confirmed in other FM cohorts, it would constitute an additional argument to consider FM as a neuropsychiatric disorder.

## ACKNOWLEDGMENTS

Acknowledgements for their technical collaboration to *Anna Arias, Juan Sanchez, Rosa Badia, Frank Ramirez, Pilar Souviron, Karola Benning, and Pilar Muñoz.* This work was supported by a CENIT contract (2008-2011), the Fundación FF and by the Lilly foundation to X. Estivill. E. Docampo is supported by the Spanish Ministry of health (FIS-PFIS).R. Rabionet is supported by the Ramon y Cajal program and the ESGI European consortium. The authors report no conflict of interests.

**REFERENCES**


[1] Arnold LM, Fan J, Russell IJ, Yunus MB, Khan MA, Kushner I, Olson JM, Iyengar SK. The fibromyalgia family study: a genome-wide linkage scan study. Arthritis Rheum 2013;65(4):1122-1128.

[2] Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21(2):263-265.

[3] Branco JC, Bannwarth B, Failde I, Abello Carbonell J, Blotman F, Spaeth M, Saraiva F, Nacci F, Thomas E, Caubere JP, Le Lay K, Taieb C, Matucci-Cerinic M. Prevalence of fibromyalgia: a survey in five European countries. Semin Arthritis Rheum 2010;39(6):448-453.

[4] Buskila D, Neumann L. Fibromyalgia syndrome (FM) and nonarticular tenderness in relatives of patients with FM. J Rheumatol 1997;24(5):941-944.

[5] Buskila D, Neumann L, Hazanov I, Carmi R. Familial aggregation in the fibromyalgia syndrome. Semin Arthritis Rheum 1996;26(3):605-611.

[6] Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, Aerts J, Andrews TD, Barnes C, Campbell P, Fitzgerald T, Hu M, Ihm CH, Kristiansson K, Macarthur DG, Macdonald JR, Onyiah I, Pang AW, Robson S, Stirrups K, Valsesia A, Walter K, Wei J, Tyler-Smith C, Carter NP, Lee C, Scherer SW, Hurles ME. Origins and functional impact of copy number variation in the human genome. Nature 2010.

[7] Craig AM, Kang Y. Neurexin-neuroligin signaling in synapse development. Curr Opin Neurobiol 2007;17(1):43-52.

[8] de Cid R, Riveira-Munoz E, Zeeuwen PL, Robarge J, Liao W, Dannhauser EN, Giardina E, Stuart PE, Nair R, Helms C, Escaramis G, Ballana E, Martin-Ezquerra G, den Heijer M, Kamsteeg M, Joosten I, Eichler EE, Lazaro C, Pujol RM, Armengol L, Abecasis G, Elder JT, Novelli G, Armour JA, Kwok PY, Bowcock A, Schalkwijk J, Estivill X. Deletion of the late cornified envelope LCE3B and LCE3C genes as a susceptibility factor for psoriasis. Nat Genet 2009;41(2):211-215.

[9] Docampo  E CA, Escaramis G,Carbonell J,Rivera J, Redondo J, Alegre J, Rabionet R and Estivill X Cluster analysis identifies fibromyalgia subgroups. PLoS One 2013;8(9) e74873.

[10] Docampo E, Ribases M, Gratacos M, Bruguera E, Cabezas C, Sanchez-Mora C, Nieva G, Puente D, Argimon-Pallas JM, Casas M, Rabionet R, Estivill X. Association of Neurexin 3 polymorphisms with smoking behavior. Genes Brain Behav 2012;11(6):704-711.

[11] Fitzcharles MA, Yunus MB. The clinical concept of fibromyalgia as a changing paradigm in the past 20 years. Pain Res Treat 2012;2012:184835.

[12] Ge D, Zhang K, Need AC, Martin O, Fellay J, Urban TJ, Telenti A, Goldstein DB. WGAViewer: software for genomic annotation of whole genome association studies. Genome Res 2008;18(4):640-643.

[13] Gracely RH, Ambrose KR. Neuroimaging of fibromyalgia. Best Pract Res Clin Rheumatol 2011;25(2):271-284.

[14] Harris RE, Clauw DJ, Scott DJ, McLean SA, Gracely RH, Zubieta JK. Decreased central mu-opioid receptor availability in fibromyalgia. J Neurosci 2007;27(37):10000-10006.

[15] Heard-Costa NL, Zillikens MC, Monda KL, Johansson A, Harris TB, Fu M, Haritunians T, Feitosa MF, Aspelund T, Eiriksdottir G, Garcia M, Launer LJ, Smith AV, Mitchell BD, McArdle PF, Shuldiner AR, Bielinski SJ, Boerwinkle E, Brancati F, Demerath EW, Pankow JS, Arnold AM, Chen YD, Glazer NL, McKnight B, Psaty BM, Rotter JI, Amin N, Campbell H, Gyllensten U, Pattaro C, Pramstaller PP, Rudan I, Struchalin M, Vitart V, Gao X, Kraja A, Province MA, Zhang Q, Atwood LD, Dupuis J, Hirschhorn JN,

Jaquish CE, O'Donnell CJ, Vasan RS, White CC, Aulchenko YS, Estrada K, Hofman A, Rivadeneira F, Uitterlinden AG, Witteman JC, Oostra BA, Kaplan RC, Gudnason V, O'Connell JR, Borecki IB, van Duijn CM, Cupples LA, Fox CS, North KE. NRXN3 is a novel locus for waist circumference: a genome-wide association study from the CHARGE Consortium. PLoS Genet 2009;5(6):e1000539.
[16] Hishimoto A, Liu QR, Drgon T, Pletnikova O, Walther D, Zhu XG, Troncoso JC, Uhl GR. Neurexin 3 polymorphisms are associated with alcohol dependence and altered expression of specific isoforms. Hum Mol Genet 2007;16(23):2880-2891.
[17] Horven S, Stiles TC, Holst A, Moen T. HLA antigens in primary fibromyalgia syndrome. J Rheumatol 1992;19(8):1269-1270.
[18] Kato K, Sullivan PF, Evengard B, Pedersen NL. Importance of genetic influences on chronic widespread pain. Arthritis Rheum 2006;54(5):1682-1686.
[19] Kuchinad A, Schweinhardt P, Seminowicz DA, Wood PB, Chizh BA, Bushnell MC. Accelerated brain gray matter loss in fibromyalgia patients: premature aging of the brain? J Neurosci 2007;27(15):4004-4007.
[20] Lachman HM, Fann CS, Bartzis M, Evgrafov OV, Rosenthal RN, Nunes EV, Miner C, Santana M, Gaffney J, Riddick A, Hsu CL, Knowles JA. Genomewide suggestive linkage of opioid dependence to chromosome 14q. Hum Mol Genet 2007;16(11):1327-1334.
[21] Lee YH, Choi SJ, Ji JD, Song GG. Candidate gene studies of fibromyalgia: a systematic review and meta-analysis. Rheumatol Int 2012;32(2):417-426.
[22] Li MY, Lai FJ, Hsu LJ, Lo CP, Cheng CL, Lin SR, Lee MH, Chang JY, Subhan D, Tsai MS, Sze CI, Pugazhenthi S, Chang NS, Chen ST. Dramatic co-activation of WWOX/WOX1 with CREB and NF-kappaB in delayed loss of small dorsal root ganglion neurons upon sciatic nerve transection in rats. PLoS One 2009;4(11):e7820.
[23] Lise MF, Wong TP, Trinh A, Hines RM, Liu L, Kang R, Hines DJ, Lu J, Goldenring JR, Wang YT, El-Husseini A. Involvement of myosin Vb in glutamate receptor trafficking. J Biol Chem 2006;281(6):3669-3678.
[24] Lukkahatai N, Majors B, Reddy S, Walitt B, Saligan LN. Gene expression profiles of fatigued fibromyalgia patients with different categories of pain and catastrophizing: A preliminary report. Nursing outlook 2013;61(4):216-224 e212.
[25] Meyer KJ, Axelsen MS, Sheffield VC, Patil SR, Wassink TH. Germline mosaic transmission of a novel duplication of PXDN and MYT1L to two male half-siblings with autism. Psychiatr Genet 2012;22(3):137-140.
[26] Missler M, Zhang W, Rohlmann A, Kattenstroth G, Hammer RE, Gottmann K, Sudhof TC. Alpha-neurexins couple Ca2+ channels to synaptic vesicle exocytosis. Nature 2003;423(6943):939-948.
[27] Mogil JS. Pain genetics: past, present and future. Trends Genet 2012;28(6):258-266.
[28] Norbury TA, MacGregor AJ, Urwin J, Spector TD, McMahon SB. Heritability of responses to painful stimuli in women: a classical twin study. Brain 2007;130(Pt 11):3041-3049.
[29] Perlis RH, Ruderfer D, Maussion G, Chambert K, Gallagher P, Turecki G, Ernst C. Bipolar Disorder and a History of Suicide Attempts With a Duplication in 5HTR1A. Am J Psychiatry 2012;169(11):1213-1214.
[30] Peters MJ, Broer L, Willemen HL, Eiriksdottir G, Hocking LJ, Holliday KL, Horan MA, Meulenbelt I, Neogi T, Popham M, Schmidt CO, Soni A, Valdes AM, Amin N, Dennison EM, Eijkelkamp N, Harris TB, Hart DJ, Hofman A, Huygen FJ, Jameson KA, Jones GT, Launer LJ, Kerkhof HJ, de Kruijf M, McBeth J, Kloppenburg M, Ollier WE, Oostra B, Payton A, Rivadeneira F, Smith BH, Smith AV, Stolk L, Teumer A, Thomson W, Uitterlinden AG, Wang K, van Wingerden SH, Arden NK, Cooper C, Felson D, Gudnason V, Macfarlane GJ, Pendleton N, Slagboom PE, Spector TD, Volzke H, Kavelaars A, van Duijn CM, Williams FM, van Meurs JB. Genome-wide association

