View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by UCL Discovery

Effects of multiple genetic loci on age at onset in frontotemporal dementia

Raffaele Ferrari*[a], Mario Grassi*[b], Francesca Graziano [b], Fernando Palluzzi [b], Silvana Archetti [c], Elisa Bonomi[d], Amalia C. Bruni[e], Raffaele G. Maletta [e], Livia Bernardi [e], Chiara Cupidi [e], Rosanna Colao [e], Innocenzo Rainero [f], Elisa Rubino [f], Lorenzo Pinessi [f], Daniela Galimberti [g], Elio Scarpini [g], Maria Serpente [g], Benedetta Nacmias [h], Irene Piaceri [h], Silvia Bagnoli [h], Giacomina Rossi [i], Giorgio Giaccone [i], Fabrizio Tagliavini [i], Luisa Benussi [l], Giuliano Binetti [l,m], Roberta Ghidoni [l], Andrew Singleton [n], John Hardy [o], Parastoo Momeni [p], Alessandro Padovani [d], and Barbara Borroni [d]

* These authors equally contributed to this work.

[a] Department of Molecular Neuroscience, Institute of Neurology, UCL, London, UK

[b] Department of Brain and Behavioural Sciences, Medical and Genomic Statistics Unit, University of Pavia, Pavia, Italy

[c] Department of Laboratories, III Laboratory of Analysis, Brescia Hospital, Brescia, Italy

[d] Neurology Unit, Department of Clinical and Experimental Science, University of Brescia, Italy

[e] Neurogenetic Regional Centre ASPCZ Lamezia Terme, Lamezia Terme, Italy

[f] Neurology I, Department of Neuroscience, University of Torino and Città della Salute e della Scienza di Torino, Turin, Italy

[g] Neurology Unit, Department of Pathophysiology and Transplantation, University of Milan, Fondazione Cà Granda, IRCCS Ospedale Policlinico, Milan, Italy

[h] Department of Neuroscience, Psychology, Drug Research and Child Health, University of Florence, Florence, Italy

[i] Division of Neurology V and Neuropathology, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milano, Italy

[l] Molecular Markers Laboratory, IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy

[m] MAC Memory Center, IRCCS Istituto Centro San Giovanni di Dio-Fatebenefratelli, Brescia, Italy

[n] Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD,

USA

[o] Department of Molecular Neuroscience, Institute of Neurology, UCL, London, UK

[p] Laboratory of Neurogenetics, Department of Internal Medicine, Texas Tech University Health Science Center, Lubbock, TX, USA

The authors have no conflict of interests to disclose.

Word counts for the text: 2230

Corresponding author:

Barbara Borroni, MD

Neurology Unit, University of Brescia

Piazza Spedali Civili 1

25100, Brescia, Italy

Email: bborroni@inwind.it

ABSTRACT

In Frontotemporal Dementia (FTD) age at disease onset (AAO) is unpredictable in both early and late-onset cases; AAO variability is found even in autosomal dominant FTD.

The present study was aimed at identifying genetic modifiers modulating AAO in a large cohort of Italian FTD patients.

We conducted an association analysis on 411 FTD patients, belonging to 7 Italian Centers, and for whom AAO was available. Population structure was evaluated by principal component analysis (PCA) to infer continuous axes of genetic variation, and single linear regression models were applied. A genetic Score (GS) was calculated on the basis of suggestive Single Nucleotide Polymorphisms (SNPs) found by association analyses.

GS showed genome-wide significant slope decrease by -3.86 (95%CI: -4.64 to -3.07, $p<2x10^{-16}$) per standard deviation of the GS for 6 SNPs mapping to genes involved in neuronal development and signaling, axonal myelinization and glutamatergic/GABA neurotransmission. An increase of the GS was associated with a decrease of the AAO.

Our data indicate that there is indeed a genetic component that underpins and modulates up to 14.5% of variability of AAO in Italian FTD. Future studies on genetic modifiers in FTD are warranted.

Key words: GWAS; age at onset; Frontotemporal Dementia; polymorphism

INTRODUCTION

Frontotemporal Dementia (FTD) is a neurodegenerative disorder characterized by behavioural abnormalities, impairment of executive functions and language deficits [1]. Mutations in the Granulin *(GRN)*, *C9orf72* or *Microtuble Associated Protein Tau (MAPT)*, and, recently, the TANK binding kinase 1 (*TBK1*) genes drive up to ~40% of Mendelian cases [2]; for the remaining cases, multiple loci and genes appear to influence FTD risk with rather small effect size [3,4].

The clinical, pathological and genetic profiles, as well as the age at onset (AAO), are highly heterogeneous in FTD. FTD is considered the most common neurodegenerative dementia in the young adulthood along with Alzheimer Disease with a mean onset age between late 50s and early 60s [5]; however, recent reports suggested an increased span in the range of disease onset up to >70 years of age [6,7]. Similarly, variability of AAO is found in autosomal dominant FTD [8-11].

