Letter to the Editor

# Penetrance of pathogenic mutations in haploinsufficient genes for intellectual disability and related disorders

ABSTRACT

*De novo* loss of function (LOF) mutations in the *ASXL3* gene cause Bainbridge-Ropers syndrome, a severe form of intellectual disability (ID) and developmental delay, but there is evidence that they also occur in healthy individuals. This has prompted us to look for non-pathogenic LOF variants in other ID genes. Heterozygous LOF mutations in *ASXL1*, a paralog of *ASXL3*, are known to cause Bohring-Opitz syndrome (BOS), and benign LOF mutations in this gene have not been published to date. Therefore, we were surprised to find 56 *ASXL1* LOF variants in the ExAC database (http://exac.broadinstitute.org), comprising exomes from 60,706 individuals who had been selected to exclude severe genetic childhood disorders. 4 of these variants have been described as disease-causing in patients with BOS, which rules out the possibility that pathogenic and clinically neutral LOF variants in this gene are functionally distinct. Apparently benign LOF variants were also detected in several other genes for ID and related disorders, including CDH15, KATNAL2, DEPDC5, ARID1B and AUTS2, both in the ExAC database and in the 6,500 exomes of the Exome Variant Server (http://evs.gs.washington.edu/EVS/). These observations argue for low penetrance of LOF mutations in *ASXL1* and other genes for ID and related disorders, which could have far-reaching implications for genetic counseling and research.

© 2015 Elsevier Masson SAS. All rights reserved.

Recently we described a novel intellectual disability (ID) syndrome (Bainbridge et al., 2013) (OMIM #615485) which is due to *de novo* loss of function (LOF) mutations in the *ASXL3* gene. ASXL3 is a paralog of ASXL1, a negative regulator of histone H3K27 methylation, which had been implicated previously in the clinically related Bohring-Opitz syndrome (BOS) (Hoischen et al., 2011). Apart from 4 disease-causing mutations in the N-terminal part of the *ASXL3* gene, we had also found several apparently benign LOF variants in the Exome Variant Database (URL: http://evs.gs.washington.edu/EVS/), which may have escaped from nonsense-mediated RNA decay due to their C-terminal location. More recently, however, we and others (Dinwiddie et al., 2013; Kuechler et al., 2015 and own unpublished observations) have identified additional disease-causing and benign *ASXL3* mutations that seem to refute this simple interpretation, thereby re-opening the question why some LOF mutations in the *ASXL3* gene are pathogenic while others are not.

This has prompted us to look for non-pathogenic LOF mutations in other ID genes. In the related *ASXL1* gene, benign LOF mutations have not been reported to date. Therefore, we were surprised to find 56 different LOF mutation in the ExAC database (Exome Aggregation Consortium (ExAC), Cambridge, MA, URL: http://exac.broadinstitute.org [accessed 12 May 2015]) which comprises 60,706 exomes, selected from >90,000 exomes to exclude

individuals with severe genetic childhood disorders. Of note, 4 of these 56 apparently benign LOF mutations have been described as disease-causing in patients with BOS (Hoischen et al., 2011; Russell et al., 2015; Magini et al., 2012) (see Fig. 1).

These findings and the similar spatial distribution of LOF mutations in affected and healthy individuals argue against the possibility that disease-causing and clinically neutral LOF variants in the *ASXL1* gene are functionally distinct. Therefore, it is likely that affected and unaffected carriers of protein-truncating *ASXL1* mutations differ in other, hitherto unknown ways. Our observations suggest that almost one in 1000 individuals of the ExAC cohort carries a truncating LOF mutation in *ASXL1*. Since BOS is a severe and clinically recognizable disorder, most heterozygous carriers of inactivating *ASXL1* variants must be clinically inconspicuous.

*ASXL1* LOF variants have been observed in 17 out of the 23 BOS cases studied to date, and the origin of the mutation could be investigated in 11 of these (see Table 1).

