denovo-db version 1.5

http://denovo-db.gs.washington.edu/denovo-db/

May 2, 2017

1 Introduction

This is a release PDF for denovo-db version 1.5. This document describes studies by phenotype, lists studies previously included but now removed, discusses studies considered but not included, and provides an appendix describing where the variants were pulled from each paper. Please see our Data Usage and Release page at http://denovo-db.gs.washington.edu/denovo-db/Usage.jsp. We invite you to email our mailing list at denovo-db@uw.edu with any comments or questions.

2 Description of studies

2.1 Studies by phenotype

In the following subsections we will describe the studies in denovo-db 1.5 in alphabetical order based on their primary phenotype.

2.1.1 Acromelic frontonasal dysostosis

Smith et al. 2014 is the only study in the database on acromelic frontonasal dysostosis. It contains data from four probands of which **three** were whole-exome sequenced.

2.1.2 Amyotrophic lateral sclerosis

There are two studies (Table 1) in denovo-db on amyotrophic lateral sclerosis (ALS). As far as we know, there are no overlaps between the two cohorts.

Table 1: Unique denominators for performing aggregate ALS analysis

Studies	Count
Steinberg2015	44
Chesi2015	47
Total	91

2.1.3 Anophthalmia/Microphthalmia

There is only one study (Slavotinek2015) on anophthalmia/microphthalmia in denovo-db and it contains **25** probands.

2.1.4 Autism

There are a number of autism whole-exome, whole-genome, and targeted sequencing studies. Table 2 contains a list of all of these studies. The data come from different cohorts studied by many researchers. These include the Simons Simplex Collection (SSC), The Autism Simplex Collection (TASC), the Autism Sequencing Consortium (ASC), MSSNG, and the National Institute of Mental Health (NIMH) as well as clinical cohorts. For aggregate studies of the autism data it is important to know the sample overlaps and what to consider for analysis (Table 3).

First, we highly recommend that the targeted sequencing studies (ASD1_2, ASD3) be ignored except for individuals from the SSC (include .p1 in their name) who are present there because of their validation status. denovo-db picks variant representation in studies based on validation status so SSC variants were detected in exome but **validated** by targeted sequencing studies.

Second, the Michaelson et al. 2012 study is a set of 10 monozygotic twin pairs and for aggregate statistics only one set of the variants should be represented in analysis.

Third, although the overall denominator reported is 2,270 in de Rubeis et al. 2014 there were 825 SSC probands included so we report the difference of 1,445 in denovo-db since the SSC proband results represented in that study are present in other studies in the database.

Fourth, the Yuen2016 and Yuen2017 studies are both in the MSSNG dataset and therefore the denominator should be adjusted to Yuen2017's 1,625 total number. Also, the Yuen studies contain related affected children which should not matter for *de novo* variants except for the less common case of germline mosaicism.

In sum there are a total of 5,624 unique individuals (Table 3) with autism from exome and genome sequencing studies.

Table 2: Autism studies in denovo-db 1.5

Phenotype	Name	PMID	Probands	Controls	Type
autism	Iossifov	25363768	2508	1911	exome
	Krumm	25961944	2377	1786	exome
	ASD1_2	23160955	2446	0	targeted
	ASD3	25418537	3486	2493	targeted
	Turner2016	26749308	53	40	genome
	DeRubeis2014	25363760	2270	0	exome
	Michaelson2012	23260136	10	0	genome
	Lee2014	24501278	1	0	exome
	Tavassoli2014	24650168	1	0	exome
	Yuen2016	27525107	200	0	genome
	Yuen2017	28263302	1625	2	genome
	Hashimoto2015	26582266	30	0	exome
	Moreno-Ramos2015	26352270	4	0	exome

2.1.5 Bipolar disorder

There is only one study on bipolar disorder (Kataoka2016) in denovo-db and it contains **79** probands.

