



Universitat
de les Illes Balears

TREBALL DE FI DE GRAU

CONTRIBUCIÓ DE LA VARIABILITAT GENÈTICA COMÚ A LA MANIFESTACIÓ DE LES MALALTIES MENTALS GREUS EN FAMÍLIES AMB ELEVADA PREVALENCIA MITJANÇANT QUANTIFICACIÓ DE LA PUNTUACIÓ DE RISC POLIGÈNIC (PRS, POLYGENIC RISK SCORE)

Toni Capó Quetglas

Biologia

Facultat de Ciències

Any acadèmic 2020-21

CONTRIBUCIÓ DE LA VARIABILITAT GENÈTICA COMÚ A LA MANIFESTACIÓ DE LES MALALTIES MENTALS GREUS EN FAMÍLIES AMB ELEVADA PREVALENCIA MITJANÇANT QUANTIFICACIÓ DE LA PUNTUACIÓ DE RISC POLIGÈNIC (PRS, POLYGENIC RISK SCORE)

Toni Capó Quetglas

Treball de Fi de Grau

Facultat de Ciències

Universitat de les Illes Balears

Any acadèmic 2020-2021

Paraules clau del treball:

Malalties mentals greus, polimorfismes d'un sol nucleòtid (SNPs), Polygenic Risk Score, heritabilitat genètica.

Nom del tutor / la tutora del treball: Cristòfol Vives Bauzà, Grup de Neurobiologia de l'Institut d'Investigació Sanitària de les Illes Balears (IdISBa), Son Espases

Autoritz la Universitat a incloure aquest treball en el repositori institucional per consultar-lo en accés obert i difondre'l en línia, amb finalitats exclusivament acadèmiques i d'investigació

Autor/a		Tutor/a	
Sí	No	Sí	No
<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

ÍNDEX

RESUM	4
1. INTRODUCCIÓ	5
1.1 Les malalties mentals greus	5
1.2 Incidència i descripció de les malalties mentals greus	5
1.3 Arquitectura genètica de les malalties mentals greus	7
1.4 Polygenic Risk Score	8
2. OBJECTIU EXPERIMENTAL	9
3. MATERIALS I MÈTODES	10
3.1 Recol·lecció de mostres	10
3.2 Extracció d'ADN	11
3.3 Quantificació d'ADN	12
3.4 Genotipat	13
3.5 Preparació de les "dades objectiu"	14
3.6 Anàlisi de les puntuacions de risc poligènic.....	16
4. RESULTATS	18
5. DISCUSSIÓ	22
6. CONCLUSIONS	25
BIBLIOGRAFIA	26
ANNEXAS	32

RESUM

Es creu que els trastorns psiquiàtrics tenen una etiologia genètica complexa que consisteix en la interacció de variants comunes i rares. En aquest treball s'ha aprofundit en l'estudi de la contribució de la variabilitat genètica comú analitzant les puntuacions de risc poligènic de trastorns psiquiàtrics (SZ, BD i TDM) en dues famílies (Família 1, n=12 i Família 2, n=31) a través del programa informàtic PRSice. Mitjançant extracció i quantificació d'ADN, SNP array i duguent a terme diferents controls de qualitat, s'han preparat les dades objectiu de les nostres famílies, les quals han estat comparades amb diferents GWAS per a SZ, BD i SZ+BD amb l'objectiu de trobar diferències en la càrrega de variants genètiques comunes entre els afectats i els controls. Així s'ha demostrat que es pot observar la correlació entre PRS i la severitat del fenotip per a MMG en famílies amb alta prevalença de MMG. A partir d'aquesta prova pilot, es podria continuar amb l'estudi del PRS com a possible biomarcador per a millorar el diagnòstic i predir el risc de patir una MMG en l'àmbit familiar.

ABSTRACT

Psychiatric disorders are thought to have a complex genetic pathology consisting of an interplay between common and rare variation. In this work, we have determined the contribution of the common genetic variability by analyzing the polygenic risk scores (PRS) of psychiatric disorders (schizophrenia, bipolar disorder and major depression) in two families (Family 1, n = 12 and Family 2, n = 31). PRS was calculated using the PRSice software from SNP array genotyping. We compared common variant contribution to phenotype with the most recent GWAS data obtained from schizophrenia and bipolar disorder large cohorts with the aim to determine differences in the burden of common genetic variants between those affected subjects and their healthy relatives. Our results have shown a correlation between PRS and the severity of the phenotype in families with a high prevalence of major mental disorders. Based on this pilot study, we conclude that the PRS could be a potential biomarker to predict the risk of suffering a major mental disorder in those subjects from families with high prevalence.

1. INTRODUCCIÓ

1.1. Les malalties mentals greus

Les malalties mental greus (MMG), que inclouen el trastorn depressiu major (TDM), l'esquizofrènia (SZ), el trastorn esquizoafectiu i el trastorn bipolar (BD)¹, es troben entre les malalties més incapacitants i costoses². Els mecanismes causals subjacents a les MMG poden amagar les claus per a una prevenció i tractament eficaços, però continuen sent poc coneguts. La manca d'èxit és atribuïble, en part, a la complexitat intrínseca del cervell. Els trastorns psiquiàtrics tenen efecte damunt funcions corticals més elevades (estat d'ànim, comportament, percepció i cognició), que són molt més difícils de localitzar, quantificar i modelar que les funcions neurològiques més bàsiques³. A més, el repte es veu ampliat per la complexitat del fenotip, el comportament, l'arquitectura genètica, la poligenicitat i la multiplicitat de factors ambientals implicats⁴.

1.2. Incidència i descripció de les malalties mentals greus

Si mirem la incidència i la prevalença dels MMG, veim que afecten a més del 25% de la població i són una de les principals causes de discapacitat mundial⁵. De fet, s'estima que aproximadament un de cada 150 individus rebrà un diagnòstic de trastorn de psicosis en algun moment de la seva vida⁶. A Espanya, un de cada 10 adults i un de cada 100 nens tenen un problema de salut mental, on la depressió és la malaltia mental més freqüent⁷. Segons l'Organització Mundial de la Salut (OMS), la depressió afecta uns 121 milions de persones al món, dels quals menys del 25% tenen accés a tractaments efectius i adverteix que una de cada cinc persones arribarà a desenvolupar un quadre depressiu en la seva vida, augmentant aquest nombre en presència d'altres factors com la comorbiditat o situacions d'estrès⁸.

La depressió es presenta com un conjunt de símptomes de predomini afectiu (tristesa patològica, apatia, anhedònia, desesperança, irritabilitat, sensació subjectiva de malestar i impotència davant de les exigències de la vida) encara que, en major o menor grau, es podria parlar d'una afectació global psíquica i física, fent especial èmfasi en

l'esfera afectiva⁹. Així, presenta un gran impacte sobre la funcionalitat i la productivitat a l'àmbit social, plasmat en els costos associats a la depressió. A més, la depressió és un dels factors més associats a la conducta suïcida. S'ha estimat que el risc de suïcidi és quatre vegades més gran en les persones amb depressió en comparació amb la població general i 20 vegades més gran en el cas de la depressió major⁵.

BD i SZ són trastorns psiquiàtrics greus que presenten components genètics significatius amb estimacions d'heretabilitat que oscil·len entre el 60% i el 80%¹⁰. Estudis genètics i epidemiològics recents han demostrat una superposició substancial entre aquests dos trastorns, amb una correlació genètica¹¹. Malgrat la genètica i la simptomatologia compartides, els sistemes de diagnòstic actuals descriuen la BD i la SZ com a malalties independents i diferenciades en funció de la seva presentació clínica¹². Per tant, la investigació i identificació de components genètics que contribueixen a ambdós trastorns proporcionarà informació sobre la biologia subjacent als símptomes compartits dels trastorns.

Les manifestacions clíniques de la SZ es caracteritzen per símptomes positius (per exemple, al·lucinacions, deliris i conducta desorganitzada) i negatius (per exemple, dèficit expressiu emocional, amotivació social, abstinència social i dificultat per experimentar plaer)¹³. La SZ, és una de les principals causes de morbiditat i consumeix molta atenció mèdica i social a llarg termini¹⁴. A més, mostra una prevalença al voltant de l'1%¹⁵ i representa una enorme càrrega assistencial, amb uns costos associats anuals molt elevats, arribant a superar els 150.000 milions de dòlars de costos associats per any als Estats Units¹⁶. El fet que una malaltia que afecta a un 1% de la població suposi tants de costos, es deu al fet que l'aparició típica es produeix a principis de l'edat adulta, suposant deficiències a llarg termini de la funció social i ocupacional de les persones afectades¹⁷. El trastorn també s'associa amb una reducció de l'esperança de vida: persones amb SZ tenen una esperança de vida mitjana uns 15 anys inferior a la població general i un risc del 5% al 10% de mort per suïcidi¹⁸. Per tant, es tracta d'una causa important de morbiditat i consumeix una gran quantitat d'atenció mèdica i social a llarg termini.

El BD es situa com la dissetena causa principal de discapacitat entre totes les malalties a tot el món¹⁹. Les persones amb BD presenten taxes elevades de condicions psiquiàtriques coexistents, inclosa l'ansietat (estimada en el 71% de les persones amb BD), l'ús de substàncies estupefaents (en un 56%), trastorns de personalitat (en un 36%) i dèficit d'atenció – trastorn de hiperactivitat (entre un 10 i un 20%). Quan es troben aquests problemes addicionals, augmenta la càrrega de la malaltia i empitjoren el pronòstic²⁰. La Iniciativa de l'Enquesta Mundial de Salut Mental va informar estimacions de prevalença per als BD del 2,4%²¹. Atès que el BD es manifesta normalment per primera vegada durant el primers anys del desenvolupament d'infants i adolescents, l'assoliment de fites educatives i ocupacionals sovint es veu afectada negativament. La disfunció cognitiva i psicosocial durant els episodis aguts agrava el problema²². A més, aproximadament el 6-7% de les persones amb BD es suïciden; hi ha proves que les taxes de suïcidi entre les persones amb BD són 20 a 30 vegades més altes que les taxes de la població general²³. Tot i que s'han suggerit intervencions específiques per a aquestes afeccions simultànies²⁴, les taxes elevades de trastorns mentals coexistents també poden reflectir el fracàs dels sistemes de diagnòstic actuals a l'hora d'identificar la salut mental general d'un pacient²⁵.

1.3. Arquitectura genètica de les malalties mentals greus

Tot i el gran repte que suposa, en la darrera dècada ha tingut lloc un notable progrés a l'hora d'aclarir els fonaments genètics dels trastorns psiquiàtrics, aconseguint grans avanços en el coneixement de l'«arquitectura genètica» d'aquests trastorns. Cal destacar que la majoria d'aquests estudis es basen en relacions familiars conegudes en lloc de mesurar directament els marcadors genètics. De fet, tots els tipus de MMG tenen tendència a manifestar-se en famílies i el risc de desenvolupar una malaltia s'associa amb el grau de relació biològica (material genòmic compartit) amb l'individu afectat²⁶. Estudis en bessons demostren constantment que els bessons monozigòtics que comparteixen el 100% del seu ADN nuclear són més propensos a presentar el mateix trastorn que els bessons dizigòtics que comparteixen el 50% del seu material genètic²⁷. Aquesta diferència suggereix que la causa de la malaltia mental és atribuïble en gran mesura a factors genètics. A més, la tecnologia genètica molecular ha aportat les eines per estudiar les variants genètiques involucrades en la transmissió de MMG cap a la

descendència. El tipus de variants més comunes, els polimorfismes d'un sol nucleòtid (SNP, per les sigles en anglès *Single Nucleotide Polymorphism*), són un tipus de variant diferencial entre diferents cromosomes en una base concreta de la seqüència d'ADN²⁸. Per exemple, alguns cromosomes d'una població poden tenir una citosina en aquest lloc (l'al·lel C), mentre que d'altres tenen una timina (l'al·lel T). S'ha estimat que, el humans tenim prop de 10 milions de SNPs (és a dir, una variant cada 300 bases de mitjana) amb freqüències al·lèliques majors al 1% i que aquests 10 milions de SNPs constitueixen el 90% de la variació en la població de caracterització²⁹. El 10% restant es deu a una àmplia gamma de variants que són rares a la població. La variabilitat genòmica en forma d'SNP es pot genotipar actualment gràcies a la tecnologia dels arrays de SNPs²⁸.

1.4 Polygenic Risk Score

Les malalties complexes són degudes a la interacció de molts de gens amb factors ambientals. Els estudis de GWAS (GWAS, per les sigles en anglès de *Genomic-Wide Association Studies*) han permès desvetllar molta variabilitat comú en forma de SNP associada al risc a patir SZ³⁰, BD³¹ o TDM³². El PRS permet quantificar, a partir dels SNPs d'associació coneguts, el risc d'un determinat individu a partir alguna d'aquestes malalties³³. No obstant això, aquestes variants solen tenir un efecte petit i correspondre a una petita fracció de variants realment associades, el que significa que tenen un poder predictiu limitat³⁴. Tot i això, Yang et al. va demostrar que gran part de l'heretabilitat en MMG es pot explicar avaluant simultàniament els efectes de tots els SNP, ja que s'ha demostrat que contribueixen entorn a un 20% de l'heretabilitat de les MMG³⁵. Els PRS ofereixen la possibilitat d'investigar la superposició genètica entre trastorns i a provar la seva associació amb múltiples fenotips³⁶. A més, estudis amb PRS demostren constantment que la predicció de MMG millora mitjançant la inclusió de variants genètiques més dèbilment associades, cosa que suggereix que molts milers de variants genètiques estan implicades en la configuració del risc de la majoria dels trastorns mentals³⁷. Aquests impliquen tant SNPs com variants estructurals rares, com ara delecions i insercions d'ADN³⁸. Si bé els SNP individuals tenen un efecte molt discret sobre el risc de malaltia, el fet de sumar la dosi al·lèlica ponderada en tots els SNPs i la creació d'una única puntuació de risc poligènic ha permès explorar com el risc genètic es manifesta directament en individus de diferents poblacions³⁷. Així, la estimació del

PRS permet estudiar el risc de patir la malaltia sense conèixer els gens específics implicats en dita malaltia complexa³⁹. Tot i que s'han proposat molts biomarcadors per les MMGs, fins al moment no se n'ha trobat cap que sigui d'utilitat clínica. A causa de la seva elevada prevalença i la magnitud de discapacitat produïda per aquestes patologies, seria molt important poder comptar amb biomarcadors que ens ajudin amb el diagnòstic, permetent diferenciar entre els subtipus de depressió, contribuint així a l'orientació terapèutica.

2. OBJECTIU EXPERIMENTAL

Objectiu Principal (OP) i Objectius específics (OE):

OP. Determinar la correlació entre PRS i severitat del fenotip clínic i estudiar la viabilitat del PRS com a biomarcador predictor de la severitat del fenotip de les malalties mentals greus en famílies amb elevada prevalença

OE.1. Genotipar per SNP-array 49 individus de dues famílies amb elevada prevalença de MMG

OE.2. Establir correlacions lineals entre puntuacions PRS i els diferents fenotips clínics de MMG de les dues famílies estudiades.

Davant l'actual situació de pandèmia, han sorgit una sèrie d'obstacles relacionats amb els permisos per accedir a l'IdISBa, des d'on s'havia de dur a terme aquest Treball de Fi de Grau, amb el que de les 12 famílies inicialment previstes d'estudiar en el marc d'aquest TFG, només se n'han pogut estudiar dues.

3. MATERIALS I MÈTODES

3.1 Recol·lecció de mostres

Tots els subjectes inclosos en aquest treball, van firmar consentiment per escrit. I tots els procediments van ser aprovats pel comitè ètic d'investigació clínica de les Illes Balears (CEIC). Les mostres de sang foren recollides en tubs de sang BD Vacutainer (que contenen EDTA) i es van guardar a la gelera fins al moment de l'extracció del ADN.

3.1.1 Famílies

Família 1:

Aquesta família prové originalment de dos petits pobles de Mallorca. Inicialment van ser reclutats després d'una sessió amb el psiquiatra. La provant, que es tracta del subjecte que permet la recol·lecció dels demás individus, és una mare diagnosticada de trastorn esquizoafectiu (subjecte 17) caracteritzat per episodis depressius i maníacs acompanyats de deliris de tipus paranoide. La provant té tres fills varons afectes i dues filles bassones sanes. Entre els fills afectes hi ha el subjecte 27, que presenta des dels 19 anys; el subjecte 28 que va debutar amb símptomes psicòtics quan tenia 16 anys, diagnosticat posteriorment de SZ paranoica, i va morir per suïcidi en 2014. I el subjecte 29 va començar a evidenciar trastorns de comportament, ansietat i símptomes somàtics als 14 anys i després fou diagnosticat amb SZ. A més a més, té dues filles bessones que no presenten símptomes psiquiàtrics (subjectes 30 i 31). La història familiar de la provant inclou un germà que va morir de cirrosi per alcoholisme (subjecte 19), un oncle amb autisme i psicosi (subjecte 5), un altre oncle amb transtorn depressiu major i epilèpsia (subjecte 6) i un cosí amb ansietat (subjecte 11) . El subjecte 11 és la mare d'una filla afectada per BD (subjecte 24). El seu pare (subjecte 12) presenta trastorn depressiu major.

Família 2:

La segona família és una gran família multigeneracional amb una alta prevalença de trastorns mentals de la regió granadina de "Las Alpujarras", coneguda per ser una regió amb elevada prevalença de SZ⁴⁰. Part de la família viu a Mallorca i va ser tractada en el Servei de Psiquiatria de l'Hospital Universitari Son Espases (HUSE) de les Illes Balears. La resta de la família va ser reclutada gràcies a la col·laboració amb el Servei de Salut Mental d'El Ejido, Almería, i del Servei de Salut Mental de l'Hospital Universitario Virgen de las Nieves, Granada). En total, es van obtenir n = 34 mostres d'ADN de la família (n = 9 corresponents a pacients psicòtics, n = 10 a pacients amb trastorn mental no psicòtic, n = 14 a individus sans i n = 1 a un fenotip indeterminat). Tots els participants van ser sotmesos a entrevistes semiestructurades per un psiquiatra capacitada, utilitzant la versió en espanyol de l'Entrevista Clínica Estructurada per al Manual Diagnòstic i Estadístic de Trastorns Mentals (DSM-IV) Eix I (SCID-I), l'Escala de Síndrome Positiu i Negatiu (PANSS) i l'Entrevista Diagnòstica per a Estudis Genètics (DIGS).

3.1.2 Disseny dels pedigrees amb el programa Progeny Genetics

La informació sobre l'estructura familiar s'ha completat entrevistant el psiquiatra que va tractar amb la provant en el cas de la Família 1 i a partir dels registres familiars en el cas de la Família 2. A partir d'aquesta informació, s'ha emprat el programa online *Progenygenetics.com*⁴¹ per tal d'estructurar els pedigrees familiars (Veure Annexa 1). Tots els membres de les famílies, dels quals va ser possible obtenir mostres de sang per poder enviar les mostres a genotipar posteriorment, van ser avaluats per un psiquiatra.

3.2 Extracció d'ADN

Es va realitzar una extracció de sang amb tubs BD Vacutainer (amb EDTA) dels diferents pacients estudiats i es van guardar les mostres a la gelera. Tot seguit, es va extreure l'ADN utilitzant el kit d'extracció *MasterPure™ DNA Purification Kit for Blood Version II (MB711400, Lucigen)*, seguint el protocol de Buffy Coat, a través del qual s'extreu l'ADN de les mostres a partir de la capa leucocítica.

3.2.1 Protocol Buffy Coat

Aquesta va ser extreta mitjançant una centrifugació a 10.621 G (10.000 rpm, rotor FA-45-30-11 - *Centrifuge 5430*) durant 10 minuts i posterior aïllament de la capa cel·lular intermitja entre el plasma i la capa d'eritròcits (*buffy coat*). D'aquesta manera, es va partir d'un volum concret de mostra amb cèl·lules leucocítiques, sobre les quals es va aplicar el reactiu *Red Cell Lysis* per tal de rompre les seves membranes cel·lulars. Seguidament es va afegir *Tissue Cell Lysis* a la solució amb l'objectiu d'acabar de rompre l'estructura cel·lular, fent més accessible l'ADN. Posteriorment, es va afegir un altre reactiu: MPC, per tal de degradar les proteïnes de la mostra i evitar interferències en la quantificació del material genètic. Aquesta solució es va centrifugar i el sobrenadant es va resuspendre amb isopropanol amb l'objectiu que l'ADN precipitès. A continuació es va realitzar una centrifugació a 10.621 G (10.000 rpm, rotor FA-45-30-11 - *Centrifuge 5430*) a 4 °C, de la qual es va eliminar el sobrenadant i es va guardar el pelet. Posteriorment es van realitzar una sèrie de rentats amb etanol 70%. Finalment, la mostra d'ADN va ser resuspesa amb *TE Buffer* (10 mM Tris, 1 mM EDTA, pH 8.0 amb HCl). La funció d'aquest tampó és solubilitzar l'ADN, protegint-lo de la degradació.

3.3 Quantificació d'ADN

Després de l'extracció, les mostres d'ADN es van quantificar mitjançant Nanodrop assegurant una relació A260 / A280 proper a 1.8, ja que un valor A260 / 280 < 1.6 indicaria una possible contaminació per compostos aromàtics com fenols i proteïnes mentre que un rati A260 / 280 > 2.1 podria ser degut a la presència d'ARN en la mostra. També es va valorar que la quantificació mostràs unes concentracions de DNA superiors a 50 ng/µl. Posteriorment, les mostres es van emmagatzemar al congelador a -20°C. A més, ens vàrem assegurar que totes les mostres arribassin a una concentració mínima de 50 ng/µl en un volum mínim de 20 µl. Tots els identificadors de mostra es van codificar de manera que només els metges poguessin accedir a la informació personal. En total, es van recol·lectar 13 mostres d'ADN de la Família 1 (Veure Annexa 3, Figura 5 A) i 34 de la Família 2.

3.4 Genotipat

3.4.1 Tècnica SNP array

La tècnica de SNP *array* (matriu de SNPs) serveix per estudiar la variació genètica en molts centenars de milers de llocs específics del genoma ⁴². S'ha demostrat que és un mètode excel·lent per estudiar la variació genètica comú, que es pot utilitzar per avaluar l'ascendència ⁴³, així com la predisposició a moltes malalties multifactorials complexes ⁴⁴. Així, per tal de conèixer la variabilitat genètica entre les nostres mostres, es van enviar a Helmholtz Zentrum München (Alemanya) per genotipar-les mitjançant la matriu GSAMD-24 v1-0. Els passos bàsics de la prova són els següents: amplificació i fragmentació de l'ADN, precipitació i resuspensió de l'ADN, hibridació, extensió, tinció i imatge ⁴⁵. Les mostres 14 i 15 no van ser genotipades per SNP array (veure Annexa 1, Família 1).

El primer pas del procés de genotipatge és amplificar les mostres d'ADN i incubar-les durant la nit. La qualitat de l'ADN és crucial en aquest pas d'amplificació. L'endemà, les mostres es fragmenten enzimàticament per fer més accessibles les regions d'ADN, i així assegurar una major eficiència d'hibridació. Després, les mostres d'ADN es precipiten i es resuspenen en un tampó especialment dissenyat per facilitar la hibridació. En aquest moment, les mostres estan preparades per hibridar-se amb el BeadChip (soport físic que conté la matriu amb totes les sondes formades per petits oligonucleòtids que contenen els SNPs). La hibridació consisteix en la unió dels fragments d'ADN genòmic a les seqüències complementàries anomenades sondes, presents a les microesferes. Després de la hibridació, cada locus SNP és "puntuat" mitjançant un pas d'extensió de base única (SBE) en presència de dideoxinucleòtids (ddNTP) marcats amb dinitrofenol per a C / G i marcats amb biotina per A / T. El nucleòtid incorporat a la sonda serà C / G, A / T en funció de l'al·lel present a l'ADN complementari. Finalment, després de l'extensió es visualitzen les diferents emissions lumíniques (verd en presència de C o G i vermell en presència de A o T). El sistema iScan capta les intensitats de fluorescència de les microesferes, basant-se en un sistema de làser confocal de dos colors amb una resolució de 0,8 µm, que genera fitxers de dades d'intensitat.

3.5 Preparació de les “dades objectiu”

Les dades objectiu consisteixen en dades genotip-fenotip a nivell individual generades en base a la informació obtinguda després de genotipar les mostres, adjuntar la informació relativa als pacients i dur a terme els diferents controls de qualitat pertinents. Aquests controls de qualitat s’han realitzat abans de computar els diferents PRS per als subjectes de la Família 1 i la Família 2 per evitar resultats no fiables causats per dades de genotipatge de baixa qualitat, seguint les pautes generals de control de qualitat recomanades per als estudis de PRS i d’associació³⁹ a través del programa informàtic PLINK 1.7⁴⁶.

3.5.1 Programa informàtic PLINK.7

PLINK.7 és un conjunt d’eines d’anàlisi d’associació de tot el genoma gratuït, fàcil d’utilitzar, que permet múltiples funcions relacionades amb anàlisis genètics. Presenta cinc tipus de funció diferents: gestió de dades (permet manipular enormes fitxers de dades), resum de diferents estadístiques (genotipatge i taxes d’heterozigositat, proves d’equilibri de Hardy-Weinberg, freqüències d’al·lels i genotips, resums d’errors mendelians per a dades familiars, etc.), estratificació de la població (anàlisi de components principals), grau d’ascendència compartida recent i anàlisi d’associació. PLINK.7 presenta un format que inclou dos fitxers separats: un fitxer de mapa (*map file*) i un fitxer de pedigrí (*ped file*):

- ❖ **Plink.ped:** Les 6 columnes inicials són: identificació familiar (columna 1), identificació del subjecte (columna 2), identificació pare (columna 3), identificació mare (columna 4), sexe - 1 = home i 2 = dona - (columna 5) i afecte - 1 = no afectat i 2 = afectat – (columna 6). Els valors no disponibles es codifiquen amb el valor “-9”. La resta de columnes (milers de columnes) corresponen als diferents SNPs genotipats (al·lels paterns i materns).
- ❖ **Plink.map:** el *map file* és un fitxer de quatre columnes que inclou informació relativa a: el cromosoma (primera columna), el nom del SNP (segona

columna), la distància genètica en centimorgan⁴⁷ (tercera columna) i la posició del parell de bases (quarta columna) per a tots els SNPs genotipats.

