The 22q11 PRODH/DGCR6 deletion is frequent in hyperprolinemic subjects but is not a strong risk factor for ASD.

Anne-Claire Richard, Anne Rovelet-Lecrux, Elsa Delaby, Camille Charbonnier, Bhooma Thiruvahindrapuram, Eli Hatchwell, Peggy S Eis, Alexandra Afenjar, Brigitte Gilbert Dussardier, Stephen W Scherer, et al.

HAL Id: inserm-01289348

https://inserm.hal.science/inserm-01289348

Submitted on 16 Mar 2016

The 22q11 *PRODH/DGCR6* deletion is frequent in hyperprolinemic subjects but is not a strong risk factor for ASD

Anne Claire Richard[1], Anne Rovelet-Lecrux[1], Elsa Delaby[2,3,4], Camille Charbonnier[1], Bhooma Thiruvahindrapuram[5], Eli Hatchwell[6], Peggy S. Eis[6], Alexandra Afenjar[7], Brigitte Gilbert Dussardier[8], Stephen W. Scherer[5,9], Catalina Betancur[2,3,4], Dominique Campion[1, 10]*

[1]Inserm U1079, Institute for Research and Innovation in Biomedicine, University of Rouen, France
[2]Inserm U1130, Neuroscience Paris Seine, Paris, France
[3]CNRS UMR 8246, Neuroscience Paris Seine, Paris, France
[4]Sorbonne Universités, UPMC Univ Paris 6, Institut de Biologie Paris-Seine, Paris, France
[5]The Centre for Applied Genomics, and Program in Genetics and Genome Biology, the Hospital for Sick Children, Toronto, Ontario, Canada
[6]Population Diagnostics, Inc., Melville, New York, USA
[7]Unité de neuropédiatrie et pathologie du développement, CHU Paris Est - Hôpital d'Enfants Armand-Trousseau, Paris, France
[8]Service de génétique médicale, CHU de Poitiers, France
[9]McLaughlin Centre and Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada
[10]Centre Hospitalier du Rouvray, Sotteville-lès-Rouen, France

*Correspondence to: Dominique Campion, Inserm UMR1079, Faculté de Médecine, 22 boulevard Gambetta, 76183 Rouen, France. Email: dominique.campion@univ-rouen.fr

Abstract

The proline dehydrogenase *(PRODH)* gene maps to 22q11.2 in the region deleted in the velo-cardio-facial syndrome (VCFS). A moderate to severe reduction (>50%) in PRODH activity resulting from recessive deletions and/or missense mutations has been shown to cause type 1 hyperprolinemia (HPI). Autistic features have been reported as a common clinical manifestation of HPI. Here we studied the frequency of a recurrent small 22q11.2 deletion encompassing *PRODH* and the neighboring *DGCR6* gene in three case-control studies, one comprising HPI patients (n=83), and the other two comprising autism spectrum disorder (ASD) patients (total of n=2800), analyzed with high-resolution microarrays. We found that the *PRODH* deletion is a strong risk factor for HPI (OR = 50.7; 95% CI = 7.5-2147) but not for ASD (p=0.4, OR= 0.6-3.3). This result indicates either that the suggested association between ASD and HPI is spurious and results from a bias leading to the preferential inclusion of ASD patients in HPI series, or that HPI is present in only a very small subset of ASD patients. In this latter case, a very large sample size would be required to detect an association between the *PRODH* deletion and ASD in a case-control study.

Key words: Proline dehydrogenase, copy number variant, 22q11.2 deletion, autism, hyperprolinemia

INTRODUCTION

The *PRODH* gene encodes proline dehydrogenase, a mitochondrial enzyme that catalyzes the conversion of proline into Δ1-pyrroline-5-carboxylate (P5C). It is located in the 22q11.2 chromosomal region, near the segmental duplication that defines the proximal boundary of the 3 Mb segment commonly deleted in velo-cardio-facial syndrome (VCFS) [Bittel et al. 2009]. Typical 22q11.2 deletions in VCFS patients include ~30 genes. We previously identified a smaller 22q11.2 deletion encompassing only *PRODH* and the neighboring *DGCR6* gene in a schizoaffective patient without VCFS clinical features [Jacquet et al. 2002] (Figure S1). The frequency of this copy number variant (CNV) was subsequently reported to be 1/250 individuals in a Japanese population, making it 10-fold more prevalent than the VCFS-associated 22q11.2 deletion [Ohtsuki et al. 2004]. The reciprocal duplication was even more common, with an estimated frequency of 2.5/100 individuals [Guilmatre et al. 2010]. The *PRODH/DGCR6* deletion is frequently observed in patients with type 1 hyperprolinemia (HPI) [Jacquet et al. 2003; Guilmatre et al. 2010]. Since HPI is a recessive condition, the *PRODH* deletion must be combined with another hit (usually a missense mutation reducing enzymatic activity) on the second allele to cause an increase in plasma proline levels [Guilmatre et al. 2010]. The clinical phenotype of HPI is still not well characterized. While some HPI subjects have intellectual disability, seizures, language delay, autistic features, schizophrenia or bipolar disorder, others are asymptomatic [Afenjar et al. 2007; Di Rosa et al. 2008; Guilmatre et al. 2010; Di Rosa et al. 2014]. In particular, autism spectrum disorder (ASD) has been reported as a prevalent clinical feature of HPI, although the diagnosis often does not rely on formal criteria [Afenjar et al. 2007; Di Rosa et al. 2008; Guilmatre et al. 2010]. The involvement of HPI in ASD determinism is supported by the fact that PRODH deficiency in mice results in increased neurotransmitter release at glutamatergic synapses [Paterlini et al. 2005] and glutamate is possibly involved in ASD pathology [reviewed in Choudhury et al. 2012].