2.1.6 Cantú syndrome

There is only one study (vanBon2012) on Cantú syndrome in denovo-db and it contains four probands with **one** that was exome sequenced.

Table 3: Unique exome/genome denominators for performing aggregate autism analysis

Cohort	Studies	Count
SSC	Iossifov, Krumm, Turner2016 ASD1_2/ASD3 (.p1 only)	2508
ASC	DeRubeis2014	1445
MSSNG	Yuen2016, Yuen2017	1625
NIMH	Michaelson2012	10
Clinical	Lee2014	1
	Tavassoli2014	1
Other	Hashimoto2015	30
	Moreno-Ramos2015	4
Total		5624

2.1.7 Congenital diaphragmatic hernia

There is only one study (Yu2015) on congenital diaphragmatic hernia in denovo-db and it contains **39** probands.

2.1.8 Congenital heart disease

There are three studies (Lifton, Homsy2015, Sifrim2016) on congenital heart disease in denovo-db. There are overlaps between these studies. By removing all sample overlaps that we could identify, there were a total of **2,072** probands.

2.1.9 Controls

There are a few studies where there were either only controls sequenced or cases and controls. These studies are shown in Table 4. The data from these studies come from different cohorts studied by many researchers. These include the SSC, MSSNG, Genome of the Netherlands (GoNL), and others. For aggregate studies of the control data, it is important to know the overlaps and what to consider for analysis (Table 5).

We highly recommend that the targeted sequencing study (ASD3) be ignored except for individuals from the SSC (include .s1 in their name) who are present there because of their validation status.

As an additional note, Conrad2011 and Ramu2013 both study only one sample (NA12878).

In sum up there are a total of 2,278 unique control individuals (Table 5) from exome and genome sequencing studies.

Table 4: Controls in denovo-db 1.5

Phenotype	Name	PMID	Cases	Controls	Type
control	Iossifov	25363768	2508	1911	exome
	Krumm	25961944	2377	1786	exome
	ASD3	25418537	3486	2493	targeted
	Turner2016	26749308	53	40	genome
	Yuen2017	28263302	1625	2	genome
	GoNL	24974849	0	250	genome
	Besenbacher2014	25597990	0	10	genome
	Conrad2011	21666693	0	1	genome
	Ramu2013	23975140	0	1	genome
	Rauch2012	23020937	51	20	exome
	Gulsuner2013	23911319	105	84	genome

Table 5: Unique exome/genome denominators for performing aggregate control analysis

Cohort	Studies	Count
SSC	Iossifov, Krumm, Turner2016 ASD3 (.s1 only)	1911
MSSNG	Yuen2017	2
Other	GoNL	250
	Besenbacher2014	10
	Conrad2011/Ramu2013	1
	Rauch2012	20
	Gulsuner2013	84
Total		2278

2.1.10 Developmental disorder

There is only one study on developmental disorder (DDD2017) in denovo-db and it contains **4,293** probands.

2.1.11 Early onset Alzheimer's disease

There is only one study (Rovelet-Lecrux2015) on early onset Alzheimer's disease in denovo-db and it contains **12** probands.

2.1.12 Early onset Parkinson's disease

There is only one study (KunRodrigues2015) on early onset Parkinson's disease in denovo-db and it contains **21** probands.

2.1.13 Epilepsy

There are five studies on epilepsy with a total of **532** probands. The studies are shown in Table 6.

Table 6: Unique denominators for performing aggregate epilepsy analysis

Studies	Count
epi4k2013	264
Veeramah2012	1
Barcia2012	3
Veeramah2013	10
Helbig2016	254
Total	532

2.1.14 Intellectual disability

There are four studies on intellectual disability with a total of **1,010** probands. The studies are shown in Table 7.

2.1.15 Neural tube defects

There is only one study (Lemay2015) on neural tube defects in denovo-db and it contains **43** probands.