study meta-analysis of chronic widespread pain: evidence for involvement of the 5p15.2 region. Annals of the rheumatic diseases 2013;72(3):427-436.
[31] Pfisterer U, Kirkeby A, Torper O, Wood J, Nelander J, Dufour A, Bjorklund A, Lindvall O, Jakobsson J, Parmar M. Direct conversion of human fibroblasts to dopaminergic neurons. Proc Natl Acad Sci U S A 2011;108(25):10343-10348.
[32] Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81(3):559-575.
[33] Rivera J AC, Ballina FJ, Carbonell J, Carmona L, Castel B, Collado A, Esteve JJ, Martínez FG, Tornero J, Vallejo MA, Vidal J. Consensus document on fibromyalgia of the Spanish Society of Rheumatology  Reumatología Clínica 2006;2S1:55-66.
[34] Sklar P, Smoller JW, Fan J, Ferreira MA, Perlis RH, Chambert K, Nimgaonkar VL, McQueen MB, Faraone SV, Kirby A, de Bakker PI, Ogdie MN, Thase ME, Sachs GS, Todd-Brown K, Gabriel SB, Sougnez C, Gates C, Blumenstiel B, Defelice M, Ardlie KG, Franklin J, Muir WJ, McGhee KA, MacIntyre DJ, McLean A, VanBeck M, McQuillin A, Bass NJ, Robinson M, Lawrence J, Anjorin A, Curtis D, Scolnick EM, Daly MJ, Blackwood DH, Gurling HM, Purcell SM. Whole-genome association study of bipolar disorder. Mol Psychiatry 2008;13(6):558-569.
[35] Smith SB, Maixner DW, Fillingim RB, Slade G, Gracely RH, Ambrose K, Zaykin DV, Hyde C, John S, Tan K, Maixner W, Diatchenko L. Large candidate gene association study reveals genetic risk factors and therapeutic targets for fibromyalgia. Arthritis Rheum 2012;64(2):584-593.
[36] Staud R, Nagel S, Robinson ME, Price DD. Enhanced central pain processing of fibromyalgia patients is maintained by muscle afferent input: a randomized, double-blind, placebo-controlled study. Pain 2009;145(1-2):96-104.
[37] Sudhof TC. Neuroligins and neurexins link synaptic function to cognitive disease. Nature 2008;455(7215):903-911.
[38] Takahashi H, Katayama K, Sohya K, Miyamoto H, Prasad T, Matsumoto Y, Ota M, Yasuda H, Tsumoto T, Aruga J, Craig AM. Selective control of inhibitory synapse development by Slitrk3-PTPdelta trans-synaptic interaction. Nat Neurosci 2012;15(3):389-398, S381-382.
[39] Uceyler N, Zeller D, Kahn AK, Kewenig S, Kittel-Schneider S, Schmid A, Casanova-Molla J, Reiners K, Sommer C. Small fibre pathology in patients with fibromyalgia syndrome. Brain 2013.
[40] Vierbuchen T, Ostermeier A, Pang ZP, Kokubu Y, Sudhof TC, Wernig M. Direct conversion of fibroblasts to functional neurons by defined factors. Nature 2010;463(7284):1035-1041.
[41] Vrijenhoek T, Buizer-Voskamp JE, van der Stelt I, Strengman E, Sabatti C, Geurts van Kessel A, Brunner HG, Ophoff RA, Veltman JA. Recurrent CNVs disrupt three candidate genes in schizophrenia patients. Am J Hum Genet 2008;83(4):504-510.
[42] Wang T, Zeng Z, Li T, Liu J, Li J, Li Y, Zhao Q, Wei Z, Wang Y, Li B, Feng G, He L, Shi Y. Common SNPs in myelin transcription factor 1-like (MYT1L): association with major depressive disorder in the Chinese Han population. PLoS One 2010;5(10):e13662.
[43] Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Katz RS, Mease P, Russell AS, Russell IJ, Winfield JB, Yunus MB. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res (Hoboken) 2010;62(5):600-610.
[44] Wolfe F, Smythe HA, Yunus MB, Bennett RM, Bombardier C, Goldenberg DL, Tugwell P, Campbell SM, Abeles M, Clark P, et al. The American College of Rheumatology 1990 Criteria for the Classification of Fibromyalgia. Report of the Multicenter Criteria Committee. Arthritis Rheum 1990;33(2):160-172.

[45] Wood PB, Patterson JC, 2nd, Sunderland JJ, Tainter KH, Glabus MF, Lilien DL. Reduced presynaptic dopamine activity in fibromyalgia syndrome demonstrated with positron emission tomography: a pilot study. J Pain 2007;8(1):51-58.
[46] Yamakawa H, Oyama S, Mitsuhashi H, Sasagawa N, Uchino S, Kohsaka S, Ishiura S. Neuroligins 3 and 4X interact with syntrophin-gamma2, and the interactions are affected by autism-related mutations. Biochem Biophys Res Commun 2007;355(1):41-46.
[47] Yunus MB, Khan MA, Rawlings KK, Green JR, Olson JM, Shah S. Genetic linkage analysis of multicase families with fibromyalgia syndrome. J Rheumatol 1999;26(2):408-412.

**FIGURE LEGENDS:**

**Fig. 1**: QQ plot of fibromyalgia GWAS. Observed p-values are plotted against expected p-values in an association study of 500 k SNPs. The almost perfect correlation between observed and expected values was indicative of absence of population stratification as proven by a genomic inflation (λ) value of 1.013. Figure obtained with WGA viewer software.

**Fig. 2**: Manhattan plot representing the results of the fibromyalgia GWAS. The negative $LOG_{10}$ p-values of all SNPs are plotted against their chromosomal positions. Chromosomes are represented by different colors.