Together with the frequency of *ASXL1* LOF-mutations in ExAC controls (i.e., 56/60,706), and assuming that the prevalence of BOS is 1/1,000,000 (OrphaNet, ORPHA97297, http://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=DE&Expert=97297), this leads to an estimated penetrance of 0.08%, using Bayes' theorem (Vassos et al., 2010). The odds ratio (OR) of the underlying fourfold table is $3.08 \times 10^3$, and the diagnostic likelihood ratio (LR) is (17/
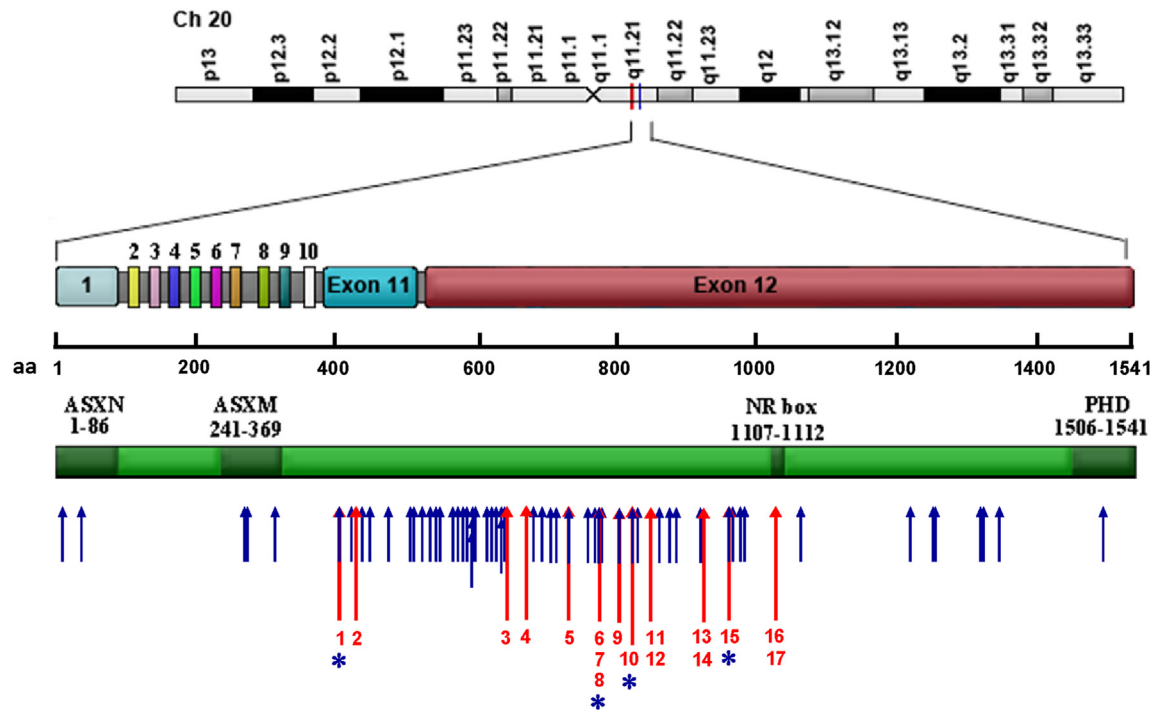
**Fig. 1. Loss-of-function (LOF) mutations in the ASXL1 gene.** Red arrows: disease-causing LOF mutations identified in 17 patients with Bohring-Opitz syndrome (see (Hoischen et al., 2011; Russell et al., 2015; Magini et al., 2012)). Blue arrows: 56 heterozygous frame shift, stop gained or splice site mutations in apparently healthy individuals (from ExAC database, http://exac.broadinstitute.org/). At 4 positions (Arg404X, Gln733X, Leu823X and Arg965X, see asterisks), red and blue arrows overlap.