Genetic variants may influence and/or modulate specific features of a disorder, including disease onset and progression. Indeed, AAO, which is rather unpredictable in FTD, might associate with a genetic structure that contributes to an earlier or later disease development.

We recently carried out a Genome-Wide Association Study (GWAS) in FTD patients from Italy, and identified an extended number of genetic risk factors associated with disease in the Italian population [3]. In the present work, we sought to explore whether such risk factors might influence disease onset in the Italian FTD population by searching for genetic variants associated with age at disease onset, and computing a genetic risk score (GS) to stratify age of disease susceptibility.

METHODS

Subjects. Genotyping data of DNA samples diagnosed with FTD were available to us from the FTD-GWAS dataset (Ferrari et al., 2014). Our association analysis was carried out on 411 FTD patients, belonging to 7 Italian Centres (see Supporting **Table S1**), already included in our previous study [3] and for whom AAO was available. Diagnosis of FTD, and FTD subtypes (i.e., behavioral variant FTD and Primary Progressive Aphasias), was made according to current clinical criteria [12,13]. Estimated AAO was referred by the proxy caregiver by a semi-structured interview, and carefully recorded.
Informed consent and Ethic Committee approval were obtained at the individual sites and each centre provided consent for their use for the purposes of this study. The cases were collected and genotyped at the University College London (UCL) by means of Illumina human 660K-Quad Beadchips assayed on the Illumina Infinium platform (Illumina, San Diego, CA, USA); 657231 single nucleotide polymorphisms (SNPs) were genotyped.

Quality Control. Quality control (QC) was performed at both markers (per-marker) and samples (per-individual) level, as previously illustrated by Anderson et al [14]. Specifically, per-marker QC of the data consisted of three different steps, namely 1) the identification and removal of SNPs with an excessive missing genotype (>0.1); 2) the identification and removal of SNPs with significant deviation from Hardy-Weinberg equilibrium (HWE) (p<= 0.001); and 3) the removal of all makers with a very low minor allele frequency (MAF) (MAF< 0.05). Per-individual QC of the data consisted of identification and removal of 1) the individuals with discordant sex information; 2) the individuals with outlying missing genotype (genotype failure rate >= 0.03); and 3) heterozygosity rate ± 3 standard deviations from the mean.

Principal Component Analysis. Population structure was evaluated by principal component analysis (PCA) to infer continuous axes of genetic variation. PCA was processed on the similarity matrix (*n x n*) between any two individuals (*i, j),* based on the genome-wide average proportion of alleles sharing identical by state (IBS) and using for the SNPs the additive coding (0,1, 2 minor frequency alleles). We excluded outlier

samples, defined as individuals exceeding a default number of standard deviations (6.0) from the whole sample. The 657,231 SNPs were reduced to 125,902 SNPs for IBS computation, by trimming the dataset so that no pair of SNPs (within a given number of base-pairs) had Linkage Disequilibrium (LD, measured by r2) greater than 0.2. PC axes were considered significant according to Tracy-Widom statistic [15].

Linear Regression Model. We used markers of highest quality to impute with Minimac [16] approximately 2.5 million SNPs, based on European-ancestry haplotype reference (HapMap II CEU population build 36, release 22). Imputed SNPs with a MAF< 0.01, and with low imputation quality (Rsq< 0.50) were removed (for Minimac protocol details, see http://genome.sph.umich.edu/wiki/Minimac:_1000_Genomes_Imputation_Cookbook). We then tested each available SNP (observed and imputed) to assess the association with AAO by using p-values (two-tailed) from a linear regression model (LRM), with an additive coding per SNP, and adjusting for sex and significant principal components. The fixed threshold for genome-wide significance was set at the consensus level of $p<5\times10^{-8}$. We defined SNPs as suggestive if $5\times10^{-5}> p> 5\times10^{-8}$. We performed this analysis in presence and absence of apolipoprotein E (*APOE*) ε4 homozygous carriers and also used *APOE* ε4 as a covariate to assess and/or correct for any potential influence of this genetic marker on final results.

Genetic Score computation. The suggestive SNPs obtained by association analysis were used to calculate a weighted genetic risk score (GS). Specifically, GS was defined as sum of the number of minor frequency alleles (0 (not present), 1 (heterozygous), 2 (homozygous)) per suggestive SNPs weighted by their log OR, i.e. the odds ratio (OR) estimate derived from our previous case-control study [4]. Thus, the weights were determined independently from the genetic-AAO association of this work, eluding over fitting biases. We defined three trimming rules for inclusion of SNPs in the GS: 1) SNPs with moderate to high case-control signal (logORs >log(1.2)), recoding all negative logOR into positive logOR on the basis of the risk allele; 2) SNPs observed or imputed with high quality (Rsq>0.9); and 3) SNPs in Linkage Equilibrium (LE). For SNPs in Linkage Disequilibrium, the SNP with the lowest AAO p-value was selected.