3.5.2 Controls de qualitat

A partir dels *map file* i el *ped file* s'han pogut realitzar una sèrie de correccions amb PLINK.7 per tal de refinar les nostres mostres i aconseguir una millor fiabilitat abans d'aplicar els PRS. Primer de tot s'ha tengut en compte la qualitat del genotipat, allà on s'han considerat diferents aspectes correctors: el *Missing Call rate* (tasa de falta d'informació), el *LogRRatio* (LRR) i el *BAlleleFreq* (BAF). El *Missing Call Rate* és la tasa de SNPs absents d'una mostra sobre el nombre total d'SNP genotipats o el percentatge d'individus genotipats que no presentaven un SNP determinat. És a dir, que amb PLINK.7 s'han eliminat aquells SNPs que no es trobaven presents en almenys un 10% dels individus genotipats i també aquells individus que presentaven una mancança del 10% dels SNPs respecte el total⁴⁸. A més, PLINK.7 també ha permès avaluar els possibles errors en la determinació del sexe dels pacients a través de comparacions entre la columna 1 del *map file* amb la columna 5 del *ped file*, comprovant que la dotació cromosòmica coincideix amb la sexe registrat. Els SNPs rars que tan sols estan presents a un 0.05% dels individus de les dades objectiu també s'han descartat. Finalment, s'han eliminat els SNPs que violin l'equilibri Hardy Weinberg - és a dir, aquells SNPs que puguin resultar un factor que alteri les freqüències al·lèliques de la població (ja que en estudis d'associació es considera que han d'estar en equilibri) -³⁹.

En la Família 1 es van mantenir totes les mostres per a calcular el PRS excepte la mostra "29" (Veure pedigree de la Família 1 al 3.1.2) degut a la baixa qualitat en el genotipat. En la Família 2 també es van descartar 3 mostres més a causa de la mala qualitat del el genotipat.

3.6 Anàlisi de les puntuacions de risc poligènic (PRS, Polygenic Risk Score)

3.6.1 Software emprat: PRSice

PRSice és un programa dissenyat per calcular PRS⁴⁹, el qual requereix dos conjunts de dades: el conjunt de “dades base” (són el conjunt de SNPs de referència provinents dels estudis de GWAS, Veure 3.5.2) i el conjunt de “dades objectiu” (són els SNPs del individu a avaluar, Veure 3.3). Consta de dos passos: primer es calculen els PRS per a diferents llindars de p-valor (o *thresholds*) del conjunt de dades base (provinents de GWAS) i després, es realitza un anàlisi de regressió per esbrinar sota quin llindars el PRS explica millor el fenotip (Veure 3.6.3). Per poder executar les funcions en el PRSice, és necessari realitzar una conversió de les dades en format PLINK.ped i PLINK.map a format binari, on les mostres es disposen en 3 formats: “bim”, “bed” i “fam”. “Bim” inclou les 4 columnes del *map file*; “fam” inclou les 6 primeres columnes del *ped file* i “bed” inclou totes les columnes restants del *ped file* amb els diferents SNPs analitzats.

3.6.2 Dades base

El primer pas en les anàlisis de la puntuació de risc poligènic (PRS) és generar o obtenir les dades base (dades estadístiques resumides de GWAS) que es corresponguessin amb els resultats GWAS més potents disponibles en el fenotip objecte d'estudi. Les nostres “dades base” es van obtenir del Consorci de Genòmica Psiquiàtrica (PGC), des d'on es van descarregar tres fitxers de resultats GWAS: (I) *BDSCZvsCONT.sumstats.gz*, un resum dels resultats d'un GWAS que conjuga casos de 20.129 pacients bipolars amb 33.426 casos de SZ i que es comparen amb 54.065 controls; (II) *BDvsCONT.sumstats.gz*, un resum dels resultats d'un GWAS que compara 20.129 casos de bipolaritat contra 21.524 controls; (III) i *SCZvsCONT.sumstats.gz*, un resum dels resultats d'un GWAS que compara 33.426 casos de SZ amb 32.541 controls. Aquests resultats provenien de Wang et al., 2017.

Els fitxers s'estructuraven de la següent manera: **CHR** -cromosoma-, **SNP** -Variant determinada-, **BP** -Posició genòmica, de l'anglès *base pair*-, **A1** -Al·lel d'efecte-, **A2** -Al·lel sense efecte-, **FQR_A** -Freqüència al·lèlica en casos o malalts)-, **FQR_U** -

Freqüència al·lèlica en controls-, **INFO** -informació de puntuació que representa la qualitat de la imputació-, **OR** -Ods Ratio basada en A1) - **SE** (Error estàndard) - **P** (p-valor) **Direction** (Direcció de l'efecte) - **HetPVa** (Heterogeneïtat del p-valor entre les cohorts incloses).

3.6.3 Puntuació del PRS

Els PRS es van calcular amb PRSice per a cada membre de la família mitjançant una puntuació estandarditzada ⁵¹ a través del següent Script:

```
Rscript PRSice.R -dir. -Price ./PRSice_linux -base data_gwas -target data_SQ -fil 1 -stat O -score std -.
```

$$PRS = \frac{\sum_i (S_i \times G_i) - \text{Mean (PRS)}}{SD (PRS)}$$

A Reininghaus 2016, s'informa que "S" explica la mida de l'efecte de l'al·lel d'efecte i "G" és el nombre d'al·lels d'efecte observats per a un fenotip concret. PRSice tracta aquest PRS per a diferents llindars de significació que després es comparen mitjançant anàlisis de regressió per seleccionar el llindar que discrimina millor entre casos i controls. Així, els llindars de PRS que expliquen la variància més alta es van utilitzar per calcular estadístiques descriptives per comparar afectats amb individus sans.

El valors "mean" i "SD" (mitjana i desviació estàndard, respectivament) ens serveixen per a normalitzar els valors. Una vegada obtingut el valor de PRS per a cada individu, es poden realitzar comparatives entre diferents grups amb l'objectiu de comprovar si els casos de malaltia mental presenten més càrrega de variants amb efecte sobre el fenotip d'interés respecte el grup control. En funció del llindar que el PRSice seleccioni, es contemplaran variants prèviament associades a malalties mentals greus en estudis de GWAS (si el p-valor <0.05) o també s'inclouran altres variants que no han estat prèviament associades a aquell fenotip (si p-valor>0.05).

4. RESULTATS

A través de PRS s'ha estudiat la contribució de variants comunes al fenotip psicòtic tant a la Família 1 com a la Família 2. Els resultats obtinguts han estat analitzats des d'un punt de vista quantitatiu, tenint en compte que pot existir un biaix associat a la petita mostra analitzada en aquest estudi, pel que s'ha de ser molt curós a l'hora d'establir certes assumpcions categòriques. D'aquesta manera, s'ha duit a terme una anàlisi estadística dels resultats mitjançant comparacions múltiples entre casos i controls tant a la Família 1 com a la Família 2 amb el programa informàtic RStudio. Primer s'han estudiat les característiques de les dades (homogeneïtat de variàncies i normalitat) i posteriorment s'ha aplicat el test estadístic paramètric *t.test*. A més, s'han realitzat correccions de p-valor per a comparacions múltiples (p-valor / n^o de comparacions) tant en la Família 1 com en la Família 2. Així, les diferències entre grups s'han considerat significatives en p-valors < $8,33 \cdot 10^{-3}$ en la Família 1 i la combinació de la Família 1+Família 2 i p-valors < $4,17 \cdot 10^{-3}$ en la Família 2. Interessantment, gairebé tots els llocals de PRS han inclòs variants prèviament associades significativament al fenotip.

4.1 Família 1

Cal destacar que per aquesta família no s'ha obtingut cap diferència significativa entre les diferents comparatives de casos contra controls. Tot i així, s'observen una sèrie de diferències que ens permeten analitzar a nivell qualitatiu els resultats obtinguts per a les diferents comparatives. Emprant el fenotip SZ com a conjunt de dades base (PRS SZ) per calcular el PRS (lloc de significació de PRS = 0,0029), tots els membres psicòtics (inclou tots els subjectes SZ i BD) del pedigree han obtingut un resultat PRS SZ positiu (fenotip psicòtic, n = 4, mitjana = 0,6730, SD = 0,6436; Veure Annexa 3, Figura 1A) mentre que la majoria dels els subjectes sans presenten PRS SZ negatiu (sans, n= 8, mitjana= -0,3364, SD= 1,0027) -Veure Annexa 3, Figura 1A i Annexa 2, la Taula 1-. El subjecte 28 és el pacient amb la major contribució de variants comunes al fenotip (PRS= 1.5916; Veure Annexa 2, Taula 1). Aquestes diferències també s'observen amb el fenotip ample (inclou subjectes psicòtics i malalts mentals no psicòtics amb TDM), tot i que

menys pronunciats (ample n= 7, mitjana= 0,3320, SD= 0,6264 vs. sans, n= 5, mitjana= -0,4648, SD = 1,3031; Veure Annexa 3, Veure Annexa 3, Figura 1B).

Quant a PRS BD (llindar de significació de PRS= 0,0299), el subjecte 24 és l'únic que ha obtingut un PRS BD positiu entre individus psicòtics (PRS BD= 0,0124; Veure Annexa 2, Taula 1). De fet, els individus psicòtics mostren un resultat PRS BD menor que els sans (fenotip psicòtic n= 4, mitjana= -0,7925, SD= 0,8764 vs. sans, n= 8, mitjana= 0,3962, SD= 0,8389; Veure Annexa 3, Figura 1C), probablement a causa de les puntuacions altes de subjectes sans 15 (PRSD = 1,1654) i 23 (PRS BD= 0,483) i subjectes del fenotip ample 16 (PRS BD= 0,6981) i 11 (PRS BD= 1,9398) -Veure Annexa 2, la Taula 1-. Aquestes diferències pràcticament desapareixen en analitzar el fenotip ample, però tot i així encara mostren PRS BD lleugerament inferior als sans (fenotip ample, n= 7, mitjana= -0,1197, SD= 1,2288 vs. sans, n= 5, mitjana= 0,1667, SD= 0,6517; Veure Annexa 3, Figura 1D).

Finalment, pel que fa a PRS SZ+BD (llindar de significació de PRS= 0,0407), de nou, el subjecte 24 és l'únic subjecte que ha obtingut resultats positius entre els afectats del fenotip psicòtic (PRS SZ + BD= 0,3634) -Veure Annexa 2, la Taula 1-. Igual que en els anàlisis anteriors, els individus psicòtics han mostrat un PRS SZ+BD generalment més baix que els sans (fenotip psicòtic, n= 4, mitjana= -0,7616, SD= 0,9286 vs sans n= 8, mitjana= 0,3808, SD= 0,8393; Veure Annexa 3, Figura 1E). Aquestes diferències gairebé desapareixen en analitzar el fenotip ample (ampli n= 7, mitjana= -0,0199, SD= 1,2551 vs. saludable, n= 5, mitjana= 0,0369, SD= 0,6303; Veure Annexa 3, Figura 1F) degut a les altes puntuacions de individus afectats 16 (PRS SZ+BD= 0,4085), 12 (PRS SZ+BD = 0,455) i 11 (PRS SZ+BD= 2,0429) - Veure Annexa 2, la Taula 1-.

4.2 Família 2

Emprant el fenotip SZ com a conjunt de dades base, s'han trobat les puntuacions PRS SZ més altes (llindar de significació PRS= 0,01575) entre els subjectes psicòtics (fenotip psicòtic; n= 8, mitjana= 1,1253, SD= 0,6393; Veure Annexa 3, Figura 2A) en comparació amb els subjectes sans, que mostren puntuacions negatives (subjectes sans; n= 22,

mitjana= -0,4560, SD= 0,7357; Veure Annexa 3, Figura 2A). El pacient amb la major contribució de variants comunes associades a MMG ha sigut el pacient 41* (PRS SZ= 2.3379; Veure Annexa 2, Taula 2). El mateix s'observa en considerar el fenotip ampli, tot i que les diferències no es mostren de forma tan pronunciada (fenotip ampli; n= 17, mitjana= 0,5245, SD= 0,8723 vs. sans; n= 13, mitjana= -0,7653, SD= 0,6101; Veure Annexa 3, Figura 2B). En comparar fenotips segons el diagnòstic de la malaltia (BD, SZ, TDM), els subjectes SZ presenten puntuacions PRS SZ més altes que els pacients BD (casos SZ n= 4, mitjana= 1,3080, SD= 0,6902; casos BD n= 3 , mitjana = 0,9599, SD = 0,7618; Veure Annexa 3, Figura 3A). Els subjectes TDM han obtingut puntuacions PRS SZ properes a zero (n= 9, mitjana= -0,0094, SD= 0,6948; Veure Annexa 3, Figura 3A) i els individus sans mostren puntuacions negatives (n= 13, mitjana= -0,7653, SD= 0,6101; Veure Annexa 3, Figura 3A). Interessantment, ha estat possible diferenciar dos grups entre els pacients amb TDM; un mostrant puntuacions més altes (TDM A) ja sigui a PRS SZ, PRS BD i PRS SZ + BD que l'altre (TDM B). TDM A mostra la tercera puntuació PRS SZ més alta després dels subjectes BD (TDM A; n= 5, mitjana= 0,55554, SD= 0,1636; Veure Annexa 3, Figura 4A) mentre que TDM B presenten puntuacions negatives, semblants a la dels individus sans (TDM B; n= 4, mitjana= -0,7155, SD= 0,2347; Veure Annexa 3, Figura 4A).

Quant al fenotip BD com a conjunt de dades base, de nou es troben les puntuacions PRS BD més altes (llindar de significació PRS = 0,0073) entre els subjectes psicòtics (fenotip psicòtic; n= 8, mitjana= 0,6914, SD= 1,1583; Veure Annexa 3, Figura 2C) en comparació amb individus sans que han obtingut puntuacions negatives (sans ; n= 22, mitjana= -0,3152, SD= 0,7753; Veure Annexa 3, Figura 2C). En aquest cas, les diferències són més fortes en analitzar el fenotip ampli (fenotip ampli; n= 17, mitjana= 0,3680, SD= 0,9828; Veure Annexa 3, Figura 2D) vs. sans (n= 13, mitjana = -0,5892, SD = 0,6926; Veure Annexa 3, Figura 2D) ja que l'anàlisi estadístic indica una significació moderada en les diferències observades (p -valor= $5,5 \cdot 10^{-3}$, lleugerament major que $4,17 \cdot 10^{-3}$). Com era d'esperar, en l'anàlisi de PRS agrupats per fenotip les puntuacions PRS BD més altes es troben entre individus BD (n= 3, mitjana= 0,9668, SD= 0,8699; Veure Annexa 3, Figura 3B) on dues de les puntuacions PRS BD més altes van correspondre a dos casos de BD 68* (PRS BD= 2.2005) i 87* (PRS BD= 1,4662) – Veure Annexa 2, la Taula 2-. De fet, en PRS BD existeixen diferències moderadament significatives entre BD vs. sans (p -

valor= $7,8 \cdot 10^{-3}$). Els casos de SZ també mostren resultats PRS BD elevats ($n= 4$, mitjana= 0,4509, SD= 1,2947; Veure Annexa 3, Figura 3B) excepte en els subjectes 36* (PRS BD= -1,4486) i 103* (PRS BD= -0,4063) –Veure Annexa 2, la Taula 2-. Pel que fa als individus TDM en conjunt, presenten puntuacions PRS BD properes a zero ($n= 9$, mitjana= 0,0805, SD= 0,7484; Veure Annexa 3, Figura 3B). No obstant això, a l'hora d'analitzar els dos subgrups pel fenotip TDM, TDM A ($n= 5$, mitjana= 0,5108, SD= 0,5643) és el segon amb les puntuacions PRS BD més altes després dels subjectes amb BD mentre que TDM B ha obtingut puntuacions negatives ($n= 4$, mitjana= -0,4574, SD= 0,6121) properes a la dels individus sans ($n= 13$, mitjana= -0,5892, SD= 0,6926; Veure Annexa 3, Figura 4B).

Finalment, combinant SNPs prèviament associats amb SZ i BD per calcular PRS SZ+BD (llindar de significació de PRS = 0,13655) s'han obtingut resultats similars: els subjectes psicòtics mostren les puntuacions significativament més altes de PRS SZ+BD (p -valor= $6,1 \cdot 10^{-4}$, $n= 8$, mitjana= 0,9682, SD= 0,7293; Veure Annexa 3, Figura 2E) en comparació amb individus sans ($n= 22$, mitjana= -0,3701, SD= 0,8620; Veure Annexa 3, Figura 2E). En l'anàlisi del fenotip ampli, aquestes diferències es redueixen lleugerament (fenotip ampli; $n= 17$, mitjana= 0,4746, SD= 0,9605 vs. sans; $n= 13$, mitjana= -0,6512, SD= 0,6914; Veure Annexa 3, Figura 2F) però es mantenen significatives (p -valor= $9,9 \cdot 10^{-4}$). L'anàlisi dels fenotips segons el diagnòstic torna a mostrar que els pacients amb SZ obtenen les puntuacions més altes de PRS SZ+BD ($n= 4$, mitjana= 1,1258, SD= 0,8363) seguits de pacients amb BD ($n= 3$, mitjana= 0,8804, SD= 0,870) i subjectes TDM ($n= 9$, mitjana= 0,0359, SD= 0,9589; Veure Annexa 3, Figura 3C). Els individus sans mostren les puntuacions PRS SZ+BD més baixes ($n= 13$, mitjana= -0,6512, SD= 0,6914; Veure Annexa 3, Figura 3C). De fet, SZ presenta puntuacions de PRS significativament més altes que els sans (p -valor= $4,5 \cdot 10^{-4}$) i que TDM B (p -valor= $1,56 \cdot 10^{-3}$). Interessantment, l'anàlisi dels subgrups de pacients amb TDM il·lustren que TDM A presenta el tercer PRS SZ + BD més alt ($n= 5$, mitjana= 0,6633, SD= 0,6452) mentre que TDM B ($n= 4$, mitjana= -0,7482, SD= 0,6914) presenta el PRS més baix (Veure Annexa 3, Figura 4C). De fet, TDM A presenta puntuacions significativament més altes que TDM B (p -valor= $6,2 \cdot 10^{-3}$) i que els sans (p -valor= $2,67 \cdot 10^{-3}$).

Per concloure l'estudi, i amb la finalitat d'augmentar la n analitzada, es va fer un anàlisi conjunt d'ambdues famílies (n=43), on es pogué observar que el PRS també és capaç de separar els grups de subjectes segons el fenotip clínic (Veure Annexa 3, Figura 5). Així, veiem com el fenotip psicòtic o el fenotip ample mostra sempre majors puntuacions de PRS que el grup control. Tot i així, en cap dels tres subconjunts de dades s'han obtingut diferències significatives entre el grup afectat i el grup control.

5. DISCUSSIÓ

Durant l'última dècada, la tecnologia genètica molecular ha permès estudiar les variants genètiques responsables de la transmissió de malalties mentals dels pares a la descendència. Els estudis d'associació a tot el genoma han identificat més d'un centenar de variants associades a malalties severes⁵². L'anàlisi de puntuacions de risc poligènic demostren que la predicció de malalties mentals millora mitjançant la inclusió de milers de variants genètiques comunes implicades en la configuració del risc de trastorns mentals^{53,54}. Tot i així, hi ha evidències que les estimacions d'heretabilitat estudiades en bessons no es tradueixen en efectes directes de les variants moleculars genètiques⁵⁵ degut a que existeixen molts de factors a tenir en compte que també contribueixen al risc de patir MMG. Els mateixos estudis sobre bessons que van confirmar que les malalties mentals són heretables també han demostrat que l'entorn és important, suggerint fins i tot que els factors genètics i ambientals contribueixen per igual a la causació de malalties mentals⁵⁶.

Davant la complexa imatge que sorgeix respecte l'etiologia de les MMG, aquest estudi és un dels primers en investigar patrons de variació genètica comuna enfocat en l'àmbit familiar, amb l'objectiu d'il·lustrar les possibilitats d'aquest enfocament investigant un pedigree multigeneracional afectat amb MMG (SZ, TDM i BP) i per tant millorar la comprensió del risc genètic comú de trastorns psiquiàtrics. D'aquesta manera, hem estat capaços de comprovar que existeix una correlació entre la puntuació de PRS i la severitat del fenotip en famílies amb alta prevalença de MMG. La viabilitat d'aquest enfocament radica en gran part a l'estructura i composició de les nostres famílies d'estudi.

L'aparellament assortatiu (no aleatori) és un fenomen comú en què les persones emparellades són fenotípicament més similars entre sí per a una característica determinada del que s'esperaria que fossin per casualitat⁵⁷. Aquesta és una de les característiques de les nostres famílies d'estudi, on molts individus casats es troben més afectats per alguna malaltia mental que la població general. Aquest fet permet explicar com variants que són comunes a la població general presenten un grau tan alt de responsabilitat genètica en la transmissió de les MMG, ja que amb el pas de les generacions es van acumulant variants genètiques comunes que es troben associades a MMG. Això demostra les possibilitats de les aplicacions PRS en un context familiar, on es permet il·luminar fenòmens dins de la família que estan fora de l'abast dels estudis de casos / control tradicionals.

Emprant diferents estudis d'associació per a MMG (tant per SZ, BD i SZ+BD) s'ha conseguit discriminar entre les diferents puntuacions de PRS afectats per diferents fenotips de MMG i controls, de la mateixa manera que s'havia aconseguit en estudis previs^{58,59}. A més, tot i les diferències de PRS observades entre els diferents grups amb els PRS SZ, PRS BD i PRS SZ+BD, tan sols s'han trobat diferències estadísticament significatives en el PRS SZ+BD, el que confirma la predicció que va comentar Boies et., al 2018 quan explicava que es podrien aconseguir uns millors resultats de PRS a través de la conjunció de PRS SZ i PRS BD. Cal destacar que després d'anitzar els PRS dels diferents subfenotips (SZ, BD, TDM i sans) hem observat PRS BD associat amb SZ en comparació amb els sans i, al contrari, PRS SZ associat amb BD en comparació amb els sans. Aquest punt en comú dóna suport a que nombroses variants de risc de SZ també estiguin associades amb BD i que nombroses variants de risc de BD estiguin alhora associades amb SZ, com s'ha demostrat anteriorment en estudis poblacionals, majoritàriament no familiars (Consortium C-DGotPG, 2013; Purcell et al., 2009)⁶⁰.

Interessantment, vàrem observar que els resultats obtinguts en TDM mostraven evidència de clares puntuacions de PRS distintives entre dos subgrups de subjectes depressius. Això ens va portar a realitzar una anàlisi on dividíem TDM en TDMA i TDMB en base al valor PRS obtingut on es veu com TDMA mostra puntuacions de PRS properes als fenotips més greus (SZ i BD) tant en PRS SZ com en PRS BD i PRS SZ+B. Això confirma que TDM (almenys un subgrup d'aquests subjectes) presenta també

nombroses variants associades a les malalties com SZ i BD que els diferencia dels controls, com ja s'ha vist en estudis previs^{37,61}. A més, s'han trobat diferències estadísticament significatives entre TDMA vs. TDMB, el que explica l'existència de subfenotips dins la depressió en els subjectes de la Família 2. De fet, Power et., al (2017) va demostrar mitjançant puntuacions poligèniques que el TDM pot presentar diferent solapament genètic en funció de l'edat d'aparició de la malaltia, suggerint que la susceptibilitat genètica al TDM difereix entre el TDM d'aparició adulta i el d'aparició primerenca, presentant una major superposició genètica amb SZ i BD en l'aparició primerenca envers el TDM d'aparició adulta³⁷. Per contra, els nostres resultats il·lustren una idea oposada a la que Power et, al. van explicar, ja que són els subjectes d'aparició adulta de la malaltia els que presenten una superposició genètica més elevada amb els fenotips més severs -Veure Taula 3-. En front d'aquests resultats sorgeixen multitud d'hipòtesis possibles que explicarien aquests resultats, però l'explicació més probable és el fet que s'estiguin avaluant diferents subtipus del TDM, ja que es tracta d'una malaltia molt heterogènia^{37,62}. D'aquesta manera, és possible que cada subtipus de depressió integrat dins el grup TDM mostri càrregues de variants genètiques comunes diferents entre sí, el que explicaria els dos subgrups obtinguts.

Tot i els resultats prometedors obtinguts en aquest estudi, cal tenir en compte una sèrie de limitacions d'aquest a l'hora d'extreure conclusions clares respecte els objectius plantejats. Un dels principals punts febles d'aquest treball és la petita mida mostral (n=43), en part degut a les dificultats que suposa reclutar famílies tan nombroses amb alta prevalença de MMG, ja que nombrosos estudis expliquen la importància d'una mida mostral gran en estudis de PRS per tal d'aconseguir nivells útils de predicció^{53,32, 39}. De fet, si comparem els resultats obtinguts en la família 1 (n=12) i la família 2 (n=31), veiem com la segona (més gran) pareix explicar millor la relació entre els casos i els controls ja que mostra uns resultats més fidels als obtinguts en estudis anteriors^{63,54}. Tot i així, quan hem estudiat les dades conjuntes d'ambdues famílies (n=43), no s'han trobat diferències significatives entre els fenotips clínics i els grups control tot i que els grups es distribueixen de la mateixa manera que en la Família 2. Per tant, és necessari ser prudents a l'hora d'interpretar i confirmar els resultats obtinguts en aquest estudi. De la mateixa manera, hem de tenir en compte que de cara a futures investigacions seria important incrementar la mida mostral per tal de refinar el poder predictor del PRS en

famílies amb elevada prevalença de MMG. A més, per poder establir un bon criteri de predicció de risc de patir una MMG és important tenir en compte l'efecte de molts altres factors com pugui ser l'efecte de les variants rares associades a MMG, la presència de malalties associades a comorbiditat (com puguin ser les malalties cardiovasculars o canceroses)^{24,2} i multitud de factors ambientals que ens poden predisposar a desenvolupar una MMG. En un futur, tenint en compte tots aquests factors de forma conjunta, persones que es realitzin les anàlisis de PRS podran beneficiar-se de discussions sobre aquesta puntuació de risc poligènic amb els seus metges i assessors genètics per a avaluacions sanitàries posteriors.