**Table 1**
ASXL1 loss-of-function mutations in patients with BOS.

| Sources | Total BOS cases | ASXL1 Mutation | Origin of mutation | |
|---|---|---|---|---|
| | | | Investigated | de novo |
| Hoischen et al. (2011) | 13 | 7 | 7 | 7 |
| Magini et al. (2012) | 2 | 2 | 2 | 2 |
| Russell et al. (2015) | 8 | 8 | 2 | 2 |
| Total | 23 | 17 | 11 | 11 |

$23)/(56/60,706) = 801.2$. The large difference between OR and LR can be explained by the relatively high 'detectance' of ASXL1 LOF-mutations in individuals with BOS.[1]

Both parents could be investigated in 11 of these patients. All of these carried *de novo* ASXL1 LOF-mutations; so far, no single BOS patient with an inherited *ASXL1* mutation has been described. This suggests that the lower 99 percent confidence limit for the proportion of BOS patients carrying a *de novo* mutation is 62%; conversely, at most 38% of these patients may carry an inherited mutation. These observations are difficult to reconcile with the very low penetrance of *ASXL1* LOF mutations inferred from the ExAC data, even under the assumption of simple digenic inheritance or epistasis.

As to the incomplete penetrance of LOF mutations, *ASXL1* may be an extreme case, but it is apparently not unique, as suggested by the evidence, presented above, that ASXL3 LOF mutations do not always cause disease. LOF variants in other haploinsufficient genes causing ID or autism spectrum disorder are sometimes

transmitted by healthy parents (e.g., see (Jiang et al., 2013; Yuen et al., 2014)), and incomplete penetrance is also a hallmark of many pathogenic copy number variants (CNVs). Since the clinical consequences of many CNVs can be ascribed to aberrant dosage of single genes (e.g., see (Slager et al., 2003; Zollino et al., 2012)), it is plausible to assume that other 'disease-causing' gene defects are not fully penetrant either. So far, reliable penetrance estimates exist only for a minority of the known pathogenic CNVs (Cooper et al., 2011; Rosenfeld et al., 2013; Tropeano et al., 2013). Even less is known about the penetrance of *de novo* LOF mutations in genes that have been identified as major cause of ID and developmental delay.

Because of its unprecedented depth, the ExAC database is a unique source of information about the penetrance of X-linked recessive and autosomal dominant LOF mutations in haploinsufficient disease genes. In accordance with the Fort Lauderdale agreement, which encourages 'the use and publication of frequency data for specific targeted sets of variants', but does not allow publication of 'global (genome-wide) analyses of these data or of large gene sets' (see http://exac.broadinstitute.org), we have limited our analysis to only a few other established genes for ID/ASD and related disorders.

Significantly elevated numbers of apparently benign truncating variants were also found for *CDH15* (n = 31), *KATNAL2* (n = 23), *DEPDC5* (n = 21), *ARID1B* (n = 16) and *AUTS2* (n = 11), suggesting that for these genes, the penetrance of LOF mutations is also low or at least incomplete. Apart from DEPDC5 which is associated with autosomal dominant focal epilepsy (e.g., see (Ishida et al., 2013; Dibbens et al., 2013) and OMIM *614191), these genes have been firmly implicated in dominant forms of ID or autism spectrum disorder (ASD), mostly by trio sequencing ((Bhalla et al., 2008; Wright et al., 2015; Sanders et al., 2012; O'Roak et al., 2011; de Rubeis et al., 2014; Beunders et al., 2013; Hoyer et al., 2012) and references therein; see also OMIM *614556 and *607270).

---

[1] Detectance is defined as Prob(Genotype|Phenotype); thus, it is the inverse of penetrance, Prob(Phenotype|Genotype), where event and condition are interchanged (Weiss and Terwilliger, 2000). Detectance equals the genotype case frequency; reduced values may be indicative of genetic heterogeneity.