Then, GS was standardized as Z=(GS-mean(GS))/standard deviation (GS), and LRM of AAO on the standardized GS was performed adjusting for the following covariates: i) sex, ii) population structure (measured by geographical zone such as North, Centre, South and/or by significant principal components), iii) type of FTD (5 subtypes: bvFTD, SD, PNFA, FTD-MND and FTD spectrum), and iv) APOE (0=e23e23, 1=e23e4, 2=e4e4).

QC analysis and statistical analyses were performed by free software PLINK [17] and R [18], respectively. Gene annotation on top-associated markers was performed by R package NCBI2R [19].

RESULTS

AAO was available for 411 out of 532 FTD patients that underwent Italian GWAS (Ferrari et al., 2015). The mean AAO was 63.7 ± 20.7 (range, 36.0 – 82.0) with male/female ratio 187/224 (male 45.5%). Two hundred fifty patients came from Northern Italy, 49 patients from the Centre and 112 from Southern Italy.

Out of 411 FTD patients, 307 had bvFTD, 82 PPA, 20 FTD-MND and 2 belonged to the FTD spectrum. A minority of cases, for which FTD genes had been screened, carried variants in *MAPT* (n=4), *GRN* (n=38) and *C9orf72* (n=11). Two *MAPT* and one *GRN* variants are now reported to be pathogenic, whilst the remaining ones are common polymorphisms; considering that the mutations have very small frequency in the current sample set, thus they are negligible, and that, together with the *C9orf72* positive cases they were kept in the previous studies (Ferrari et al., 2014; Ferrari et al., 2015), we also kept and included them in our current analysis.

Population structure through PCA showed a genetic pattern that correlated with geographical provenience (i.e., North, Centre and South), as displayed in the scatter plot of the first two PCs (**Fig. 1**).

According to Tracy-Widom statistics, tests of additional population structure of the first 10 eigenvalues, only the first PC (TW1=5.03480, P<0.001) was deemed significant and used as covariate in AAO association analyses to correct for population stratification.

By single linear regression analysis, we identified suggestive evidence ($p<5x10^{-5}$) of an association with AAO for 127 SNPs (using sex and first PC as covariates), while none achieved genome-wide significance (see Manhattan plot in **Fig. 2**).

As defined in the method section, log OR weighted genetic score (GS) was computed after three trimming rules identifying 6 SNPs mapping to the following loci: intergenic region on chromosome (chr) 1 encompassing the *G protein-coupled receptor 137B* (*GPR137B*) and the *Endoplasmic Reticulum Oxidoreductase Beta (ERO1B*) genes; chr 2, intronic to the *Hypocalcin-like 1 (HPCAL1*) gene and *LOC730100* (which is proximal to Neurexin 1 (*NRXN1*)); chr 9, intronic to the *Protein Tyrosine Phosphatase, Receptor Type D (PTPRD*) gene, and; chr 13, in *LOC101926897*, *LOC105370219* and *LOC105370220*, and *LOC105370290* and *LINC00375* (**Table 1**).

Linear regression analysis of AAO based on the score effect of these 6 markers (GS), adjusted for covariates (see method), showed genome-wide significant slope decrease by -3.68 (95%CI: -4.52 to -2.84, $p<2x10^{-16}$) per standard deviation of the GS meaning that an increase of the GS was associated with a decrease of the AAO. When we excluded *APOE* ε4 homozygous carriers from our analysis still was a similar genome-wide significant slope decrease, and when using *APOE* ε4 as a covariate results were again not affected indicating that effect on AAO is independent from the *APOE* status. Of note, considering all the predictors (GS + covariates) we found that total contribution (R^2) to AAO variance was ~22% where only gender and PC1 were the actual significant covariates contributing to AAO variance in addition to GS of ~ 7%.
Therefore and in summary, the AAO variance explained by GS only (excluding the covariates) reached ~14.5%. The AAO-GS association is visualized and summarized in **Fig. 3 and Table 2**, respectively. For overfitting control, we also computed the prediction error (PE) by Leave-One-Out method (K=n), and K=10-fold Cross-Validation: the corresponding R^2 were 13.7%, and 14.2%, respectively. Finally, backward stepwise analyses in the GS showed that 7.8% of this signal was driven by the AAO-associated SNPs in *HPCAL1* and *PTPRD*.

DISCUSSION

FTD is a complex disorder that encompasses heterogeneous phenotypes and variable AAO. Our current study supports the notion that individual genetic background modifies age at disease onset and modulates disease presentation.