Thus, the question arises whether the *PRODH/DGCR6* deletion is associated with ASD. Supporting a putative role as a risk factor, Guilmatre et al. [2009] reported a significant association of this deletion with ASD in a small sample (9/260 ASD cases vs. 1/236 controls, p=0.02 Fisher exact test, OR= 8.4; 95% CI= 1.15-370.2). Due to its combined frequency >1% in the general population, *PRODH/DGCR6* deletions and reciprocal duplications were not reported in most of the recently published large CNV case-control studies in ASD, which focused on rare variants, usually with a frequency <1% [Pinto et al. 2010; Sanders et al. 2011; Gai et al. 2012; Girirajan et al. 2013; Pinto et al. 2014]. Moreover, since these studies used different microarray platforms, this CNV might have been missed due to insufficient probe coverage or differences in CNV-calling algorithms. To further clarify the implication of the *PRODH/DGCR6* CNV in ASD risk we examined its frequency in three case/control studies, one including HPI subjects and two including ASD patients, using data generated by both CGH and SNP array platforms.

PATIENTS AND METHODS

HPI study

To constitute the HPI cohort, we reviewed all HPI cases (with fasting plasma proline > 400 µmol/L) in which a genetic analysis of the *PRODH* gene had previously been undertaken by our laboratory. Only cases where the phase of *PRODH* mutations was unambiguously established were retained (n=83). VCFS patients were not included. The mean age of the patients was 9.6 years (SD=2.8, range=2–42). All cases except five were Caucasian. The control cohort consisted of 236 adult healthy Caucasian subjects with normal plasma proline level [Guilmatre et al. 2009]. This study was approved by the Rouen University Hospital ethics committee and all patients or their parents and controls signed a written informed consent. The DNA samples were derived from whole blood. CNV analysis was performed by using the Human High-Resolution Discovery Microarray Kit 1x1 M (design ID 021529, Agilent Technologies, Santa Clara, CA, USA) following standard protocols. A non-commercial genomic DNA pool of 10 control individuals was used as a reference sample. Hybridization results were extracted with the Feature Extraction software v.10.5.1.1 (Agilent Technologies) and analyzed using the DNA Analytics software v.4.0.81 (Agilent Technologies). The data were processed by the ADM-2 algorithm, with the threshold set at 6.0 SD. CNVs were confirmed by quantitative multiplex PCR of short fluorescent fragments (QMPSF) as previously described [Guilmatre et al. 2009]. Sanger sequencing of the *PRODH* gene was performed on an ABI prism 3100 automated sequencer (Applied Biosystems, Foster City, CA, USA) using previously described primers [Raux et al. 2007].

ASD sample 1

The subjects analyzed by the international consortium Autism Genome Project (AGP) included 2,446 ASD unrelated cases (2,114 males and 332 females) and 4,768 controls (2,022 males and 2,746 females) [Pinto et al. 2014]. Affected individuals were assessed using the Autism Diagnostic Interview-Revised (ADI-R) and/or the Autism Diagnostic Observation Schedule (ADOS). Unrelated **adult** control subjects were assembled from three studies in which individuals had no obvious psychiatric history: the Study of Addiction Genetics and Environment [Bierut et al. 2010], the Ontario Colorectal Cancer Case-Control Study [Figueiredo et al. 2011], and Health, Aging, and Body Composition [Fox et al. 2012]. The ASD cases and comparison subjects were not matched for age. The Autism Genome Project cases included both children and adults, whereas the control subjects where all adults, reflecting the fact that they were not recruited specifically for the ASD study. Given that this is a genetic association study, the age of the subjects should not have any effect on the results.

DNA samples were derived from whole blood or cell lines. All samples were genotyped using Illumina 1M single or duo microarrays (Illumina, San Diego, CA, USA) and analyzed with the same quality-control procedures and CNV analysis pipeline. CNVs were detected using a multiple-algorithm approach including QuantiSNP [Colella et al. 2007], iPattern [Pinto et al. 2011], and PennCNV [Wang et al. 2007]. CNVs spanning less than five consecutive probes or less than 5 kb, or with low

confidence score were excluded. For a detailed description of quality control procedures, see Pinto et al. [2010]. Ancestry was inferred from genotyping data; among the cases, 2,147 (87.8%) were of European ancestry; in the control cohort, 2,640 (55.4%) were of European ancestry. Informed consent was obtained from all participants or their parents, and all procedures followed were in accordance with the ethical standards on human experimentation of the participating sites.