6. CONCLUSIONS

- 1) PRS es correlaciona amb la severitat del fenotip de MMG en famílies amb alta prevalença de MMG.
- 2) SZ, BD i TDM comparteixen un alt nombre de variants associades a MMG.
- 3) Tant el fenotip psicòtic (SZ + BD) com el fenotip ample (psicòtics i TDM) permeten trobar diferències de PRS entre afectats i controls.
- 4) Emprar el GWAS combinat amb variants associades a SZ + BD permet trobar més diferències significatives entre afectats i controls i entre diferents fenotips d'afectats.
- 5) TDM en la Família 2 és un grup heterogeni amb almenys 2 subtipus de TDM amb diferents puntuacions de PRS.
- 6) Es necessiten estudis amb una mida mostral més gran per seguir investigant el PRS com a possible biomarcador pel diagnòstic i predicció de les MMG en l'àmbit familiar.

BIBLIOGRAFIA

1. Correll, C. U. *et al.* Prevalence, incidence and mortality from cardiovascular disease in patients with pooled and specific severe mental illness: a large-scale meta-analysis of 3,211,768 patients and 113,383,368 controls. *World Psychiatry* **16**, 163-180 (2017).
2. Whiteford, H. A. *et al.* Global burden of disease attributable to mental and substance use disorders: Findings from the Global Burden of Disease Study 2010. *Lancet* **382**, 1575-1586 (2013).
3. Sullivan, P. F. & Geschwind, D. H. Defining the Genetic, Genomic, Cellular, and Diagnostic Architectures of Psychiatric Disorders. *Cell* **177**, 162-183 (2019).
4. Taylor, J. H., Calkins, M. E. & Gur, R. E. Markers of Psychosis Risk in the General Population. *Biol. Psychiatry* **88**, 337-348 (2020).
5. Ministerio de Sanidad Servicios Sociales e Igualdad. Guía de Práctica Clínica sobre el manejo de la depresión en el adulto. *Minist. Sanidad, Serv. Soc. e Igualdad. Agencia Evaluación Tecnol. Sanit. Galicia* **61**, 38-41 (2014).
6. Castillejos, M. C., Martín-Pérez, C. & Moreno-Küstner, B. Incidence of psychotic disorders and its association with methodological issues. A systematic review and meta-analyses. *Schizophr. Res.* **204**, 458-459 (2019).
7. Haro, J. M. *et al.* Prevalencia Trastornos Mentales España 2006. *Med Clin* **126**, 445-451[visto23.10.2020] (2006).
8. Depression. https://www.who.int/health-topics/depression#tab=tab_1.
9. Lustman, P. J., Griffith, L. S., Gavard, J. A. & Clouse, R. E. Depression in adults with diabetes. *Diabetes Care* **15**, 1631-1639 (1992).
10. Nöthen, M. M., Nieratschker, V., Cichon, S. & Rietschel, M. New findings in the genetics of major psychoses. *Dialogues Clin. Neurosci.* **12**, 85-93 (2010).

11. Lee, S. H. *et al.* Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat. Genet.* **45**, 984-994 (2013).
12. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®) - American Psychiatric Association - Google Libros.
<https://books.google.es/books?hl=es&lr=&id=-JivBAAAQBAJ&oi=fnd&pg=PT15&dq=American+Psychiatric+Association.+Diagnostic+and+statistical+manual+of+mental+disorders,+4th+ed.+Washington,+DC:+American+Psychiatric+Press,+1994.&ots=cfPM24IKua&sig=s8d8k8MQ4nBnHEEJxXZz3SfogTI#v=onepage&q&f=false>.
13. Habtewold, T. D. *et al.* A systematic review and narrative synthesis of data-driven studies in schizophrenia symptoms and cognitive deficits. *Transl. Psychiatry* **10**, (2020).
14. Craddock, N., O'Donovan, M. C. & Owen, M. J. The genetics of schizophrenia and bipolar disorder: Dissecting psychosis. *J. Med. Genet.* **42**, 193-204 (2005).
15. Saha, S., Chant, D., Welham, J. & McGrath, J. A systematic review of the prevalence of schizophrenia. *PLoS Med.* **2**, 0413-0433 (2005).
16. McCutcheon, R. A., Reis Marques, T. & Howes, O. D. Schizophrenia - An Overview. *JAMA Psychiatry* **77**, 201-210 (2020).
17. Carver, L. A., City, I., Andreasen, N. C. & Iowa, P. The lifetime trajectory of schizophrenia and the concept of neurodevelopment. *Dialogues Clin. Neurosci.* **12**, 409-415 (2010).
18. Hjorthøj, C., Stürup, A. E., McGrath, J. J. & Nordentoft, M. Years of potential life lost and life expectancy in schizophrenia: a systematic review and meta-analysis. *The Lancet Psychiatry* **4**, 295-301 (2017).
19. Vigo, D., Thornicroft, G. & Atun, R. Estimating the true global burden of mental illness. *The Lancet Psychiatry* **3**, 171-178 (2016).
20. Geddes, J. R. & Briess, D. Bipolar disorder. *BMJ Clin. Evid.* **2007**, 58-66 (2007).

21. Merikangas, K. R. *et al.* Prevalence and correlates of bipolar spectrum disorder in the World Mental Health Survey Initiative. *Arch. Gen. Psychiatry* **68**, 241-251 (2011).
22. Miskowiak, K. W. *et al.* Assessing and addressing cognitive impairment in bipolar disorder: the International Society for Bipolar Disorders Targeting Cognition Task Force recommendations for clinicians. *Bipolar Disord.* **20**, 184-194 (2018).
23. Plans, L. *et al.* Association between completed suicide and bipolar disorder: A systematic review of the literature. *J. Affect. Disord.* **242**, 111-122 (2019).
24. Krishnan, K. R. R. Psychiatric and medical comorbidities of bipolar disorder. *Psychosom. Med.* **67**, 1-8 (2005).
25. Maj, M. «Psychiatric comorbidity»: An artefact of current diagnostic systems? *Br. J. Psychiatry* **186**, 182-184 (2005).
26. Rasic, D., Hajek, T., Alda, M. & Uher, R. Risk of mental illness in offspring of parents with schizophrenia, bipolar disorder, and major depressive disorder: A meta-analysis of family high-risk studies. *Schizophr. Bull.* **40**, 28-38 (2014).
27. Polderman, T. J. C. *et al.* Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat. Genet.* **47**, 702-709 (2015).
28. Tanaka, T. International HapMap project. *Nippon rinsho. Japanese J. Clin. Med.* **63 Suppl 1**, 29-34 (2005).
29. Kruglyak, L. & Nickerson, D. A. Variation is the spice of life. *Nat. Genet.* **27**, 234-236 (2001).
30. Jonas, K. G. *et al.* Schizophrenia polygenic risk score and 20-year course of illness in psychotic disorders. *Transl. Psychiatry* **9**, (2019).
31. Lichtenstein, P. *et al.* Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* **373**, 234-239 (2009).

32. Martin, A. R., Daly, M. J., Robinson, E. B., Hyman, S. E. & Neale, B. M. Predicting Polygenic Risk of Psychiatric Disorders. *Biol. Psychiatry* **86**, 97-109 (2019).
33. Polygenic Risk Scores. <https://www.genome.gov/Health/Genomics-and-Medicine/Polygenic-risk-scores#one>.
34. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565-569 (2010).
35. Lee, S. H. *et al.* Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nat. Genet.* **44**, 247-250 (2012).
36. Kember, R. L. *et al.* Polygenic Risk of Psychiatric Disorders Exhibits Cross-trait Associations in Electronic Health Record Data From European Ancestry Individuals. *Biol. Psychiatry* **89**, 236-245 (2021).
37. Power, R. A. *et al.* Genome-wide Association for Major Depression Through Age at Onset Stratification: Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. *Biol. Psychiatry* **81**, 325-335 (2017).
38. Marshall, C. R. *et al.* Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat. Genet.* **49**, 27-35 (2017).
39. Choi, S. W., Mak, T. S. H. & O'Reilly, P. F. Tutorial: a guide to performing polygenic risk score analyses. *Nat. Protoc.* **15**, 2759-2772 (2020).
40. Moreno, B., García-Alonso, C. R., Negín Hernández, M. A., Torres-González, F. & Salvador-Carulla, L. Spatial analysis to identify hotspots of prevalence of schizophrenia. *Soc. Psychiatry Psychiatr. Epidemiol.* **43**, 782-791 (2008).
41. Progeny Free Online Pedigree Tool - 9.101.9047.00.
<https://pedigree.progenygenetics.com/>.
42. Mn, W. *et al.* Use of SNP chips to detect rare pathogenic variants: retrospective, population based diagnostic evaluation. *BMJ* **372**, n214 (2021).

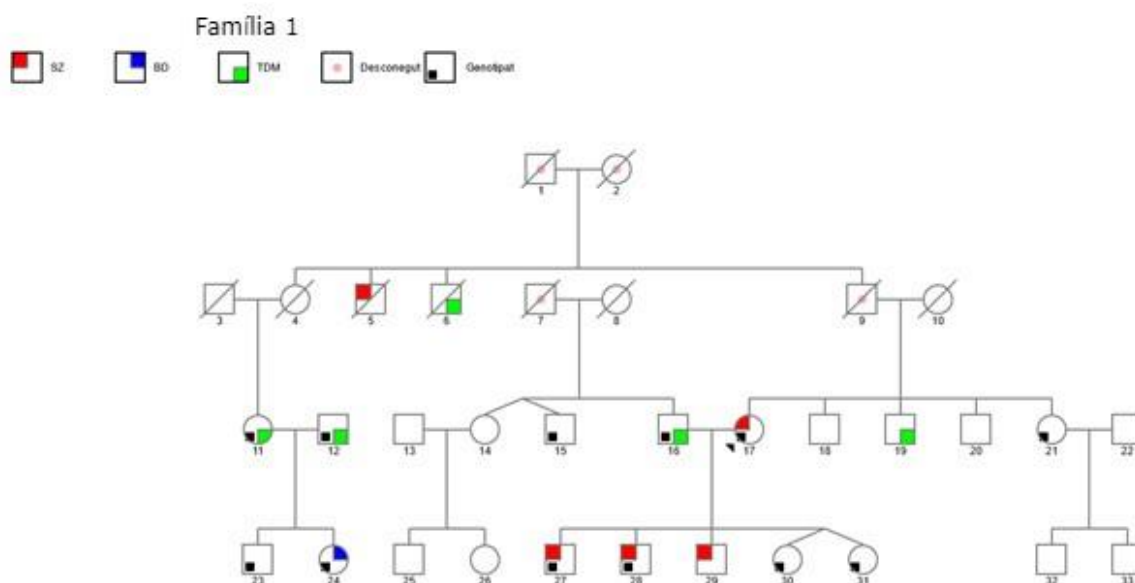
43. Smart, A., Bolnick, D. A. & Tutton, R. Health and genetic ancestry testing: Time to bridge the gap. *BMC Med. Genomics* **10**, 1-9 (2017).
44. Price, A. L., Spencer, C. C. A. & Donnelly, P. Progress and promise in understanding the genetic basis of common diseases. *Proc. R. Soc. B Biol. Sci.* **282**, (2015).
45. Steemers, F. J. & Gunderson, K. L. Whole genome genotyping technologies on the BeadArray™ platform. *Biotechnol. J.* **2**, 41-49 (2007).
46. Purcell, S. *et al.* PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559-575 (2007).
47. Centimorgan | NHGRI. <https://www.genome.gov/es/genetics-glossary/Centimorgan>.
48. Anderson, C. A. *et al.* Data quality control in genetic case-control association studies. *Nat. Protoc.* **5**, 1564-1573 (2010).
49. Euesden, J., Lewis, C. M. & O'Reilly, P. F. PRSice: Polygenic Risk Score software. *Bioinformatics* **31**, 1466-1468 (2015).
50. Wang, J., Samuels, D. C., Shyr, Y. & Guo, Y. StrandScript: Evaluation of Illumina genotyping array design and strand correction. *Bioinformatics* **33**, 2399-2401 (2017).
51. Reininghaus, U. *et al.* Evaluation of the validity and utility of a transdiagnostic psychosis dimension encompassing schizophrenia and bipolar disorder. *Br. J. Psychiatry* **209**, 107-113 (2016).
52. Ripke, S. *et al.* Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-427 (2014).
53. Dudbridge, F. Power and Predictive Accuracy of Polygenic Risk Scores. *PLoS Genet.* **9**, (2013).

54. Wray, N. R. *et al.* Research Review: Polygenic methods and their application to psychiatric traits. *J. Child Psychol. Psychiatry Allied Discip.* **55**, 1068-1087 (2014).
55. Wager, M. G. T. and J. F. S. 基因的改变 NIH Public Access. *Bone* **23**, 1-7 (2011).
56. Uher, R. & Zwickler, A. *Etiology in psychiatry: embracing the reality of poly-gene-environmental causation of mental illness.*
57. Ries Merikangas, K. & Duan, S. Assortative Mating Among in-Patients with Primary Affective Disorder. *Psychol. Med.* **12**, 753-764 (1982).
58. Boies, S., Mérette, C., Paccalet, T., Maziade, M. & Bureau, A. Polygenic risk scores distinguish patients from non-affected adult relatives and from normal controls in schizophrenia and bipolar disorder multi-affected kindreds. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **177**, 329-336 (2018).
59. Fullerton, J. M. *et al.* Assessment of first and second degree relatives of individuals with bipolar disorder shows increased genetic risk scores in both affected relatives and young At-Risk Individuals. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **168**, 617-629 (2015).
60. Smoller, J. W. *et al.* Identification of risk loci with shared effects on five major psychiatric disorders: A genome-wide analysis. *Lancet* **381**, 1371-1379 (2013).
61. Uher, R. & Zwickler, A. Etiology in psychiatry: embracing the reality of poly-gene-environmental causation of mental illness. *World Psychiatry* **16**, 121-129 (2017).
62. Schwabe, I. *et al.* Unraveling the genetic architecture of major depressive disorder: Merits and pitfalls of the approaches used in genome-wide association studies. *Psychol. Med.* **49**, 2646-2656 (2019).
63. Demirkan, A. *et al.* Genetic risk profiles for depression and anxiety in adult and elderly cohorts. *Mol. Psychiatry* **16**, 773-783 (2011).

ANNEXAS

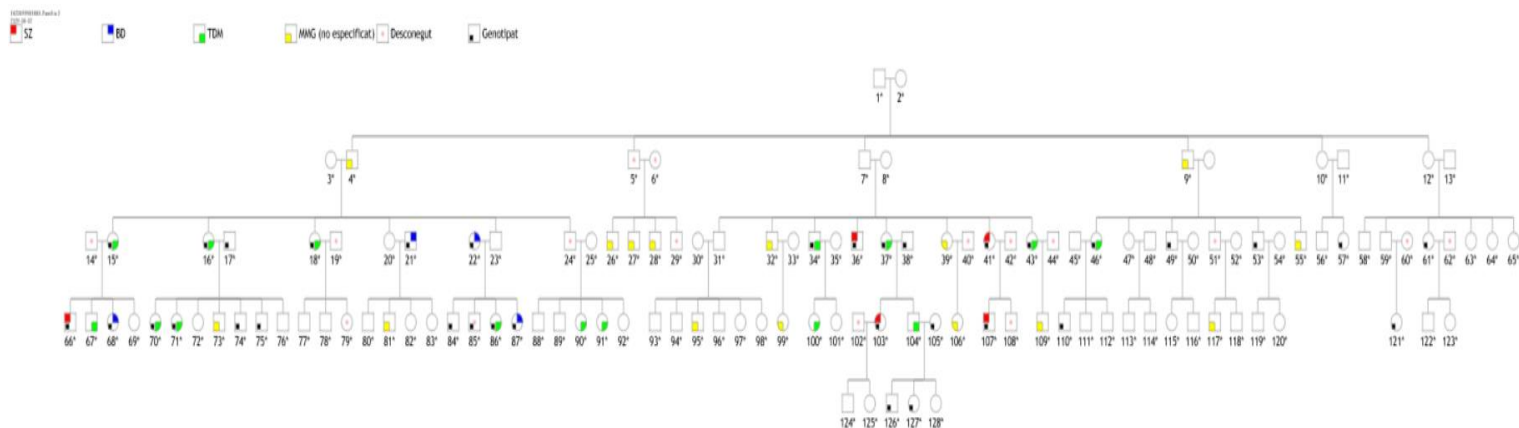
1) Pedigrees:

❖ Família 1



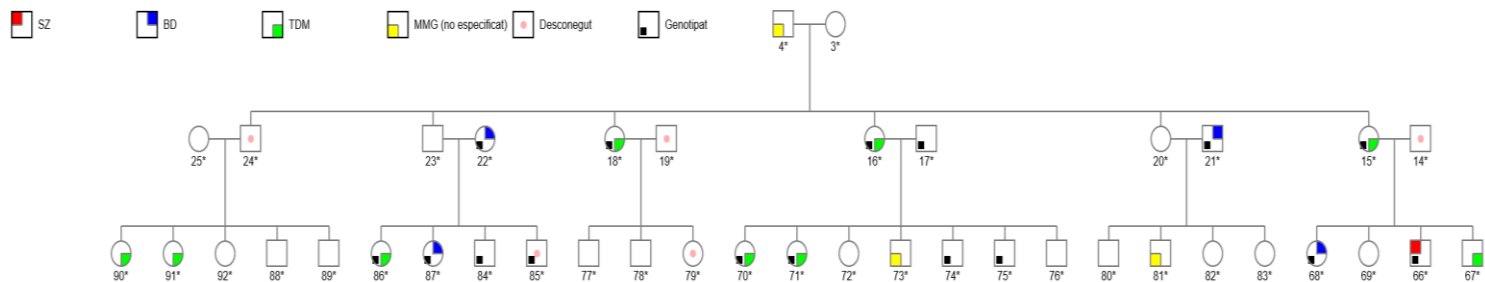
❖ Família 2

Pedigree complet:

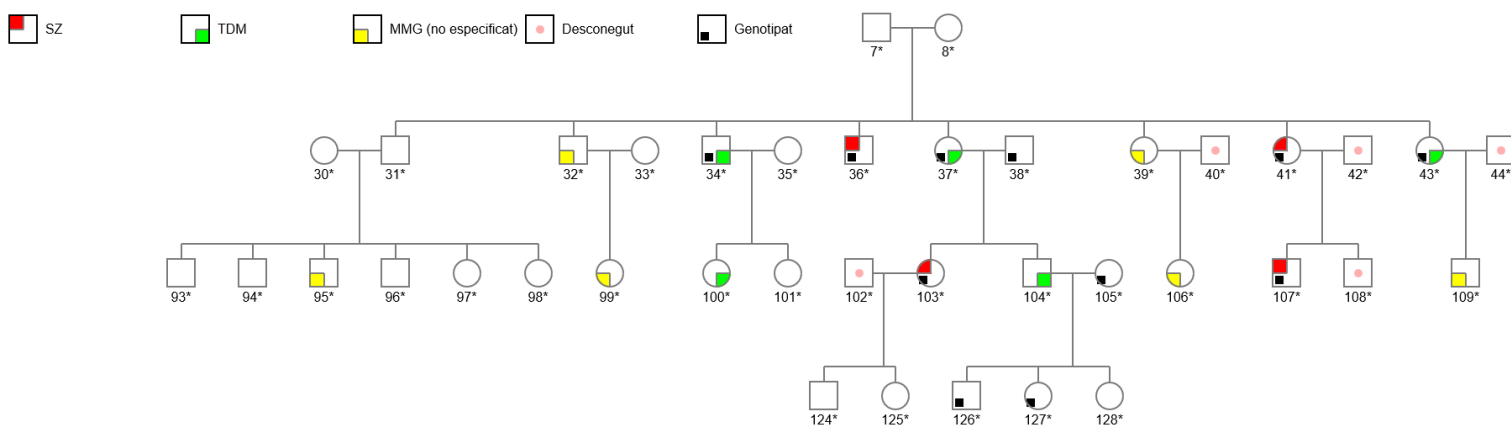


Subpedigrees:

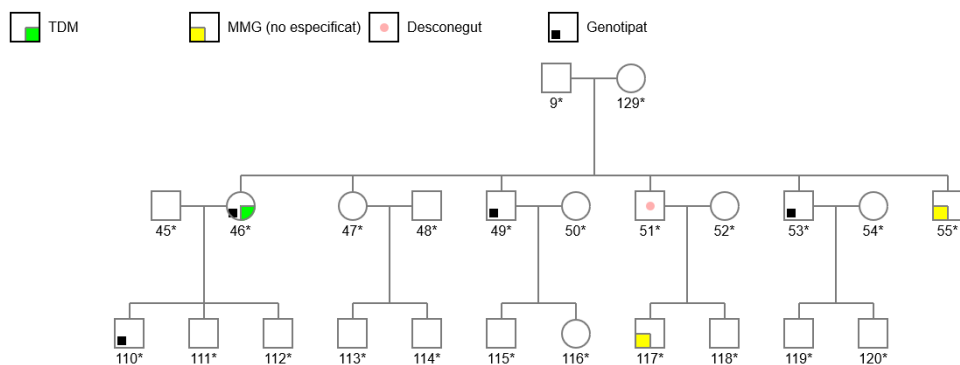
❖ Subfamília 1



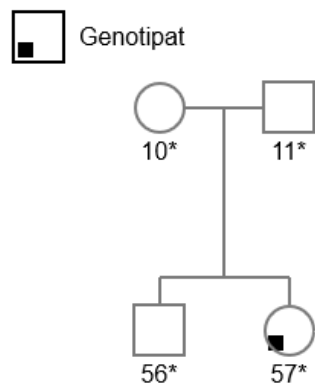
❖ Subfamília 2



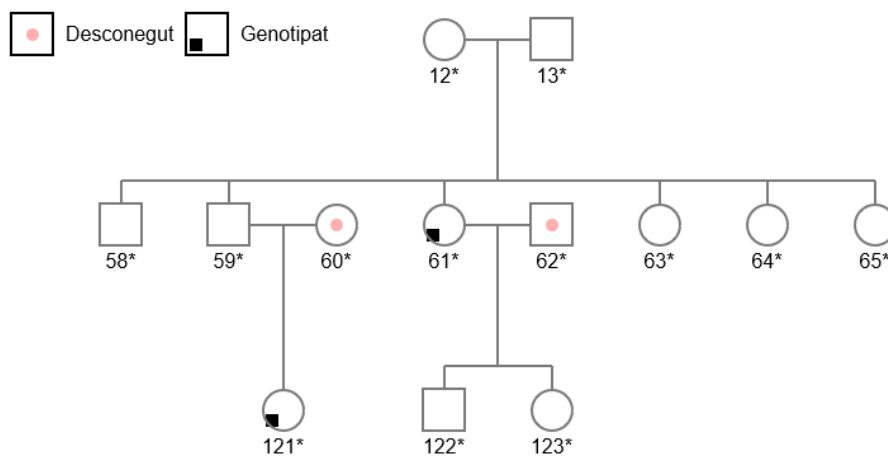
❖ Subfamília 3



❖ Subfamília 4



❖ Subfamília 5



2) Taules:

Família 1				
Individu	PRS SZ	PRS BD	PRS SZ + BD	Diagnosi
28	1,591	-0,4556	-0,7332	Esquizofrènia
17	0,1422	-2,0309	-1,9108	Esquizofrènia
27	0,6246	-0,6962	-0,7661	Esquizofrènia
23	0,3338	0,0124	0,3634	Transtorn bipolar
16	-0,0299	0,6981	0,4085	Transtorn depressiu major
12	-0,0856	-0,3057	0,4555	Transtorn depressiu major
11	-0,2522	1,9398	2,0429	Transtorn depressiu major
21	0,8904	-0,4099	0,6028	Sa
23	0,3424	0,483	0,5582	Sa
30	-0,5092	-0,2002	-0,0454	Sa
31	-0,5092	-0,2002	-0,0454	Sa
15	-2,5386	1,1654	-0,9303	Sa

Taula 1. Resultats del PRSice de la Família 1 per a tres subconjunt diferents de dades base: PRS SZ, PRS BD i PRS SZ+BD. S'inclou el fenotip diagnosticat per als subjectes d'estudi (n=12).