We computed a genetic score (GS) able to weight the genetic contribution to the variability of AAO in FTD (the higher the GS, the lower the age at disease onset). Particularly, the GS of six SNPs indicated significant genetic contribution for up to 14% of variability of AAO in FTD. This is remarkable as it suggests that the genetic component underlying variability of AAO has a robust implication in FTD comparatively, for example, to Parkinson's disease (PD) for which a similar study approach revealed that their GS only contributed to 0.6% of variability of AAO [20]. Similar observations for PD were made in other studies [21]. Conversely, a recent study in Alzheimer's disease (AD) showed that either *APOE* or a combination of nine loci contributed to 3.7% and 2.2% of variability of AAO, respectively [22], whereas in our current study the contribution of the *APOE* status was basically negligible. Taken all this together, we gather that in FTD the genetic component appears to have a stronger influence on AAO prediction than in PD and in AD [20,21].

Considering the 6 loci that we identified influencing AAO in FTD, the literature suggests, particularly for *NRXN, HPCAL1*, and *PTPRD*, a direct link with the biology of the brain through calcium-mediated neuronal signaling, development of neocortical regions as well as axonal myelination and glutamatergic/GABAergic synapsis, respectively [23-26]. It is interesting to gather that variability in these genes, and their associated functions and processes, could underpin the modulation of disease onset in FTD. In addition, it is noteworthy that *GPR137B* was previously shown to associate with modulation of drug-response in schizophrenia, particularly mediating effect of antipsychotics on working memory [27].

NRXN1 has been implicated in neocortical development and aging [23], and *NRXN1* genetic variations affect the risk of autism and schizophrenia [28,29]. Of note, we reported the locus including *NRXN1* already in our previous work highlighting novel suggestive loci for FTD in the Italian population (Ferrari et al., 2015); here, we provide further support for this association and expand on its potential implication in driving AAO. All the more, in the previous [3] and the current work, the OR associated with SNPs at the *NRXN1* locus were

among the highest, 2.5 and 1.64, respectively, suggesting that variability at this locus might associate with disease as well as AAO with a moderate effect size. Moreover, a close relationship between NRXN1 and one of the main actors of FTD pathogenesis, namely TDP-43 protein, was recently shown: a novel miRNA (miR-NID1) processed from the intron 5 of *NRXN1* is able to repress NRXN1 expression by binding to TDP-43, thus this complex (TDP-43 + miR-NID1) might hold relevance and be implicated in FTD-TDP pathology [30]. In the same view, *HPCAL1*, a gene encoding the HPCAL1 protein that modulates calcium-mediated neuronal signalling [25], has been implicated in the risk of Alzheimer disease and autism development [31,32], thus supporting a general involvement in neurodegenerative disease, including FTD, and *PTPRD* has been associated with obsessive-compulsive disorders [26]. Currently, there are no known relationships with conditions affecting the brain for *ERO1B* (involved in protein oxidative folding) [33], whilst the role of *LOC101926897*, *LOC105370219* and *LOC105370220*, and *LOC105370290* and *LINC00375* appear to be more complicated to infer, given that the latter are non-coding RNAs, which belong to a different group of cellular functional elements, rather involved in transcription and/or translation regulation. Nevertheless, this piece of data raises the awareness that non-coding functional elements are important and deserve attention, fostering the implementation of study designs and analysis strategies that focus and shed light on such regulatory elements.

Our study overall suggest and supports the relevance of factors that modulate features of FTD; this goes along with other examples of common risk variants acting as genetic modifiers in Mendelian FTD such as *TMEM106B*. The latter has been reported as a genetic modifier affecting both AAO and circulating levels of progranulin in FTD patients carrying *Granulin* mutations [34,35]. Another example of genetic modifiers of AAO is the one recently reported in *PSEN1*-related-AD [36].

In summary, we herein describe an additive effect of six SNPs and their likely implication in modulating disease onset in Italian FTD patients, for which we further considered the genetic make-up of Italian regions (i.e. North, Centre and South) as possible confounder in the analyses, when a Mendelian trait was excluded *a priori*.

Our data indicate that there is a genetic component that underpins and modulates up to 14.5% of variability of AAO in Italian FTD.

Acknowledgements

This work was supported by NeTAlz grant to A.P., by Fondazione Cassa di Risparmio di Pistoia e Pescia (grants 2011.0264 e 2013.0347), a grant for Research from the University of Florence to B.N., the Ricerca Corrente, Italy Ministry of Health to L.B., R.G., G.B., and Alzheimer's Society (grant number 284) to R.F. The authors would also like to thank Dr. Erika Salvi and Professor Daniele Cusi from the Department of Health Sciences, University of Milan at San Paolo Hospital, Milan, Italy and Hypergenes - European Network for Genetic-Epidemiological Studies (http://www.hypergenes.eu/) for previously providing control data that helped generating, filtering and scoring the current dataset.

REFERENCES

[1] van der Zee J, Van Broeckhoven C. Dementia in 2013: frontotemporal lobar degeneration-building on breakthroughs. Nat Rev Neurol 2014;10:70-72.