Table 7: Unique denominators for performing aggregate intellectual disability analysis

Studies	Count
deLigt2012	100
Rauch	51
Halvardson2016	39
Lelieveld2016	820
Total	1010

2.1.16 Schizophrenia

There are five studies on schizophrenia with a total of **800** probands. The studies are shown in Table 8.

Table 8: Unique denominators for performing aggregate schizophrenia analysis

Studies	Count
Gulsuner2013	105
McCarthy2014	57
Fromer2014	623
Smedemark-Margulies2016	1
Kranz2015	14
Total	800

2.1.17 Sporadic infantile spasm syndrome

There is only one study (Dimassi2015) on sporadic infantile spasm syndrome in denovo-db and it contains **10** probands.

2.2 Studies previously in denovo-db and removed in version 1.5

2.2.1 Jiang et al. 2013, PMID: 23849776

Jiang et al. 2013 was a whole-genome sequencing study of 32 families with autism. We realized in the preparation of denovo-db 1.5 that there were 20 pairs of identical duplicates with Yuen et al. 2016. The Yuen paper did not mention previous sequencing of these families but the senior author is the same so we believe the samples overlap. The naming convention is similar between studies. For example, one of the overlaps is at chromosome 1, position 16376358, C to G change and is in sample 2-1342_proband in the Jiang study and sample 2-1342-003 in the Yuen study. To be conservative we removed the Jiang study from version 1.5.

2.2.2 Yuen et al. 2015, PMID: 25621899

Yuen et al. 2015 was removed from denovo-db 1.5 as it had considerable overlap with Yuen et al. 2017. The two studies have the same samples but used "-" in the sample names in one paper and "_" in the other paper therefore bypassing our original pipelines for catching duplicates. Also, some of the samples in the 2017 paper had an **a** or **b** at the end of their name, which may correspond to different sequencing platforms. We removed those from denovo-db 1.5.

2.3 Studies considered for denovo-db but not included in version 1.5

2.3.1 Neale et al. 2012, PMID: 22495311

We had hoped to add the Neale et al. 2012 study to the database, which studied the exomes of 175 families with autism. Unfortunately, we were not able to disentangle the sample overlaps between this paper and the de Rubeis 2014 paper published by some of the same authors. Because of potential overlap we excluded it from this version of the database.

2.3.2 DDD_2015, PMID:25533962

Because of potential overlap with DDD 2017 we excluded DDD 2015 from this version of the database.

3 Appendix

In the appendix we describe the location of the data in the original papers.

3.0.1 ASD1_2, PMID:23160955

The data from this paper comes from Tables S11 and S14.

3.0.2 ASD3, PMID:25418537

The data from this paper comes from Supplementary Data 1 and 2.

3.0.3 Barcia2012, PMID:23086397

The data from this paper comes from Supplementary Table 2 and the number of trios exome sequenced (n = 3) is described in the main text.

3.0.4 Besenbacher2014, PMID:25597990

The data from this paper comes from Supplementary Data 3. The count of n = 10 trios is from the abstract.

3.0.5 Chesi2015, PMID:23708140

The data from this study comes from Table 1 and the trio count of n = 47 is from the abstract.

3.0.6 Conrad2011, PMID:21666693

The data from this study comes from Table S1 and only the sites that were marked as germline *de novo* in sample NA12878 were retained for denovo-db. Since the coordinates were on hg18, we lifted them over to hg19.

3.0.7 DDD_2017, PMID:28135719

The data from this paper comes from Supplementary Table 1 and the count of families (n = 4,293) comes from the abstract. We also noticed that some individuals have very high numbers of variants (up to 36 in the exome). This may be a consideration for aggregate statistical analysis.

3.0.8 deLigt2012, PMID:23033978

The data from this paper comes from Table S3 and the number of trios (n = 100) comes from the abstract.