Família 2				
Individu	PRS SZ	PRS BD	PRS SZ + BD	Diagnosi
3.11	2,3379	1,2391	2,2739	Esquizofrènia
3.31	1,0658	1,3082	1,0993	Esquizofrènia
3.12	0,9150	-1,4486	0,2907	Esquizofrènia
1.18	0,9134	0,7052	0,8394	Esquizofrènia
3.30	0,8906	-0,4063	0,6013	Esquizofrènia
1.25	0,6750	1,4662	0,6516	Transtorn Bipolar
1.19	1,8232	2,2005	1,7978	Transtorn bipolar
1.6	0,3816	0,4674	0,1920	Transtorn Bipolar
3.10	0,7766	0,0413	0,8573	Transtorn depressiu major
3.9	0,6209	0,0532	1,1943	Transtorn depressiu major
1.5	0,5914	1,2853	0,9754	Transtorn depressiu major
3.13	0,4116	0,2463	0,7409	Transtorn depressiu major
1.1	0,3768	0,9282	-0,4514	Transtorn depressiu major
1.24	-0,3894	-0,6870	0,1740	Transtorn depressiu major
1.20	-0,9281	0,1212	-1,3203	Transtorn depressiu major
1.22	-0,7116	-0,0507	-1,0344	Transtorn depressiu major
4.15	-0,833	-1,2132	-0,8121	Transtorn depressiu major
1.27	1,0316	1,404	0,3966	Transtorn depressiu major
3.33	-0,4675	-0,5872	-0,1147	Sa
4.14	0,3631	-0,1179	0,1321	Sa
1.26	-0,0121	0,7385	-0,0416	Sa
3.8	-0,6247	-0,8440	-0,7880	Sa
2.7	-0,0579	-0,6625	0,3873	Sa
2.28	-0,8851	-0,0956	-0,8329	Sa
4.16	-0,6305	-1,5320	-0,2497	Sa
4.32	-0,9617	-0,3477	-0,9285	Sa
1.23	-0,9825	-0,0431	-1,1969	Sa
1.4	-1,1031	-0,6543	-1,7319	Sa
1.21	-1,4133	0,2870	-1,9237	Sa
3.29	-1,5000	-1,7090	-0,6281	Sa
3.34	-1,6738	-1,5185	-0,5497	Sa

Taula 2. Resultats del PRSice de la Família 2 per a tres subconjunt diferents de dades base: PRS SZ, PRS BD i PRS SZ+BD. S'inclou el fenotip diagnosticat per als subjectes d'estudi (n=31).

Subjecte	Edat	PRS SZ	PRS BD	PRS SZ + BD
14*	18	0,3768	0,9282	-0,4514
70*	22	-0,9281	0,1212	-1,3203
18*	20	0,5914	1,2853	0,9754
46*	24	-0,833	-1,2132	-0,8121
71*	13	-0,7116	-0,0507	-1,0344
86*	8	-0,3894	-0,6870	0,1740
37*	70	0,6209	0,0532	1,1943
34*	41	0,7766	0,0413	0,8573
43*	>50	0,4116	0,2463	0,7409

Taula 3. Recull de resultats del PRSice de subjectes de diferents edats que pertanyen al grup TDM

3) Figures:

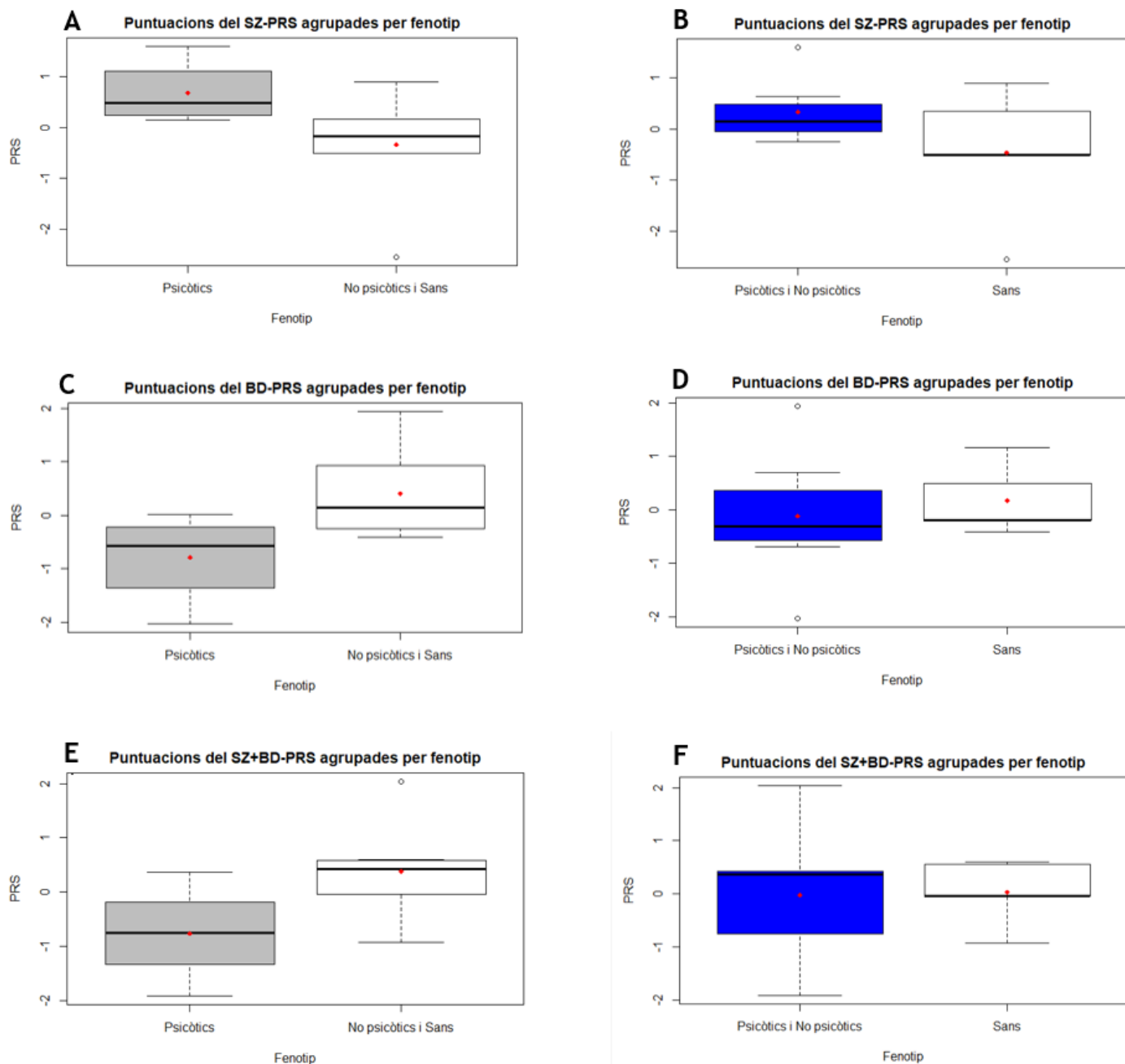


Figura 1. Resultats del PRSice de la Família 1. Es mostren les diferents puntuacions obtingudes en el programa PRSice a través de la prova PRS per als diferents grups de casos (psicòtics o malalts mentals no psicòtics) i controls (sans).

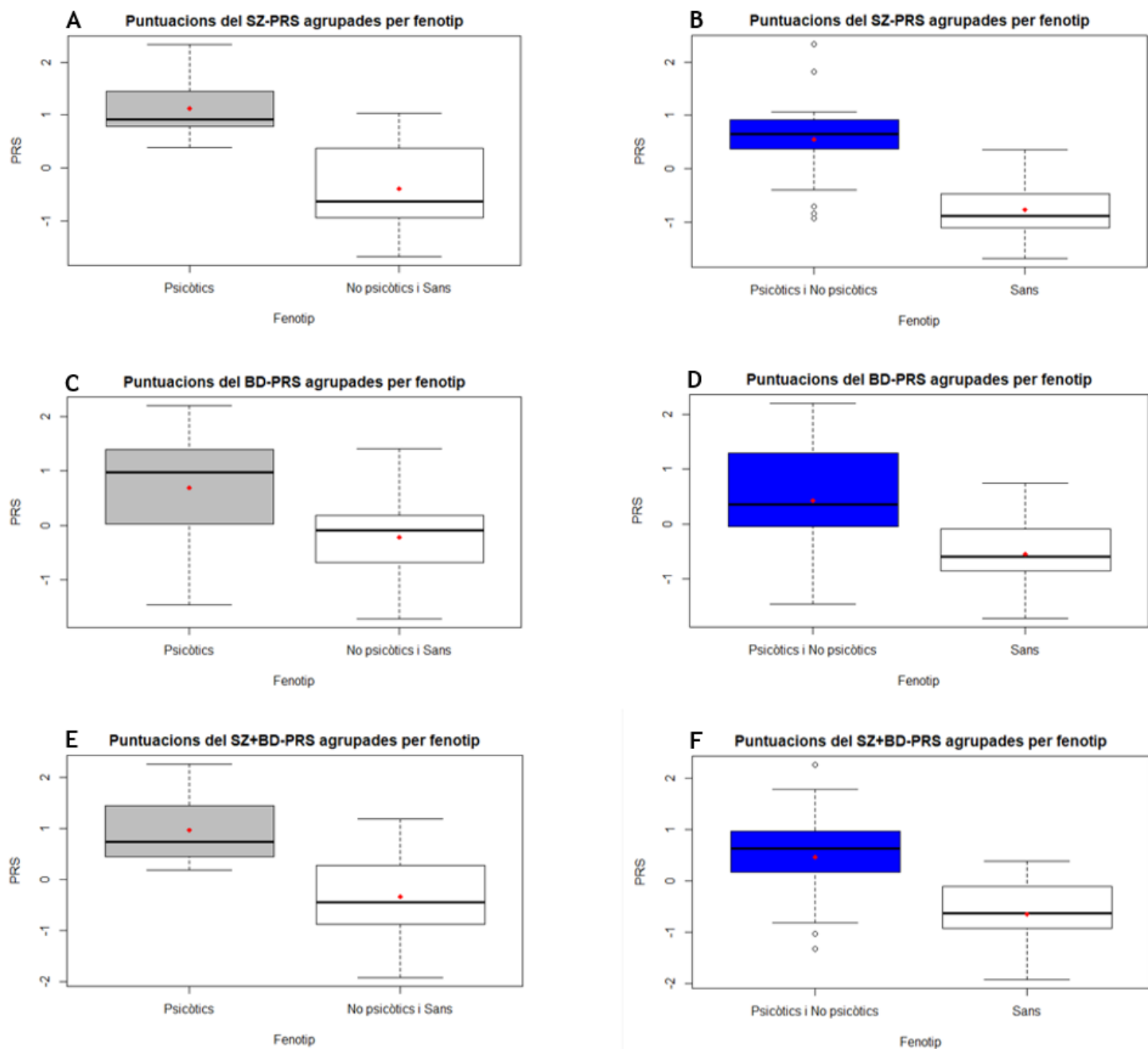


Figura 2. Resultats del PRSice de la Família 2. Es mostren les diferents puntuacions obtingudes en el programa PRSice a través de la prova PRS per als diferents grups de casos (psicòtics o malalts mentals no psicòtics) i controls (sans).

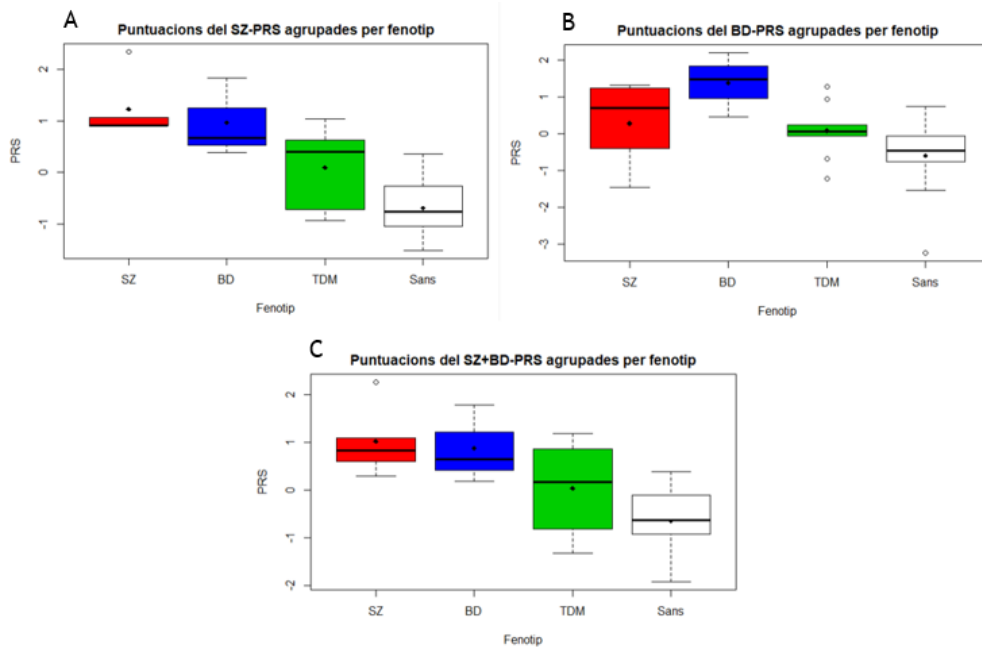


Figura 3. Resultats del PRSice de la Família 2 agrupats per fenotip. Es mostren les diferents puntuacions obtingudes en el programa PRSice a través de la prova PRS per als diferents fenotips dels subjectes malalts (SZ, BD i TDM) i controls (sans).

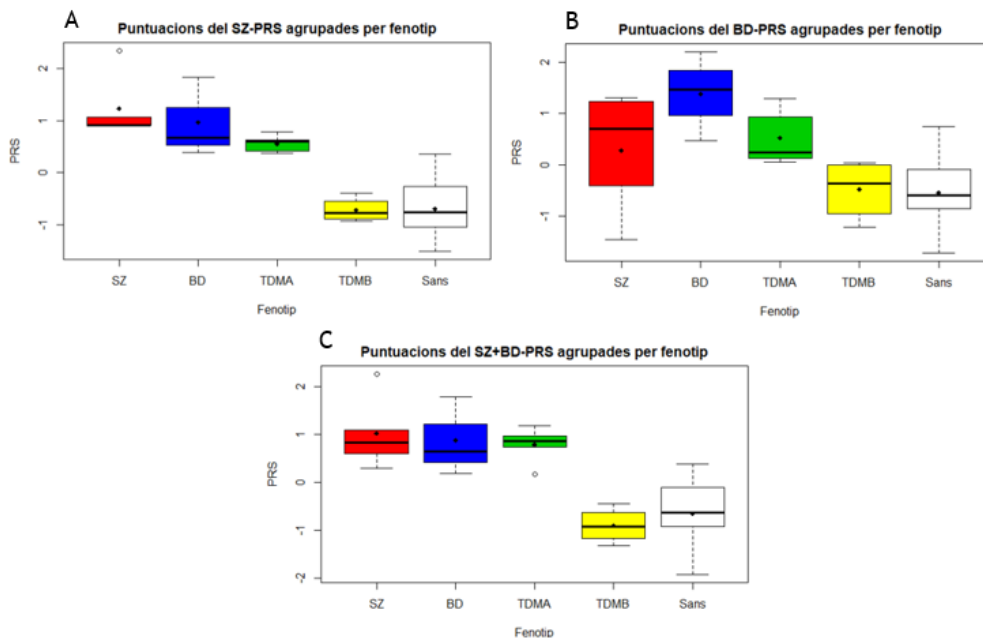


Figura 4. Resultats del PRSice de la Família 2 agrupats per fenotip. Es mostren les diferents puntuacions obtingudes en el programa PRSice a través de la prova PRS per als diferents fenotips dels subjectes malalts i controls (sans) (SZ, BD, TDMA i TDMB) i controls (sans). El fenotip TDM s'ha subdividit en dos grups TDM A (subjectes depressius amb alts valors de PRS) i TDM B (subjectes depressius amb baixos valors de PRS)

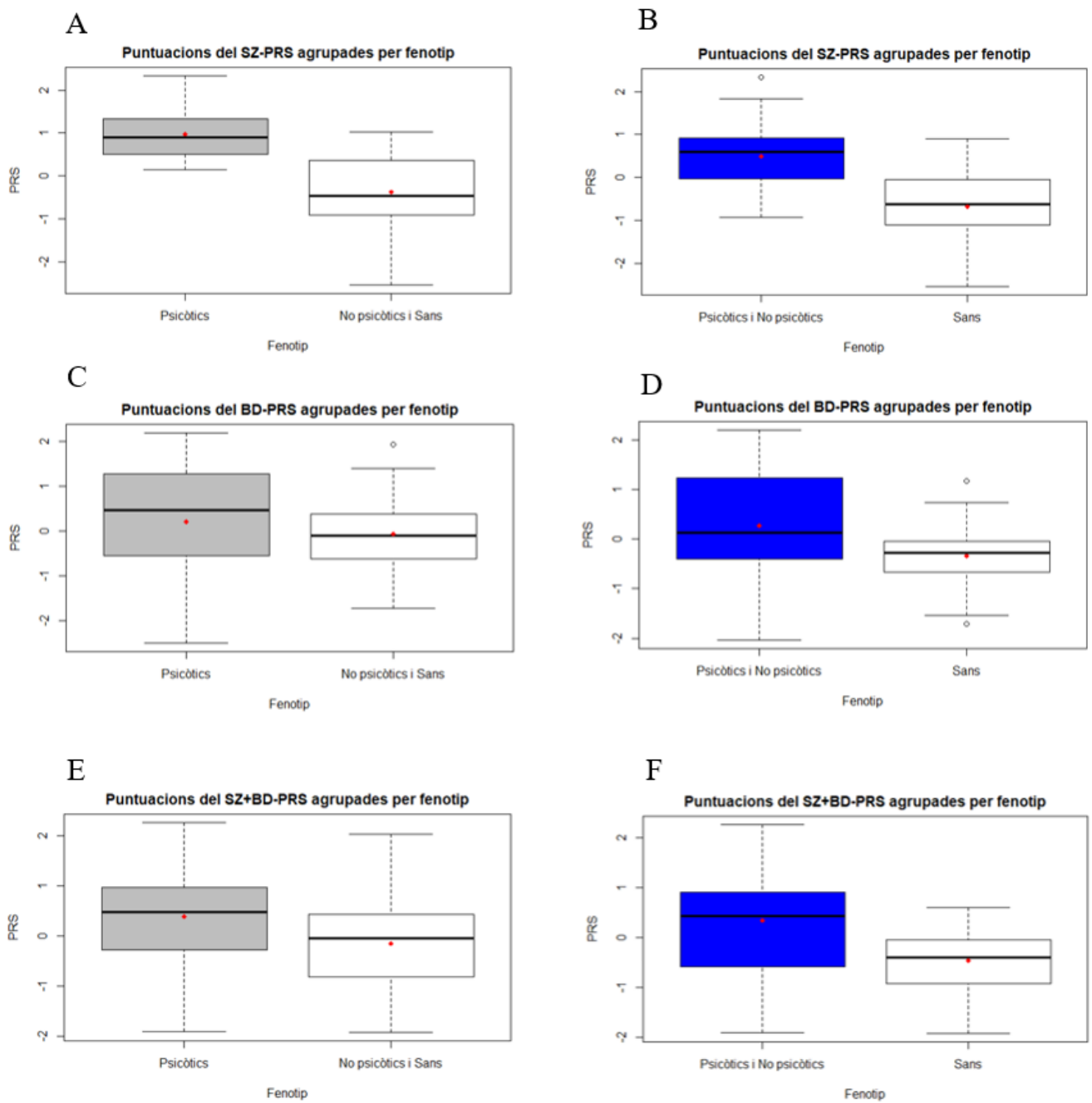


Figura 5. Resultats conjunts del PRSice de les Famílies 1 i 2. Es mostren les diferents puntuacions obtingudes en el programa PRSice a través de la prova PRS per als diferents grups de casos (psicòtics o malalts mentals no psicòtics) i controls (sans) de tots els pacients inclosos en aquest estudi.

4) Estadística:

Família 1

Toni Capó Quetglas

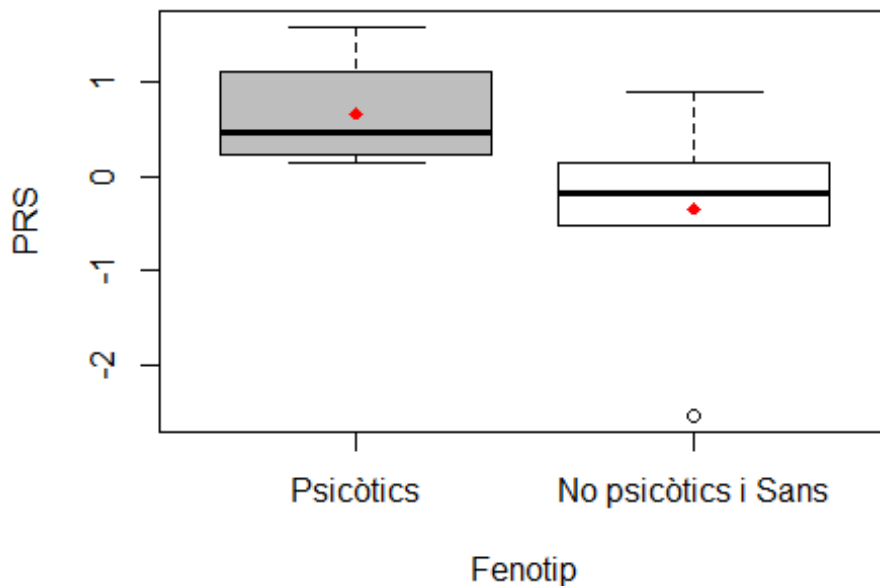
30/5/2021

PRS SZ

1) Fenotip estret: Psicòtics vs No psicòtics i Sans

```
rm(list=ls())
Fenotip_estret<-c( 1.5916,0.1422,0.6246,0.3338)
Sans<-c(-0.0299,-0.0856,-0.2522,0.8904,0.3424,-0.5092,-0.5092,-2.5386)
boxplot(Fenotip_estret,Sans, names=c("Psicòtics","No psicòtics i Sans"),
ylab="PRS", main="Puntuacions del SZ-PRS agrupades per fenotip",xlab="Fenotip",col=c("grey", "white"))
points(1,mean(Fenotip_estret),col="red",pch=18)
points(2,mean(Sans),col="red",pch=18)
```

Puntuacions del SZ-PRS agrupades per fenotip



```
rm(list=ls())
Fenotip_estret<-c( 1.5916,0.1422,0.6246,0.3338,NA,NA,NA,NA)
Sans<-c(-0.0299,-0.0856,-0.2522,0.8904,0.3424,-0.5092,-0.5092,-2.5386)
```

```
Dades_SZ_estret<-data.frame(Fenotip_estret,Sans)
shapiro.test(Fenotip_estret)
```

```
##
## Shapiro-Wilk normality test
##
## data: Fenotip_estret
## W = 0.8777, p-value = 0.3289
```

```
shapiro.test(Sans)
```

```
##
## Shapiro-Wilk normality test
##
## data: Sans
## W = 0.84402, p-value = 0.08282
```

```
stacked<-stack(Dades_SZ_estret)
bartlett.test(values~ind,data=stacked)
```

```
##
## Bartlett test of homogeneity of variances
##
## data: values by ind
## Bartlett's K-squared = 0.63926, df = 1, p-value = 0.424
```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variances, pel que aplicarem el test paramètric "t.test"

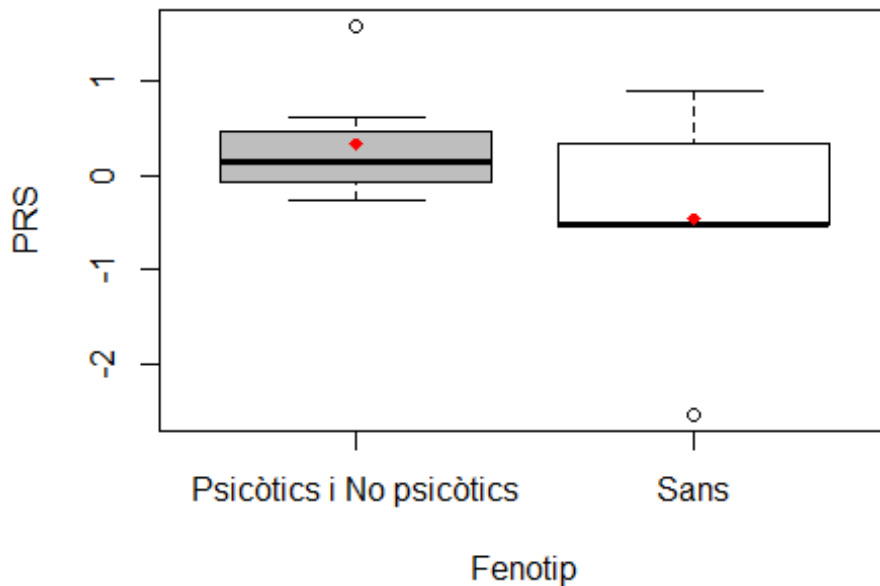
```
t.test(values~ind, data=stacked, var.equal=T)
```

```
##
## Two Sample t-test
##
## data: values by ind
## t = 1.8115, df = 10, p-value = 0.1002
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.2321966 2.2512716
## sample estimates:
## mean in group Fenotip_estret          mean in group Sans
##                0.6730500                -0.3364875
```

2) Fenotip ample: Psicòtics i No psicòtics vs Sans

```
rm(list=ls())
Fenotip_ample<-c(1.5916,0.1422,0.6246,0.3338,-0.0299,-0.0856,-0.2522)
Sans<-c(0.8904,0.3424,-0.5092,-0.5092,-2.5386)
boxplot(Fenotip_ample,Sans, names=c("Psicòtics i No psicòtics","Sans"), y
lab="PRS",main="Puntuacions del SZ-PRS agrupades per fenotip", xlab="Feno
tip",col=c("grey", "white"))
points(1,mean(Fenotip_ample),col="red",pch=18)
points(2,mean(Sans),col="red",pch=18)
```