[2] Pottier C, Ravenscroft TA, Sanchez-Contreras M, et al. Genetics of FTLD: Overview and what else we can expect from genetic studies. J Neurochem. 2016;Mar 24 doi: 10.1111/jnc.13622 in press.

[3] Ferrari R, Grassi, M, Salvi E, et al. A genome-wide screening and SNPs-to-genes approach to identify novel genetic risk factors associated with frontotemporal dementia. Neurobiol Aging 2015;36:13-26.

[4] Ferrari R, Hernandez DG, Nalls MA, et al. Frontotemporal dementia and its subtypes: a genome-wide association study. Lancet Neurol 2014;13:686-699.

[5] Lashley T, Rohrer JD, Mead S, et al. An update on clinical, genetic and pathological aspects of frontotemporal lobar degenerations. Neuropathology Appl Neurobiol 2015;41:858–881.

[6] Borroni B, Alberici A, Grassi M, et al. Prevalence and demographic features of early-onset neurodegenerative dementia in Brescia County, Italy. Alzheimer Dis Assoc Disord 2011;25:341-344.

[7] Jellinger KA. Elderly individuals with FTLD. JAMA Neurol 2013;70:412-413.

[8] Borroni B, Padovani A. Dementia: a new algorithm for molecular diagnostics in FTLD. Nat Rev Neurol 2013;9:241-242.

[9] Benussi L, Ghidoni R, Pegoiani E, et al. Progranulin Leu271LeufsX10 is one of the most common FTLD and CBS associated mutations worldwide. Neurobiol Dis 2009;33:379-385.

[10] Benussi A, Padovani A, Borroni, B. Phenotypic heterogeneity of monogenic Frontotemporal dementia. Front Aging Neurosci 2015;7:e171.

[11] Bernardi L, Frangipane F, Smirne N, et al. Epidemiology and genetics of frontotemporal dementia: a door-to-door survey in southern Italy. Neurobiol Aging 2012;33:2948.e1-2948.e10.

[12] Gorno-Tempini ML, Hillis AE, Weintraub S, et al. Classification of primary progressive aphasia and its variants. Neurology 2011;76:1006-1014.

[13] Rascovsky K, Hodges JR, Knopman D, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain 2011;134:2456-2477.

[14] Anderson, C.A., Pettersson, F.H., Clarke, G.M., et al. Data quality control in genetic case-control association studies. Nat Protoc 2010;5:1564-1573.

[15] Limpiti T, Intarapanich A, Assawamakin A, et al. Study of large and highly stratified population datasets by combining iterative pruning principal component analysis and structure. BMC Bioinformatics 2011;12: e255.

[16] Fuchsberger C, Abecasis GR, Hinds, D.A. Minimac2: faster genotype imputation. Bioinformatics 2015;31:782-784.

[17] Purcell S, Neale B, Todd-Brown K, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am J Hum Genet 2007;81:559-575.

[18] R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria 2015. URL http://www.R-project.org

[19] Melville S. NCBI2R: NCBI2R-An R package to navigate and annotate genes and SNPs. R package 2014 version 1.4.6.

[20] Lill CM, Hansen J, Olsen JH, et al. Impact of Parkinson's disease risk loci on age at onset. Mov Disor 2015;30:847-850.

[21] Davis AA, Andruska KM, Benitez BA, et al. Variants in GBA, SNCA, and MAPT influence Parkinson disease risk, age at onset, and progression. Neurobiol Aging 2016;209:e1-7.

[22] Naj AC, Jun G, Reitz C, et al. Effects of multiple genetic loci on age at onset in late-onset Alzheimer disease: a genome-wide association study. JAMA Neurol 2014;71:1394-1404.

[23] Jenkins AK, Paterson C, Wang Y, et al. Neurexin 1 (NRXN1) splice isoform expression during human neocortical development and aging. Mol Psychiatry 2015;21:701-716.

[24] Mattheisen M, Samuels JF, Wang Y, et al. Genome-wide association study in obsessive-compulsive disorder: results from the OCGAS. Mol Psychiatry 2015;20:337-344.

[25] Sharp BM, Chen H, Gong S, et al. Gene expression in accumbens GABA neurons from inbred rats with different drug-taking behavior. Genes Brain Behav 2011;10:778-788.

[26] Zhu Q, Tan Z, Zhao S, et al. Developmental expression and function analysis of protein tyrosine phosphatase receptor type D in oligodendrocyte myelination. Neuroscience 2015;308:106-114.

[27] McClay JL, Adkins DE, Aberg K, et al. Genome-wide pharmacogenomic study of neurocognition as an indicator of antipsychotic treatment response in schizophrenia. Neuropsychopharmacology 2011;36:616-626.

[28] Jiang YH, Yuen RK, Jin X, et al. Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing. Am J Hum Genet 2013;93:249-263.

[29] Todarello G, Feng N, Kolachana BS, et al. Incomplete penetrance of NRXN1 deletions in families with schizophrenia. Schizophr Res 2014;155:1-7.