3.0.9 DeRubeis2014, PMID:25363760

The data from this paper comes from Supplementary Table 3. The proband count for trios (n = 2,270) comes from the main text and the removal of the 825 SSC trios for denominator counts is based on Supplementary Table 1. This results in 1,445 individuals.

3.0.10 Dimassi2015, PMID:26138355

The data from this paper comes from Table 1 and the count of n = 10 is from the abstract.

3.0.11 epi4k2013, PMID:23934111

The data from this paper comes from Supplementary Table 2 with sample names from Supplementary Table 1. The count of n = 264 comes from the abstract.

3.0.12 Fromer2014, PMID:24463507

The data from this paper comes from Supplementary Table 1. The number of n = 623 trios comes from that supplementary table legend.

3.0.13 GoNL, PMID:24974849

The data from this paper was downloaded from the website that was in the URLs section of the paper: http://www.nlgenome.nl/ The file at this website was named GoNL_DNMs.txt The number of n = 250 trios comes from the abstract.

3.0.14 Gulsuner2013, PMID:23911319

The data from this paper comes from Table S3. The number n = 105 comes from the results section. The number n = 84 comes from the main text.

3.0.15 Halvardson2016, PMID:27334371

The data from this paper comes from Supplementary Table 2. The number of n = 39 trios comes from the abstract.

3.0.16 Hashimoto2015, PMID:26582266

The data from this paper comes from Table 1. The number of n = 30 trios comes from the abstract.

3.0.17 Helbig2016, PMID:26795593

The data from this paper comes from Tables S3 and S7. The number n = 254 comes from the "Exome sequencing strategy" section of the paper.

3.0.18 Homsy2015, PMID:26785492

The data from this paper comes from Supplementary Database S2. The number n = 1,213 trios comes from the main text.

3.0.19 Iossifov, PMID:25363768

The data from this paper comes from Supplementary Table 2. The numbers 2,508 and 1,911 come from the main text.

3.0.20 Kataoka2016, PMID:27217147

The data from this paper comes from Table S2. The number n = 79 comes from the abstract.

3.0.21 Kranz2015, PMID:26091878

The data from this paper comes from Table 1. The number n = 14 comes from the abstract.

3.0.22 Krumm, PMID:25961944

The data from this paper comes from Supplementary Table 1. The number n = 2,377 comes from the abstract and the number 1,786 comes from the methods section.

3.0.23 Kun-Rodrigues, PMID:26362251

The data from this paper comes from Table 1. The number n = 21 comes from the abstract.

3.0.24 Lee2014, PMID:24501278

The data from this paper comes from Table 1. We only retained one of the monozygotic twin pairs for the database.

3.0.25 Lelieveld2016, PMID:27479843

The data from this paper comes from corrected Supplementary Table 2. The number n = 820 comes from the main text.

3.0.26 Lemay2015, PMID:25805808

The data from this paper comes from Table 1. The number n = 43 comes from the abstract.

3.0.27 Lifton, PMID:23665959

The data from this paper comes from Table S4. The number n = 362 comes from the abstract.

3.0.28 McCarthy2014, PMID:24776741

The data from this paper comes from Supplementary Table 3. The number n = 57 comes from the abstract.

3.0.29 Michaelson2012, PMID:23260136

The data from this paper comes from Table S1 and was lifted over to hg19. The number of 10 monozygotic twin pairs comes from Figure S1. Since variants from the Michaelson paper come from a set of 10 monozygotic twin pairs, only one set of each pair should be considered if doing large aggregate *de novo* statistics.

3.0.30 Moreno-Ramos2015, PMID:26352270

The data from this study comes from Table 1.

3.0.31 Ramu2013, PMID:23975140

The data from this study comes from Supplementary Table 9. The number n = 1 comes from the main text.

3.0.32 Rauch2012, PMID:23020937

The data from this study comes from Supplementary Tables S2, S3, S6, S7, and S8. The number n = 51 and n = 20 comes from the supplement under the "Study patients" section.