Puntuacions del SZ-PRS agrupades per fenotip



```
rm(list=ls())
Fenotip_ample<-c(1.5916,0.1422,0.6246,0.3338,-0.0299,-0.0856,-0.2522)
Sans<-c(0.8904,0.3424,-0.5092,-0.5092,-2.5386,NA,NA)
Dades_SZ_ample<-data.frame(Fenotip_ample,Sans)
shapiro.test(Fenotip_ample)

##
## Shapiro-Wilk normality test
##
## data: Fenotip_ample
## W = 0.85074, p-value = 0.1248

shapiro.test(Sans)

##
## Shapiro-Wilk normality test
##
## data: Sans
## W = 0.90879, p-value = 0.4604

stacked<-stack(Dades_SZ_ample)
bartlett.test(values~ind,data=stacked)

##
## Bartlett test of homogeneity of variances
##
## data: values by ind
## Bartlett's K-squared = 2.354, df = 1, p-value = 0.125
```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variàncies, pel que aplicarem el test paramètric "t.test"

```
t.test(values~ind, data=stacked, var.equal=T)

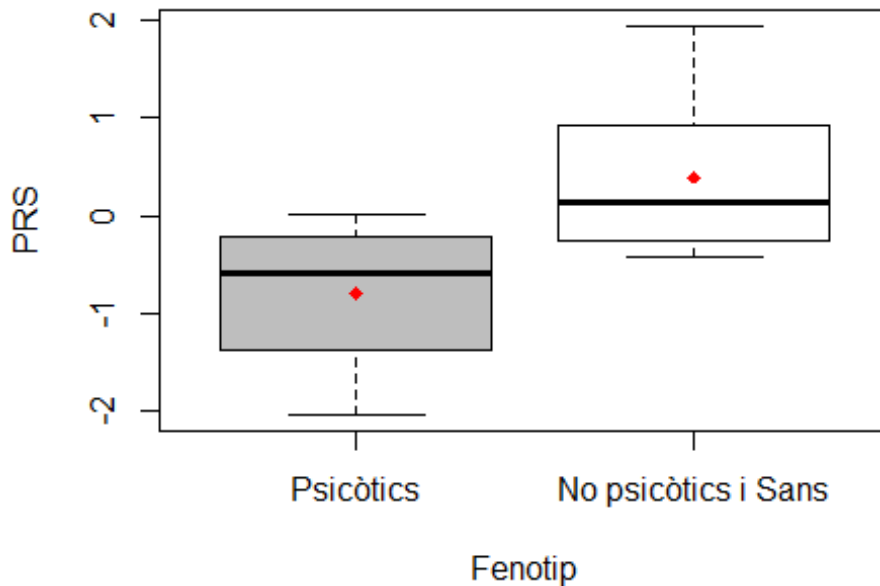
##
## Two Sample t-test
##
## data: values by ind
## t = 1.423, df = 10, p-value = 0.1852
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.4509109 2.0447337
## sample estimates:
## mean in group Fenotip_ample          mean in group Sans
##                0.3320714                -0.4648400
```

PRS BD

1) Fenotip estret: Psicòtics vs No psicòtics i Sans

```
rm(list=ls())
Fenotip_estret<-c(-0.4556,-2.0309,-0.6962,0.0124)
Sans<-c(0.6981,-0.3057,1.9398,-0.4099,0.483,-0.2002,-0.2002,1.1654)
boxplot(Fenotip_estret,Sans, names=c("Psicòtics","No psicòtics i Sans"),
ylab="PRS",main="Puntuacions del BD-PRS agrupades per fenotip", xlab="Fenotip",col=c("grey", "white"))
points(1,mean(Fenotip_estret),col="red",pch=18)
points(2,mean(Sans),col="red",pch=18)
```

Puntuacions del BD-PRS agrupades per fenotip



```
rm(list=ls())
Fenotip_estret<-c(-0.4556, -2.0309, -0.6962, 0.0124, NA, NA, NA, NA)
Sans<-c(0.6981, -0.3057, 1.9398, -0.4099, 0.483, -0.2002, -0.2002, 1.1654)
Dades_BD_estret<-data.frame(Fenotip_estret, Sans)
shapiro.test(Fenotip_estret)

##
## Shapiro-Wilk normality test
##
## data: Fenotip_estret
## W = 0.90529, p-value = 0.4578

shapiro.test(Sans)

##
## Shapiro-Wilk normality test
##
## data: Sans
## W = 0.88058, p-value = 0.1908

stacked<-stack(Dades_BD_estret)
bartlett.test(values~ind, data=stacked)

##
## Bartlett test of homogeneity of variances
##
## data: values by ind
## Bartlett's K-squared = 0.0072013, df = 1, p-value = 0.9324
```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variàncies, pel que aplicarem el test paramètric "t.test"

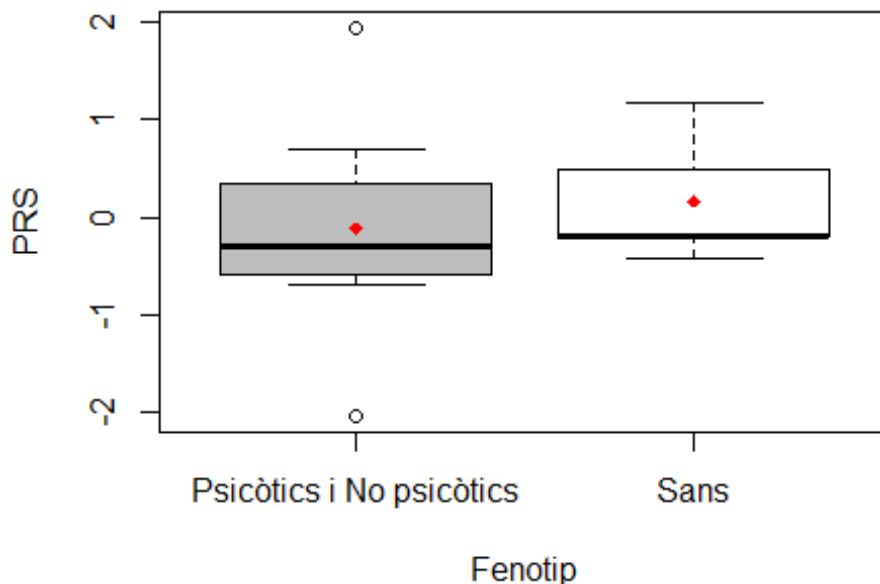
```
t.test(values~ind, data=stacked, var.equal=T)

##
## Two Sample t-test
##
## data: values by ind
## t = -2.283, df = 10, p-value = 0.04555
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -2.34914064 -0.02858436
## sample estimates:
## mean in group Fenotip_estret          mean in group Sans
##                -0.7925750                0.3962875
```

2) Fenotip ample: Psicòtics i No psicòtics vs Sans

```
rm(list=ls())
Fenotip_ample<-c(-0.4556,-2.0309,-0.6962,0.0124,0.6981,-0.3057,1.9398)
Sans<-c(-0.4099,0.483,-0.2002,-0.2002,1.1654)
boxplot(Fenotip_ample,Sans, names=c("Psicòtics i No psicòtics","Sans"), y
lab="PRS",main="Puntuacions del BD-PRS agrupades per fenotip", xlab="Feno
tip",col=c("grey", "white"))
points(1,mean(Fenotip_ample),col="red",pch=18)
points(2,mean(Sans),col="red",pch=18)
```

Puntuacions del BD-PRS agrupades per fenotip



```

rm(list=ls())
Fenotip_ample<-c(-0.4556,-2.0309,-0.6962,0.0124,0.6981,-0.3057,1.9398)
Sans<-c(-0.4099,0.483,-0.2002,-0.2002,1.1654,NA,NA)
Dades_BD_ample<-data.frame(Fenotip_ample,Sans)
shapiro.test(Fenotip_ample)

##
## Shapiro-Wilk normality test
##
## data: Fenotip_ample
## W = 0.96881, p-value = 0.8898

shapiro.test(Sans)

##
## Shapiro-Wilk normality test
##
## data: Sans
## W = 0.86438, p-value = 0.2444

stacked<-stack(Dades_BD_ample)
bartlett.test(values~ind,data=stacked)

##
## Bartlett test of homogeneity of variances
##
## data: values by ind
## Bartlett's K-squared = 1.523, df = 1, p-value = 0.2172

```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variàncies, pel que aplicarem el test paramètric “t.test”

```

t.test(values~ind, data=stacked, var.equal=T)

##
## Two Sample t-test
##
## data: values by ind
## t = -0.47311, df = 10, p-value = 0.6463
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -1.640644 1.065947
## sample estimates:
## mean in group Fenotip_ample      mean in group Sans
##          -0.1197286                0.1676200

```

PRS SZ+BD

1) Fenotip estret: Psicòtics vs No psicòtics i Sans

```

rm(list=ls())
Fenotip_estret<-c(-0.7332,-1.9108,-0.7661,0.3634)
Sans<-c(0.4085,0.4555,2.0429,0.6028,0.5582,-0.0454,-0.0454,-0.9303 )

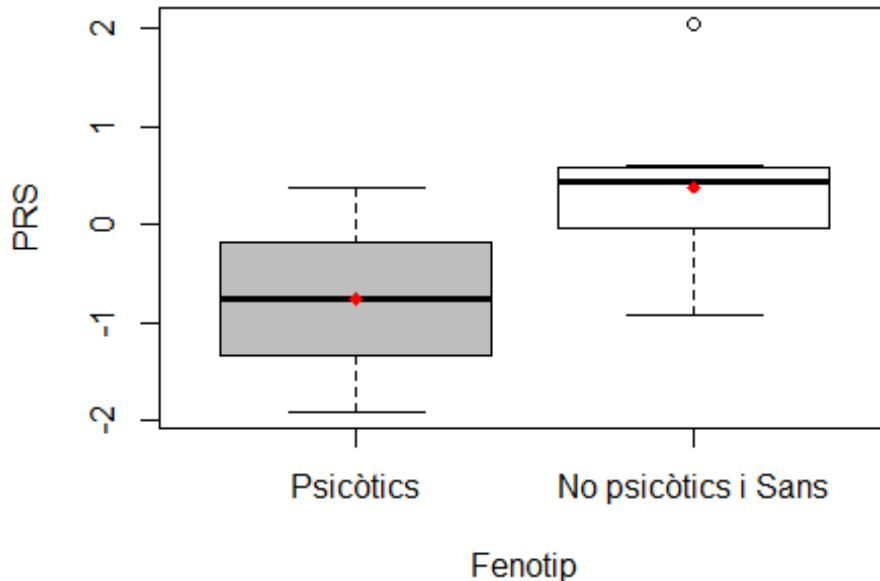
```

```

boxplot(Fenotip_estret,Sans, names=c("Psicòtics","No psicòtics i Sans"),
ylab="PRS", main="Puntuacions del SZ+BD-PRS agrupades per fenotip",xlab="
Fenotip",col=c("grey", "white"))
points(1,mean(Fenotip_estret),col="red",pch=18)
points(2,mean(Sans),col="red",pch=18)

```

Puntuacions del SZ+BD-PRS agrupades per fenotip



```

rm(list=ls())
Fenotip_estret<-c(-0.7332,-1.9108,-0.7661,0.3634,NA,NA,NA,NA)
Sans<-c(0.4085,0.4555,2.0429,0.6028,0.5582,-0.0454,-0.0454,-0.9303 )
Dades_SZBD_estret<-data.frame(Fenotip_estret,Sans)
shapiro.test(Fenotip_estret)

##
## Shapiro-Wilk normality test
##
## data: Fenotip_estret
## W = 0.95088, p-value = 0.7216

shapiro.test(Sans)

##
## Shapiro-Wilk normality test
##
## data: Sans
## W = 0.90471, p-value = 0.3183

stacked<-stack(Dades_SZBD_estret)
bartlett.test(values~ind,data=stacked)

```



```
##
## Bartlett test of homogeneity of variances
##
## data: values by ind
## Bartlett's K-squared = 0.0391, df = 1, p-value = 0.8433
```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variàncies, pel que aplicarem el test paramètric "t.test"

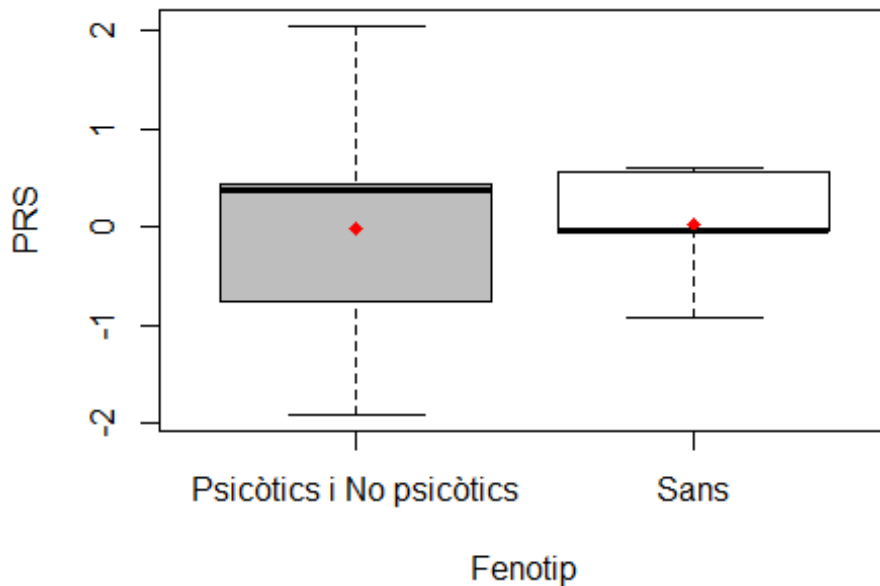
```
t.test(values~ind, data=stacked, var.equal=T)

##
## Two Sample t-test
##
## data: values by ind
## t = -2.1516, df = 10, p-value = 0.05691
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -2.32567515 0.04062515
## sample estimates:
## mean in group Fenotip_estret          mean in group Sans
##                -0.761675                0.380850
```

2) Fenotip ample: Psicòtics i No psicòtics vs Sans

```
rm(list=ls())
Fenotip_ample<-c(-0.7332,-1.9108,-0.7661,0.3634,0.4085,0.4555,2.0429)
Sans<-c(0.6028,0.5582,-0.0454,-0.0454,-0.9303)
boxplot(Fenotip_ample,Sans, names=c("Psicòtics i No psicòtics","Sans"), y
lab="PRS",main="Puntuacions del SZ+BD-PRS agrupades per fenotip", xlab="F
enotip",col=c("grey", "white"))
points(1,mean(Fenotip_ample),col="red",pch=18)
points(2,mean(Sans),col="red",pch=18)
```

Puntuacions del SZ+BD-PRS agrupades per fenotip



```
rm(list=ls())
Fenotip_ample<-c(-0.7332,-1.9108,-0.7661,0.3634,0.4085,0.4555,2.0429)
Sans<-c(0.6028,0.5582,-0.0454,-0.0454,-0.9303,NA,NA)
Dades_SZBD_ample<-data.frame(Fenotip_ample,Sans)
shapiro.test(Fenotip_ample)

##
## Shapiro-Wilk normality test
##
## data: Fenotip_ample
## W = 0.94885, p-value = 0.7193

shapiro.test(Sans)

##
## Shapiro-Wilk normality test
##
## data: Sans
## W = 0.88042, p-value = 0.3112

stacked<-stack(Dades_SZBD_ample)
bartlett.test(values~ind,data=stacked)

##
## Bartlett test of homogeneity of variances
##
## data: values by ind
## Bartlett's K-squared = 1.8418, df = 1, p-value = 0.1747
```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variàncies, pel que aplicarem el test paramètric "t.test"

```
t.test(values~ind, data=stacked, var.equal=T)

##
## Two Sample t-test
##
## data: values by ind
## t = -0.078105, df = 10, p-value = 0.9393
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -1.415880 1.319977
## sample estimates:
## mean in group Fenotip_ample          mean in group Sans
##                -0.01997143                0.02798000
```

Família 2

Toni Capó Quetglas

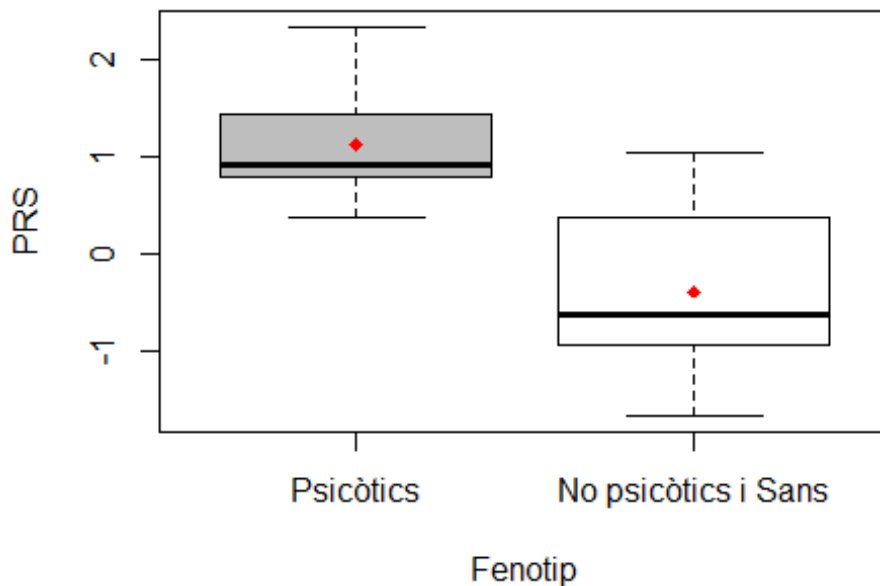
31/5/2021

PRS SZ

1) Fenotip estret: Psicòtics vs No psicòtics i Sans

```
rm(list=ls())
Fenotip_estret<-c(2.3379,1.0658,0.9150,0.9134,0.8906,0.6750,1.8232,0.3816
)
Sans<-c(0.7766,0.6209,0.5914,0.4116,-0.6305,0.3768,-0.3894,-0.4675,0.3631
,-0.0121,-0.6247,-0.0579,-0.7116,-0.833,-0.8851,-0.9281,-0.9617,-0.9825,-
1.1031,-1.4133,-1.5000,-1.6738,1.0316)
boxplot(Fenotip_estret,Sans, names=c("Psicòtics","No psicòtics i Sans"),
ylab="PRS",main="Puntuacions del SZ-PRS agrupades per fenotip", xlab="Fen
otip",col=c("grey","white"))
points(1,mean(Fenotip_estret),col="red",pch=18)
points(2,mean(Sans),col="red",pch=18)
```

Puntuacions del SZ-PRS agrupades per fenotip



```
rm(list=ls())
Fenotip_estret<-c(2.3379,1.0658,0.9150,0.9134,0.8906,0.6750,1.8232,0.3816
,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA)
Sans<-c(0.7766,0.6209,0.5914,0.4116,-0.6305,0.3768,-0.3894,-0.4675,0.3631
```

```
, -0.0121, -0.6247, -0.0579, -0.7116, -0.833, -0.8851, -0.9281, -0.9617, -0.9825, -
1.1031, -1.4133, -1.5000, -1.6738, 1.0316)
Dades_SZ_estret<-data.frame(Fenotip_estret,Sans)
shapiro.test(Fenotip_estret)

##
## Shapiro-Wilk normality test
##
## data: Fenotip_estret
## W = 0.86992, p-value = 0.1505

shapiro.test(Sans)

##
## Shapiro-Wilk normality test
##
## data: Sans
## W = 0.95089, p-value = 0.3053

stacked<-stack(Dades_SZ_estret)
bartlett.test(values~ind,data=stacked)

##
## Bartlett test of homogeneity of variances
##
## data: values by ind
## Bartlett's K-squared = 0.38548, df = 1, p-value = 0.5347
```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variances, pel que aplicarem el test paramètric “t.test”

```
t.test(values~ind, data=stacked, var.equal=T)

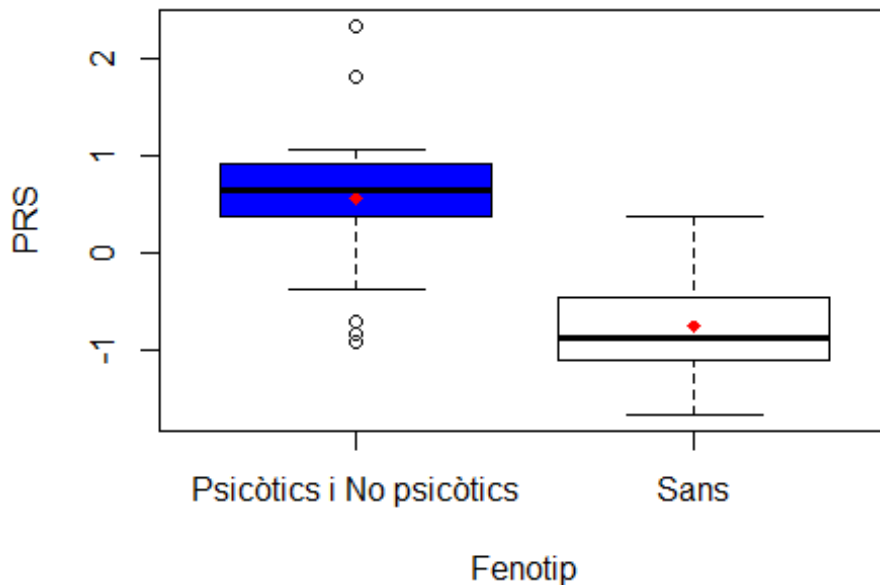
##
## Two Sample t-test
##
## data: values by ind
## t = 4.9218, df = 29, p-value = 3.152e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 0.8864578 2.1469759
## sample estimates:
## mean in group Fenotip_estret          mean in group Sans
##                1.1253125                -0.3914043
```

2) Fenotip ample: Psicòtics i No psicòtics vs Sans

```
rm(list=ls())
Fenotip_ample<-c(2.3379,1.0658,0.9150,0.9134,0.8906,0.6750,1.8232,0.3816,
0.7766,0.6209,0.5914,0.4116,0.3768,-0.3894,-0.9281,-0.7116,-0.833,1.0316)
Sans<-c(-0.4675,0.3631,-0.0121,-0.6247,-0.0579,-0.8851,-0.6305,-0.9617,-0
.9825,-1.1031,-1.4133,-1.5000,-1.6738)
boxplot(Fenotip_ample,Sans, names=c("Psicòtics i No psicòtics","Sans"), y
```

```
lab="PRS",main="Puntuacions del SZ-PRS agrupades per fenotip", xlab="Fenotip",col=c("blue", "white"))
points(1,mean(Fenotip_ample),col="red",pch=18)
points(2,mean(Sans),col="red",pch=18)
```

Puntuacions del SZ-PRS agrupades per fenotip



```
rm(list=ls())
Fenotip_ample<-c(2.3379,1.0658,0.9150,0.9134,0.8906,0.6750,1.8232,0.3816,
0.7766,0.6209,0.5914,0.4116,0.3768,-0.3894,-0.9281,-0.7116,-0.833,1.0316)
Sans<-c(-0.4675,0.3631,-0.0121,-0.6247,-0.0579,-0.8851,-0.6305,-0.9617,-0
.9825,-1.1031,-1.4133,-1.5000,-1.6738,NA,NA,NA,NA,NA)
Dades_SZ_ample<-data.frame(Fenotip_ample,Sans)
shapiro.test(Fenotip_ample)

##
## Shapiro-Wilk normality test
##
## data: Fenotip_ample
## W = 0.93117, p-value = 0.2036

shapiro.test(Sans)

##
## Shapiro-Wilk normality test
##
## data: Sans
## W = 0.96733, p-value = 0.8604
```

```

stacked<-stack(Dades_SZ_ample)
bartlett.test(values~ind,data=stacked)

##
## Bartlett test of homogeneity of variances
##
## data: values by ind
## Bartlett's K-squared = 1.4604, df = 1, p-value = 0.2269

```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variances, pel que aplicarem el test paramètric “t.test”

```

t.test(values~ind, data=stacked, var.equal=T)

##
## Two Sample t-test
##
## data: values by ind
## t = 4.7456, df = 29, p-value = 5.143e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 0.7500056 1.8861029
## sample estimates:
## mean in group Fenotip_ample          mean in group Sans
##                0.5527389                -0.7653154

```

PRS BD

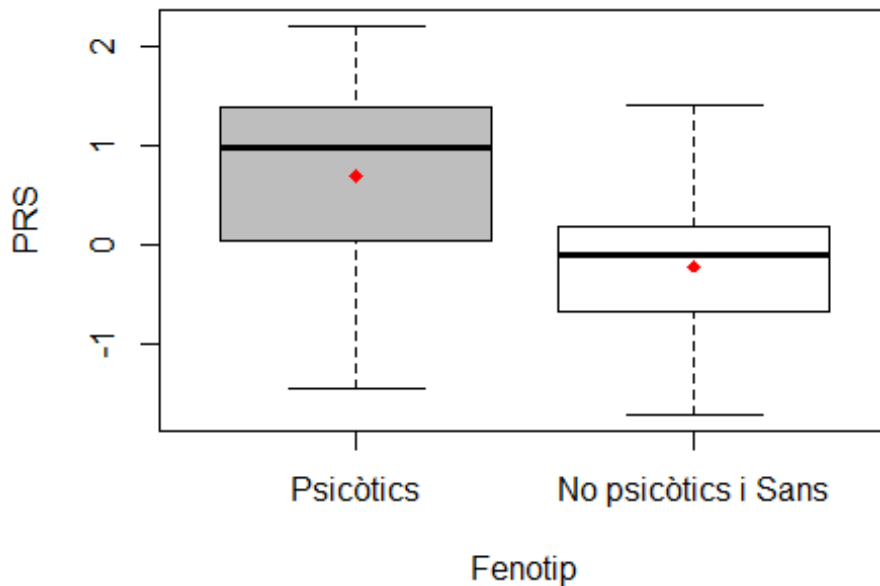
1) Fenotip estret: Psicòtics vs No psicòtics i Sans

```

rm(list=ls())
Fenotip_estret<-c(1.2391,1.3082,-1.4486,0.7052,-0.4063,1.4662,2.2005,0.4674)
Sans<-c(0.0413,0.0532,1.2853,0.2463,-1.5320,0.9282,-0.6870,-0.5872,-0.1179,0.7385,-0.8440,-0.6625,-0.0507,-1.2132,-0.0956,0.1212,-0.3477,-0.0431,-0.6543,0.2870,-1.7090,-1.5185,1.404)
boxplot(Fenotip_estret,Sans, names=c("Psicòtics","No psicòtics i Sans"),
ylab="PRS",main="Puntuacions del BD-PRS agrupades per fenotip", xlab="Fenotip",col=c("grey", "white"))
points(1,mean(Fenotip_estret),col="red",pch=18)
points(2,mean(Sans),col="red",pch=18)