A potential problem of the ExAC database is the fact that its 60,706 'healthy controls' are derived from various different cohorts, and although its authors state to 'have removed individuals affected by severe pediatric diseases', there is no guarantee that all affected individuals have been excluded in this way. Therefore it is reassuring that for all of these genes, *bona fide* LOF variants (supported by >10 sequencing reads) were also found among the 6503 thoroughly phenotyped controls from the NHLBI GO Exome Sequencing Project which focuses on heart, lung and blood disorders (see Exome Variant Server, EVS; http://evs.gs.washington.edu/EVS/database). Indeed, with 9, 6, 6, 2, 2 and 1 high-quality LOF variants listed for *ASXL1, KATNAL2, CDH15, DEPDC, ARID1B* and *AUTS2*, respectively, the frequency of inactivating mutations in these genes tends to be even higher in the EVS than in the ExAC database.

For *CDH15* and *DEPDC5*, incomplete penetrance of LOF mutations has been documented before (e.g., see (Dibbens et al., 2013; Bhalla et al., 2008) and OMIM *114019), which is in keeping with our observations and lends further credibility to the ExAC data. In contrast, evidence for incomplete penetrance of inactivating mutations is scarce or lacking for *ASXL1, KATNAL2, ARID1B* and *AUTS2*. Exonic disruptions of *AUTS2* give rise to a highly penetrant ID phenotype and are rarely inherited, but occasionally, non-manifesting carrier parents have been observed ((Beunders et al., 2013), see also OMIM *607270). Similarly, there is recent evidence that some of the 'causative' *KATNAL2* mutations reported are also inherited (Yuen et al., 2014; De Rubeis et al., 2014).

*De novo ARID1B* mutations are among the most common causes of autosomal dominant ID (e.g., see (Hoyer et al., 2012) and OMIM *614556), and the fact that incomplete penetrance has been never described before is difficult to reconcile with the presence of 16 different *ARID1B* LOF variants in the ExAC cohort, even though these findings are supported by 2 *bona fide* LOF variants in the much smaller EVS cohort. Very recently, however, an individual with an ARID1B LOF variant but without any sign of the characteristic Coffin-Siris phenotype has been reported (Johnston et al., 2015), indicating that the penetrance of inactivating ARID1B mutations is not complete either. Assuming that the prevalence of severe ID in our population is 0.5% and that in 0.9% of the affected individuals, ID is due to heterozygous inactivating ARID1B mutations (Hoyer et al., 2012), their prevalence may be around $4.5 \times 10^{-5}$ or 1/22.222. Given the 16 LOF variants in the ARID1B gene identified in 60.706 ExAC controls, this suggests that their penetrance may be close to 15%.

Although intuitively, this estimate may seem much too low, it may be premature to dismiss these results as artifacts reflecting shortcomings of the ExAC dataset, because this database is the first one that is sufficiently large to allow observations of this kind. Moreover, these results are in keeping with the frequency of LOF variants in the (much smaller) EVS dataset. Until very recently, mutation screening has been almost entirely focused on patients, and the number of healthy controls may have been simply too low to identify more than a small minority of the unaffected individuals carrying 'disease-causing' variants.

No or very few LOF mutations were detected in the ExAC database for numerous other established haploinsufficient ID/ASD genes such as *CHD8*, and these may indeed reflect 'background noise'. When selecting healthy controls, a few affected individuals may have 'slipped through'; some of the apparent LOF variants may be edited or processed upon transcription or translation, and they could be due to rare sample swaps or sequencing errors. A more comprehensive analysis will have to await the publication of the announced 'ExAC flagship paper' (see http://exac.broadinstitute.org).

Manifesting carriers of pathogenic CNVs often harbor other potentially pathogenic CNVs which may act as a necessary 'second hit' (Girirajan et al., 2010, 2012). Our observations, suggesting that in *ASXL1* and several other genes for ID or related disorders, LOF mutations are not fully penetrant, may have identified new entry points for research into the vulnerability or resilience to genetic disease. Identification of the underlying genetic, epigenetic or environmental factors promises to shed more light on the pathogenesis of many disorders and may lead to improved disease management and therapy (Friend and Schadt, 2014). Moreover, the systematic ascertainment of apparently benign LOF mutations in haploinsufficient disease genes and their validation by detailed clinical reexamination (Johnston et al., 2015) may have far-reaching implications for genetic diagnosis and counseling.