**SUMMARY**

Variants in *NRXN3* and *MYT1L* are associated with fibromyalgia. Our results point to a role for the central nervous system in susceptibility to fibromyalgia.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Figure 1
Click here to download high resolution image

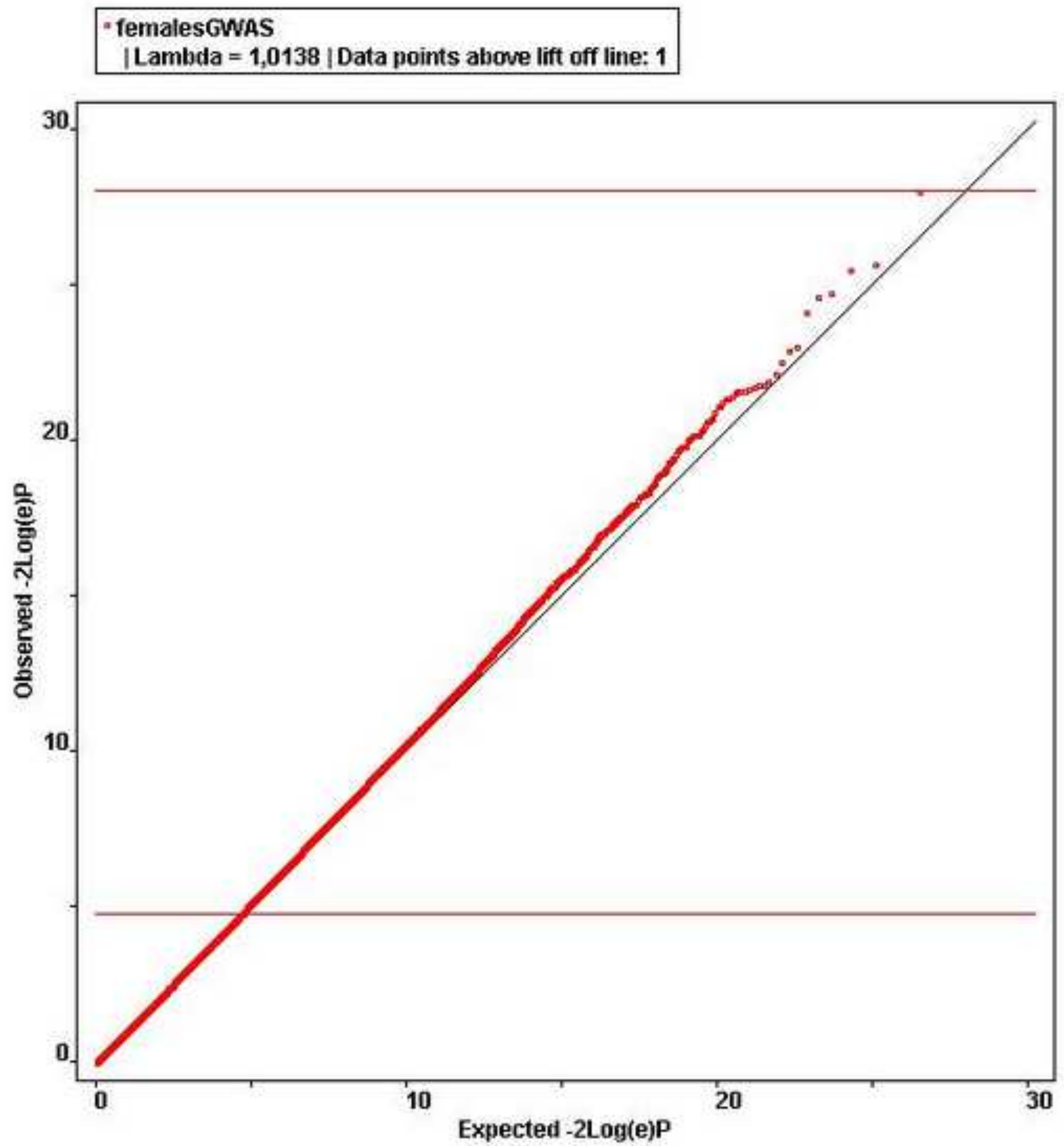
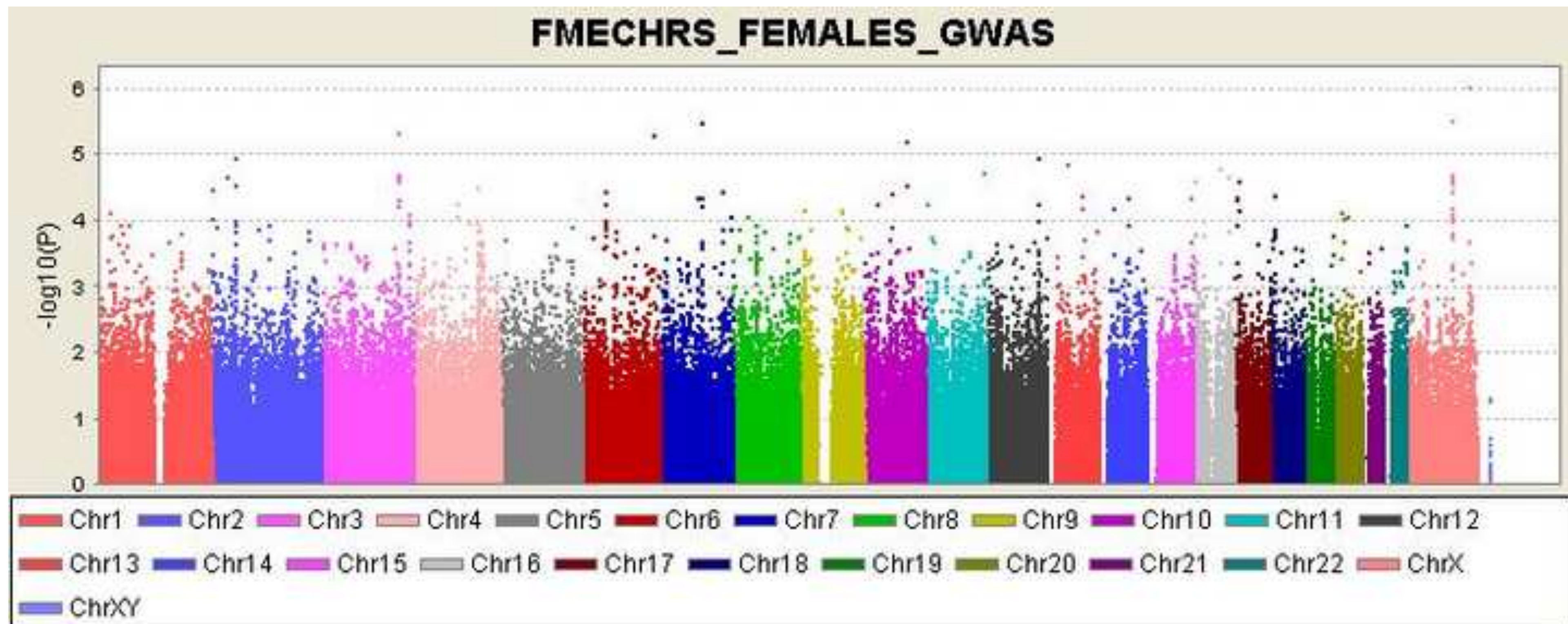