3.0.33 Rovelet-Lecrux_2015, PMID:26194182

The data from this paper comes from Table 2. The number n = 12 is from the abstract.

3.0.34 Sifrim2016, PMID:27479907

The data from this paper comes from Supplementary Tables 21, 22, and 23. The Sifrim et al. 2016 study on congenital heart disorders had overlap with other studies in denovo-db. In particular, it overlapped with the DDD 2017 and Lifton studies. We removed all samples from the Sifrim paper with a DDD in the name or a CG in the name that corresponded to these two studies. Supplementary Table 2 in Sifrim paper has the information for trios and based on that table there are 859 total trios in the Sifrim 2016 set (minus DDD and Lifton).

3.0.35 Slavotinek2015, PMID:25457163

The data from this paper were derived directly from the main text.

3.0.36 Smedemark-Margulies2016, PMID:27626066

The data from this paper comes from the main text under the "Genomic Analyses" section.

3.0.37 Smith2014, PMID:25105228

The data from this paper comes from the main text.

3.0.38 Steinberg2015, PMID:25773295

The data from this paper comes from Table 3.

3.0.39 Tavassoli2014, PMID:24650168

The data from this paper comes from the main text.

3.0.40 Turner2016, PMID:26749308

The data from this paper comes from Table S9. The number n = 53 comes from the abstract and the number n = 40 comes from the main text.

3.0.41 VanBon, PMID:22608503

The data from this paper comes from Table S3.

3.0.42 Veeramah2012, PMID:22365152

The data from this paper comes from the main text.

3.0.43 Veeramah2013, PMID:23647072

The data from this paper comes from Table 2.

3.0.44 Yu2015, PMID:26034137

The data from this paper comes from Table 1.

3.0.45 Yuen2016, PMID:27525107

The data from this paper comes from Supplementary Tables S4 and S5. The number n = 200 comes from the abstract.

3.0.46 Yuen2017, PMID:28263302

The data from this paper comes from Supplementary Tables S3, S4, and S5. The total number n = 1,627 comes from the main text. Note, there are two unaffected individuals with *de novo* variants based on Table S1.

4 References

- Allen AS, Berkovic SF, Cossette P, Delanty N, Dlugos D, Eichler EE, Epstein MP, Glauser T, Goldstein DB, Han Y et al. 2013. De novo mutations in epileptic encephalopathies. Nature 501: 217-221.
- Barcia G, Fleming MR, Deligniere A, Gazula VR, Brown MR, Langouet M, Chen H, Kronengold J, Abhyankar A, Cilio R et al. 2012. De novo gain-of-function KCNT1 channel mutations cause malignant migrating partial seizures of infancy. Nature Genetics 44: 1255-1259.
- Besenbacher S, Liu S, Izarzugaza JM, Grove J, Belling K, Bork-Jensen J, Huang S, Als TD, Li S, Yadav R et al. 2015. Novel variation and de novo mutation rates in population-wide de novo assembled Danish trios. Nature Communications 6: 5969.
- Chesi A, Staahl BT, Jovicic A, Couthouis J, Fasolino M, Raphael AR, Yamazaki T, Elias L, Polak M, Kelly C et al. 2013. Exome sequencing to identify de novo mutations in sporadic ALS trios. Nature Neuroscience 16: 851-855.
- Conrad DF, Keebler JE, DePristo MA, Lindsay SJ, Zhang Y, Casals F, Idaghdour Y, Hartl CL, Torroja C, Garimella KV et al. 2011. Variation in genome-wide mutation rates within and between human families. Nature Genetics 43: 712-714.
- DDD. Prevalence and architecture of de novo mutations in developmental disorders. Nature 542: 433-438.
- de Ligt J, Willemsen MH, van Bon BW, Kleefstra T, Yntema HG, Kroes T, Vulto-van Silfhout AT, Koolen DA, de Vries P, Gilissen C et al. 2012. Diagnostic exome sequencing in persons with severe intellectual disability. The New England Journal of Medicine 367: 1921-1929.
- De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, Cicek AE, Kou Y, Liu L, Fromer M, Walker S et al. 2014. Synaptic, transcriptional and chromatin genes disrupted in autism. Nature 515: 209-215.
- Dimassi S, Labalme A, Ville D, Calender A, Mignot C, Boutry-Kryza N, de Bellescize J, Rivier-Ringenbach C, Bourel-Ponchel E, Cheillan D et al. 2016. Whole-exome sequencing improves the diagnosis yield in sporadic infantile spasm syndrome. Clinical Genetics 89: 198-204.
- Fromer M, Pocklington AJ, Kavanagh DH, Williams HJ, Dwyer S, Gormley P, Georgieva L, Rees E, Palta P, Ruderfer DM et al. 2014. De novo mutations in schizophrenia implicate synaptic networks. Nature 506: 179-184.

