**Supplementary Information: Identifying schizophrenia patients who carry pathogenic genetic copy number variants- the importance of the neurodevelopmental history.**

Supplementary table 1. Proportions of missing data in the Irish cohort for the phenotypes included in the analysis (discovery dataset).

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | **N (Total)** | **N (Data available)** | **% Complete** |
| Early age of onset of symptoms | 1215 | 1199 | 98.7 |
| History of learning difficulties | 1215 | 1195 | 98.4 |
| History of specific learning disorder | 1215 | 1192 | 98.1 |
| History of school supports | 1215 | 1195 | 98.4 |
| Primary school attainment only | 1215 | 1204 | 99.1 |
| History of developmental delay | 1215 | 1162 | 95.6 |
| Comorbid neurodevelopmental diagnosis | 1215 | 1195 | 98.4 |
| Family history of neurodevelopmental disorder | 1215 | 1087 | 89.5 |
|  |

Supplementary table 2. Proportions of missing data in the Cardiff cohort for the phenotypes included in the analysis (replication dataset).

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | **N (Total)** | **N (Data available)** | **% Complete** |
| History of developmental delay | 479 | 337 | 70.4 |
| Comorbid neurodevelopmental diagnosis | 479 | 417 | 87.1 |
|  |  |  |  |

Supplementary table 3. Schizophrenia-associated risk CNVs.

|  |  |
| --- | --- |
| **Locus** | **Position in Mb** |
| 1q21.1 del | chr1:146,57-147,39 |
| 1q21.1 dup | chr1:146,57-147,39 |
| NRXN1 del | chr2:50,15-51,26 |
| 3q29 del | chr3:195,73-197,34 |
| WBS dup | chr7:72,74-74,14 |
| VIPR2 dup | chr7:158,82-158,94 |
| 15q11.2 del | chr15:22,80-23,09 |
| AS/PWS dup | chr15:24,82-28,43 |
| 15q13.3 del | chr15:31,13-32,48 |
| 16p13.11 dup | chr16:15,51-16,30 |
| 16p11.2 distal del | chr16:28,82-29,05 |
| 16p11.2 dup | chr16:29,64-30,20 |
| 17p12 del | chr17:14,16-15,43 |
| 17q12 del | chr17:34,81-36,20 |
| 22q11.2 del | chr22:19,02-20,26 |
| Note: Copy number variant loci listed in Rees et al. (1) with previous evidence for associations with schizophrenia. Copy number variation positions are in UCSC Build 37. Del, deletion; dup, duplication; WBS, Williams-Beuren Syndrome; AS/PWS, Angelman/Prader-Willi syndrome. |

Supplementary table 4. Copy number variants identified in individuals in discovery dataset.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Case** | **Locus** | **CNV** | **Chr** | **Position in Mb** |
|  | **Start** | **Stop** |
| Case 1 | 1q21.1 | Del | 1 | 145932468 | 147831184 |
| Case 2 | 1q21.1 | Del | 1 | 146330584 | 147825662 |
| Case 3 | 1q21.1 | Del | 1 | 146717564 | 146970946 |
| Case 4 | NRXN1 | Del | 2 | 50870373 | 50971478 |
| Case 5 | 15q11.2 | Del | 15 | 22751094 | 23487547 |
| Case 6 | 15q11.2 | Del | 15 | 22770994 | 23236972 |
| Case 7 | 15q11.2 | Del | 15 | 22770994 | 23236972 |
| Case 8 | 15q13.3 | Del | 15 | 30755047 | 32489254 |
| Case 9 | 17p12 | Del | 17 | 14098277 | 15475088 |
| Case 10 | WBS | Dup | 7 | 72732833 | 74142105 |
| Case 11 | AS/PWS | Dup | 15 | 24727231 | 25107593 |
| Case 12 | 16p13.11 | Dup | 16 | 14929751 | 16396513 |
| Case 13 | 16p13.11 | Dup | 16 | 15481934 | 18314234 |
| Case 14 | 16p13.11 | Dup | 16 | 15481934 | 17296550 |
| Case 15 | 16p13.11 | Dup | 16 | 15481934 | 16340441 |
| Case 16 | 16p11.2 | Dup | 16 | 29517711 | 30177808 |
| Case 17 | 16p11.2 | Dup | 16 | 29517711 | 30306969 |
| Case 18 | 16p11.2 | Dup | 16 | 29580034 | 30191908 |
| Case 19 | 16p11.2 | Dup | 16 | 29595483 | 30017620 |
| Note: Copy number variation positions are in UCSC Build 37. Del, deletion; dup, duplication; WBS, Williams-Beuren Syndrome; AS/PWS, Angelman/Prader-Willi syndrome. |