```

Puntuacions del BD-PRS agrupades per fenotip



```
rm(list=ls())
Fenotip_estret<-c(1.2391,1.3082,-1.4486,0.7052,-0.4063,1.4662,2.2005,0.46
74,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA)
Sans<-c(0.0413,0.0532,1.2853,0.2463,-1.5320,0.9282,-0.6870,-0.5872,-0.117
9,0.7385,-0.8440,-0.6625,-0.0507,-1.2132,-0.0956,0.1212,-0.3477,-0.0431,-
0.6543,0.2870,-1.7090,-1.5185,1.404)
Dades_BD_estret<-data.frame(Fenotip_estret,Sans)
shapiro.test(Fenotip_estret)

##
## Shapiro-Wilk normality test
##
## data: Fenotip_estret
## W = 0.9439, p-value = 0.6498

shapiro.test(Sans)

##
## Shapiro-Wilk normality test
##
## data: Sans
## W = 0.968, p-value = 0.6413

stacked<-stack(Dades_BD_estret)
bartlett.test(values~ind,data=stacked)

##
## Bartlett test of homogeneity of variances
```



```
##  
## data: values by ind  
## Bartlett's K-squared = 1.1074, df = 1, p-value = 0.2926
```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variàncies, pel que aplicarem el test paramètric “t.test”

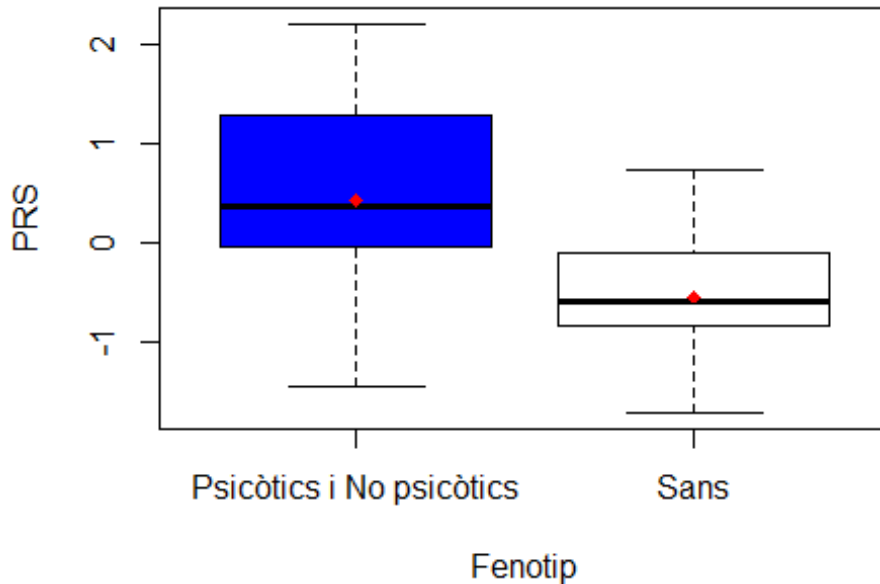
```
t.test(values~ind, data=stacked, var.equal=T)
```

```
##  
## Two Sample t-test  
##  
## data: values by ind  
## t = 2.3749, df = 29, p-value = 0.02438  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## 0.1259092 1.6881201  
## sample estimates:  
## mean in group Fenotip_estret          mean in group Sans  
##                0.6914625                -0.2155522
```

2) Fenotip ample: Psicòtics i No psicòtics vs Sans

```
rm(list=ls())  
Fenotip_ample<-c(1.2391,1.3082,-1.4486,0.7052,-0.4063,1.4662,2.2005,0.467  
4,0.0413,0.0532,1.2853,0.2463,0.9282,-0.6870,0.1212,-0.0507,-1.2132,1.404  
)  
Sans<-c(-0.5872,-0.1179,0.7385,-0.8440,-0.6625,-0.0956,-1.5320,-0.3477,-0  
.0431,-0.6543,0.2870,-1.7090,-1.5185)  
boxplot(Fenotip_ample,Sans, names=c("Psicòtics i No psicòtics","Sans"), y  
lab="PRS",main="Puntuacions del BD-PRS agrupades per fenotip", xlab="Feno  
tip",col=c("blue", "white"))  
points(1,mean(Fenotip_ample),col="red",pch=18)  
points(2,mean(Sans),col="red",pch=18)
```

Puntuacions del BD-PRS agrupades per fenotip



```
rm(list=ls())
Fenotip_ample<-c(1.2391,1.3082,-1.4486,0.7052,-0.4063,1.4662,2.2005,0.467
4,0.0413,0.0532,1.2853,0.2463,0.9282,-0.6870,0.1212,-0.0507,-1.2132,1.404
)
Sans<-c(-0.5872,-0.1179,0.7385,-0.8440,-0.6625,-0.0956,-1.5320,-0.3477,-0
.0431,-0.6543,0.2870,-1.7090,-1.5185,NA,NA,NA,NA,NA)
Dades_BD_ample<-data.frame(Fenotip_ample,Sans)
shapiro.test(Fenotip_ample)

##
## Shapiro-Wilk normality test
##
## data: Fenotip_ample
## W = 0.97007, p-value = 0.7989

shapiro.test(Sans)

##
## Shapiro-Wilk normality test
##
## data: Sans
## W = 0.95683, p-value = 0.7042

stacked<-stack(Dades_BD_ample)
bartlett.test(values~ind,data=stacked)

##
## Bartlett test of homogeneity of variances
```

```
##  
## data: values by ind  
## Bartlett's K-squared = 1.1485, df = 1, p-value = 0.2839
```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variàncies, pel que aplicarem el test paramètric "t.test"

```
t.test(values~ind, data=stacked, var.equal=T)
```

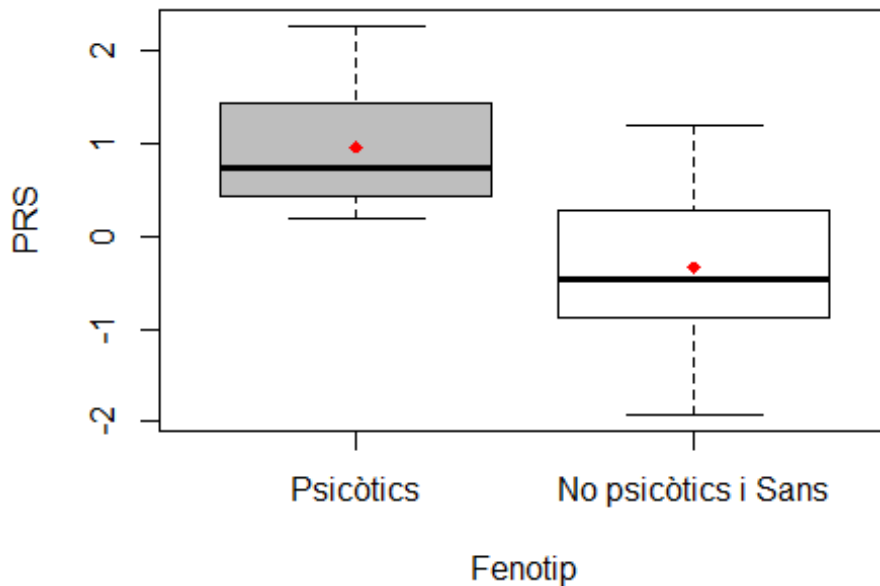
```
##  
## Two Sample t-test  
##  
## data: values by ind  
## t = 3.0026, df = 29, p-value = 0.005464  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## 0.3094909 1.6318536  
## sample estimates:  
## mean in group Fenotip_ample          mean in group Sans  
##                0.4255722                -0.5451000
```

PRS SZ+BD

1) Fenotip estret: Psicòtics vs No psicòtics i Sans

```
rm(list=ls())  
Fenotip_estret<-c(2.2739,1.0993,0.2907,0.8394,0.6013,0.6516,1.7978,0.1920  
)  
Sans<-c(0.8573,1.1943,0.9754,0.7409,-0.2497,-0.4514,0.1740,-0.1147,0.1321  
, -0.0416,-0.7880,0.3873,-1.0344,-0.8121,-0.8329,-1.3203,-0.9285,-1.1969,-  
1.7319,-1.9237,-0.6281,-0.5497,0.3966)  
boxplot(Fenotip_estret,Sans, names=c("Psicòtics","No psicòtics i Sans"),  
ylab="PRS", main="Puntuacions del SZ+BD-PRS agrupades per fenotip",xlab="  
Fenotip",col=c("grey", "white"))  
points(1,mean(Fenotip_estret),col="red",pch=18)  
points(2,mean(Sans),col="red",pch=18)
```

Puntuacions del SZ+BD-PRS agrupades per fenotip



```
rm(list=ls())
Fenotip_estret<-c(2.2739,1.0993,0.2907,0.8394,0.6013,0.6516,1.7978,0.1920
,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA)
Sans<-c(0.8573,1.1943,0.9754,0.7409,-0.2497,-0.4514,0.1740,-0.1147,0.1321
,-0.0416,-0.7880,0.3873,-1.0344,-0.8121,-0.8329,-1.3203,-0.9285,-1.1969,-
1.7319,-1.9237,-0.6281,-0.5497,0.3966)
Dades_SZBD_estret<-data.frame(Fenotip_estret,Sans)
shapiro.test(Fenotip_estret)

##
## Shapiro-Wilk normality test
##
## data: Fenotip_estret
## W = 0.90359, p-value = 0.3111

shapiro.test(Sans)

##
## Shapiro-Wilk normality test
##
## data: Sans
## W = 0.97877, p-value = 0.8842

stacked<-stack(Dades_SZBD_estret)
bartlett.test(values~ind,data=stacked)

##
## Bartlett test of homogeneity of variances
```

```
##  
## data: values by ind  
## Bartlett's K-squared = 0.24914, df = 1, p-value = 0.6177
```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variàncies, pel que aplicarem el test paramètric "t.test"

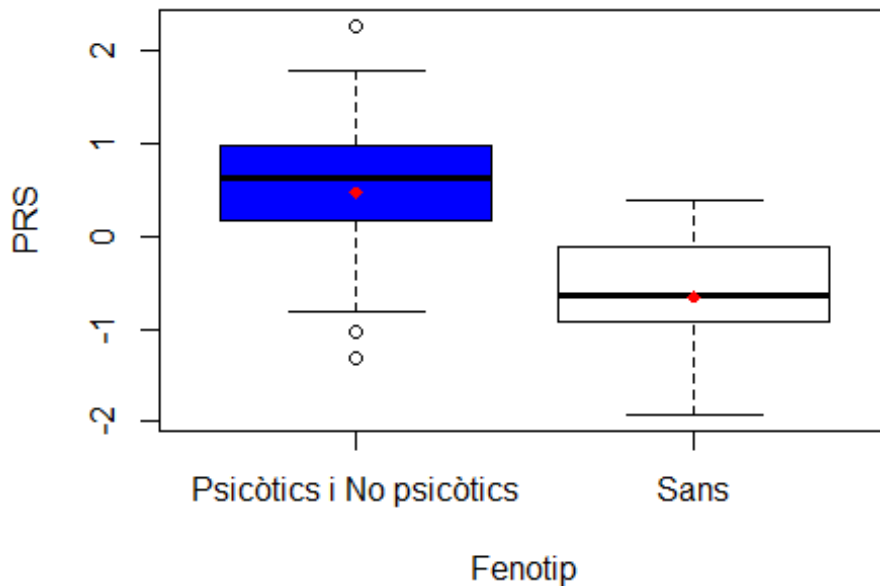
```
t.test(values~ind, data=stacked, var.equal=T)
```

```
##  
## Two Sample t-test  
##  
## data: values by ind  
## t = 3.839, df = 29, p-value = 0.0006183  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## 0.6097794 2.0002858  
## sample estimates:  
## mean in group Fenotip_estret          mean in group Sans  
##                0.9682500                -0.3367826
```

2) Fenotip ample: Psicòtics i No psicòtics vs Sans

```
rm(list=ls())  
Fenotip_ample<-c(2.2739,1.0993,0.2907,0.8394,0.6013,0.6516,1.7978,0.1920,  
0.8573,1.1943,0.9754,0.7409,-0.4514,0.1740,-1.3203,-1.0344,-0.8121,0.3966  
)  
Sans<-c(-0.1147,0.1321,-0.0416,-0.7880,0.3873,-0.8329,-0.2497,-0.9285,-1.  
1969,-1.7319,-1.9237,-0.6281,-0.5497)  
boxplot(Fenotip_ample,Sans, names=c("Psicòtics i No psicòtics","Sans"), y  
lab="PRS",main="Puntuacions del SZ+BD-PRS agrupades per fenotip", xlab="F  
enotip",col=c("blue", "white"))  
points(1,mean(Fenotip_ample),col="red",pch=18)  
points(2,mean(Sans),col="red",pch=18)
```

Puntuacions del SZ+BD-PRS agrupades per fenotip



```
rm(list=ls())
Fenotip_ample<-c(2.2739,1.0993,0.2907,0.8394,0.6013,0.6516,1.7978,0.1920,
0.8573,1.1943,0.9754,0.7409,-0.4514,0.1740,-1.3203,-1.0344,-0.8121,0.3966
)
Sans<-c(-0.1147,0.1321,-0.0416,-0.7880,0.3873,-0.8329,-0.2497,-0.9285,-1.
1969,-1.7319,-1.9237,-0.6281,-0.5497,NA,NA,NA,NA,NA)
Dades_SZBD_ample<-data.frame(Fenotip_ample,Sans)
shapiro.test(Fenotip_ample)

##
## Shapiro-Wilk normality test
##
## data: Fenotip_ample
## W = 0.96475, p-value = 0.6952

shapiro.test(Sans)

##
## Shapiro-Wilk normality test
##
## data: Sans
## W = 0.96483, p-value = 0.8259

stacked<-stack(Dades_SZBD_ample)
bartlett.test(values~ind,data=stacked)

##
## Bartlett test of homogeneity of variances
```

```
##  
## data: values by ind  
## Bartlett's K-squared = 1.1544, df = 1, p-value = 0.2826
```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variàncies, pel que aplicarem el test paramètric “t.test”

```
t.test(values~ind, data=stacked, var.equal=T)
```

```
##  
## Two Sample t-test  
##  
## data: values by ind  
## t = 3.6648, df = 29, p-value = 0.0009858  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## 0.4956613 1.7475464  
## sample estimates:  
## mean in group Fenotip_ample          mean in group Sans  
##                0.4703500                -0.6512538
```

Família 2. PRS agrupats per fenotip

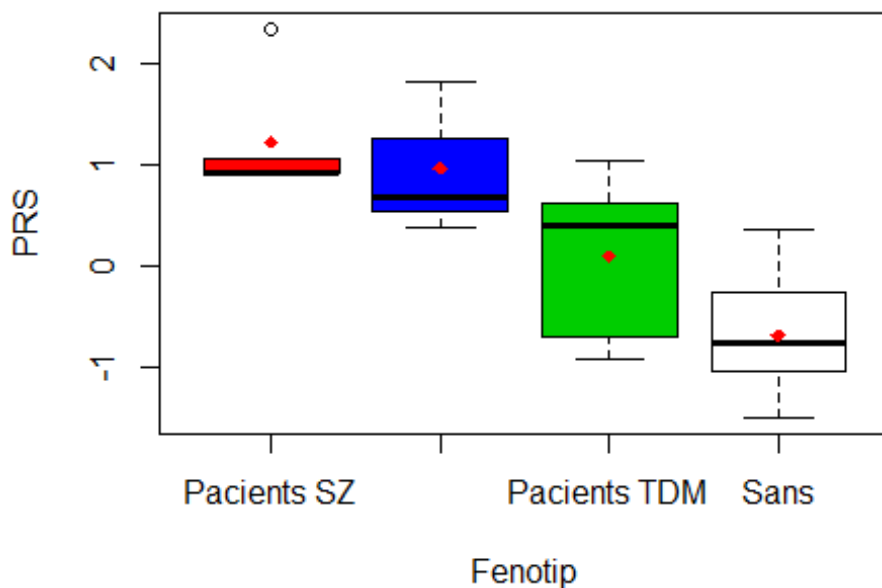
Toni Capó Quetglas

31/5/2021

PRS SZ

```
rm(list=ls())
SZ<-c(2.3379,1.0658,0.9150,0.9134,0.8906)
BD<-c(0.6750,1.8232,0.3816)
TDM<-c(0.7766,0.6209,0.5914,0.4116,0.3768,-0.3894,-0.9281,-0.7116,-0.8330,1.0316)
Sans<-c(-0.4675,0.3631,-0.0121,-0.6247,-0.0579,-0.8851,-0.6305,-0.9617,-0.9825,-1.1031,-1.4133,-1.5000)
boxplot(SZ,BD,TDM,Sans, names=c("Pacients SZ","Pacients BD", "Pacients TDM", "Sans"), ylab="PRS",main="Puntuacions del SZ-PRS agrupades per fenotip", xlab="Fenotip",col=c("red","blue","green3","white"))
points(1,mean(SZ),col="red",pch=18)
points(2,mean(BD),col="red",pch=18)
points(3,mean(TDM),col="red",pch=18)
points(4,mean(Sans),col="red",pch=18)
```

Puntuacions del SZ-PRS agrupades per fenotip



```
Valors=c(2.3379,1.0658,0.9150,0.9134,0.8906,0.6750,1.8232,0.3816,0.7766,0.6209,0.5914,0.4116,0.3768,-0.3894,-0.9281,-0.7116,-0.8330,-0.4675,0.3631
```



```

, -0.0121, -0.6247, -0.0579, -0.8851, -0.6305, -0.9617, -0.9825, -1.1031, -1.4133,
-1.5000, -1.6738)
Fenotip=c(rep("SZ",5),rep("BD",3), rep("TDM",9), rep("Sans",13))
Dades_SZ=data.frame(Valors,Fenotip)
dim (Dades_SZ)

## [1] 30 2

summary(Dades_SZ)

##      Valors      Fenotip
## Min.   :-1.67380 Length:30
## 1st Qu.:-0.87208 Class :character
## Median :-0.03500 Mode  :character
## Mean   :-0.03438
## 3rd Qu.: 0.66148
## Max.   : 2.33790

SZ <- subset(Dades_SZ, Fenotip=="SZ")
BD <- subset(Dades_SZ, Fenotip=="BD")
TDM <- subset(Dades_SZ, Fenotip=="TDM")
Sans <- subset(Dades_SZ, Fenotip=="Sans")

shapiro.test(SZ$Valors)

##
## Shapiro-Wilk normality test
##
## data:  SZ$Valors
## W = 0.63575, p-value = 0.001795

shapiro.test(BD$Valors)

##
## Shapiro-Wilk normality test
##
## data:  BD$Valors
## W = 0.8951, p-value = 0.3701

shapiro.test(TDM$Valors)

##
## Shapiro-Wilk normality test
##
## data:  TDM$Valors
## W = 0.84764, p-value = 0.07018

shapiro.test(Sans$Valors)

##
## Shapiro-Wilk normality test
##

```

```
## data: Sans$Valors
## W = 0.96033, p-value = 0.3158

bartlett.test(Valors~Fenotip, data=Dades_SZ)

##
## Bartlett test of homogeneity of variances
##
## data: Valors by Fenotip
## Bartlett's K-squared = 0.26254, df = 3, p-value = 0.9669
```

Els valors es presenten amb una distribució normal menys els valors de PRS SZ i tots presenten heterogeneïtat de variances, així que aplicarem el test no paramètric "kluskal.test"

```
kruskal.test(Valors~Fenotip, data= Dades_SZ)

##
## Kruskal-Wallis rank sum test
##
## data: Valors by Fenotip
## Kruskal-Wallis chi-squared = 19.231, df = 3, p-value = 0.000245

library("dunn.test")
dunn.test (Dades_SZ$Valors, g=Dades_SZ$Fenotip, method="bonferroni")

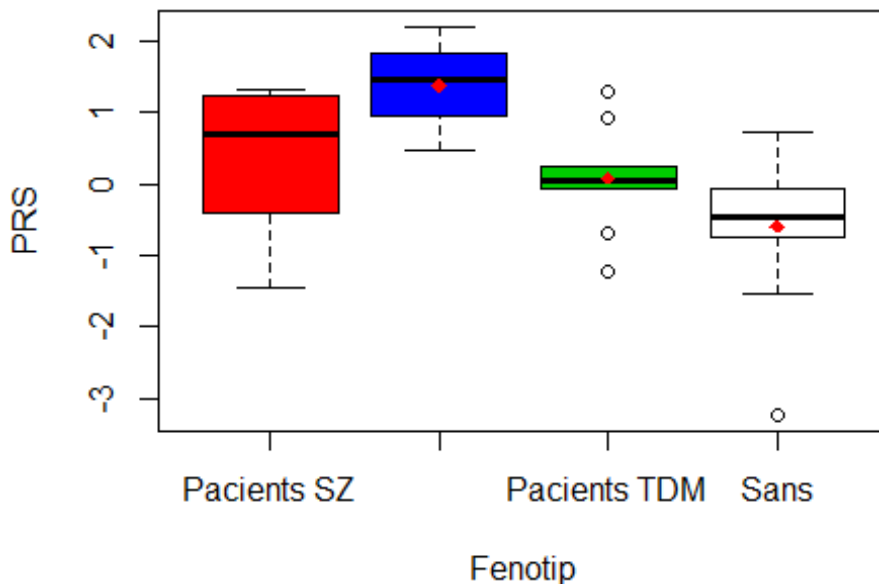
## Kruskal-Wallis rank sum test
##
## data: x and group
## Kruskal-Wallis chi-squared = 19.2306, df = 3, p-value = 0
##
##
## Comparison of x by group
## (Bonferroni)
## Col Mean-|
## Row Mean |          BD          Sans          SZ
##          +
## Sans | 2.655643
##       | 0.0237*
##       |
## SZ | -0.549584 -3.995046
##       | 1.0000 0.0002*
##       |
## TDM | 1.287380 -1.943410 2.258288
##       | 0.5939 0.1559 0.0718
##
## alpha = 0.05
## Reject Ho if p <= alpha/2
```

Existeixen diferències significatives entre SZ vs Sans i BD vs Sans. El grup TDM difereix bastant del grup Sans però no lo suficient com per que existeixi significació.

PRS BD

```
rm(list=ls())
SZ<-c(1.2391,1.3082,-1.4486,0.7052,-0.4063)
BD<-c(1.4662,2.2005,0.4674)
TDM<-c(0.0413,0.0532,1.2853,0.2463,0.9282,-0.6870,0.1212,-0.0507,-1.2132)
Sans<-c(-0.5872,-0.1179,0.7385,-0.8440,-0.6625,-0.0956,-1.5320,-0.3477,-0.0431,-0.6543,0.2870,-1.7090,-1.5185)
boxplot(SZ,BD,TDM,Sans, names=c("Pacients SZ","Pacients BD", "Pacients TDM", "Sans"), ylab="PRS", main="Puntuacions del BD-PRS agrupades per fenotip", xlab="Fenotip", col=c("red","blue","green3","white"))
points(1,mean(SZ),col="red",pch=18)
points(2,mean(BD),col="red",pch=18)
points(3,mean(TDM),col="red",pch=18)
points(4,mean(Sans),col="red",pch=18)
```

Puntuacions del BD-PRS agrupades per fenotip



```
Valors=c(1.2391,1.3082,-1.4486,0.7052,-0.4063,1.4662,2.2005,0.4674,0.0413,0.0532,1.2853,0.2463,0.9282,-0.6870,0.1212,-0.0507,-1.2132,-0.5872,-0.1179,0.7385,-0.8440,-0.6625,-0.0956,-1.5320,-0.3477,-0.0431,-0.6543,0.2870,-1.7090,-1.5185)
Fenotip=c(rep("SZ",5),rep("BD",3), rep("TDM",9), rep("Sans",13))
Dades_BD=data.frame(Valors,Fenotip)
dim (Dades_BD)

## [1] 30 2

summary(Dades_BD)
```

```

##      Valors      Fenotip
## Min.   :-1.70900 Length:30
## 1st Qu.:-0.66045 Class :character
## Median :-0.04690 Mode  :character
## Mean   :-0.02767
## 3rd Qu.: 0.64575
## Max.    : 2.20050

SZ <- subset(Dades_BD, Fenotip=="SZ")
BD <- subset(Dades_BD, Fenotip=="BD")
TDM <- subset(Dades_BD, Fenotip=="TDM")
Sans <- subset(Dades_BD, Fenotip=="Sans")

shapiro.test(SZ$Valors)

##
## Shapiro-Wilk normality test
##
## data:  SZ$Valors
## W = 0.88487, p-value = 0.332

shapiro.test(BD$Valors)

##
## Shapiro-Wilk normality test
##
## data:  BD$Valors
## W = 0.9923, p-value = 0.8321

shapiro.test(TDM$Valors)

##
## Shapiro-Wilk normality test
##
## data:  TDM$Valors
## W = 0.94785, p-value = 0.6665

shapiro.test(Sans$Valors)

##
## Shapiro-Wilk normality test
##
## data:  Sans$Valors
## W = 0.97648, p-value = 0.7262

bartlett.test(Valors~Fenotip, data=Dades_BD)

##
## Bartlett test of homogeneity of variances
##
## data:  Valors by Fenotip
## Bartlett's K-squared = 1.6615, df = 3, p-value = 0.6455

```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variàncies, pel que aplicarem el test paramètric “anova d’1 factor”

```
anova(lm(Valors~Fenotip, data= Dades_BD))

## Analysis of Variance Table
##
## Response: Valors
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Fenotip     3  9.9857   3.3286   4.8025 0.008592 **
## Residuals  26 18.0204   0.6931
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

pairwise.t.test(Dades_BD$Valors,Dades_BD$Fenotip,p.adjust.method="bonferroni")

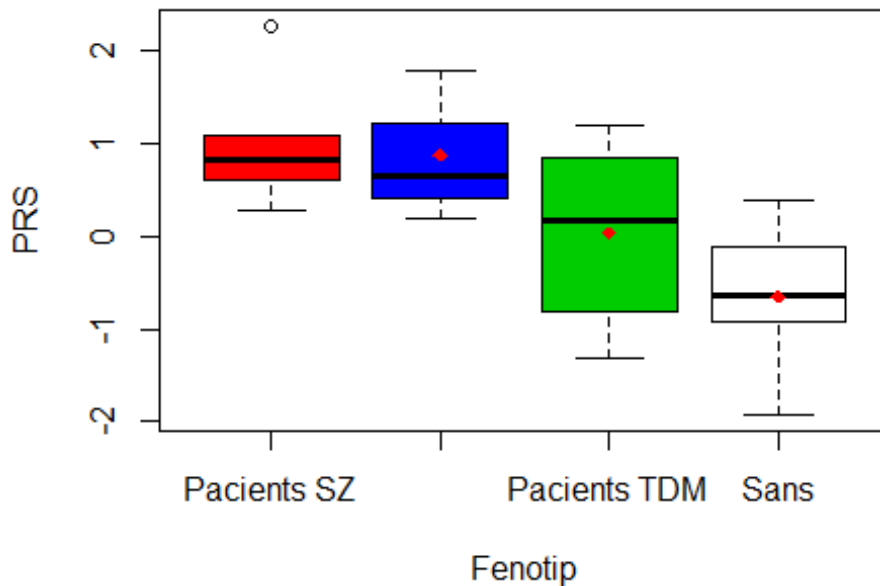
##
## Pairwise comparisons using t tests with pooled SD
##
## data:  Dades_BD$Valors and Dades_BD$Fenotip
##
##          BD      Sans      SZ
## Sans 0.0078 -        -
## SZ   0.4943 0.4262 -
## TDM  0.1642 0.5697 1.0000
##
## P value adjustment method: bonferroni
```

Existeixen diferències significatives entre Bd vs Sans únicament.