Figure 2

Click here to download high resolution image



Table

**Table 2**. rs11127292 allelic association in the different fibromyalgia clusters in GWAS, replication and joint cohorts

| rs11127292 | F_FM | F_CONTROLS | P-value | OR (95%CI) |
|---|---|---|---|---|
| Discovery (300 FM vs. 203 C) | 0.051 | 0.125 | $2.6x10^{-5}$ | 0.37 (0.23-0.60) |
| Replication (940 FM vs. 592 C) | 0.091 | 0.114 | 0.03 | 0.76 (0.60-0.97) |
| Discovery + Replication (1240 FM vs. 795 C) | 0.081 | 0.117 | $1.76x10^{-4}$ | 0.67 (0.54-0.82) |
| Discovery cl3 (196 FM vs. 203 C) | 0.045 | 0.125 | $6.21x10^{-5}$ | 0.33 (0.19-0.58) |
| Replication cl3 (450 FM vs. 592 C) | 0.083 | 0.114 | 0.019 | 0.70 (0.52-0.94) |
| Discovery cl1 + Replication cl1 (240 FM vs. 795 C) | 0.085 | 0.117 | 0.05 | 0.70 (0.49-1.01) |
| Discovery cl2 + Replication cl2 (304FM vs. 795 C) | 0.092 | 0.117 | 0.09 | 0.76 (0.55-1.04) |
| Discovery cl3 + Replication cl3 (646 FM vs. 795 C) | 0.071 | 0.117 | $4.28x10^{-5}$ | 0.58 (0.44-0.75) |
| Discovery cl1_3 + Replication cl1_3 (886 FM vs. 795 C) | 0.075 | 0.117 | $4.03x10^{-5}$ | 0.61 (0.48-0.77) |

F: frequency of the effect allele (minor allele; A or T in rs1112792); Discovery: FM_discovery vs CON_discovery; Replication: FM_replication vs CON_VH + CON_SAL; FM: fibromyalgia individuals; C: controls; cl1, cl2, cl3, cl1_3: clusters 1, 2, 3, and 1 and 3, respectively.

Table

**Table 3.** Genotype distribution of NRXN3_DEL among cases and controls in the studied cohorts

| GENOTYPES | CON = 445<br>N (%) | FM = 1358<br>N (%) | FM cl1 = 268<br>N (%) | FM cl2 = 330<br>N (%) | FM cl3 = 712<br>N (%) | FM cl1_3 = 980<br>N (%) |
|---|---|---|---|---|---|---|
| Del/Del | 48 (10.8) | 204 (15.0) | 50 (17.5) | 497(14.2) | 103 (14.5) | 150 (15.3) |
| Del/NoDel | 207 (46.5) | 640 (47.1) | 132 (47.4) | 146 (44.2) | 349 (49.0) | 476 (48.6) |
| NoDel/NoDel | 190 (42.7) | 514 (37.8) | 95 (35.1) | 137 (41.5) | 260 (36.5) | 354 (36.1) |

Del: Deleted; NoDel: Non-deleted; CON: Control; FM: Fibromialgia (FM_discovery + FM_replication); FM cl1, FM cl2, FM cl3, FM cl1_3: Fibromyalgia cases corresponding to clusters 1, 2, 3, and 1 and 3, respectively. 48 FM samples could not be assigned to a cluster.

**Table**

**Table 1**. SNPs selected for replication in our GWAS study of fibromyalgia

| SNP | Rank | Type | Gene/Region | P GWAS | OR (95%CI) | P replication | OR (95%CI) | Combined P | OR (95%CI) |
|---|---|---|---|---|---|---|---|---|---|
| rs12556003 | 1 | intron | MCF2 | $2.14 \times 10^{-6}$ | 0.34 (0.21-0.54) | 0.132 | 0.81 (0.62-1.06) | $1.98 \times 10^{-4}$ | 0.65 (0.52-0.81) |
| rs12704506 | 2 | intergenic | STEAP1/STEAP2 | $3.20 \times 10^{-6}$ | 2.19 (1.56-3.07) | 0.010 | 0.81 (0.67-0.96) | 0.9 | 1.00 (0.86-1.17) |
| rs11923054 | 3 | intron | ZBBX | $3.52 \times 10^{-6}$ | 0.54 (0.42-0.70) | 0.580 | 0.93 (0.80-1.08) | 0.0047 | 0.83 (0.73-0.94) |
| rs2858166 | 4 | 5 kb_down | ARMCX6 | $3.61 \times 10^{-6}$ | 1.83 (1.42-2.38) | 0.695 | 0.99 (0.86-1.15) | 0.008 | 1.18 (1.04-1.34) |
| rs10782344 | 5 | intergenic | RP11-518I13.1 | $3.63 \times 10^{-6}$ | 2.22 (1.57-3.12) | 0.526 | 1.07 (0.89-1.28) | 0.004 | 1.26 (1.07-1.48) |
| rs1998709 | 6 | intron | PLCE1 | $7.62 \times 10^{-6}$ | 1.80 (1.39-2.33) | 0.14 | 0.87 (0.75-1.02) | 0.3397 | 1.06 (0.93-1.21) |
| rs2901761 | IM | Intron | PLCE1 | $5.11 \times 10^{-7}$ | 0.50 (0.38-0.66) | ‡ | ‡ | ‡ | ‡ |
| rs2194390 | 7 | intron | NRXN1 | $8.06 \times 10^{-6}$ | 0.41 (0.21-0.67) | 0.702 | 1.07 (0.84-1.35) | 0.049 | 0.81 (0.67-0.99) |
| rs2701106 | 8 | intergenic | TBX5 | $8.94 \times 10^{-6}$ | 1.84 (1.40-2.42) | 0.287 | 0.90 (0.77-1.05) | 0.194 | 1.09 (0.95-1.24) |
| rs7963168 | IM | intergenic | TBX5 | $1.49 \times 10^{-8}$ | 0.52(0.40-0.67) | 0.08 | 1.17 (1.01-1.36) | 0.28 | 1.07 (0.94-1.22) |
| rs17512210 | 15 | intron | SHISA6 | $2.21 \times 10^{-5}$ | 0.56(0.43-0.73) | 0.851 | 1.01 (0.86-1.17) | 0.049 | 0.87 (0.76-0.99) |
| rs9381682 | 18 | intergenic | - | $2.48 \times 10^{-5}$ | 0.45 (0.31-0.65) | 0.156 | 0.87 (0.69-1.09) | $7.38 \times 10^{-4}$ | 0.72 (0.59-0.87) |
| rs11127292 | 20 | intron | MYT1L | $2.60 \times 10^{-5}$ | 0.37 (0.23-0.60) | 0.039 | 0.76 (0.60-0.97) | $1.76 \times 10^{-4}$ | 0.67 (0.54-0.82) |
| rs12770855 | 32 | intergenic | ZNF438 | $4.05 \times 10^{-5}$ | 0.41 (0.27-0.63) | 0.149 | 0.85 (0.67-1.07) | 0.001 | 0.71 (0.58-0.87) |
| rs10821659 | 33 | intron | ANK3 | $4.06 \times 10^{-5}$ | 0.58 (0.44-0.75) | 0.110 | 0.88 (0.73-1.03) | $6.22 \times 10^{-4}$ | 0.79 (0.70-0.90) |
| rs265015 | 34 | intron | UNC5C | $4.12 \times 10^{-5}$ | 0.41 (0.26-0.63) | 0.016 | 1.41 (1.06-1.88) | 0.852 | 0.97 (0.77-1.23) |
| rs9565180 | 37 | intron | LMO7 | $4.53 \times 10^{-5}$ | 0.54 (0.41-0.73) | 0.242 | 1.13 (0.94-1.35) | 0.267 | 0.91 (0.79-1.06) |
| rs6043433 | 43 | intron | MACROD2 | $5.00 \times 10^{-5}$ | 0.52 (0.38-0.71) | 0.288 | 1.11 (0.91-1.36) | 0.899 | 0.89 (0.76-1.06) |
| rs6131711 | IM | intron | MACROD2 | $9.92 \times 10^{-8}$ | 1.87 (1.45-2.83) | 0.596 | 1.05 (0.91-1.23) | $9 \times 10^{-4}$ | 1.25 (1.09-1.42) |
| rs11602757 | 57 | intron | LRG_164 | $7.15 \times 10^{-5}$ | 2.77 (1.64-4.68) | 0.795 | 0.95 (0.74-1.23) | 0.077 | 1.22 (0.97-1.53) |
| rs981524 | 67 | intron | AKAP6 | $7.94 \times 10^{-5}$ | 1.98 (1.40-2.81) | 0.139 | 1.18 (0.99-1.42) | 0.001 | 1.30 (1.11-15.52) |

SNPs are listed in ascending order according to allelic association GWAS p-value (with the exception of the three imputed SNPs). Type indicates the relative position of the SNP with respect to the nearest gene, and gene provides the gene in which the SNP is located or the nearest gene in a 500 kb window. IM: imputed SNP. Type: SNP location upon the gene according to WGA viewer classification; ‡ rs2901761 was not in HWE in controls (p-value = 0.005).