ASD sample 2

This study included 676 unrelated ASD cases and 1005 controls that passed array quality control. The ASD samples (560 males and 116 females) came from three Canadian sites: Hospital for Sick Children in Toronto, Ontario; McMaster University, Hamilton, Ontario; and Memorial University of Newfoundland, St. John's, Newfoundland [Prasad et al. 2012]. All the cases met the criteria for autism on one or both diagnostic measures – ADI-R and ADOS. CNV analysis was performed on genomic DNA derived from blood or lymphoblastoid cell lines. Data for the control cohort of 1005 DNA samples were provided by Population Diagnostics, Inc. (Melville, NY, USA) [Prasad et al. 2012]. Briefly, the cohort (referred to as PDx controls) comprised reportedly healthy donors of European ancestry (500 females and 505 males, >45 years of age) with no known history of disease, including psychiatric disorders. All donors were consented and deidentified *via* a protocol approved by the institutional review board. The DNA samples were all derived from whole blood. All case and control samples were run on the Agilent 1M CGH array (Agilent Technologies, Santa Clara, CA, USA) and used CNV-calling methods as previously described [Prasad et al. 2012]. The reference DNA for cases was a pool of 50 sex-matched European ancestry control samples; for controls, the same sex-matched sample (one female or one male) was used. Ancestry analysis in 615 out of 676 cases with SNP microarray data available indicated that 505 cases were of European ancestry. Of the 676 ASD cases studied with Agilent 1M CGH arrays, 322 were also included in the AGP and were studied with SNP arrays [Pinto et al. 2010; Pinto et al. 2014], thus allowing a direct comparison between the two microarray platforms.

RESULTS

HPI cohort

Among the 83 HPI subjects, 15 (18%) carried the *PRODH/DGCR6* deletion (Table I). In four patients the deletion was homozygous. CGH array analysis indicated that *PRODH* deletions and duplications had similar breakpoints, with an approximate size of 120 kb (Figure S2). The CNV is immediately preceded by a 237 kb-long segmental duplication (also termed low-copy repeat, LCR) region, referred to as LCR22-2, which corresponds to the proximal breakpoint of the classical VCFS deletions (Figure S1). Of note, the presence of this LCR is characterized by a gap in probe coverage on the CGH array, which makes the exact determination of the breakpoint uncertain. The clinical phenotype of HPI subjects included ASD in 7/15 subjects (Table I). In the control group (Table II), only one subject carried the *PRODH/DGCR6* deletion. His plasma proline level (337 μmol/L) was at the upper limit of

normal values and his complete *PRODH* genotype was Del/Q19P+T275N. The frequency of deletion carriers (either at the homozygous or heterozygous state) significantly differed between cases and controls (p= 8.6×10^{-9}, Fisher exact test), and carrying at least one deleted allele was a strong risk factor for HPI (OR = 50.7; 95% CI = 7.5-2,147).

ASD samples

In ASD study 1 (AGP), the SNP array revealed that 11/2,446 ASD subjects carried the deletion (10 males, 1 female), compared to 15/4,768 control individuals (8 males, 7 females) (p = 0.4, OR = 0.6-3.3) (Table II). Two ASD cases with a larger deletion corresponding to VCFS were not included in these counts. In ASD study 2, array CGH analysis identified 6 *PRODH/DGCR6* deletion carriers among the 676 ASD individuals (4 deletions in 354 unique individuals not included in the AGP sample), and 9 deletion carriers among the 1005 control individuals (p = 1; OR = 0.3-3.2, Fisher exact test). Analyses restricted to European cases and controls yielded similar negative findings (Table II). Genotype calls were concordant among the 322 ASD patients assessed on both Agilent 1M CGH and SNP array platforms. The two deletion carriers present in this subsample were found in both analyses, indicating that two microarray platforms seem equally suitable to detect the *PRODH/DGCR6* deletion.

There were 45 duplications in the AGP cases (1.8%, 41 males, 4 females; 41 in Europeans) - excluding 5 duplications of the VCFS region - and 102 in controls (2.1%, 45 males, 57 females; 57 in Europeans, 11 mixed ethnicity, 34 other ethnicity). The overall frequency of the duplication in controls and ASD patients was consistent with that previously reported (2.5%) [Guilmatre et al. 2010], and is also comparable to what was observed in the PDx control cohort (18 duplications corresponding to 1.8% frequency, 7 males, 11 females; all European).

DISCUSSION

Despite a possible role of hyperprolinemia in ASD, and the strong involvement of the *PRODH/DGCR6* deletion in the genetic determinism of hyperprolinemia, we did not find any significant association between this deletion and ASD in two large case/control samples.