**Supplementary table 5. Proportion of individuals with positive history of phenotypic variables**

|  |  |  |
| --- | --- | --- |
|  | **Discovery Sample** | **Replication Sample** |
| **Phenotypic Variable****(References)** | CNV(n = 19) | No CNV(n = 1196) | Total(n = 1215) | CNV(n = 8) | No CNV(n = 471) | Total(n = 479) |
| Early onset of symptoms(2, 3) | 6 (31.6%) | 270 (22.6%) | 276 (22.7%) | - | - | - |
| History of learning difficulties(4-6) | 8 (42.1%) | 181 (15.1%) | 189 (15.6%) | - | - | - |
| Specific learning disorder(4-6) | 2 (10.5%) | 15 (1.3%) | 17 (1.4%) | - | - | - |
| Remedial school support(4-6) | 1 (5.3%) | 36 (3.0%) | 37 (3.0%) | - | - | - |
| Low educational attainment(4-6) | 4 (21.1%) | 150 (12.5%) | 154 (12.7%) | - | - | - |
| History of developmental delay(7-9) | 5 (26.3%) | 77 (6.4%) | 82 (6.7%) | 3 (37.5%) | 92 (19.5%) | 95 (19.8%) |
| Comorbid neurodevelopmental diagnosis(7-9) | 3 (15.8%) | 43 (3.6%) | 46 (3.8%) | 4 (50.0%) | 40 (8.5%) | 44 (9.2%) |
| Family history of Neurodevelopmental disorder(10, 11) | 5 (26.3%) | 343 (28.7%) | 348 (28.6%) | - | - | - |
| Note: Proportions of individuals with positive history of phenotypic variables and SCZ-associated pathogenic CNV status. |

Supplementary table 6. Multiple logistic regression model using two phenotypic variables to model SCZ-associated CNV carrier status in Irish data set.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable** | **β** | **SE** | **Wald statistic (Z)** | **P** | **OR (95% CI)** |
| Comorbid neurodevelopmental disorder | 1.76 | 0.67 | 2.63 | 0.009 | 5.81 (1.28-19.29) |
| History of developmental delay | 1.72 | 0.55 | 3.13 | 0.002 | 5.60 (1.73-15.74) |
| Note: Predictor coefficients were tested using Wald tests and confidence intervals obtained using the Wald method. Nagelkerke pseudo R Square = 0.172). |

Supplementary table 7. ROC curve results for modelling SCZ-associated CNV status using two phenotypic variables in the Irish dataset.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Cutoff Value** | **Sensitivity (%)****(95% CI)** | **Specificity (%)****(95% CI)** | **AUC (%)****(95% CI)** | **PPV (%)****(95% CI)** | **NPV (%)****(95% CI)** |
| 0.05 | 47.1 (23.0-72.2) | 90.1 (88.3-91.8) | 68.6 (56.3-80.9) | 6.7 (5.5-17.3) | 99.1 (97.4-99.3) |
| Note: Optimal Cutoff Value, Sensitivity, Specificity, Area Under The Curve, and Predictive Values using two independent variables (“History of developmental delay”, “Comorbid neurodevelopmental disorder”) to model for SCZ-associated CNV status in Irish cohort (discovery dataset).CI, confidence interval; AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value |