- GoNL. Whole-genome sequence variation, population structure and demographic history of the Dutch population. Nature Genetics 46: 818-825.
- Gulsuner S, Walsh T, Watts AC, Lee MK, Thornton AM, Casadei S, Rippey C, Shahin H, Nimgaonkar VL, Go RC et al. 2013. Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. Cell 154: 518-529.
- Halvardson J, Zhao JJ, Zaghlool A, Wentzel C, Georgii-Hemming P, Mansson E, Ederth Savmarker H, Brandberg G, Soussi Zander C, Thuresson AC et al. 2016. Mutations in HECW2 are associated with intellectual disability and epilepsy. Journal of Medical Genetics 53: 697-704.
- Hashimoto R, Nakazawa T, Tsurusaki Y, Yasuda Y, Nagayasu K, Matsumura K, Kawashima H, Yamamori H, Fujimoto M, Ohi K et al. 2016. Whole-exome sequencing and neurite outgrowth analysis in autism spectrum disorder. Journal of Human Genetics 61: 199-206.
- Helbig KL, Farwell Hagman KD, Shinde DN, Mroske C, Powis Z, Li S, Tang S, Helbig I. 2016. Diagnostic exome sequencing provides a molecular diagnosis for a significant proportion of patients with epilepsy. Genetics in medicine : official journal of the American College of Medical Genetics doi:10.1038/gim.2015.186.
- Homsy J, Zaidi S, Shen Y, Ware JS, Samocha KE, Karczewski KJ, DePalma SR, McKean D, Wakimoto H, Gorham J et al. 2015. De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. Science (New York, NY) 350: 1262-1266.
- Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, Stessman HA, Witherspoon KT, Vives L, Patterson KE et al. 2014. The contribution of de novo coding mutations to autism spectrum disorder. Nature 515: 216-221.
- Kataoka M, Matoba N, Sawada T, Kazuno AA, Ishiwata M, Fujii K, Matsuo K, Takata A, Kato T. 2016. Exome sequencing for bipolar disorder points to roles of de novo loss-of-function and protein-altering mutations. Molecular Psychiatry 21: 885-893.
- Kranz TM, Harroch S, Manor O, Lichtenberg P, Friedlander Y, Seandel M, Harkavy-Friedman J, Walsh-Messinger J, Dolgalev I, Heguy A et al. 2015. De novo mutations from sporadic schizophrenia cases highlight important signaling genes in an independent sample. Schizophrenia Research 166: 119-124.
- Krumm N, Turner TN, Baker C, Vives L, Mohajeri K, Witherspoon K, Raja A, Coe BP, Stessman HA, He ZX et al. 2015. Excess of rare, inherited truncating mutations in autism. Nature Genetics 47: 582-588.
- Kun-Rodrigues C, Ganos C, Guerreiro R, Schneider SA, Schulte C, Lesage S, Darwent L, Holmans P, Singleton A, Bhatia K et al. 2015. A systematic screening to identify de novo mutations causing sporadic early-onset Parkinson's disease. Human Molecular Genetics 24: 6711-6720.
- Lee H, Lin MC, Kornblum HI, Papazian DM, Nelson SF. 2014. Exome sequencing identifies de novo gain of function missense mutation in KCND2 in identical twins with autism and seizures that slows potassium channel inactivation. Human Molecular Genetics 23: 3481-3489.
- Lelieveld SH, Reijnders MR, Pfundt R, Yntema HG, Kamsteeg EJ, de Vries P, de Vries BB, Willemsen MH, Kleefstra T, Lohner K et al. 2016. Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. Nature Neuroscience 19: 1194-1196.
- Lemay P, Guyot MC, Tremblay E, Dionne-Laporte A, Spiegelman D, Henrion E, Diallo O, De Marco P, Merello E, Massicotte C et al. 2015. Loss-of-function de novo mutations play an important role in severe human neural tube defects. Journal of Medical Genetics 52: 493-497.