PRS SZ+BD

```
rm(list=ls())
SZ<-c(2.2739,1.0993,0.2907,0.8394,0.6013)
BD<-c(0.6516,1.7978,0.1920)
TDM<-c(0.8573,1.1943,0.9754,0.7409,-0.4514,0.1740,-1.3203,-1.0344,-0.8121)
Sans<-c(-0.1147,0.1321,-0.0416,-0.7880,0.3873,-0.8329,-0.2497,-0.9285,-1.1969,-1.7319,-1.9237,-0.6281,-0.5497)
boxplot(SZ,BD,TDM,Sans, names=c("Pacients SZ","Pacients BD", "Pacients TDM", "Sans"), ylab="PRS",main="Puntuacions del SZ+BD-PRS agrupades per fenotip", xlab="Fenotip",col=c("red","blue","green3","white"))
points(1,mean(SZ),col="red",pch=18)
points(2,mean(BD),col="red",pch=18)
points(3,mean(TDM),col="red",pch=18)
points(4,mean(Sans),col="red",pch=18)
```

Puntuacions del SZ+BD-PRS agrupades per fenotip



```
Valors=c(2.2739,1.0993,0.2907,0.8394,0.6013,0.6516,1.7978,0.1920,0.8573,1
.1943,0.9754,0.7409,-0.4514,0.1740,-1.3203,-1.0344,-0.8121,-0.1147,0.1321
,-0.0416,-0.7880,0.3873,-0.8329,-0.2497,-0.9285,-1.1969,-1.7319,-1.9237,-
0.6281,-0.5497)
Fenotip=c(rep("SZ",5),rep("BD",3), rep("TDM",9), rep("Sans",13))
Dades_SZBD=data.frame(Valors,Fenotip)
dim (Dades_SZBD)

## [1] 30 2

summary(Dades_SZBD)

##      Valors      Fenotip
## Min.   :-1.92370 Length:30
## 1st Qu.:-0.80607 Class :character
## Median : 0.04525 Mode  :character
## Mean   :-0.01322
## 3rd Qu.: 0.71857
## Max.    : 2.27390

SZ <- subset(Dades_SZBD, Fenotip=="SZ")
BD <- subset(Dades_SZBD, Fenotip=="BD")
TDM <- subset(Dades_SZBD, Fenotip=="TDM")
Sans <- subset(Dades_SZBD, Fenotip="Sans")

shapiro.test(SZ$Valors)
```

```

##
## Shapiro-Wilk normality test
##
## data:  SZ$Valors
## W = 0.89233, p-value = 0.369

shapiro.test(BD$Valors)

##
## Shapiro-Wilk normality test
##
## data:  BD$Valors
## W = 0.94256, p-value = 0.5378

shapiro.test(TDM$Valors)

##
## Shapiro-Wilk normality test
##
## data:  TDM$Valors
## W = 0.90215, p-value = 0.2647

shapiro.test(Sans$Valors)

##
## Shapiro-Wilk normality test
##
## data:  Sans$Valors
## W = 0.98713, p-value = 0.9677

bartlett.test(Valors~Fenotip, data=Dades_SZBD)

##
## Bartlett test of homogeneity of variances
##
## data:  Valors by Fenotip
## Bartlett's K-squared = 0.98169, df = 3, p-value = 0.8057

```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variàncies, pel que aplicarem el test paramètric “anova d’1 factor”

```

anova(lm(Valors~Fenotip, data= Dades_SZBD))

## Analysis of Variance Table
##
## Response: Valors
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Fenotip     3  13.057   4.3524   6.7438 0.001627 **
## Residuals  26  16.780   0.6454
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

pairwise.t.test(Dades_SZBD$Valors,Dades_SZBD$Fenotip,p.adjust.method="bonferroni")

##
## Pairwise comparisons using t tests with pooled SD
##
## data: Dades_SZBD$Valors and Dades_SZBD$Fenotip
##
##      BD      Sans   SZ
## Sans 0.0374 -      -
## SZ   1.0000 0.0032 -
## TDM  0.7616 0.3555 0.2223
##
## P value adjustment method: bonferroni

```

Existeixen diferències significatives entre SZ vs Sans i BD vs Sans.

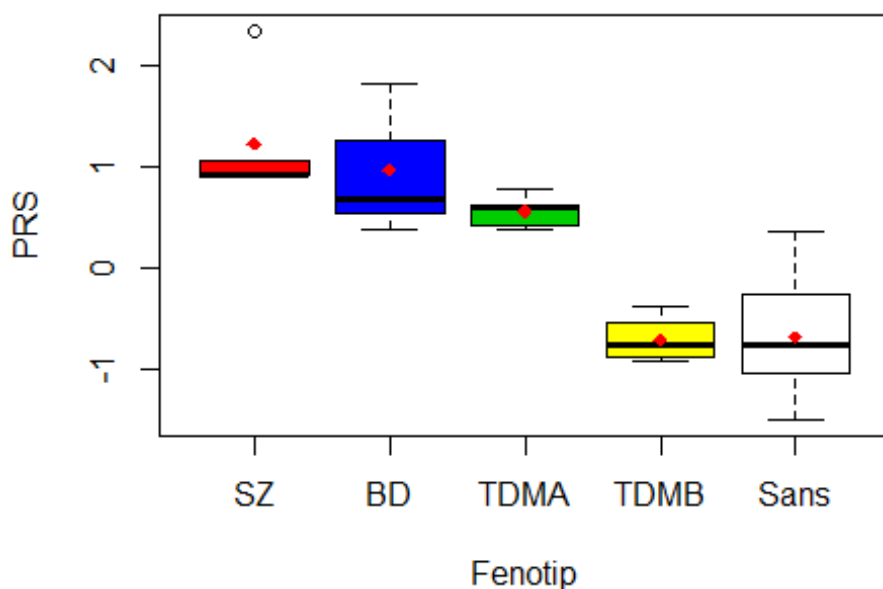
PRS SZ

```

rm(list=ls())
SZ<-c(2.3379,1.0658,0.9150,0.9134,0.8906)
BD<-c(0.6750,1.8232,0.3816)
TDM_A<-c(0.7766,0.6209,0.5914,0.4116,0.3768)
TDM_B<-c(-0.3894,-0.9281,-0.7116,-0.8330)
Sans<-c(-0.4675,0.3631,-0.0121,-0.6247,-0.0579,-0.8851,-0.6305,-0.9617,-0.9825,-1.1031,-1.4133,-1.5000)
boxplot(SZ,BD,TDM_A,TDM_B,Sans, names=c("SZ","BD","TDMA","TDMB","Sans"),
, ylab="PRS",main="Puntuacions del SZ-PRS agrupades per fenotip", xlab="Fenotip",col=c("red","blue","green3","yellow","white"))
points(1,mean(SZ),col="red",pch=18)
points(2,mean(BD),col="red",pch=18)
points(3,mean(TDM_A),col="red",pch=18)
points(4,mean(TDM_B),col="red",pch=18)
points(5,mean(Sans),col="red",pch=18)

```


Puntuacions del SZ-PRS agrupades per fenotip



```
Valors=c(2.3379,1.0658,0.9150,0.9134,0.8906,0.6750,1.8232,0.3816,0.7766,0.6209,0.5914,0.4116,0.3768,-0.3894,-0.9281,-0.7116,-0.8330,-0.4675,0.3631,-0.0121,-0.6247,-0.0579,-0.8851,-0.6305,-0.9617,-0.9825,-1.1031,-1.4133,-1.5000,-1.6738)
```

```
Fenotip=c(rep("SZ",5),rep("BD",3), rep("TDMA",5),rep("TDMB",4), rep("Sans",13))
```

```
Dades_SZ=data.frame(Valors,Fenotip)
```

```
dim (Dades_SZ)
```

```
## [1] 30 2
```

```
summary(Dades_SZ)
```

```
##      Valors      Fenotip
## Min. :-1.67380 Length:30
## 1st Qu.:-0.87208 Class :character
## Median :-0.03500 Mode  :character
## Mean  :-0.03438
## 3rd Qu.: 0.66148
## Max.   : 2.33790
```

```
SZ <- subset(Dades_SZ, Fenotip=="SZ")
```

```
BD <- subset(Dades_SZ, Fenotip=="BD")
```

```
TDM_A <- subset(Dades_SZ, Fenotip=="TDMA")
```

```
TDM_B <- subset(Dades_SZ, Fenotip=="TDMB")
```

```
Sans <- subset(Dades_SZ, Fenotip=="Sans")
```

```
shapiro.test(SZ$Valors)
```

```

##
## Shapiro-Wilk normality test
##
## data:  SZ$Valors
## W = 0.63575, p-value = 0.001795

shapiro.test(BD$Valors)

##
## Shapiro-Wilk normality test
##
## data:  BD$Valors
## W = 0.8951, p-value = 0.3701

shapiro.test(TDM_A$Valors)

##
## Shapiro-Wilk normality test
##
## data:  TDM_A$Valors
## W = 0.93384, p-value = 0.6228

shapiro.test(TDM_B$Valors)

##
## Shapiro-Wilk normality test
##
## data:  TDM_B$Valors
## W = 0.9218, p-value = 0.5471

shapiro.test(Sans$Valors)

##
## Shapiro-Wilk normality test
##
## data:  Sans$Valors
## W = 0.96033, p-value = 0.3158

bartlett.test(Valors~Fenotip, data=Dades_SZ)

##
## Bartlett test of homogeneity of variances
##
## data:  Valors by Fenotip
## Bartlett's K-squared = 8.6693, df = 4, p-value = 0.06992

```

Els valors es presenten amb una distribució normal menys els valors de PRS SZ i tots presenten heterogeneïtat de variances, així que aplicarem el test no paramètric “kluskal.test”

```
kruskal.test(Valors~Fenotip, data= Dades_SZ)
```

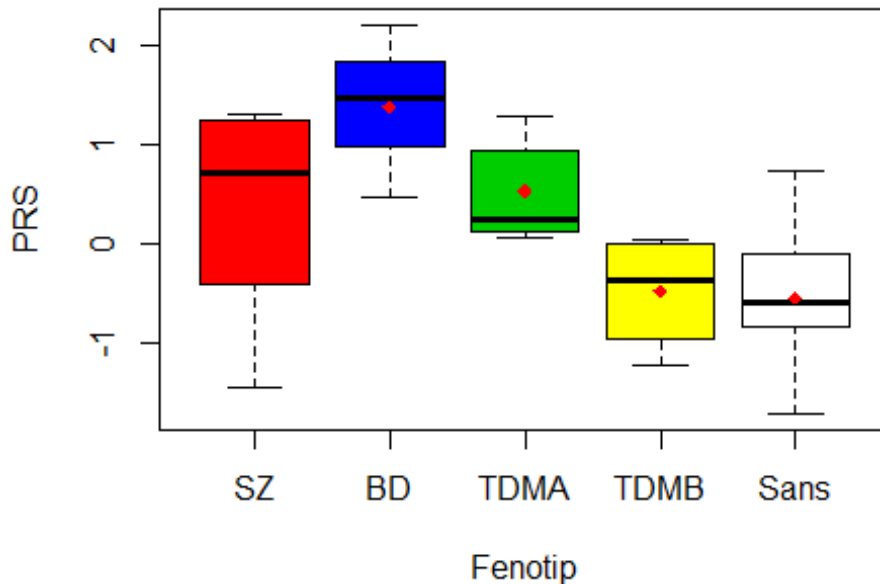
```
##
## Kruskal-Wallis rank sum test
##
## data: Valors by Fenotip
## Kruskal-Wallis chi-squared = 22.7, df = 4, p-value = 0.0001453
library("dunn.test")
dunn.test (Dades_SZ$Valors, g=Dades_SZ$Fenotip, method="bonferroni")
## Kruskal-Wallis rank sum test
##
## data: x and group
## Kruskal-Wallis chi-squared = 22.7002, df = 4, p-value = 0
##
##
## Comparison of x by group
## (Bonferroni)
## Col Mean- |
## Row Mean |          BD          Sans          SZ          TDMA
## -----|-----
## Sans |      2.655643
##      |      0.0396
## SZ | -0.549584 -3.995046
##      |      1.0000      0.0003*
## TDMA |      0.414780 -2.656722      1.113552
##      |      1.0000      0.0395      1.0000
## TDMB |      2.032606 -0.259795      2.912536      1.862668
##      |      0.2105      1.0000      0.0179*      0.3125
##
## alpha = 0.05
## Reject Ho if p <= alpha/2
```

PRS BD

```
rm(list=ls())
SZ<-c(1.2391,1.3082,-1.4486,0.7052,-0.4063)
BD<-c(1.4662,2.2005,0.4674)
TDM_A<-c(0.0532,1.2853,0.2463,0.9282,0.1212)
TDM_B<-c(-0.0507,-1.2132,-0.6870,0.0413)
Sans<-c(-0.5872,-0.1179,0.7385,-0.8440,-0.6625,-0.0956,-1.5320,-0.3477,-0.0431,-0.6543,0.2870,-1.7090,-1.5185)
boxplot(SZ,BD,TDM_A,TDM_B,Sans, names=c("SZ","BD","TDMA","TDMB","Sans")
, ylab="PRS",main="Puntuacions del BD-PRS agrupades per fenotip", xlab="Fenotip",col=c("red","blue","green3","yellow","white"))
points(1,mean(SZ),col="red",pch=18)
points(2,mean(BD),col="red",pch=18)
points(3,mean(TDM_A),col="red",pch=18)
```

```
points(4,mean(TDM_B),col="red",pch=18)
points(5,mean(Sans),col="red",pch=18)
```

Puntuacions del BD-PRS agrupades per fenotip



```
Valors=c(1.2391,1.3082,-1.4486,0.7052,-0.4063,1.4662,2.2005,0.4674,0.0532
,1.2853,0.2463,0.9282,0.1212,-0.0507,-1.2132,-0.6870,0.0413,-0.5872,-0.11
79,0.7385,-0.8440,-0.6625,-0.0956,-1.5320,-0.3477,-0.0431,-0.6543,0.2870,
-1.7090,-1.5185)
Fenotip=c(rep("SZ",5),rep("BD",3), rep("TDMA",5),rep("TDMB",4), rep("Sans
",13))
Dades_BD=data.frame(Valors,Fenotip)
dim (Dades_BD)

## [1] 30 2

summary(Dades_BD)

##      Valors      Fenotip
## Min.   :-1.70900   Length:30
## 1st Qu.:-0.66045   Class :character
## Median :-0.04690   Mode  :character
## Mean   :-0.02767
## 3rd Qu.: 0.64575
## Max.   : 2.20050

SZ <- subset(Dades_BD, Fenotip=="SZ")
BD <- subset(Dades_BD, Fenotip=="BD")
TDM_A <- subset(Dades_BD, Fenotip=="TDMA")
```

```
TDM_B <- subset(Dades_BD, Fenotip=="TDMB")
Sans <- subset(Dades_BD, Fenotip="Sans")

shapiro.test(SZ$Valors)

##
## Shapiro-Wilk normality test
##
## data:  SZ$Valors
## W = 0.88487, p-value = 0.332

shapiro.test(BD$Valors)

##
## Shapiro-Wilk normality test
##
## data:  BD$Valors
## W = 0.9923, p-value = 0.8321

shapiro.test(TDM_A$Valors)

##
## Shapiro-Wilk normality test
##
## data:  TDM_A$Valors
## W = 0.85349, p-value = 0.2058

shapiro.test(TDM_B$Valors)

##
## Shapiro-Wilk normality test
##
## data:  TDM_B$Valors
## W = 0.90409, p-value = 0.4517

shapiro.test(Sans$Valors)

##
## Shapiro-Wilk normality test
##
## data:  Sans$Valors
## W = 0.97648, p-value = 0.7262

bartlett.test(Valors~Fenotip, data=Dades_BD)

##
## Bartlett test of homogeneity of variances
##
## data:  Valors by Fenotip
## Bartlett's K-squared = 2.8829, df = 4, p-value = 0.5776

anova(lm(Valors~Fenotip, data= Dades_BD))
```

```

## Analysis of Variance Table
##
## Response: Valors
##           Df Sum Sq Mean Sq F value Pr(>F)
## Fenotip    4 12.227  3.05670   4.8429 0.00496 **
## Residuals 25 15.779  0.63117
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

pairwise.t.test(Dades_BD$Valors,Dades_BD$Fenotip,p.adjust.method="bonferroni")

##
## Pairwise comparisons using t tests with pooled SD
##
## data:  Dades_BD$Valors and Dades_BD$Fenotip
##
##      BD      Sans      SZ      TDMA
## Sans 0.0087 -        -        -
## SZ   0.6994 0.5972 -        -
## TDMA 1.0000 0.1674 1.0000 -
## TDMB 0.0525 1.0000 1.0000 0.7120
##
## P value adjustment method: bonferroni

```

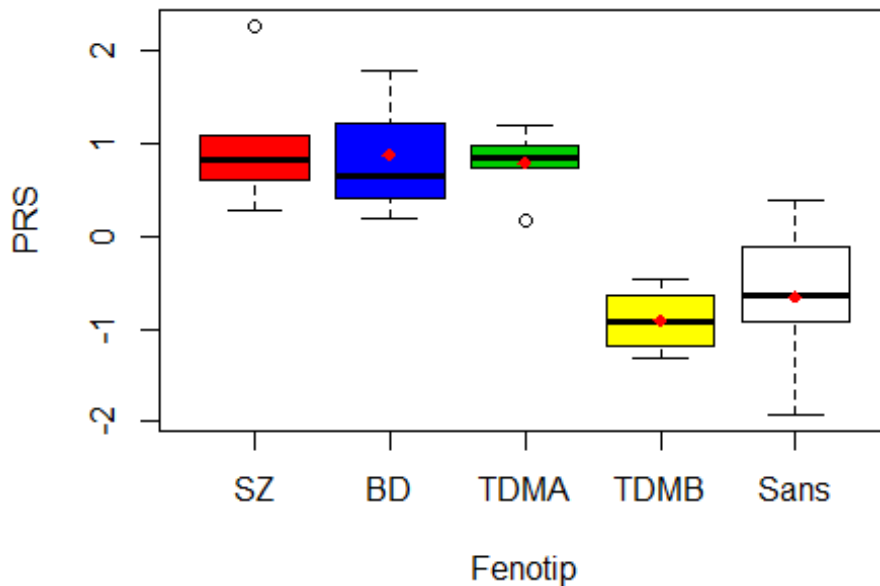
PRS SZ+BD

```

rm(list=ls())
SZ<-c(2.2739,1.0993,0.2907,0.8394,0.6013)
BD<-c(0.6516,1.7978,0.1920)
TDM_A<-c(0.8573,1.1943,0.9754,0.7409,0.1740)
TDM_B<-c(-0.4514,-1.3203,-1.0344,-0.8121)
Sans<-c(-0.1147,0.1321,-0.0416,-0.7880,0.3873,-0.8329,-0.2497,-0.9285,-1.1969,-1.7319,-1.9237,-0.6281,-0.5497)
boxplot(SZ,BD,TDM_A,TDM_B,Sans, names=c("SZ","BD","TDMA","TDMB","Sans"),
, ylab="PRS",main="Puntuacions del SZ+BD-PRS agrupades per fenotip", xlab="Fenotip",col=c("red","blue","green3","yellow","white"))
points(1,mean(SZ),col="red",pch=18)
points(2,mean(BD),col="red",pch=18)
points(3,mean(TDM_A),col="red",pch=18)
points(4,mean(TDM_B),col="red",pch=18)
points(5,mean(Sans),col="red",pch=18)

```

Puntuacions del SZ+BD-PRS agrupades per fenotip



```
Valors=c(2.2739,1.0993,0.2907,0.8394,0.6013,0.6516,1.7978,0.1920,0.8573,1.1943,0.9754,0.7409,0.1740,-0.4514,-1.3203,-1.0344,-0.8121,-0.1147,0.1321,-0.0416,-0.7880,0.3873,-0.8329,-0.2497,-0.9285,-1.1969,-1.7319,-1.9237,-0.6281,-0.5497)
```

```
Fenotip=c(rep("SZ",5),rep("BD",3), rep("TDMA",5),rep("TDMB",4), rep("Sans",13))
```

```
Dades_SZBD=data.frame(Valors,Fenotip)
```

```
dim (Dades_SZBD)
```

```
## [1] 30 2
```

```
summary(Dades_SZBD)
```

```
##      Valors      Fenotip
## Min.   :-1.92370 Length:30
## 1st Qu.:-0.80607 Class :character
## Median : 0.04525 Mode  :character
## Mean   :-0.01322
## 3rd Qu.: 0.71857
## Max.    : 2.27390
```

```
SZ <- subset(Dades_SZBD, Fenotip=="SZ")
```

```
BD <- subset(Dades_SZBD, Fenotip=="BD")
```

```
TDM_A <- subset(Dades_SZBD, Fenotip=="TDMA")
```

```
TDM_B <- subset(Dades_SZBD, Fenotip=="TDMB")
```

```
Sans <- subset(Dades_SZBD, Fenotip="Sans")
```

```
shapiro.test(SZ$Valors)
```

```

##
## Shapiro-Wilk normality test
##
## data:  SZ$Valors
## W = 0.89233, p-value = 0.369
shapiro.test(BD$Valors)

##
## Shapiro-Wilk normality test
##
## data:  BD$Valors
## W = 0.94256, p-value = 0.5378
shapiro.test(TDM_A$Valors)

##
## Shapiro-Wilk normality test
##
## data:  TDM_A$Valors
## W = 0.92402, p-value = 0.5562
shapiro.test(TDM_B$Valors)

##
## Shapiro-Wilk normality test
##
## data:  TDM_B$Valors
## W = 0.99637, p-value = 0.9873
shapiro.test(Sans$Valors)

##
## Shapiro-Wilk normality test
##
## data:  Sans$Valors
## W = 0.98713, p-value = 0.9677
bartlett.test(Valors~Fenotip, data=Dades_SZBD)

##
## Bartlett test of homogeneity of variances
##
## data:  Valors by Fenotip
## Bartlett's K-squared = 3.2384, df = 4, p-value = 0.5187
anova(lm(Valors~Fenotip, data= Dades_SZBD))

## Analysis of Variance Table
##
## Response: Valors
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Fenotip    4  19.426   4.8565  11.662 1.758e-05 ***

```



```

## Residuals 25 10.411 0.4164
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

pairwise.t.test(Dades_SZBD$Valors,Dades_SZBD$Fenotip,p.adjust.method="bonferroni")

##
## Pairwise comparisons using t tests with pooled SD
##
## data: Dades_SZBD$Valors and Dades_SZBD$Fenotip
##
##      BD      Sans      SZ      TDMA
## Sans 0.01051 -          -          -
## SZ    1.00000 0.00045 -          -
## TDMA 1.00000 0.00267 1.00000 -
## TDMB 0.01300 1.00000 0.00156 0.00623
##
## P value adjustment method: bonferroni