### Note added in proof

After this manuscript had been accepted, we have used LOFTEE, a Loss-Of-Function Transcript Effect Estimator pipeline (https://github.com/konradjk/loftee) to estimate the functional relevance of all LOF variants listed in the ExAC browser for the 6 aforementioned ID genes. To this end, we screened the current version of the ExAC database (ExAC release r03, file ExAC.r0.3.sites.vep.vcf, Jan. 13th, 2015) to rule out 'low-confidence' LOF variants located in non-canonical transcripts, variants representing the ancestral allele across primates, as well as stop-gained and frameshift variants in the terminal 5% of the transcript. LOFTEE filtering removed only two ASXL1 and one of the CDH15 LOF variants, i.e. after filtering, 54 and 30 high-confidence LOF mutations in these genes were retained, including 4 pathogenic ASXL1 mutations that had been described in BOS patients before. For KATNAL2, 16 out of 23 LOF variants survived the filtering, but only 9 out of 21 LOF variants listed for DEPDC5, and 7 out of 11 AUTS2 variants. No more than 8 out of 16 ARID1B LOF variants identified in the 60,706 ExAC controls turned out to be high-confidence LOF mutations. This means that their penetrance may be close to 25% rather than 15%, as estimated above.

### Conflicts of interest

The authors declare no conflicts of interest.

### Acknowledgments

### References

Bainbridge, M.N., Hu, H., Muzny, D., et al., 2013. De novo truncating mutations in ASXL3 are associated with a novel clinical phenotype with similarities to Bohring-Opitz syndrome. Genome Med. 5, 11.

Beunders, G., Voorhoeve, E., Golzio, C., et al., 2013. Exonic deletions in AUTS2 cause a syndromic form of intellectual disability and suggest a critical role for the C terminus. Am. J. Hum. Genet. 92, 210–222.

Bhalla, K., Luo, Y., Buchan, T., et al., 2008. Alterations in CDH15 and KIRREL3 in patients with mild to severe intellectual disability. Am. J. Hum. Genet. 83, 703–713.

Cooper, G.M., Coe, B.P., Girirajan, S., et al., 2011. A copy number variation morbidity map of developmental delay. Nat. Genet. 43, 838–846.

De Rubeis, S., He, X., Goldberg, A.P., et al., 2014. Synaptic, transcriptional and chromatin genes disrupted in autism. Nature 515, 209–215.

Dibbens, L.M., de Vries, B., Donatello, S., et al., 2013. Mutations in DEPDC5 cause familial focal epilepsy with variable foci. Nat. Genet. 45, 546–551.

Dinwiddie, D.L., Soden, S.E., Saunders, C.J., et al., 2013. De novo frameshift mutation

in ASXL3 in a patient with global developmental delay, microcephaly, and craniofacial anomalies. BMC Med. Genomics 6, 32.

Friend, S.H., Schadt, E.E., 2014. Clues from the resilient. Science 344, 970–972.

Girirajan, S., Rosenfeld, J.A., Cooper, G.M., et al., 2010. A recurrent 16p12.1 microdeletion supports a two-hit model for developmental delay. Nat. Genet. 42, 203–209.

Girirajan, S., Rosenfeld, J.A., Coe, B.P., et al., 2012. Phenotypic heterogeneity of genomic disorders and rare copy-number variants. N. Engl. J. Med. 367, 1321–1331.

Hoischen, A., van Bon, B.W.M., Rodríguez-Santiago, B., et al., 2011. *De novo* nonsense mutations in *ASXL1* cause Bohring-Opitz syndrome. Nat. Genet. 43, 729–731.