[30] Fan Z, Chen X, Chen R. Transcriptome-wide analysis of TDP-43 binding small RNAs identifies miR-NID1 (miR-8485), a novel miRNA that represses *NRXN1* expression. Genomics 2014;103:76-82.

[31] Egger G, Roetzer KM, Noor A, et al. Identification of risk genes for autism spectrum disorder through copy number variation analysis in Austrian families. Neurogenetics 2014;15:117-127.

[32] Lee JH, et al., 2011. Identification of novel loci for Alzheimer disease and replication of CLU, PICALM, and BIN1 in Caribbean Hispanic individuals. Arch Neurol 68, 320-328.

[33] Sevier CS. New insights into oxidative folding. J Cell Biol 2010;188:757-758.

[34] Premi E, Formenti A, Gazzina S, et al. Effect of TMEM106B polymorphism on functional network connectivity in asymptomatic GRN mutation carriers. JAMA Neurol 2014;71:216-221.

[35] Cruchaga C, Graff C, Chiang HH, et al. Association of TMEM106B gene polymorphism with age at onset in granulin mutation carriers and plasma granulin protein levels. Arch Neurol 2011;68:581-586.

[36] Lee JH, Cheng R, Vardarajan B, et al. Genetic Modifiers of Age at Onset in Carriers of the G206A Mutation in PSEN1 With Familial Alzheimer Disease Among Caribbean Hispanics. JAMA Neurol 2015;72:1043-1051.

Table 1. SNPs and related genes after the weighted Genetic Score (GS).

SNP	Chr	Allele change	Gene/nearby gene(s)	P-value	OR	Brain function	Brain-related Disease
rs10924761	chr1	C>T	**GPR137B;** **ERO1B**	1.81E-05	1.21	G protein-coupled receptor superfamily, known to be expressed in brain Oxidative folding (30)	Schizophrenia; drug-response (24) NA
rs3821200	chr2	C>T	**HPCAL1**	1.43E-05	1.27	GABA neuronal signaling (22)	Autism (28), Alzheimer Disease (29), addiction (22)
rs17863722	chr2	C>G	**LOC730100;** **NRXN1**	1.14E-05	1.64	NA Neocortical development and aging (20)	NA Autism (25), schizophrenia (26)
rs904780	chr9	A>T	**PTPRD**	1.29E-05	1.25	Axonal myelination (21), development of glutamatergic and GABAergic synapsis (23)	Obsessive –compulsive disorder (23)
rs9527836	chr13	C>T	**LOC101926897;** **LOC105370219;** **LOC105370220**	7.51E-06	1.23	NA NA NA	NA NA NA
rs9602613	chr13	C>T	**LOC105370290** **LINC00375**	2.42E-05	1.21	NA NA	NA NA

Chr: chromosome; p-value: age-onset p-value, OR: FTD odds ratio; NA: not available. Related reference between brackets.

Table 2. Mean (SD) of age at onset across Genetic score classes, after recoding the continuous Genetic Score into categorical variable.

GS category	**n**	**Mean (SD) of age at onset**
Low (min to -1SD)	114	67.6 (8.40)
Medium low (-1SD to 0 =mean)	154	63.8 (7.90)
Medium high (0=mean to +1SD)	104	62.2 (7.69)
High (+1SD to max)	39	54.9 (11.0)

SD: standard deviation

Legend to figures:

Fig. 1. Principal component analysis plot of the first two principal components of FTD patients (labeled by geographic site) derived from their SNP genotypes.

Fig. 2. Manhattan plots of genome-wide association (GWAS) analyses of age at onset. Each dot represents a single nucleotide polymorphism (SNP: black= typed, red=imputed), with the horizontal axis showing the genomic position (hg18) and the vertical axis showing the –log10(p value) for age at onset derived by linear regression model. The horizontal line indicates the less stringent threshold ($p<5*10-5$)

Fig. 3. Scatter plot of the standardized genetic score (GS) and age at onset (AAO) of FTD patients. Simple regression slope of AAO on GS and correlation coefficient (r) are also shown.

Figure 1.