- McCarthy SE, Gillis J, Kramer M, Lihm J, Yoon S, Berstein Y, Mistry M, Pavlidis P, Solomon R, Ghiban E et al. 2014. De novo mutations in schizophrenia implicate chromatin remodeling and support a genetic overlap with autism and intellectual disability. Molecular Psychiatry 19: 652-658.
- Michaelson JJ, Shi Y, Gujral M, Zheng H, Malhotra D, Jin X, Jian M, Liu G, Greer D, Bhandari A et al. 2012. Whole-genome sequencing in autism identifies hot spots for de novo germline mutation. Cell 151: 1431-1442.
- Moreno-Ramos OA, Olivares AM, Haider NB, de Autismo LC, Lattig MC. 2015. Whole-Exome Sequencing in a South American Cohort Links ALDH1A3, FOXN1 and Retinoic Acid Regulation Pathways to Autism Spectrum Disorders. PloS One 10: e0135927.
- O'Roak BJ, Stessman HA, Boyle EA, Witherspoon KT, Martin B, Lee C, Vives L, Baker C, Hiatt JB, Nickerson DA et al. 2014. Recurrent de novo mutations implicate novel genes underlying simplex autism risk. Nature Communications 5: 5595.
- O'Roak BJ, Vives L, Fu W, Egertson JD, Stanaway IB, Phelps IG, Carvill G, Kumar A, Lee C, Ankenman K et al. 2012. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. Science (New York, NY) 338: 1619-1622.
- Ramu A, Noordam MJ, Schwartz RS, Wuster A, Hurles ME, Cartwright RA, Conrad DF. 2013. DeNovoGear: de novo indel and point mutation discovery and phasing. Nature Methods 10: 985-987.
- Rauch A, Wieczorek D, Graf E, Wieland T, Endele S, Schwarzmayr T, Albrecht B, Bartholdi D, Beygo J, Di Donato N et al. 2012. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. Lancet 380: 1674-1682.
- Rovelet-Lecrux A, Charbonnier C, Wallon D, Nicolas G, Seaman MN, Pottier C, Breusegem SY, Mathur PP, Jenardhanan P, Le Guennec K et al. 2015. De novo deleterious genetic variations target a biological network centered on Abeta peptide in early-onset Alzheimer disease. Molecular Psychiatry 20: 1046-1056.
- Sifrim A, Hitz MP, Wilsdon A, Breckpot J, Turki SH, Thienpont B, McRae J, Fitzgerald TW, Singh T, Swaminathan GJ et al. 2016. Distinct genetic architectures for syndromic and nonsyndromic congenital heart defects identified by exome sequencing. Nature Genetics 48: 1060-1065.
- Slavotinek AM, Garcia ST, Chandratillake G, Bardakjian T, Ullah E, Wu D, Umeda K, Lao R, Tang PL, Wan E et al. 2015. Exome sequencing in 32 patients with anophthalmia/microphthalmia and developmental eye defects. Clinical Genetics 88: 468-473.
- Smedemark-Margulies N, Brownstein CA, Vargas S, Tembulkar SK, Towne MC, Shi J, Gonzalez-Cuevas E, Liu KX, Bilguvar K, Kleiman RJ et al. 2016. A novel de novo mutation in ATP1A3 and childhood-onset schizophrenia. Cold Spring Harbor Molecular Case Studies 2: a001008.
- Smith JD, Hing AV, Clarke CM, Johnson NM, Perez FA, Park SS, Horst JA, Mecham B, Maves L, Nickerson DA et al. 2014. Exome sequencing identifies a recurrent de novo ZSWIM6 mutation associated with acromelic frontonasal dysostosis. American Journal of Human Genetics 95: 235-240.
- Steinberg KM, Yu B, Koboldt DC, Mardis ER, Pamphlett R. 2015. Exome sequencing of case-unaffected-parents trios reveals recessive and de novo genetic variants in sporadic ALS. Scientific Reports 5: 9124.
- Tavassoli T, Kolevzon A, Wang AT, Curchack-Lichtin J, Halpern D, Schwartz L, Soffes S, Bush L, Grodberg D, Cai G et al. 2014. De novo SCN2A splice site mutation in a boy with Autism spectrum disorder. BMC Medical Genetics 15: 35.