```

Resultats Família 1 + Família 2

Toni Capó Quetglas

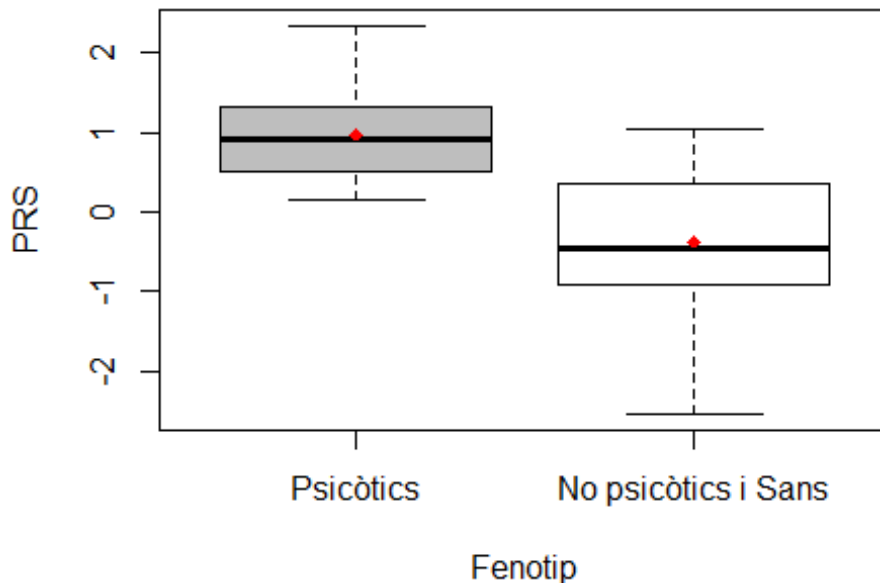
15/6/2021

PRS SZ

1) Fenotip estret: Psicòtics vs No psicòtics i Sans

```
rm(list=ls())
Fenotip_estret<-c(2.3379,1.0658,0.9150,0.9134,0.8906,1.591,0.1422,0.6246,
0.3338,0.6750,1.8232,0.3816)
Sans<-c(-0.0299,-0.0856,-0.2522,0.776,0.6209,0.5914,0.4116,0.3768,-0.3894
,-0.9281,-0.7116,-0.833,1.0316,0.8904,0.3424,-0.5092,-0.5092,-2.5386,-0.4
675,0.3631,-0.0121,-0.6247,-0.0579,-0.8851,-0.6305,-0.9617,-0.9825,-1.103
1,-1.4133,-1.5,-1.6738)
boxplot(Fenotip_estret,Sans, names=c("Psicòtics","No psicòtics i Sans"),
ylab="PRS",main="Puntuacions del SZ-PRS agrupades per fenotip", xlab="Fenotip",col=c("grey","white"))
points(1,mean(Fenotip_estret),col="red",pch=18)
points(2,mean(Sans),col="red",pch=18)
```

Puntuacions del SZ-PRS agrupades per fenotip



```
rm(list=ls())
Fenotip_estret<-c(2.3379,1.0658,0.9150,0.9134,0.8906,1.591,0.1422,0.6246,
0.3338,0.6750,1.8232,0.3816,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
```

```

NA,NA,NA,NA)
Sans<-c(-0.0299,-0.0856,-0.2522,0.776,0.6209,0.5914,0.4116,0.3768,-0.3894
,-0.9281,-0.7116,-0.833,1.0316,0.8904,0.3424,-0.5092,-0.5092,-2.5386,-0.4
675,0.3631,-0.0121,-0.6247,-0.0579,-0.8851,-0.6305,-0.9617,-0.9825,-1.103
1,-1.4133,-1.5,-1.6738)
Dades_SZ_estret<-data.frame(Fenotip_estret,Sans)
shapiro.test(Fenotip_estret)

##
## Shapiro-Wilk normality test
##
## data: Fenotip_estret
## W = 0.92526, p-value = 0.3326

shapiro.test(Sans)

##
## Shapiro-Wilk normality test
##
## data: Sans
## W = 0.97589, p-value = 0.6916

stacked<-stack(Dades_SZ_estret)
bartlett.test(values~ind,data=stacked)

##
## Bartlett test of homogeneity of variances
##
## data: values by ind
## Bartlett's K-squared = 0.83702, df = 1, p-value = 0.3603

```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variàncies, pel que aplicarem el test paramètric “t.test”

```

t.test(values~ind, data=stacked, var.equal=T)

##
## Two Sample t-test
##
## data: values by ind
## t = 5.0725, df = 41, p-value = 8.88e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 0.8135759 1.8899440
## sample estimates:
## mean in group Fenotip_estret          mean in group Sans
## 0.9745083                             -0.3772516

```

2) Fenotip ample: Psicòtics i No psicòtics vs Sans

```

rm(list=ls())
Fenotip_ample<-c(2.3379,1.0658,0.915,0.9134,0.8906,1.591,0.1422,0.6246,0.
3338,0.675,1.8232,0.3816,-0.0299,-0.0856,-0.2522,0.7766,0.6209,0.5914,0.4

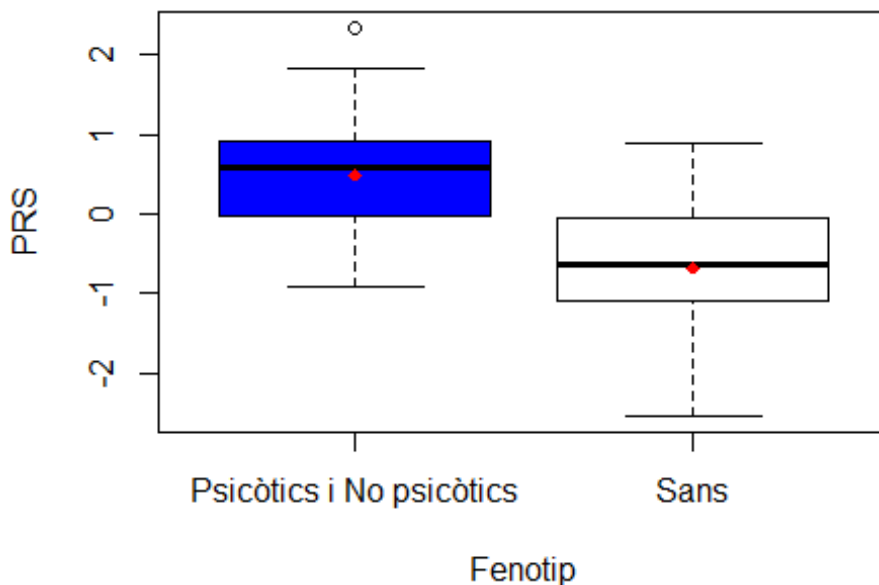
```

```

116,0.3768,-0.3894,-0.9281,-0.7116,-0.833,1.0316)
Sans<-c(0.8904,0.3424,-0.5092,-0.5092,-2.5386,-0.4675,0.3631,-0.0121,-0.6
247,-0.0579,-0.8851,-0.6305,-0.9617,-0.9825,-1.1031,-1.4133,-1.5,-1.6738)
boxplot(Fenotip_ample,Sans, names=c("Psicòtics i No psicòtics","Sans"), y
lab="PRS",main="Puntuacions del SZ-PRS agrupades per fenotip", xlab="Feno
tip",col=c("blue", "white"))
points(1,mean(Fenotip_ample),col="red",pch=18)
points(2,mean(Sans),col="red",pch=18)

```

Puntuacions del SZ-PRS agrupades per fenotip



```

rm(list=ls())
Fenotip_ample<-c(2.3379,1.0658,0.915,0.9134,0.8906,1.591,0.1422,0.6246,0.
3338,0.675,1.8232,0.3816,-0.0299,-0.0856,-0.2522,0.7766,0.6209,0.5914,0.4
116,0.3768,-0.3894,-0.9281,-0.7116,-0.833,1.0316)
Sans<-c(0.8904,0.3424,-0.5092,-0.5092,-2.5386,-0.4675,0.3631,-0.0121,-0.6
247,-0.0579,-0.8851,-0.6305,-0.9617,-0.9825,-1.1031,-1.4133,-1.5,-1.6738,
NA,NA,NA,NA,NA,NA,NA)
Dades_SZ_ample<-data.frame(Fenotip_ample,Sans)
shapiro.test(Fenotip_ample)

##
## Shapiro-Wilk normality test
##
## data: Fenotip_ample
## W = 0.97396, p-value = 0.7459

shapiro.test(Sans)

```

```

##
## Shapiro-Wilk normality test
##
## data: Sans
## W = 0.98377, p-value = 0.9804

stacked<-stack(Dades_SZ_ample)
bartlett.test(values~ind,data=stacked)

##
## Bartlett test of homogeneity of variances
##
## data: values by ind
## Bartlett's K-squared = 0.035488, df = 1, p-value = 0.8506

t.test(values~ind, data=stacked, var.equal=T)

##
## Two Sample t-test
##
## data: values by ind
## t = 4.7098, df = 41, p-value = 2.838e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 0.6698985 1.6756575
## sample estimates:
## mean in group Fenotip_ample          mean in group Sans
##                0.490928                -0.681850

```

PRS BD

1) Fenotip estret: Psicòtics vs No psicòtics i Sans

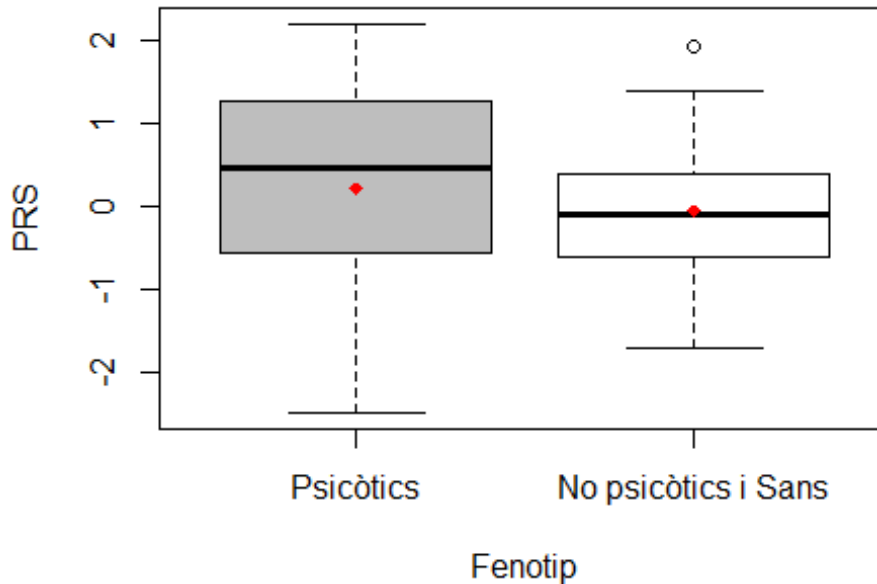
```

rm(list=ls())
Fenotip_estret<-c(1.2391,1.3082,-1.4486,0.7052,-0.4063,-0.4556-2.0309,-0.6962,0.0124,1.4662,2.2005,0.4674)
Sans<-c(0.6981,-0.3057,1.9398,0.0413,0.0532,1.2853,0.2463,0.9282,-0.687,0.1212,-0.0507,-1.2132,1.404,-0.4099,0.483,-0.2002,-0.2002,1.1654,-0.5872,-0.1179,0.7385,-0.844,-0.6625,-0.0956,-1.532,-0.3477,-0.0431,-0.6543,0.287,-1.709,-1.5185)

boxplot(Fenotip_estret,Sans, names=c("Psicòtics","No psicòtics i Sans"),
ylab="PRS",main="Puntuacions del BD-PRS agrupades per fenotip", xlab="Fenotip",col=c("grey", "white"))
points(1,mean(Fenotip_estret),col="red",pch=18)
points(2,mean(Sans),col="red",pch=18)

```

Puntuacions del BD-PRS agrupades per fenotip



```
rm(list=ls())
Fenotip_estret<-c(1.2391,1.3082,-1.4486,0.7052,-0.4063,-0.4556-2.0309,-0.
6962,0.0124,1.4662,2.2005,0.4674,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA)
Sans<-c(0.6981,-0.3057,1.9398,0.0413,0.0532,1.2853,0.2463,0.9282,-0.687,0
.1212,-0.0507,-1.2132,1.404,-0.4099,0.483,-0.2002,-0.2002,1.1654,-0.5872,
-0.1179,0.7385,-0.844,-0.6625,-0.0956,-1.532,-0.3477,-0.0431,-0.6543,0.28
7,-1.709,-1.5185)
Dades_BD_estret<-data.frame(Fenotip_estret,Sans)
shapiro.test(Fenotip_estret)

##
## Shapiro-Wilk normality test
##
## data: Fenotip_estret
## W = 0.96671, p-value = 0.8515

shapiro.test(Sans)

##
## Shapiro-Wilk normality test
##
## data: Sans
## W = 0.97952, p-value = 0.7991

stacked<-stack(Dades_BD_estret)
bartlett.test(values~ind,data=stacked)
```

```
##
## Bartlett test of homogeneity of variances
##
## data: values by ind
## Bartlett's K-squared = 3.5862, df = 1, p-value = 0.05826
```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variàncies, pel que aplicarem el test paramètric "t.test"

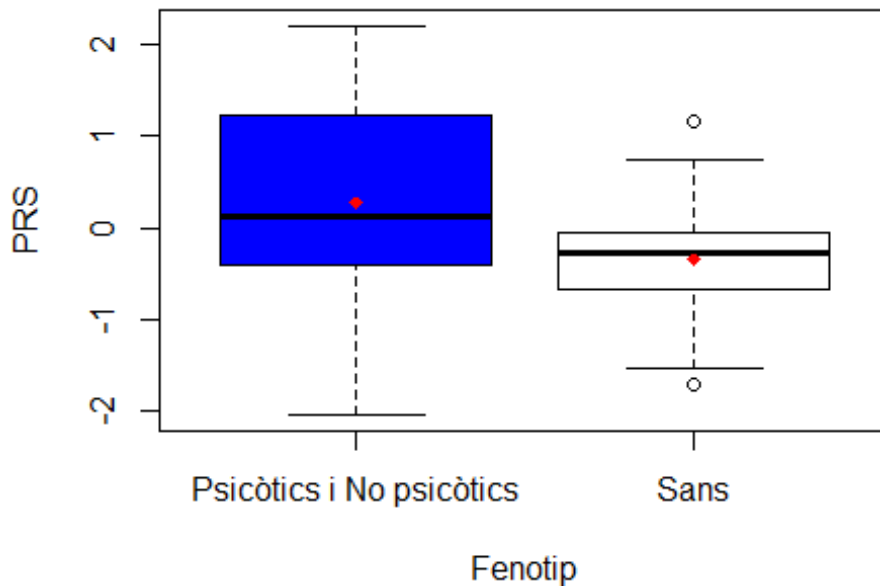
```
t.test(values~ind, data=stacked, var.equal=T)

##
## Two Sample t-test
##
## data: values by ind
## t = 0.75511, df = 40, p-value = 0.4546
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.4565684 1.0012300
## sample estimates:
## mean in group Fenotip_estret          mean in group Sans
##                0.21467273                -0.05765806
```

2) Fenotip ample: Psicòtics i No psicòtics vs Sans

```
rm(list=ls())
Fenotip_ample<-c(1.2391,1.3082,-1.4486,0.7052,-0.4063,-0.4556,-2.0309,-0.6962,0.0124,1.4662,2.2005,0.4674,0.6981,-0.3057,1.9398,0.0413,0.0532,1.2853,0.2463,0.9282,-0.687,0.1212,-0.0507,-1.2132,1.404)
Sans<-c(-0.4099,0.483,-0.2002,-0.2002,1.1654,-0.5872,-0.1179,0.7385,-0.844,-0.6625,-0.0956,-1.532,-0.3477,-0.0431,-0.6543,0.287,-1.709,-1.5185)
boxplot(Fenotip_ample,Sans, names=c("Psicòtics i No psicòtics","Sans"), ylab="PRS",main="Puntuacions del BD-PRS agrupades per fenotip", xlab="Fenotip",col=c("blue", "white"))
points(1,mean(Fenotip_ample),col="red",pch=18)
points(2,mean(Sans),col="red",pch=18)
```

Puntuacions del BD-PRS agrupades per fenotip



```
rm(list=ls())
Fenotip_ample<-c(1.2391,1.3082,-1.4486,0.7052,-0.4063,-0.4556,-2.0309,-0.
6962,0.0124,1.4662,2.2005,0.4674,0.6981,-0.3057,1.9398,0.0413,0.0532,1.28
53,0.2463,0.9282,-0.687,0.1212,-0.0507,-1.2132,1.404)
Sans<-c(-0.4099,0.483,-0.2002,-0.2002,1.1654,-0.5872,-0.1179,0.7385,-0.84
4,-0.6625,-0.0956,-1.532,-0.3477,-0.0431,-0.6543,0.287,-1.709,-1.5185,NA,
NA,NA,NA,NA,NA)
Dades_BD_ample<-data.frame(Fenotip_ample,Sans)
shapiro.test(Fenotip_ample)

##
## Shapiro-Wilk normality test
##
## data: Fenotip_ample
## W = 0.98321, p-value = 0.9405

shapiro.test(Sans)

##
## Shapiro-Wilk normality test
##
## data: Sans
## W = 0.96664, p-value = 0.7327

stacked<-stack(Dades_BD_ample)
bartlett.test(values~ind,data=stacked)
```



```
##
## Bartlett test of homogeneity of variances
##
## data: values by ind
## Bartlett's K-squared = 1.9757, df = 1, p-value = 0.1598
```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variàncies, pel que aplicarem el test paramètric "t.test"

```
t.test(values~ind, data=stacked, var.equal=T)

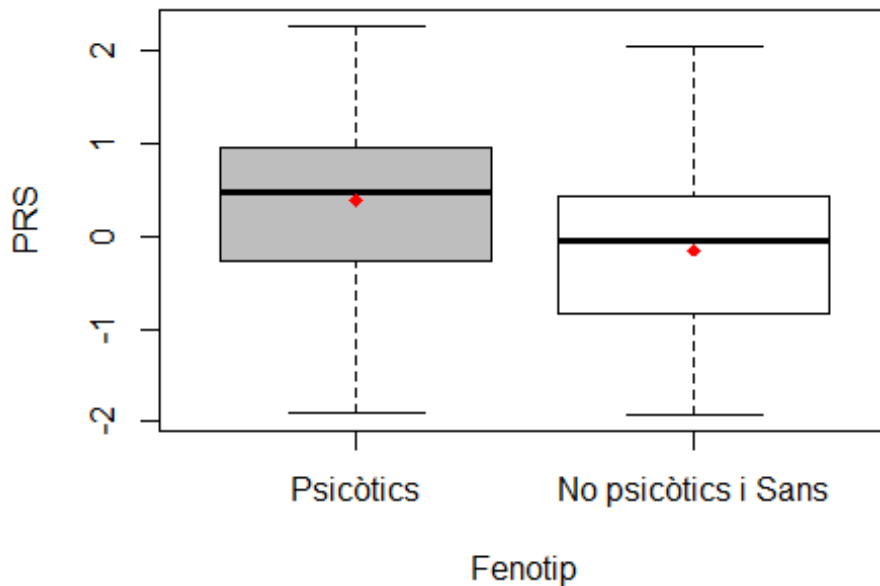
##
## Two Sample t-test
##
## data: values by ind
## t = 2.1122, df = 41, p-value = 0.0408
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 0.02720372 1.21281673
## sample estimates:
## mean in group Fenotip_ample          mean in group Sans
##                0.2728880                -0.3471222
```

PRS SZ+BD

1) Fenotip estret: Psicòtics vs No psicòtics i Sans

```
rm(list=ls())
Fenotip_estret<-c(2.2739,1.0993,0.2907,0.8394,0.6013,-0.7332,-1.9108,-0.7661,0.3634,0.6516,1.7978,0.192)
Sans<-c(0.4085,0.4555,2.0429,0.8573,1.1943,0.9754,0.7409,-0.4514,0.174,-1.3203,-1.0344,-0.8121,0.3966,0.6028,0.5582,-0.0454,-0.0454,-0.9303,-0.1147,0.1321,-0.0416,-0.788,0.3873,-0.8329,-0.2497,-0.9285,-1.1969,-1.7319,-1.9237,-0.6281,-0.5497)
boxplot(Fenotip_estret,Sans, names=c("Psicòtics","No psicòtics i Sans"),
ylab="PRS", main="Puntuacions del SZ+BD-PRS agrupades per fenotip",xlab="Fenotip",col=c("grey", "white"))
points(1,mean(Fenotip_estret),col="red",pch=18)
points(2,mean(Sans),col="red",pch=18)
```

Puntuacions del SZ+BD-PRS agrupades per fenotip



```
rm(list=ls())
Fenotip_estret<-c(2.2739,1.0993,0.2907,0.8394,0.6013,-0.7332,-1.9108,-0.7
661,0.3634,0.6516,1.7978,0.192,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA)
Sans<-c(0.4085,0.4555,2.0429,0.8573,1.1943,0.9754,0.7409,-0.4514,0.174,-1
.3203,-1.0344,-0.8121,0.3966,0.6028,0.5582,-0.0454,-0.0454,-0.9303,-0.114
7,0.1321,-0.0416,-0.788,0.3873,-0.8329,-0.2497,-0.9285,-1.1969,-1.7319,-1
.9237,-0.6281,-0.5497)
Dades_SZBD_estret<-data.frame(Fenotip_estret,Sans)
shapiro.test(Fenotip_estret)

##
## Shapiro-Wilk normality test
##
## data: Fenotip_estret
## W = 0.96667, p-value = 0.8729

shapiro.test(Sans)

##
## Shapiro-Wilk normality test
##
## data: Sans
## W = 0.98713, p-value = 0.9641

stacked<-stack(Dades_SZBD_estret)
bartlett.test(values~ind,data=stacked)
```

```
##
## Bartlett test of homogeneity of variances
##
## data: values by ind
## Bartlett's K-squared = 0.95302, df = 1, p-value = 0.329
```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variàncies, pel que aplicarem el test paramètric “t.test”

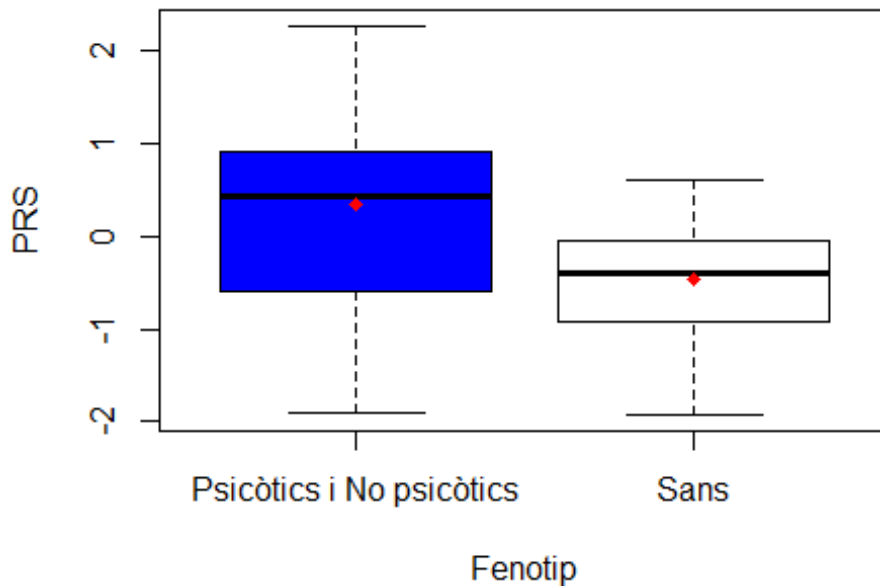
```
t.test(values~ind, data=stacked, var.equal=T)
```

```
##
## Two Sample t-test
##
## data: values by ind
## t = 1.6499, df = 41, p-value = 0.1066
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.1217036 1.2080944
## sample estimates:
## mean in group Fenotip_estret          mean in group Sans
##                0.3916083                -0.1515871
```

2) Fenotip ample: Psicòtics i No psicòtics vs Sans

```
rm(list=ls())
Fenotip_ample<-c(2.2739,1.0993,0.2907,0.8394,0.6013,-0.7332,-1.9108,-0.76
61,0.3634,0.6516,1.7978,0.192,0.4085,0.4555,2.0429,0.8573,1.1943,0.9754,0
.7409,-0.4514,0.174-1.3203,-1.0344,-0.8121,0.3966)
Sans<-c(0.6028,0.5582,-0.0454,-0.0454,-0.9303,-0.1147,0.1321,-0.0416,-0.7
88,0.3873,-0.8329,-0.2497,-0.9285,-1.1969,-1.7319,-1.9237,-0.6281,-0.5497
)
boxplot(Fenotip_ample,Sans, names=c("Psicòtics i No psicòtics","Sans"), y
lab="PRS",main="Puntuacions del SZ+BD-PRS agrupades per fenotip", xlab="F
enotip",col=c("blue", "white"))
points(1,mean(Fenotip_ample),col="red",pch=18)
points(2,mean(Sans),col="red",pch=18)
```

Puntuacions del SZ+BD-PRS agrupades per fenotip



```
rm(list=ls())
Fenotip_ample<-c(2.2739,1.0993,0.2907,0.8394,0.6013,-0.7332,-1.9108,-0.76
61,0.3634,0.6516,1.7978,0.192,0.4085,0.4555,2.0429,0.8573,1.1943,0.9754,0
.7409,-0.4514,0.174-1.3203,-1.0344,-0.8121,0.3966)
Sans<-c(0.6028,0.5582,-0.0454,-0.0454,-0.9303,-0.1147,0.1321,-0.0416,-0.7
88,0.3873,-0.8329,-0.2497,-0.9285,-1.1969,-1.7319,-1.9237,-0.6281,-0.5497
,NA,NA,NA,NA,NA,NA)
Dades_SZBD_ample<-data.frame(Fenotip_ample,Sans)
shapiro.test(Fenotip_ample)

##
## Shapiro-Wilk normality test
##
## data: Fenotip_ample
## W = 0.96708, p-value = 0.5957

shapiro.test(Sans)

##
## Shapiro-Wilk normality test
##
## data: Sans
## W = 0.9578, p-value = 0.5598

stacked<-stack(Dades_SZBD_ample)
bartlett.test(values~ind,data=stacked)
```

```
##  
## Bartlett test of homogeneity of variances  
##  
## data: values by ind  
## Bartlett's K-squared = 2.3363, df = 1, p-value = 0.1264
```

Els valors es presenten amb una distribució normal i amb heterogeneïtat de variàncies, pel que aplicarem el test paramètric “t.test”

```
t.test(values~ind, data=stacked, var.equal=T)
```

```
##  
## Two Sample t-test  
##  
## data: values by ind  
## t = 2.8243, df = 40, p-value = 0.007353  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## 0.230229 1.388802  
## sample estimates:  
## mean in group Fenotip_ample          mean in group Sans  
##                0.3469375                -0.4625778
```