This negative finding could indicate that the suggested enrichment of ASD cases in HPI patients is spurious and reflects an inclusion bias (i.e., clinical geneticists preferentially refer HPI patients with an ASD diagnosis for genetic analysis). A limitation of this study concerning HPI patients is that we relied on the clinical description of patients made by the clinician who referred blood samples to our laboratory and that structured interviews allowing a formal diagnosis were rarely used. Alternatively, our results could be explained by the low prevalence of HPI among ASD patients. For example, we recently measured plasma proline level in a consecutive series of 230 ASD patients (unpublished data) and found only one HPI case (95% CI [0.028-0.0002]). Thus, only approximately one HPI in 230 ASD cases might contribute to increase the load of *PRODH/DGCR6* deletions compared to the baseline. Based on an estimated frequency of HPI among ASD subjects of 1/230 and assuming that

only 18% of these HPI subjects will be expected to harbor the *PRODH/DGCR6* deletion (this study), the deletion frequency would be expected to increase from 0.4% (i.e., the average frequency observed in the two control samples) to 0.48%. A sample size of about 107,000 patients and a similar number of controls would be required to reach an 80% chance of detecting such an association between ASD and the *PRODH/DGCR6* deletion at the 5% level (http://www.stat.ubc.ca/~rollin/stats/ ssize/b2.html). At best, if we consider the upper limit of the 95% confidence interval for the frequency of HPI among ASD, 2.8%, the deletion frequency would be expected to increase from 0.4% in controls to 0.9% in ASD cases, which would require a sample size of about 4,000 each. We conclude that the involvement of HPI in the determinism of ASD remains controversial and is not supported by this CNV study. Given the extreme genetic heterogeneity of ASD and the fact that all disease-causing variants identified to date are very rare [Betancur 2011], the risk conferred by HPI, if any, is likely to be small. To definitely settle this issue there is an urgent need to measure fasting proline levels in large unbiased cohorts of ASD subjects and controls.

ACKNOWLEDGEMENTS

REFERENCES

Afenjar A, Moutard ML, Doummar D, Guet A, Rabier D, Vermersch AI, Mignot C, Burglen L, Heron D, Thioulouse E, de Villemeur TB, Campion D, Rodriguez D. 2007. Early neurological phenotype in 4 children with biallelic PRODH mutations. Brain Dev 29:547-552.

Betancur C. 2011. Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. Brain Res 1380:42-77.

Bierut LJ, Agrawal A, Bucholz KK, Doheny KF, Laurie C, Pugh E, Fisher S, Fox L, Howells W, Bertelsen S, Hinrichs AL, Almasy L, Breslau N, Culverhouse RC, Dick DM, Edenberg HJ, Foroud T, Grucza RA, Hatsukami D, Hesselbrock V, Johnson EO, Kramer J, Krueger RF, Kuperman S, Lynskey M, Mann K, Neuman RJ, Nothen MM, Nurnberger JI, Jr., Porjesz B, Ridinger M, Saccone NL, Saccone SF, Schuckit MA, Tischfield JA, Wang JC, Rietschel M, Goate AM, Rice JP, Gene EASC. 2010. A genome-wide association study of alcohol dependence. Proc Natl Acad Sci USA 107:5082-5087.

Bittel DC, Yu S, Newkirk H, Kibiryeva N, Holt A, 3rd, Butler MG, Cooley LD. 2009. Refining the 22q11.2 deletion breakpoints in DiGeorge syndrome by aCGH. Cytogenet Genome Res 124:113-120.

Choudhury PR, Lahiri S, Rajamma U. 2012. Glutamate mediated signaling in the pathophysiology of autism spectrum disorders. Pharmacol Biochem Behav 100:841-849.

Colella S, Yau C, Taylor JM, Mirza G, Butler H, Clouston P, Bassett AS, Seller A, Holmes CC, Ragoussis J. 2007. QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data. Nucleic Acids Res 35:2013-2025.

Di Rosa G, Nicotera AG, Lenzo P, Spano M, Tortorella G. 2014. Long-term neuropsychiatric follow-up in hyperprolinemia type I. Psychiatr Genet 24:172-175.

Di Rosa G, Pustorino G, Spano M, Campion D, Calabro M, Aguennouz M, Caccamo D, Legallic S, Sgro DL, Bonsignore M, Tortorella G. 2008. Type I hyperprolinemia and proline dehydrogenase (PRODH) mutations in four Italian children with epilepsy and mental retardation. Psychiatr Genet 18:40-42.

Figueiredo JC, Lewinger JP, Song C, Campbell PT, Conti DV, Edlund CK, Duggan DJ, Rangrej J, Lemire M, Hudson T, Zanke B, Cotterchio M, Gallinger S, Jenkins M, Hopper J, Haile R, Newcomb P, Potter J, Baron JA, Le Marchand L, Casey G. 2011. Genotype-environment interactions in microsatellite stable/microsatellite instability-low colorectal cancer: results from a genome-wide association study. Cancer Epidemiol Biomarkers Prev 20:758-766.

Fox CS, Liu Y, White CC, Feitosa M, Smith AV, Heard-Costa N, Lohman K, Consortium G, Consortium M, Consortium G, Johnson AD, Foster MC, Greenawalt DM, Griffin P, Ding J, Newman AB, Tylavsky F, Miljkovic I, Kritchevsky SB, Launer L, Garcia M, Eiriksdottir G, Carr JJ, Gudnason V, Harris TB, Cupples LA, Borecki IB. 2012. Genome-wide association for abdominal subcutaneous and visceral adipose reveals a novel locus for visceral fat in women. PLoS Genet 8:e1002695.