Hoyer, J., Ekici, A.B., Endele, S., et al., 2012. Haploinsufficiency of ARID1B, a member of the SWI/SNF-A chromatin-remodeling complex, is a frequent cause of intellectual disability. Am. J. Hum. Genet. 90, 565–572.

Ishida, S., Picard, F., Rudolf, G., et al., 2013. Mutations of DEPDC5 cause autosomal dominant focal epilepsies. Nat. Genet. 45, 552–555.

Jiang, Y.H., Yuen, R.K.C., Jin, X., et al., 2013. Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing. Am. J. Hum. Genet. 93, 249–263.

Johnston, J.J., Lewis, K.L., Ng, D., 2015. Individualized iterative phenotyping for genome-wide analysis of loss-of-function mutations. Am. J. Hum. Genet. 96, 913–925.

Kuechler, A., Czeschik, J.C., Sperl, W., et al., 2015. Bainbridge-Ropers syndrome caused by mutations in ASXL3-a recognizable condition? Med. Genet. 1, 85.

Magini, P., della Monica, M., Giovannucci Uzielli, M.L., et al., 2012. Two novel patients with Bohring-Opitz syndrome caused by de novo ASXL1 mutations. Am. J. Med. Genet. Part A 158A, 917–921.

O'Roak, B.J., Deriziotis, P., Lee, C., et al., 2011. Exome sequencing in sporadic autism spectrum disorders identifies severe *de novo* mutations. Nat. Genet. 43, 585–589.

Rosenfeld, J.A., Coe, B.P., Eichler, E.E., et al., 2013. Estimates of penetrance for recurrent pathogenic copy-number variations. Genet. Med. 15, 478–481.

Russell, B., Johnston, J.J., Biesecker, L.G., et al., 2015. Clinical management of patients with ASXL1 mutations and Bohring-Opitz syndrome, emphasizing the need for Wilms tumor surveillance. Am. J. Med. Genet. Part A 167A, 2122–2131.

Sanders, S.J., Murtha, M.T., Gupta, A.R., et al., 2012. *De novo* mutations revealed by whole-exome sequencing are strongly associated with autism. Nature 485, 237–241.

Slager, R.E., Newton, T.L., Vlangos, C.N., et al., 2003. Mutations in *RAI1* associated with Smith–Magenis syndrome. Nat. Genet. 33, 466–468.

Tropeano, M., Ahn, J.W., Dobson, R.J.B., et al., 2013. Male-biased autosomal effect of 16p13.11 copy number variation in neuro developmental disorders. PLoS One 8 (4), e61365.

Vassos, E., Collier, D.A., Holden, S., et al., 2010. Penetrance for copy number variants associated with schizophrenia. Hum. Mol. Genet. 19, 3477–3481.

Weiss, K.M., Terwilliger, J.D., 2000. How many diseases does it take to map a gene with SNPs? Nat. Genet. 26, 151–157.

Wright, C.F., Fitzgerald, T.W., Jones, J.F., et al., 2015. Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. Lancet 385, 1305–1314.

Yuen, R.K.C., Thiruvahindrapuram, B., Merico, D., et al., 2014. Whole-genome sequencing of quartet families with autism spectrum disorder. Nat. Med. 21, 185–191.

Zollino, M., Orteschi, D., Murdolo, D., 2012. Mutations in *KANSL1* cause the 17q21.31 microdeletion syndrome phenotype. Nat. Genet. 44, 636–638.

H. Hilger Ropers[*]
*Institute for Human Genetics, University Medicine, Langenbeckstrasse 1, Building 601, 55131 Mainz, Germany*

*Max Planck Institute for Molecular Genetics, Ihnestrasse 73, 14195 Berlin, Germany*

Thomas Wienker
*Max Planck Institute for Molecular Genetics, Ihnestrasse 73, 14195 Berlin, Germany*

[*] Corresponding author.
*E-mail address:* ropers@molgen.mpg.de (H.H. Ropers).

15 October 2015
Available online 24 October 2015