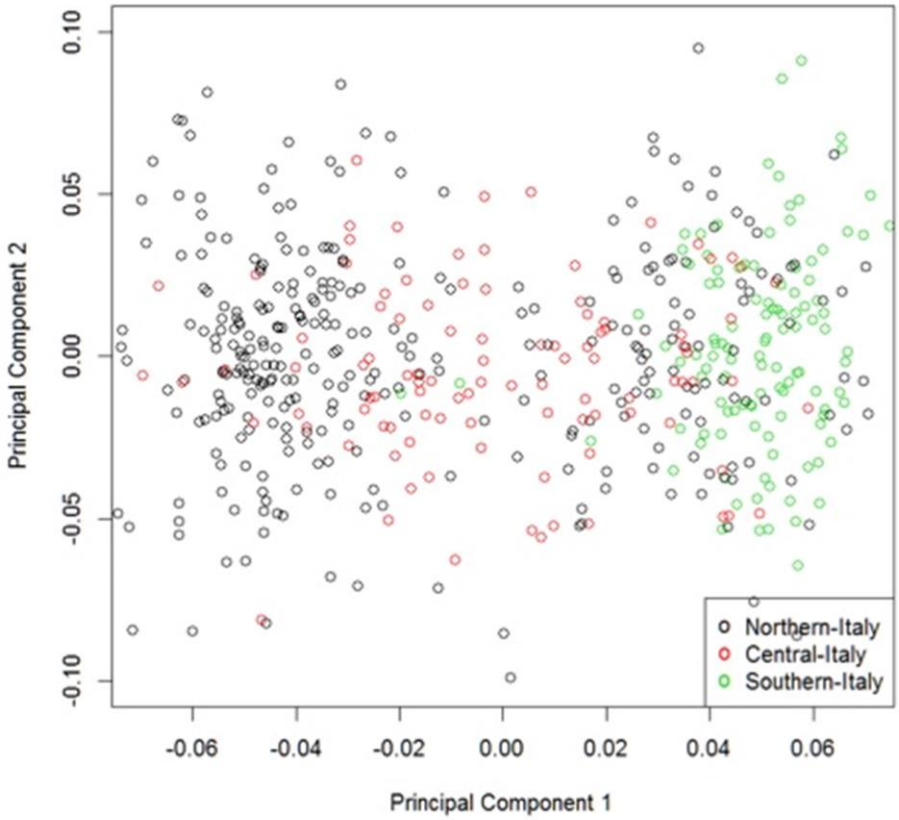

Principal Component 2
0.10
0.05
0.00
-0.05
-0.10
-0.06
-0.04
-0.02
0.00
0.02
0.04
0.06
Principal Component 1
Northern-Italy
Central-Italy
Southern-Italy

Figure 2.