**Table**

**Table 4.** Allele distribution of NRXN3_DEL among cases and controls in the studied cohorts.

| ALLELES | CON (%) | FM (%) | FM (cl1+cl3) (%) |
|---|---|---|---|
| NoDel | 587 (65.9) | 1668(61.4) | 1184 (60.4) |
| Del | 303 (34.1) | 1048(38.6) | 776 (39.6) |

Del: deleted; NoDel: non-deleted; CON: con_VH + con_SAL; FM: FM_discovery + FM_replication; Cl1: cluster 1; Cl3: cluster 3.

Supplementary Materials: figures, tables
Click here to download Supplementary Materials: figures, tables: Supplementary info.docx

# SUPPLEMENTARY MATERIAL

## SUPPLEMENTARY METHODS

### Genome-wide association study

*Quality control*

Quality control (QC) was performed with PLINK software. In order to minimize platform bias, QC steps were executed separately in the FM and control datasets, taking only into account the SNPs that overlapped between the two datasets (582,892 SNPs). Samples were checked for origin, heterozygosity, genotyping rate, inbreeding and gender. Sample filtered datasets were then filtered at the SNP level considering genotyping rate (≥96%), Hardy Weinberg Equilibrium (p>0.0001) and minimum allele frequency (MAF) ≥5%. After QC at the sample and SNP levels, the FM dataset included 308 cases (300 females and 8 males) and 513,897 SNPs, and the control dataset 395 controls (203 females and 192 males) and 512,615 SNPs. Then, FM and control datasets were merged with 505,454 overlapping SNPs. Since 97% of the FM samples were females we only included female controls to have a gender matched control set.

*Imputation*

For GWAS regions showing positive signals, we performed imputation in a window span of 100kb with Impute v2 (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html) (Supplementary methods), considering an interval buffer region of 500 kb and using as reference panels 1000Genomes and Hapmap3 CEU data. Association analysis of imputed data was performed with SNPTEST (http://mathgen.stats.ox.ac.uk/genetics), which allows dealing with genotype uncertainty due to the imputation process. For this we used the score approach based on missing data likelihood score test. Only SNPs with an Infoscore >0.8 were taken into consideration.

### Cluster analysis

A full description of the cluster analysis is provided in ref. 9. Briefly, the full 1500 fibromyalgia samples cohort was randomly divided in two sets. The cluster analysis of the clinical variables was performed in one of the sample sets and validated in the second set. Then, two sets of variables were used to score all samples, which were a priori divided into three clusters by K-means analysis. A post-hoc labeling of the clusters based on their clinical characteristics was applied. A few individuals could not be classified into clusters as some of the necessary variables were not available. The cluster distribution of the samplesfrom each cohort is given below:

| | Cluster1 | Cluster2 | Cluster3 |
|---|---|---|---|
| **FM_discovery (291)** | 66(22.7%) | 29(10%) | 196 (67.3%) |
| **FM_Replication(1135)** | 217(19.1%) | 328(28.9%) | 590 (52%) |
| **Joint (1398)** | 283(20.2%) | 357(25.6%) | 758 (54.2) |

**CNV analysis: WWOX Breakpoints detection**

*PCR conditions*

Primers (Table S1) were designed with Primer3 software taking into account aCGH positive probes coordinates and overlapping CNVs described in the Database of Genomic Variants. We took advantage of data available in Conrad et al work and included, as positive controls, HapMap samples with known genotypes for the CNV (Two copies: NA07019, NA06994; One copy: NA12056, NA12145, NA12864; Zero copies: NA06991; NA12761).

**Table S1.** Primers used for WWOX breakpoints detection.

| CNV | Primers (5'-3') | Product size (bp) |
|---|---|---|
| *WWOX* | **WWOX_F1**: TGGGTAGGAATCCTGCAGAC<br>**WWOX_R1**: TGCCTAAAAGCACACACTGC<br>**WWOX_R2**: GGGCATCCCAGTTTTCTACC<br>**WWOX_R3**: CCTGCTTCCTGAACATTCCT | **Depending on primers combinations** |

PCR reaction was performed following the PCR program below:

2' 94ºC and 30 cycles: 30" 94ºC; 30" 60ºC; 45" 68ºC; 7' 68ºC

Mix components:

50 ng DNA; 10x Roche® PCR reaction buffer + $Mg^{+2}$; 0.2 mM dNTPs; 0.4 pM/µl each primer 0.1 U/µl *Taq* Polymerase; $H_2O$ to reach a final volume of 25 µl.

*Removal of PCR primers and reagents*

5 µl PCR product of homozygous deleted samples were cleaned up with 2 µl of USB® Exo-SAP-IT® following an incubation of 15' at 37ºC plus 15' at 80ºC.

*Sequencing PCR*

1 µl of PCR product after Exosap, was added to a mix of 1 µl Big Dye Terminator® v3.1 (Applied Biosystems), 1.5 µl 5X Buffer, 0.5 µl of either reverse or forward primer (10 µM) and 6 µl $H_2O$,

and a sequencing PCR reaction was performed following the PCR program below (30 cycles): 30” 95ºC; 30” 50ºC; 3’ 60ºC

*Purification sequencing PCR*

Sequencing PCR products were purified with sepharose (Sephadex®-G50) columns. Briefly, 800 µl of sepharose were pipetted into a column and centrifuged at 1000 g for one minute for sepharose compactation. Flow-through was discarded and 10 µl of water were added to the sepharose column and centrifuged for one minute again at 1000g. Finally, the column was introduced in a new eppendorf, the sequencing PCR loaded into the column and centrifuged for 1 min at 1000g. The purified PCR was run in a capillary sequenced (3730XL Applied Biosystems).

*Sequence analysis and blast to human genome*

Sequencing results were analyzed with CLC workbench with standard settings. Only clean sequences were selected for blast analysis in UCSC Genome Browser.

**CNVs genotyping**

*PCR conditions*

PCR conditions and primers for CNVs genotyping are summarized in Tables S2 and S3 below.