Gai X, Xie HM, Perin JC, Takahashi N, Murphy K, Wenocur AS, D'Arcy M, O'Hara RJ, Goldmuntz E, Grice DE, Shaikh TH, Hakonarson H, Buxbaum JD, Elia J, White PS. 2012. Rare structural variation of synapse and neurotransmission genes in autism. Mol Psychiatry 17:402-411.

Girirajan S, Dennis MY, Baker C, Malig M, Coe BP, Campbell CD, Mark K, Vu TH, Alkan C, Cheng Z, Biesecker LG, Bernier R, Eichler EE. 2013. Refinement and discovery of new hotspots of copy-number variation associated with autism spectrum disorder. Am J Hum Genet 92:221-237.

Guilmatre A, Dubourg C, Mosca AL, Legallic S, Goldenberg A, Drouin-Garraud V, Layet V, Rosier A, Briault S, Bonnet-Brilhault F, Laumonnier F, Odent S, Le Vacon G, Joly-Helas G, David V, Bendavid C, Pinoit JM, Henry C, Impallomeni C, Germano E, Tortorella G, Di Rosa G, Barthelemy C, Andres C, Faivre L, Frebourg T, Saugier Veber P, Campion D. 2009. Recurrent rearrangements in synaptic and neurodevelopmental genes and shared biologic pathways in schizophrenia, autism, and mental retardation. Arch Gen Psychiatry 66:947-956.

Guilmatre A, Legallic S, Steel G, Willis A, Di Rosa G, Goldenberg A, Drouin-Garraud V, Guet A, Mignot C, Des Portes V, Valayannopoulos V, Van Maldergem L, Hoffman JD, Izzi C, Espil-Taris C, Orcesi S, Bonafe L, Le Galloudec E, Maurey H, Ioos C, Afenjar A, Blanchet P, Echenne B, Roubertie A, Frebourg T, Valle D, Campion D. 2010. Type I hyperprolinemia: genotype/phenotype correlations. Hum Mutat 31:961-965.

Jacquet H, Berthelot J, Bonnemains C, Simard G, Saugier-Veber P, Raux G, Campion D, Bonneau D, Frebourg T. 2003. The severe form of type I hyperprolinaemia results from homozygous inactivation of the PRODH gene. J Med Genet 40:e7.

Jacquet H, Raux G, Thibaut F, Hecketsweiler B, Houy E, Demilly C, Haouzir S, Allio G, Fouldrin G, Drouin V, Bou J, Petit M, Campion D, Frebourg T. 2002. PRODH mutations and hyperprolinemia in a subset of schizophrenic patients. Hum Mol Genet 11:2243-2249.

Ohtsuki T, Tanaka S, Ishiguro H, Noguchi E, Arinami T, Tanabe E, Yara K, Okubo T, Takahashi S, Matsuura M, Sakai T, Muto M, Kojima T, Matsushima E, Toru M, Inada T. 2004. Failure to find association between PRODH deletion and schizophrenia. Schizophr Res 67:111-113.

Paterlini M, Zakharenko SS, Lai WS, Qin J, Zhang H, Mukai J, Westphal KG, Olivier B, Sulzer D, Pavlidis P, Siegelbaum SA, Karayiorgou M, Gogos JA. 2005. Transcriptional and behavioral interaction between 22q11.2 orthologs modulates schizophrenia-related phenotypes in mice. Nat Neurosci 8:1586-1594.

Pinto D, Darvishi K, Shi X, Rajan D, Rigler D, Fitzgerald T, Lionel AC, Thiruvahindrapuram B, Macdonald JR, Mills R, Prasad A, Noonan K, Gribble S, Prigmore E, Donahoe PK, Smith RS, Park JH, Hurles ME, Carter NP, Lee C, Scherer SW, Feuk L. 2011. Comprehensive assessment of array-based platforms and calling algorithms for detection of copy number variants. Nat Biotechnol 29:512-520.

Pinto D, Delaby E, Merico D, Barbosa M, Merikangas A, Klei L, Thiruvahindrapuram B, Xu X, Ziman R, Wang Z, Vorstman JA, Thompson A, Regan R, Pilorge M, Pellecchia G, Pagnamenta AT, Oliveira B, Marshall CR, Magalhaes TR, Lowe JK, Howe JL, Griswold AJ, Gilbert J, Duketis E, Dombroski BA, De Jonge MV, Cuccaro M, Crawford EL, Correia CT, Conroy J, Conceicao IC, Chiocchetti AG, Casey JP, Cai G, Cabrol C, Bolshakova N, Bacchelli E, Anney R, Gallinger S, Cotterchio M, Casey G, Zwaigenbaum L, Wittemeyer K, Wing K, Wallace S, van Engeland H, Tryfon A, Thomson S, Soorya L, Roge B, Roberts W, Poustka F, Mouga S, Minshew N, McInnes LA, McGrew SG, Lord C, Leboyer M, Le Couteur AS, Kolevzon A, Jimenez Gonzalez P, Jacob S, Holt R, Guter S, Green J, Green A, Gillberg C, Fernandez BA, Duque F, Delorme R, Dawson G, Chaste P, Cafe C, Brennan S, Bourgeron T, Bolton PF, Bolte S, Bernier R, Baird G, Bailey AJ, Anagnostou E, Almeida J, Wijsman EM, Vieland VJ, Vicente AM, Schellenberg GD, Pericak-Vance M, Paterson AD, Parr JR, Oliveira G, Nurnberger JI, Monaco AP, Maestrini E, Klauck SM, Hakonarson H, Haines JL, Geschwind DH, Freitag CM, Folstein SE, Ennis S, Coon H, Battaglia A, Szatmari P, Sutcliffe JS, Hallmayer J, Gill M, Cook EH, Buxbaum JD, Devlin B, Gallagher L, Betancur C, Scherer SW. 2014. Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. Am J Hum Genet 94:677-694.

Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, Regan R, Conroy J, Magalhaes TR, Correia C, Abrahams BS, Almeida J, Bacchelli E, Bader GD, Bailey AJ, Baird G, Battaglia A, Berney T, Bolshakova N, Bolte S, Bolton PF, Bourgeron T, Brennan S, Brian J, Bryson SE, Carson AR, Casallo G, Casey J, Chung BH, Cochrane L, Corsello C, Crawford EL, Crossett A, Cytrynbaum C, Dawson G, de Jonge M, Delorme R, Drmic I, Duketis E, Duque F, Estes A, Farrar P, Fernandez BA, Folstein SE, Fombonne E, Freitag CM, Gilbert J, Gillberg C, Glessner JT, Goldberg J, Green A, Green J, Guter SJ, Hakonarson H, Heron EA, Hill M, Holt R, Howe JL, Hughes G, Hus V,

Igliozzi R, Kim C, Klauck SM, Kolevzon A, Korvatska O, Kustanovich V, Lajonchere CM, Lamb JA, Laskawiec M, Leboyer M, Le Couteur A, Leventhal BL, Lionel AC, Liu XQ, Lord C, Lotspeich L, Lund SC, Maestrini E, Mahoney W, Mantoulan C, Marshall CR, McConachie H, McDougle CJ, McGrath J, McMahon WM, Merikangas A, Migita O, Minshew NJ, Mirza GK, Munson J, Nelson SF, Noakes C, Noor A, Nygren G, Oliveira G, Papanikolaou K, Parr JR, Parrini B, Paton T, Pickles A, Pilorge M, Piven J, Ponting CP, Posey DJ, Poustka A, Poustka F, Prasad A, Ragoussis J, Renshaw K, Rickaby J, Roberts W, Roeder K, Roge B, Rutter ML, Bierut LJ, Rice JP, Salt J, Sansom K, Sato D, Segurado R, Sequeira AF, Senman L, Shah N, Sheffield VC, Soorya L, Sousa I, Stein O, Sykes N, Stoppioni V, Strawbridge C, Tancredi R, Tansey K, Thiruvahindrapduram B, Thompson AP, Thomson S, Tryfon A, Tsiantis J, Van Engeland H, Vincent JB, Volkmar F, Wallace S, Wang K, Wang Z, Wassink TH, Webber C, Weksberg R, Wing K, Wittemeyer K, Wood S, Wu J, Yaspan BL, Zurawiecki D, Zwaigenbaum L, Buxbaum JD, Cantor RM, Cook EH, Coon H, Cuccaro ML, Devlin B, Ennis S, Gallagher L, Geschwind DH, Gill M, Haines JL, Hallmayer J, Miller J, Monaco AP, Nurnberger JI, Jr., Paterson AD, Pericak-Vance MA, Schellenberg GD, Szatmari P, Vicente AM, Vieland VJ, Wijsman EM, Scherer SW, Sutcliffe JS, Betancur C. 2010. Functional impact of global rare copy number variation in autism spectrum disorders. Nature 466:368-372.

Prasad A, Merico D, Thiruvahindrapuram B, Wei J, Lionel AC, Sato D, Rickaby J, Lu C, Szatmari P, Roberts W, Fernandez BA, Marshall CR, Hatchwell E, Eis PS, Scherer SW. 2012. A discovery resource of rare copy number variations in individuals with autism spectrum disorder. G3 (Bethesda) 2:1665-1685.

Raux G, Bumsel E, Hecketsweiler B, van Amelsvoort T, Zinkstok J, Manouvrier-Hanu S, Fantini C, Breviere GM, Di Rosa G, Pustorino G, Vogels A, Swillen A, Legallic S, Bou J, Opolczynski G, Drouin-Garraud V, Lemarchand M, Philip N, Gerard-Desplanches A, Carlier M, Philippe A, Nolen MC, Heron D, Sarda P, Lacombe D, Coizet C, Alembik Y, Layet V, Afenjar A, Hannequin D, Demily C, Petit M, Thibaut F, Frebourg T, Campion D. 2007. Involvement of hyperprolinemia in cognitive and psychiatric features of the 22q11 deletion syndrome. Hum Mol Genet 16:83-91.