- Turner TN, Hormozdiari F, Duyzend MH, McClymont SA, Hook PW, Iossifov I, Raja A, Baker C, Hoekzema K, Stessman HA et al. 2016. Genome Sequencing of Autism-Affected Families Reveals Disruption of Putative Noncoding Regulatory DNA. American Journal of Human Genetics 98: 58-74.

- van Bon BW, Gilissen C, Grange DK, Hennekam RC, Kayserili H, Engels H, Reutter H, Ostergaard JR, Morava E, Tsiakas K et al. 2012. Cantu syndrome is caused by mutations in ABCC9. American Journal of Human Genetics 90: 1094-1101.

- Veeramah KR, Johnstone L, Karafet TM, Wolf D, Sprissler R, Salogiannis J, Barth-Maron A, Greenberg ME, Stuhlmann T, Weinert S et al. 2013. Exome sequencing reveals new causal mutations in children with epileptic encephalopathies. Epilepsia 54: 1270-1281.

- Veeramah KR, O'Brien JE, Meisler MH, Cheng X, Dib-Hajj SD, Waxman SG, Talwar D, Girirajan S, Eichler EE, Restifo LL et al. 2012. De novo pathogenic SCN8A mutation identified by whole-genome sequencing of a family quartet affected by infantile epileptic encephalopathy and SUDEP. American Journal of Human Genetics 90: 502-510.

- Yu L, Sawle AD, Wynn J, Aspelund G, Stolar CJ, Arkovitz MS, Potoka D, Azarow KS, Mychaliska GB, Shen Y et al. 2015. Increased burden of de novo predicted deleterious variants in complex congenital diaphragmatic hernia. Human Molecular Genetics 24: 4764-4773.

- Yuen RK, Merico D, Cao H, Pellecchia G, Alipanahi B, Thiruvahindrapuram B, Tong X, Sun Y, Cao D, Zhang T et al. 2016. Genome-wide characteristics of de novo mutations in autism. NPJ Genomic Medicine 1: 160271-1602710.

- Yuen RK, Merico D, Bookman M, J LH, Thiruvahindrapuram B, Patel RV, Whitney J, Deflaux N, Bingham J, Wang Z et al. 2017. Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. Nature Neuroscience 20: 602-611.

- Zaidi S, Choi M, Wakimoto H, Ma L, Jiang J, Overton JD, Romano-Adesman A, Bjornson RD, Breitbart RE, Brown KK et al. 2013. De novo mutations in histone-modifying genes in congenital heart disease. Nature 498: 220-223.