**Table S2**. PCR conditions for CNV genotyping. Grey cells correspond to shared components in between reactions

| | *ACACA* | *GALNTL6* | *WWOX* | *MYO5B* | *PTPRD* | *NRXN3* |
|---|---|---|---|---|---|---|
| PCR program<br>(30 cycles) | 2’ 94ºC<br>30” 94ºC<br>30” 60ºC<br>1’ 72ºC<br>7’ 72ºC | 2’ 94ºC<br>30” 94ºC<br>30” 63ºC<br>30” 72ºC<br>25’ 72ºC | 2’ 94ºC<br>30” 94ºC<br>30” 61ºC<br>30” 72ºC<br>7’ 72ºC | 2’ 94ºC<br>30” 94ºC<br>30” 62ºC<br>30” 72ºC<br>7’ 72ºC | 2’ 94ºC<br>30” 94ºC<br>30” 60ºC<br>30” 72ºC<br>7’ 72ºC | 2’ 94ºC<br>30” 94ºC<br>30” 60ºC<br>30” 72ºC<br>7’ 72ºC |
| Mix components | | | | | | |
| 50 ng DNA | - | 75 ng | - | - | - | - |
| 10x Roche® PCR reaction buffer+$Mg^{+2}$ | - | - | - | - | - | - |
| dNTPs (mM) | 0.125 | 0.06 | 0.125 | 0.2 | 0.2 | 0.15 |
| Del Primers (pM) | 0.40 | 0.02 | 0.4 | 0.4 | 0.24 | 0.4 |
| Non-Del primers (pM) | 0.24 | 0.032 | 0.4 | 0.4 | 0.24 | 0.2 |
| Taq Polymerase (U) | 0.1 | 0.06 | 0.1 | 0.06 | 0.06 | 0.02 |
| $H_2O$ to volume of 25 µl | - | - | - | - | - | 15µl |

**Table S3**. Primers for CNV genotyping PCR reactions

| Allele | Primers (5'-3') | Product size (bp) |
|---|---|---|
| **ACAC Del** | **ACACA_DelF**: GGCCTCCTCTTCTAGCTGTTG | **1164** |
| | **ACACA_DelR**: AACAGGTGCCCAATAAATGC | |
| **ACACA Non-Del** | **ACACA_F**: GAGCCCATTAATCCAGAAAGG | **449** |
| | **ACACA_R**: TGACTTAGTGCCCATTCAAGG | |
| **GALNTL6 Del** | **FF_Mi_A:** [6FAM]GCAAGTAATGCCCAAGGAAA | **250** |
| | **RF_Mi_del2:** AGAGCATAAACCTCACAGGAC | |
| **GALNTL6 Non-Del** | **FF_Mi_wt:** [6FAM]TGGTAATGAGCAGAGGAAAGG | **283** |
| | **RF_Mi_wt5:** TGAGCACTTACCCTGTCTGC | |
| **WWOX Del** | **WWOX_DelF**: ATCTGGCCATGTCCTCATTT | **192** |
| | **WWOX_DelR**: TGTGACCTGATAACCGCTGA | |
| **WWOX Non-Del** | **WWOX_F**: AATGGGAATCTTTGCCTGTG | **217** |
| | **WWOX_R**: ATGGCAACTGACTTGGGAAG | |
| **MYO5B Del** | **MYO5B_DelF**: AACAGGCTGTCTCTTCCATGA | **234** |
| | **MYO5B_DelR:** CAGGGGTGGTTAGAATGAGG | |
| **MYO5B Non-Del** | **MYO5B_F:** GAATGCATTTTGTCCAGCAGT | **201** |
| | **MYO5B_R:** CTCATAGAGGCGGTGTTCTTG | |
| **PTPRD Del** | **PTPRD_DelF**: GGGTGGTGGAAGGTGGTTAT | **450 Del** |
| | **PTPRD_DelR:** GGTCTGGCATTTTGACATGA | **980 Non-Del** |
| **PTPRD Non-Del** | **PTPRD _F:** GCCAATTTCAGATCCTCAGC | **219** |
| | **PTPRD _R:** TTAGTGGCGTTCACACATGG | |
| **NRXN3 Del** | **NRXN3_FDel**: CAGTCTTGACTGCTGGGTGAAC | **466** |
| | **NRXN3_R**: [6FAM]GTGACTGCTGATGAGCCACGC | |
| **NRXN3 Non-Del** | **NRXN3_FNodel**: GTGAGCACTCGATCCAGCATAA | **350** |
| | **NRXN3_R**: [6FAM]GTGACTGCTGATGAGCCACGC | |

*PCR products detection*

PTPRD PCR products were loaded in a 2% agarose gel and visualized with a UV transilluminator (Gel Doc® (Bio Rad)). WWOX and MYO5B PCR products were loaded in a 3% agarose gel and visualized with a UV transilluminator (Gel Doc® (Bio Rad)).

*GALNTL6* CNV and NRXN3_DEL were genotyped by multiplex PCR with 5' FAM modification, followed by capillary electrophoresis in a 3730XL automatic sequencer and analysis with the Gene Mapper package (Applied Biosystems, Foster City, CA). Analysis was performed with the Gene Mapper package (Applied Biosystems). Samples showing peak intensities below 1000 fluorescent units or ratios of deleted allele to non-deleted allele <0.2 or >5 were not considered for analysis. For capillary detection, NRXN3_DEL PCR reactions were diluted at 1:15, and 1 µl of PCR dilution was then added to 9 µl of a formamide/ROX mixture (950 µl + 20 µl per 100 samples), and samples were loaded into 3730XL. GALNTL6 PCR products were not diluted: 1 µl of the PCR was added to the formamide/ROX mixture.

NRXN3_del was also assessed with 3 SNPs included in a Veracode assay. Two of the SNPs were located within the deleted region (rs12894142 and rs12100748), and we designed a third SNP assay (NRXN3del) with each of its extension probes flanking the breakpoints of the CNV. A combination of the results for these SNPs was used to assess the genotype. A sample was considered as homozygous deleted when failing in both SNPs included in the deleted region and amplyfing in the breakpoints SNP; an heterozygous sample for the CNV was defined by presenting genotype for the three SNPs (the two inside the deletion having to be mandatorily homozygous); the homozygous non deleted samples were characterized by the failure of the breakpoints SNP and presenting genotype at the SNPs inside the CNV region (being either homozygous or heterozygous).