Sanders SJ, Ercan-Sencicek AG, Hus V, Luo R, Murtha MT, Moreno-De-Luca D, Chu SH, Moreau MP, Gupta AR, Thomson SA, Mason CE, Bilguvar K, Celestino-Soper PB, Choi M, Crawford EL, Davis L, Wright NR, Dhodapkar RM, DiCola M, DiLullo NM, Fernandez TV, Fielding-Singh V, Fishman DO, Frahm S, Garagaloyan R, Goh GS, Kammela S, Klei L, Lowe JK, Lund SC, McGrew AD, Meyer KA, Moffat WJ, Murdoch JD, O'Roak BJ, Ober GT, Pottenger RS, Raubeson MJ, Song Y, Wang Q, Yaspan BL, Yu TW, Yurkiewicz IR, Beaudet AL, Cantor RM, Curland M, Grice DE, Gunel M, Lifton RP, Mane SM, Martin DM, Shaw CA, Sheldon M, Tischfield JA, Walsh CA, Morrow EM, Ledbetter DH, Fombonne E, Lord C, Martin CL, Brooks AI, Sutcliffe JS, Cook EH, Jr., Geschwind D, Roeder K, Devlin B, State MW. 2011. Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. Neuron 70:863-885.

Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SF, Hakonarson H, Bucan M. 2007. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res 17:1665-1674.

Table I. Clinical features and genotype at the *PRODH* locus in HPI subjects.

Subject	Age (y)	Sex	Clinical features	Plasma proline (μmol/l)	Genotype
144	42	F	ID, SZ AFF	538	del / R453C
45435	31	F	ASD	487	del / R185W + Q19P
TF1	4	M	ID, Epi	2246	del / del
BP1	39	F	BP	470	del / P406L
001	2	F	ID, ASD, Epi	1515-2186	del / del
117	11	F	ASD	488	del / T275N + V427M
Si30	5	M	ID, ASD, Epi	422-1883	del / R453C
12363	10	M	ID	512	del / R185W
107	14	F	LD	692-878	del / del
105	7	M	ID, ASD	599-899	del / T275N
201*	13	M	ID, ASD, Epi	1008	del /del
202*	7	F	Behav D, Epi	1700	del / R453C + T466M
203*	13	F	ID, Behav D	411	del / R185W
204*	15	F	ID, Epi	556	del / Q19P + R431H
205*	11	M	ASD	481	del / R453C + R185W

*Previously unreported. ASD, autism spectrum disorder; Behav D, behavioral disorder (e.g., **aggression**, hyperactivity); BP, bipolar disorder; Epi, epilepsy; F, female; ID, intellectual disability; LD, learning disability; M, male; SZ AFF, schizoaffective disorder.

Table II. *PRODH* deletions in the three cohorts included in this study.

Study	Ancestry	Cases	Controls	p-value[a]	Odds ratio (95% confidence interval)	Array platform
HPI	All ancestries	15/83 (18%)	1/236 (0.4%)	p = 8.6x10-9	OR = 50.7 (7.5-2,147)	Agilent 1M CGH array
	European	12/78 (15%)	1/236 (0.4%)	p = 2.9x10-7	OR = 42.2 (6.0-1,815)	
	Other	3/5	—			
ASD 1 (AGP)	All ancestries	11/2,446 (0.4%)	15/4,768 (0.3%)	p = 0.4	OR = 1.4 (0.6-3.3)	Illumina 1M SNP array
	European	9/2,147 (0.4%)	12/2,640 (0.5%)	p = 1	OR = 0.92 (0.3-2.39)	
	Other	2/299 (0.7%)	3/2,128 (0.1%)			
ASD 2	All ancestries	6/676[b,c] (354) (0.9%)	9/1005 (0.9%)	p = 1	OR = 0.99 (0.3-3.2)	Agilent 1M CGH array
	European	4/505 (0.8%)	9/1005 (0.9%)	p = 1	OR = 0.88 (0.2-3.2)	
	Other	1/110 (0.9%)	—			

[a]Fisher exact test

[b]Among the 676 ASD subjects from Canada tested with CGH array, 322 were also tested with Illumina 1M SNP array as part of the AGP study, leaving 354 unique patients. Among these 354 patients, there were 4 deletion carriers (1.1%).