**Schizophrenia Working Group of the Psychiatric Genomics Consortium**

Rolf Adolfsson, Ingrid Agartz, Esben Agerbo, Margot Albus, Madeline Alexander, Farooq Amin, Ole A Andreassen, Silviu A Bacanu, Martin Begemann, Richard A Belliveau Jr, Judit Bene, Sarah E Bergen, Elizabeth Bevilacqua, Tim B Bigdeli, Donald W Black, Douglas HR Blackwood, Anders D Børglum, Elvira Bramon, Richard Bruggeman, Nancy G Buccola, Randy L Buckner, Brendan Bulik-Sullivan, Joseph D Buxbaum, William Byerley, Wiepke Cahn, Guiqing Cai, Murray J Cairns, Dominique Campion, Rita M Cantor, Vaughan J Carr, Noa Carrera, Stanley V Catts, Kimberley D Chambert, Raymond CK Chan, Eric YH Chen, Ronald YL Chen, Wei Cheng, Eric FC Cheung, Siow Ann Chong, Sven Cichon, C Robert Cloninger, David Cohen, Nadine Cohen, David A Collier, Paul Cormican, Aiden Corvin, Nick Craddock, Benedicto Crespo-Facorro, James J Crowley, David Curtis, Mark J Daly, Ariel Darvasi, Michael Davidson, Kenneth L Davis, Franziska Degenhardt, Jurgen Del Favero, Lynn E DeLisi, Ditte Demontis, Dimitris Dikeos, Timothy Dinan, Srdjan Djurovic, Enrico Domenici, Gary Donohoe , Elodie Drapeau, Jubao Duan, Frank Dudbridge, Hannelore Ehrenreich, Peter Eichhammer, Johan Eriksson, Valentina Escott-Price, Tõnu Esko, Laurent Essioux, Ayman H Fanous, Kai-How Farh, Martilias S Farrell, Josef Frank, Lude Franke, Robert Freedman, Nelson B Freimer, Joseph I Friedman, Menachem Fromer, Pablo V Gejman, Giulio Genovese, Elliot S Gershon, Ina Giegling, Michael Gill , Paola Giusti-Rodríguez, Stephanie Godard, Jacqueline I Goldstein, Srihari Gopal, Jacob Gratten, Lieuwe de Haan, Christian Hammer, Marian L Hamshere, Mark Hansen, Thomas Hansen, Vahram Haroutunian, Annette M Hartmann, Frans A Henskens, Stefan Herms, Joel N Hirschhorn, Per Hoffmann, Andrea Hofman, Mads V Hollegaard, Peter A Holmans, David M Hougaard, Hailiang Huang, Christina M Hultman, Masashi Ikeda, Nakao Iwata, Assen V Jablensky, Inge Joa , Erik G Jönsson, Antonio Julià, Anna K Kähler, René S Kahn, Luba Kalaydjieva, Sena Karachanak-Yankova, Juha Karjalainen, David Kavanagh, Matthew C Keller, Brian J Kelly, Kenneth S Kendler, James L Kennedy, Andrey Khrunin, Yunjung Kim, George Kirov, Janis Klovins, Jo Knight, James A Knowles, Bettina Konte, Vaidutis Kucinskas, Zita Ausrele Kucinskiene, Hana Kuzelova-Ptackova, Max Lam, Claudine Laurent, Phil Lee, S Hong Lee, Jimmy Lee Chee Keong, Sophie E Legge, Todd Lencz, Bernard Lerer, Douglas F Levinson, Miaoxin Li, Tao Li, Qingqin S Li, Kung-Yee Liang, Jeffrey Lieberman, Svetlana Limborska, Jianjun Liu, Jouko Lönnqvist, Carmel M Loughland, Jan Lubinski, Milan Macek Jr, Patrik KE Magnusson, Brion S Maher, Wolfgang Maier, Anil K Malhotra, Jacques Mallet, Sara Marsal, Manuel Mattheisen, Morten Mattingsdal, Robert W McCarley, Steven A McCarroll, Colm McDonald, Andrew M McIntosh, Andrew McQuillin, Sandra Meier, Carin J Meijer, Bela Melegh, Ingrid Melle, Raquelle I Mesholam-Gately, Andres Metspalu, Patricia T Michie, Lili Milani, Vihra Milanova, Younes Mokrab, Jennifer L Moran, Derek W Morris, Ole Mors, Preben B Mortensen, Bryan J Mowry, Bertram Müller-Myhsok, Kieran C Murphy, Robin M Murray, Inez Myin-Germeys, Benjamin M Neale, Mari Nelis, Igor Nenadic, Deborah A Nertney, Gerald Nestadt, Kristin K Nicodemus, Liene Nikitina-Zake, Laura Nisenbaum, Annelie Nordin, Nina Norgren, Markus M Nöthen, Eadbhard O'Callaghan, Michael C O'Donovan, Colm O'Dushlaine, F Anthony O'Neill, Sang-Yun Oh, Ann Olincy, Line Olsen, Roel A Ophoff, Jim Van Os, Michael J Owen, Sara A Paciga, Aarno Palotie, Christos Pantelis, George N Papadimitriou, Sergi Papiol, Elena Parkhomenko, Michele T Pato, Carlos N Pato, Tiina Paunio, Diana O Perkins, Tune H Pers, Tracey L Petryshen, Olli Pietiläinen, Jonathan Pimm, Andrew J Pocklington, Danielle Posthuma, John Powell, Alkes Price, Ann E Pulver, Shaun M Purcell, Digby Quested, Henrik B Rasmussen, Abraham Reichenberg, Mark A Reimers, Alexander L Richards, Marcella Rietschel, Brien P Riley, Stephan Ripke, Joshua L Roffman, Panos Roussos, Douglas M Ruderfer, Dan Rujescu, Veikko Salomaa, Alan R Sanders, Ulrich Schall, Thomas G Schulze, Sibylle G Schwab, Edward M Scolnick, Rodney J Scott, Larry J Seidman, Pak C Sham, Jianxin Shi, Jeremy M Silverman, Kang Sim, Pamela Sklar, Petr Slominsky, Jordan W Smoller, Hon-Cheong So, Erik Söderman, Chris C A Spencer, David St Clair, Eli A Stahl, Elisabeth Stogmann , Richard E Straub, Eric Strengman, Jana Strohmaier, T Scott Stroup, Mythily Subramaniam, Patrick F Sullivan, Jaana Suvisaari, Dragan M Svrakic, Jin P Szatkiewicz, Srinivas Thirumalai, Draga Toncheva, Paul A Tooney, Sarah Tosato, Peter M Visscher, John Waddington, Dermot Walsh, James TR Walters, Dai Wang, Qiang Wang, Bradley T Webb, Daniel R Weinberger, Mark Weiser, Thomas Werge, Dieter B Wildenauer, Nigel M Williams, Stephanie Williams, Stephanie H Witt, Aaron R Wolen, Emily HM Wong, Brandon K Wormley, Naomi R Wray, Jing Qin Wu, Hualin Simon Xi, Clement C Zai, Xuebin Zheng, Fritz Zimprich.