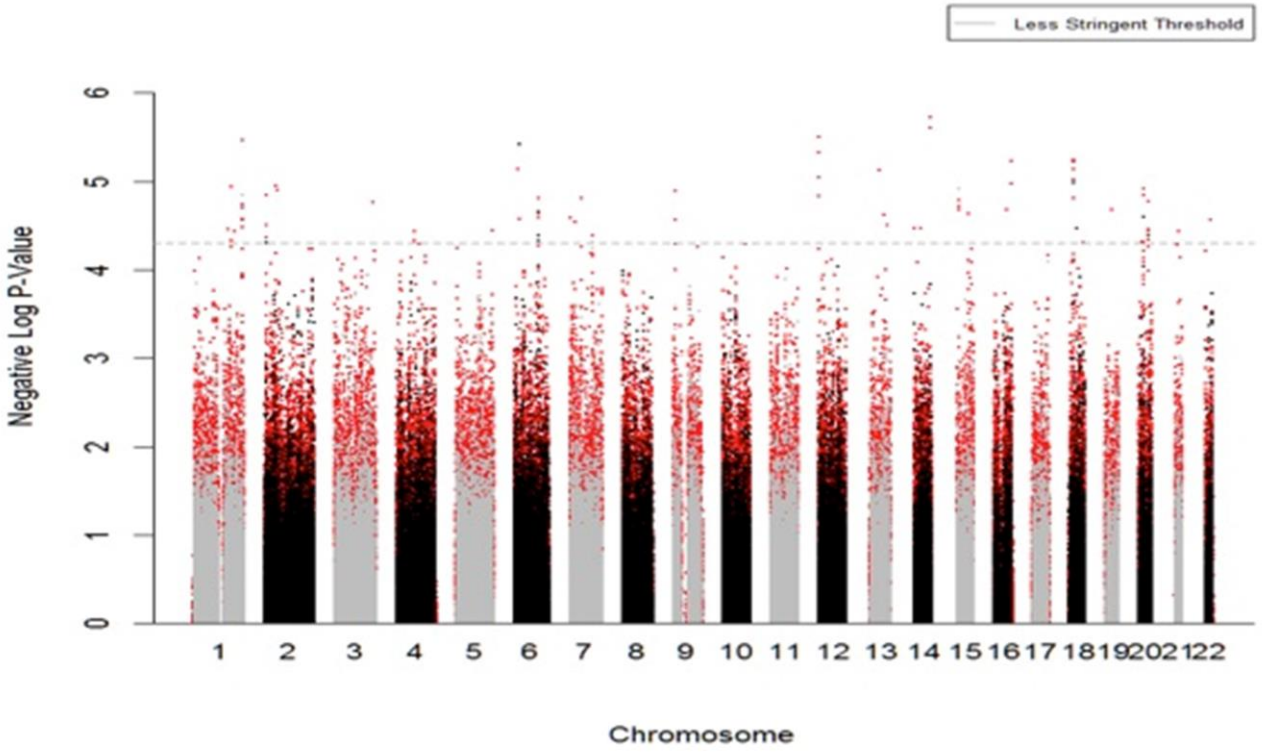
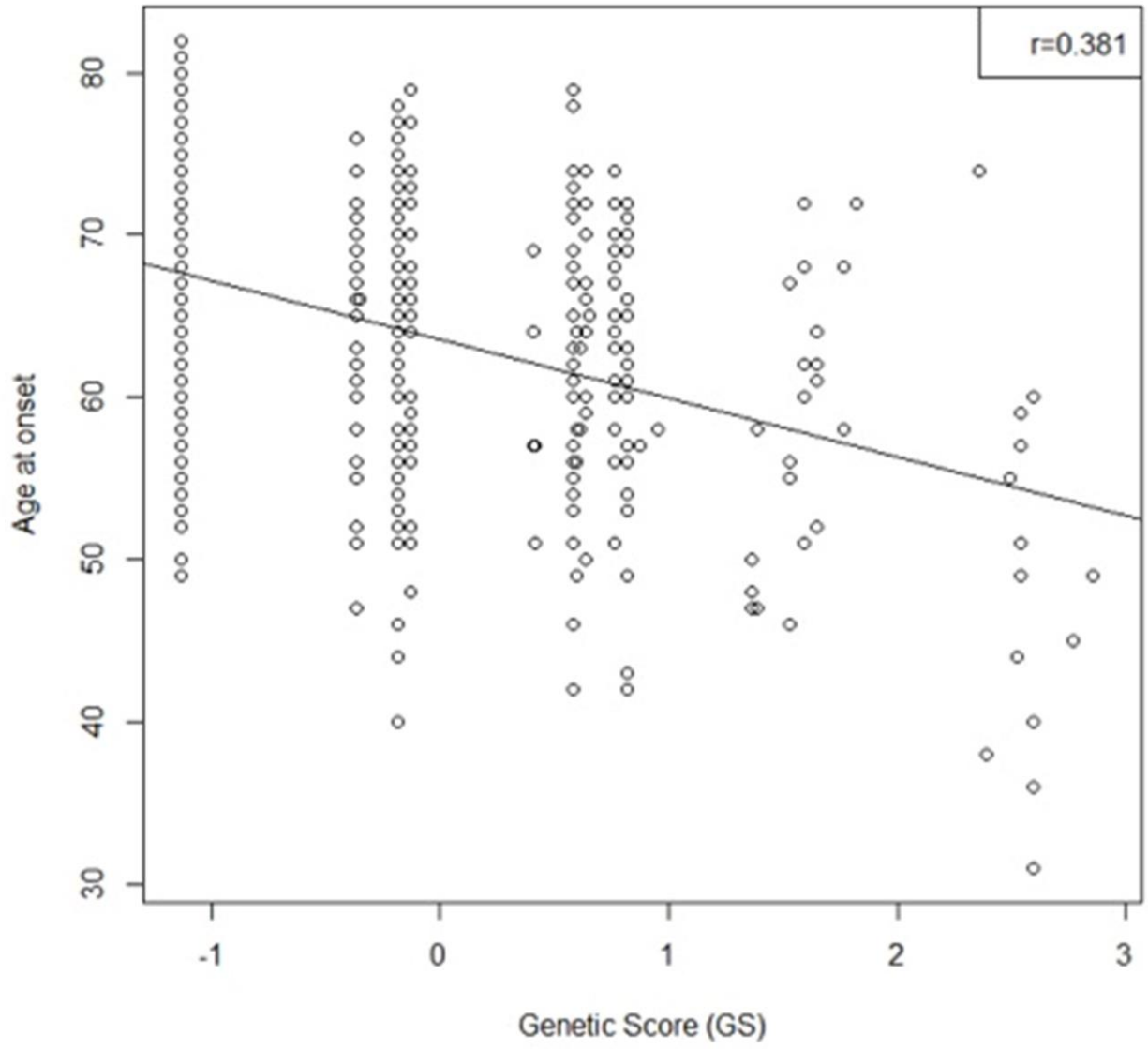

Less Stringent Threshold
Negative Log P-Value
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
Chromosome

Figure 3.



Supporting Table S1. Italian Centres and number of FTD patients

Centre	**Principal Investigator**	**number**
University of Brescia, Brescia	Barbara Borroni	77
IRCCS Fatebenefratelli, Brescia	Luisa Benussi	12
University of Turin, Turin	Innocenzo Rainiero	72
IRCCS Policlinico, Milan	Daniela Galimberti	48
IRCCS Carlo Besta, Milan	Fabrizio Tagliavini	41
University of Florence, Florence	Benedetta Nacmias	49
Lamezia Terme, Regional Genetic Centre	Amalia Bruni	112