[c]Ancestry was determined in 615 ASD cases that were genotyped previously with SNP microarrays; no ancestry information was available for the remaining 61 subjects.
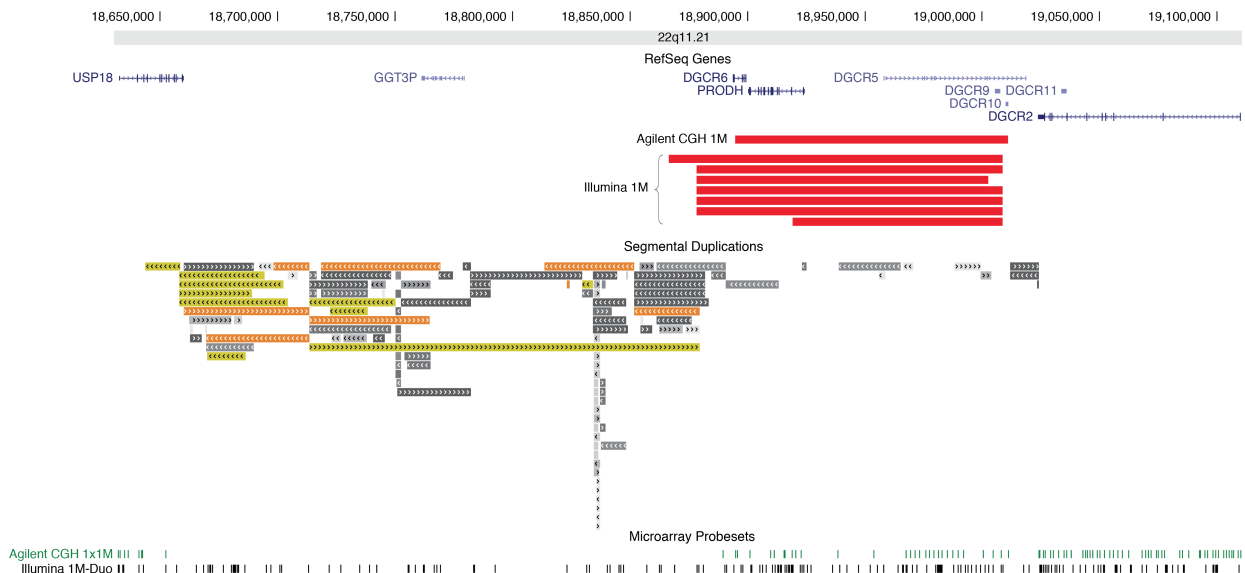
Figure S1. 22q11.21 genomic region involved in the *PRODH* recurrent imbalances. Map of the genomic region 18.63-19.11 Mb of chromosome 22 displayed using the UCSC genome browser (GRCh37/hg19). Representative deletions identified with Agilent 1M CGH arrays and with Illumina 1M SNP arrays are shown in red. Non-coding genes are shown in light blue. The centromeric breakpoint of the *PRODH/DGCR6* rearrangement falls within a large region of segmental duplications, which is not covered by Agilent CGH probes. The telomeric breakpoint is located between the non-coding *DGCR10* gene and the *DGCR2* gene.
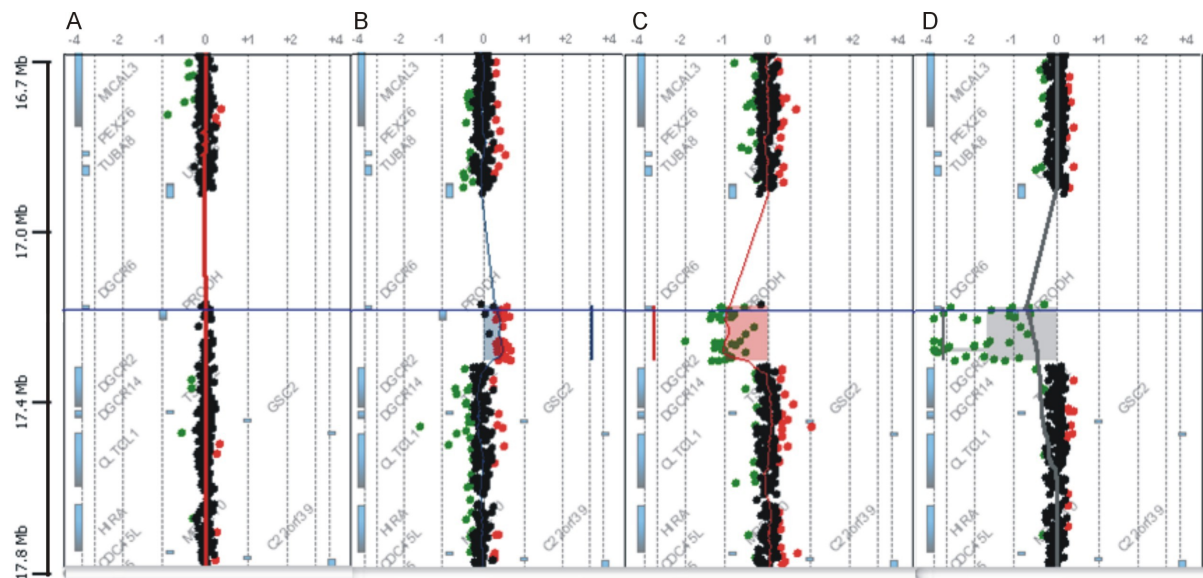
Figure S2. Detailed view of *PRODH* rearrangements detected by array-CGH. Note that the region located upstream of *PRODH* is not covered by array-CGH probes, due to the presence of an LCR. A. Control individual showing no *PRODH* rearrangement. B. Individual carrying a heterozygous duplication of *PRODH*. C. Individual carrying a heterozygous deletion of *PRODH*. D. Individual carrying a homozygous deletion of *PRODH*.