## SUPPLEMENTARY RESULTS

**Table S4**. SNPs showing the strongest allelic associations (p-value <10-4)

| SNP | P-value | Chromosome | Coordinate (Hg18) | Gene/Region |
|---|---|---|---|---|
| **rs12556003** | 2,14x10-6 | X | 138743267 | *MCF2* |
| **rs12704506** | 3,20x10-6 | 7 | 89621311 | *STEAP1/STEAP2* |
| **rs11923054** | 3,52x10-6 | 3 | 167051769 | *ZBBX* |
| **rs2858166** | 3,61x10-6 | X | 100875273 | *ARMCX6* |
| **rs10782344** | 3,63x10-6 | 6 | 156778660 | RP11-518I13.1 |
| **rs1998709** | 7,62x10-6 | 10 | 95884574 | *PLCE1* |
| **rs2194390** | 8,06x10-6 | 2 | 50902931 | *NRXN1* |
| **rs2701106** | 8,94x10-6 | 12 | 114697547 | *TBX5* |
| rs9525923 | 1,11x10-5 | 13 | 44783715 | RP11-478K15.2 |
| rs1347532 | 1,12x10-5 | 16 | 60615455 | RP11-51O6.1 |
| rs12486010 | 1,53x10-5 | 3 | 166942627 | *ZBBX* |
| rs7616572 | 1,61x10-5 | 3 | 167046536 | *ZBBX* |
| rs10894241 | 1,75x10-5 | 11 | 130635852 | AP003486.1 |
| rs11925091 | 2,00x10-5 | 3 | 166944651 | *ZBBX* |
| **rs17512210** | 2,21x10-5 | 17 | 11230466 | *SHISA6* |
| rs5951332 | 2,35x10-5 | X | 100743826 | *ARMCX4* |
| rs7060491 | 2,35x10-5 | X | 100754149 | ARMCX4 |
| **rs9381682** | 2,48x10-5 | 6 | 48620238 | AL391538.1 |
| rs17689185 | 2,50x10-5 | 16 | 77525081 | AC025284.1 |
| **rs11127292** | 2,60x10-5 | 2 | 2029943 | *MYT1L* |
| rs6523526 | 2,64x10-5 | X | 100917910 | *ARMCX2* |
| rs3784820 | 2,76x10-5 | 16 | 1569252 | *IFT140* |
| rs6621083 | 2,81x10-5 | X | 100760626 | OTTHUMG00000022030 |
| rs10432656 | 2,82x10-5 | 2 | 33375032 | OTTHUMG00000152118 |
| rs2071222 | 3,00x10-5 | X | 100617372 | LRG_128 |
| rs11971008 | 3,05x10-5 | 7 | 82136025 | *CACNA2D1* |
| rs858939 | 3,10x10-5 | 2 | 50971951 | *NRXN1* |
| rs11187789 | 3,23x10-5 | 10 | 95871655 | RP11-162K11.4 |
| rs13068321 | 3,46x10-5 | 3 | 167013777 | *ZBBX* |
| rs963618 | 3,79x10-5 | X | 100743037 | *ARMCX4* |
| rs9296606 | 3,91x10-5 | 6 | 48640714 | AL391538.1 |
| **rs12770855** | 4,05x10-5 | 10 | 31120198 | *ZNF438* |
| **rs10821659** | 4,06x10-5 | 10 | 61793424 | *ANK3* |
| **rs265015** | 4,12x10-5 | 4 | 96360796 | *UNC5C* |
| rs882847 | 4,39x10-5 | 17 | 4382729 | *SPNS3* |
| rs10507243 | 4,43x10-5 | 12 | 114708798 | *TBX5* |
| **rs9565180** | 4,53x10-5 | 13 | 76231470 | OTTHUMG00000017093 |
| rs4680657 | 4,61x10-5 | 3 | 166894537 | AC112501.2 |
| rs1994979 | 4,68x10-5 | 17 | 4350990 | SPNS3 |
| rs259154 | 4,72x10-5 | 7 | 89626822 | OTTHUMG00000065036 |
| rs13238853 | 4,73x10-5 | 7 | 135959346 | OTTHUMG00000155618 |
| rs4148965 | 4,75x10-5 | 18 | 9109484 | NDUFV2 |
| **rs6043433** | 5,00x10-5 | 20 | 15659486 | MACROD2 |
| rs6537129 | 5,49x10-5 | 4 | 143779613 | INPP4B |
| rs9299090 | 5,68x10-5 | 9 | 9264932 | AL353733.1 |
| rs5951269 | 5,99x10-5 | X | 100778274 | ARMCX4 |
| rs11869601 | 6,19x10-5 | 17 | 11234035 | SHISA6 |
| rs9410632 | 6,27x10-5 | 9 | 90400909 | CTSL3 |
| rs259152 | 6,42x10-5 | 7 | 89626611 | OTTHUMG00000065036 |
| rs8034595 | 6,43x10-5 | 15 | 96719229 | AC016251.2 |
| rs4986649 | 6,43x10-5 | X | 100736761 | ARMCX4 |
| rs265018 | 6,52x10-5 | 4 | 96362497 | UNC5C |
| rs7022749 | 6,61x10-5 | 9 | 90405495 | CTSL3 |
| rs309853 | 6,62x10-5 | 8 | 29873603 | OTTHUMG00000163815 |
| rs10507833 | 6,97x10-5 | 13 | 76226139 | OTTHUMG00000017093 |

| | | | | |
|---|---|---|---|---|
| rs4910595 | 7,15x10-5 | 11 | 4049129 | *STIM1* |
| **rs11602757** | 7,15x10-5 | 11 | 4053881 | *STIM1* |
| rs2920137 | 7,15x10-5 | 11 | 4079318 | *STIM1* |
| rs6719219 | 7,19x10-5 | 2 | 2010779 | *MYT1L* |
| rs1938204 | 7,24x10-5 | 6 | 48787015 | AL391538.1 |
| rs6083017 | 7,28x10-5 | 20 | 23119766 | RP4-737E23.4 |
| rs12588013 | 7,29x10-5 | 14 | 62724837 | AL390816.1 |
| rs9643612 | 7,39x10-5 | 8 | 50430756 | RP11-738G5.1 |
| rs2065703 | 7,53x10-5 | 20 | 31966698 | *CDK5RAP1* |
| rs7314743 | 7,55x10-5 | 12 | 114718647 | *TBX5* |
| rs2009626 | 7,84x10-5 | 3 | 187600404 | RP11-44H4.1 |
| **rs981524** | 7,94x10-5 | 14 | 33186257 | *AKAP6* |
| rs6778044 | 8,16x10-5 | 3 | 187594092 | RP11-44H4.1 |
| rs5951340 | 8,24x10-5 | X | 100771055 | *ARMCX4* |
| rs1323851 | 8,36x10-5 | 1 | 64450437 | *ROR1* |
| rs12744386 | 8,80x10-5 | 1 | 24168019 | *HMGCL* |
| rs6966421 | 9,18x10-5 | 7 | 155329530 | *CNPY1* |
| rs6556373 | 9,25x10-5 | 5 | 158359476 | *EBF1* |
| rs2997370 | 9,42x10-5 | 6 | 48778106 | AL391538.1 |
| rs10184672 | 9,42x10-5 | 2 | 11198448 | AC062028.1 |
| rs2009627 | 9,68x10-5 | 3 | 187600359 | RP11-44H4.1 |
| rs5991939 | 9,73x10-5 | X | 100712422 | *ARMCX4* |

SNPs are listed on descending order based on p-value. SNPs selected for replication appear in bold.

**TableS5**: Analysis of copy number variants (CNV) in pooled samples of fibromyalgia

| Chromosome | Cytoband | Gene | Start | End | Probes | Log2ratio_Direct | Log2ratio_DS | Pools |
|---|---|---|---|---|---|---|---|---|
| 4 | q34.1 | *GALNTL6* | 173661791 | 173666272 | 8 | -1.00 | 1.20 | FM_3 |
| 7 | q36.3 | *WDR60* | 158400565 | 158402804 | 9 | 0.48 | -0.32 | FM_3 |
| 8 | p21.2 | *DOCK5* | 25122432 | 25126488 | 6 | -1.58 | 1.21 | FM_2 |
| 9 | p23 | *PTPRD* | 10394403 | 10395130 | 3 | -2.22 | 1.83 | FM_2 |
| 14 | q31.1 | *NRXN3* | 79175885 | 79184422 | 18 | -0.54 | 0.58 | FM_3 |
| 16 | q23.1 | *WWOX* | 76929398 | 76941774 | 11 | -0.46 | 0.54 | FM_3 |
| 20 | p13 | *SIRPB1* | 1511432 | 1531941 | 28 | -0.41 | 0.47 | FM_2 |

Results are based on hybridization with the 400k-CNV aCGH; genomic locations are based on build Hg 18. DS: Dye swap.

**SUPPLEMENTARY FIGURES LEGENDS:**

**Supplementary figure 1**: Workflow indicating which sample sets were used in each analysis.

**Supplementary figure 2:** Power calculation graph. For each OR, the given power for three different minor allele frequency variants is calculated (given our sample size). OR are represented in the X axis, and power is represented in the Y axis.

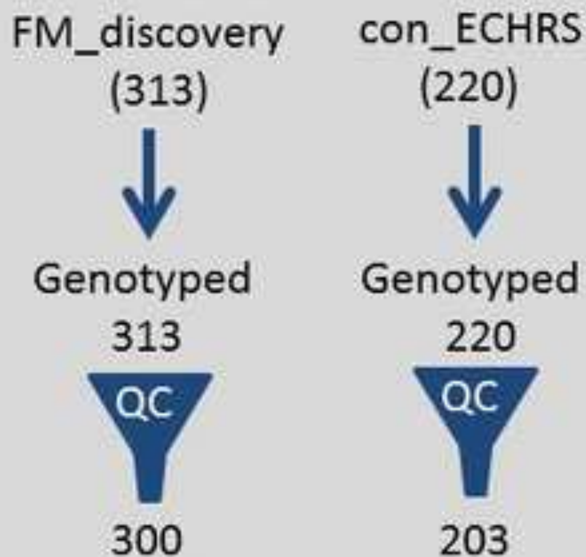
**Supplementary figure 3:** *SNTG2* expression levels (as assessed by one probe) for the three rs11127292 genotypes in twins' lymphoblastyoid cell lines (GENVAR software).