**References**

1. Rees E, Walters JTR, Georgieva L, Isles AR, Chambert KD, Richards AL, et al. Analysis of copy number variations at 15 schizophrenia-associated loci. The British Journal of Psychiatry. 2014;204(2):108-14.

2. Ahn K, Gotay N, Andersen TM, Anvari AA, Gochman P, Lee Y, et al. High rate of disease-related copy number variations in childhood onset schizophrenia. Mol Psychiatry. 2014;19(5):568-72.

3. Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science (New York, NY). 2008;320(5875):539-43.

4. Derks EM, Ayub M, Chambert K, Del Favero J, Johnstone M, MacGregor S, et al. A genome wide survey supports the involvement of large copy number variants in schizophrenia with and without intellectual disability. American journal of medical genetics Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics. 2013;162b(8):847-54.

5. Stefansson H, Meyer-Lindenberg A, Steinberg S, Magnusdottir B, Morgen K, Arnarsdottir S, et al. CNVs conferring risk of autism or schizophrenia affect cognition in controls. Nature. 2014;505(7483):361-6.

6. Kendall KM, Rees E, Escott-Price V, Einon M, Thomas R, Hewitt J, et al. Cognitive Performance Among Carriers of Pathogenic Copy Number Variants: Analysis of 152,000 UK Biobank Subjects. Biol Psychiatry. 2016.

7. Sahoo T, Theisen A, Rosenfeld JA, Lamb AN, Ravnan JB, Schultz RA, et al. Copy number variants of schizophrenia susceptibility loci are associated with a spectrum of speech and developmental delays and behavior problems. Genet Med. 2011;13(10):868-80.

8. Kirov G, Rees E, Walters JT, Escott-Price V, Georgieva L, Richards AL, et al. The penetrance of copy number variations for schizophrenia and developmental delay. Biol Psychiatry. 2014;75(5):378-85.

9. Fung WLA, Butcher NJ, Costain G, Andrade DM, Boot E, Chow EW, et al. Practical guidelines for managing adults with 22q11. 2 deletion syndrome. Genetics in Medicine. 2015;17(8):599.

10. Wilson NK, Lee Y, Long R, Hermetz K, Rudd MK, Miller R, et al. A Novel Microduplication in the Neurodevelopmental Gene SRGAP3 That Segregates with Psychotic Illness in the Family of a COS Proband. Case reports in genetics. 2011;2011:585893.

11. Costain G, Lionel AC, Fu F, Stavropoulos DJ, Gazzellone MJ, Marshall CR, et al. Adult neuropsychiatric expression and familial segregation of 2q13 duplications. American journal of medical genetics Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics. 2014;165b(4):337-44.