Supplementary Materials: figures, tables

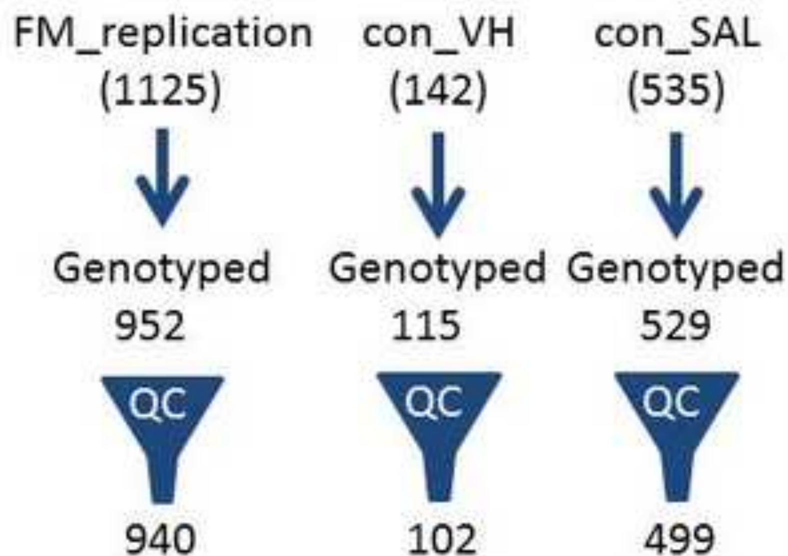
Click here to download high resolution image

SNPs

**GWAS**

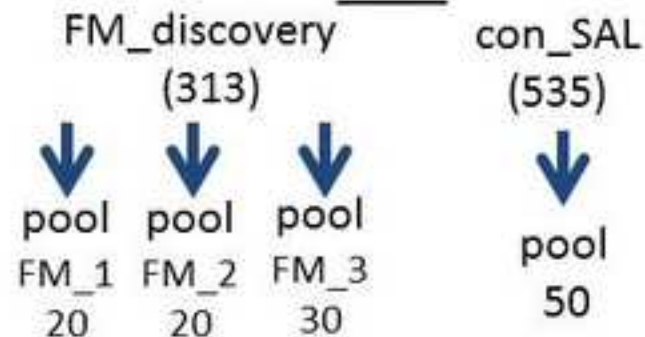



**Replication**



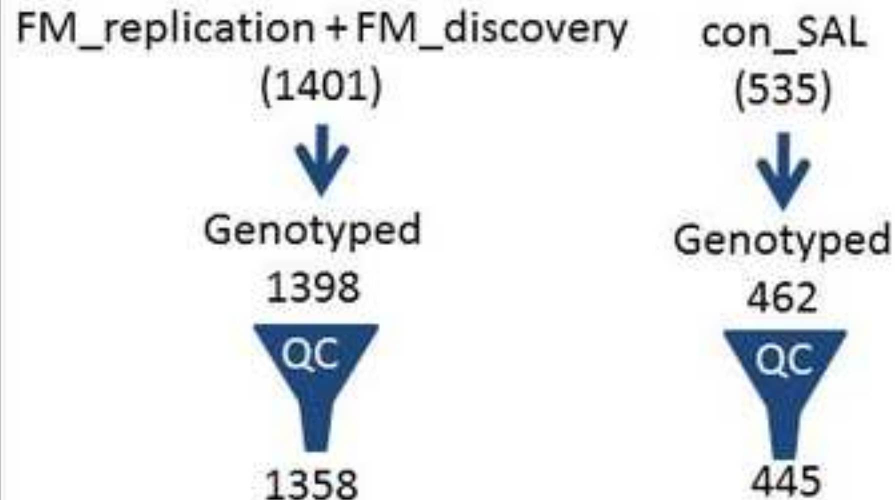

CNVs

**aCGH**

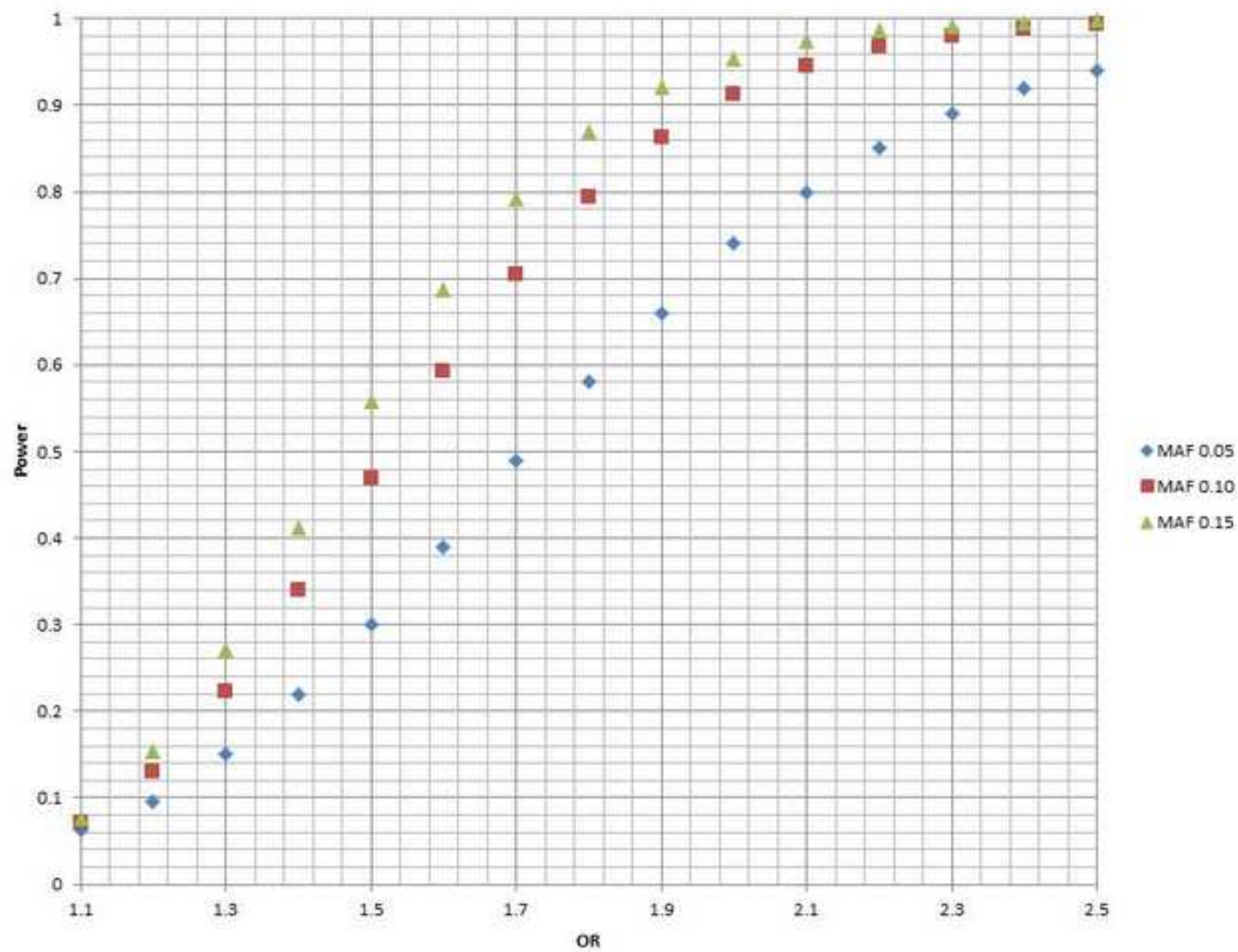



**Validation (all candidates)**

**Replication (NRXN3)**



Supplementary Materials: figures, tables
Click here to download high resolution image



Supplementary Materials: figures, tables
Click here to download high resolution image