Zebrafish homologs of genes within 16p11.2, a genomic region associated with brain disorders, are active during brain development, and include two deletion dosage sensor genes

Alicia Blaker-Lee1,*, Sunny Gupta1,*, Jasmine M. McCammon1,*, Gianluca De Rienzo1 and Hazel Sive1,2,‡

SUMMARY

Deletion or duplication of one copy of the human 16p11.2 interval is tightly associated with impaired brain function, including autism spectrum disorders (ASDs), intellectual disability disorder (IDD) and other phenotypes, indicating the importance of gene dosage in this copy number variant region (CNV). The core of this CNV includes 25 genes; however, the number of genes that contribute to these phenotypes is not known. Furthermore, genes whose functional levels change with deletion or duplication (termed dosage sensors), which can associate the CNV with pathologies, have not been identified in this region. Using the zebrafish as a tool, a set of 16p11.2 homologs was identified, primarily on chromosomes 3 and 12. Use of 11 phenotypic assays, spanning the first 5 days of development, demonstrated that this set of genes is highly active, such that 21 out of the 22 homologs tested showed loss-of-function phenotypes. Most genes in this region were required for nervous system development – impacting brain morphology, eye development, axonal density or organization, and motor response. In general, human genes were able to substitute for the fish homolog, demonstrating orthology and suggesting conserved molecular pathways. In a screen for 16p11.2 genes whose function is sensitive to hemizygosity, the aldolase a (aldoa) and kinesin family member 22 (kif22) genes were identified as giving clear phenotypes when RNA levels were reduced by ~50%, suggesting that these genes are dosage deletion sensors. This study leads to two major findings. The first is that the 16p11.2 region comprises a highly active set of genes, which could present a large genetic target and might explain why multiple brain function, and other, phenotypes are associated with this interval. The second major finding is that there are (at least) two genes with deletion dosage sensor properties among the 16p11.2 set, and these could link this CNV to brain disorders such as ASD and IDD.

INTRODUCTION

Copy number variant regions (CNVs), intervals of the genome that are deleted or duplicated, have been associated with multiple human diseases, including infectious, autoimmune and neuropsychiatric disorders (Fanciulli et al., 2010). The 16p11.2 CNV comprises two direct repeats of 143 kbp and a central core of 593 kbp, which includes 25 putative protein-coding genes (Ghebranious et al., 2007; Sebat et al., 2007), and is associated with a multitude of disorders, most commonly with speech and/or developmental delay [often characteristics of intellectual disability disorder (IDD)] [85% of deletion carriers] and autism spectrum disorders (ASDs; 19-28% of deletion carriers) (Ciuladaite et al., 2011; The Simons VIP Consortium, 2012). The deletion or duplication of 16p11.2 is also associated with other disorders in addition to IDD (Bijlsma et al., 2009) and ASD (Kumar et al., 2008; Marshall et al., 2008; Weiss et al., 2008), including seizure disorder (Ghebranious et al., 2007), obesity or being overweight (Jacquemont et al., 2011), macro- or microcephaly (Shinawi et al., 2010), schizophrenia (McCarthy et al., 2009), ADHD (Lionel et al., 2011), eye anomalies (Bardakjian et al., 2010), heart disorders (Puvabanditsin et al., 2010), vertebral abnormalities (Shimojima et al., 2009) and severe combined immunodeficiency (Shiow et al., 2009). Smaller deletions within this region are also associated with pathologies, including ASD (Crepel et al., 2011) and abnormal sexual development (Tannour-Louet et al., 2010). 16p11.2 copy number changes generally occur de novo, during meiosis (Sebat et al., 2007), and the paucity of inherited changes indicates its importance for normal health and reproduction (Levy et al., 2011; Weiss et al., 2008). Consistently, mice that are hemizygous for the 16p11.2 homologs exhibit a 50% rate of neonatal lethality (Horev et al., 2011).

A key general question is how do genes in a CNV contribute to an associated disorder? In the most direct case, duplication or deletion of a gene would increase or decrease cognate RNA and protein levels proportionally, and this change would be pivotal in the development of the pathology (Nord et al., 2011). We term genes with such properties ‘dosage sensors’. In other cases, structural effects caused by the chromosomal rearrangement might contribute to a phenotype (Ricard et al., 2010). For the 16p11.2 region, a mouse model in the syntenic region shows that levels of brain expression in 79% of genes are affected by deletion or duplication, as predicted...
by gene dosage (Horev et al., 2011). Further supporting the notion of dosage sensitivity, in humans, 16p11.2 duplication or deletion is sometimes correlated with reciprocal phenotypes, such that obesity and macrocephaly are associated with 16p11.2 deficiency, and being underweight and microcephaly are associated with 16p11.2 duplication (Jacquemont et al., 2011; Shinawi et al., 2010). Similarly, opposite phenotypes exist in the hemizygous 16p11.2 region mice compared with mice containing the duplication, including brain volume and certain behaviors (Horev et al., 2011).

In this study, we used the zebrafish to analyze activity of 16p11.2 CNV genes, and to identify gene dosage sensors within this region. Although many abnormalities are linked to 16p11.2 CNVs, we were interested in the high prevalence of associated brain disorders, with particular interest in this region’s connection to ASD. Specifically, the 16p11.2 CNV is by far the most prevalent CNV to be associated with ASD (Sanders et al., 2011), contributing to ~1% of ASD cases. Zebrafish do not have the same behavioral repertoire as humans, and therefore have limitations in comparable assays for behaviors associated with human disorders. However, the zebrafish genome is similar to that of the human, and molecular pathways engaged by homologous mammalian and fish genes are conserved. We therefore termed the zebrafish a ‘tool’, rather than a phenotypic model, for analysis of brain disorders (Sive, 2011). The zebrafish allows rapid functional analysis of many genes, at a rate unprecedented in the mouse, owing to the ability to obtain many embryos and to inhibit gene function in the whole embryo by injection of antisense oligonucleotides. There are indications that many functional brain disorders, including ASD and IDD, are developmental in nature, because they present soon after birth (Konopka et al., 2012; Ploeger et al., 2010). However, defective maintenance of brain structure, or faulty brain function, after birth may play key roles in the etiology of these disorders (Okado et al., 2001; Trembath, 1994). The most accessible time window for analysis in the zebrafish is the first 5 days of development, which covers the equivalent of several weeks in mice and a couple of years in humans, potentially addressing both developmental and later gene function. Thus, results from zebrafish assays can provide useful data to suggest more targeted mammalian studies.

We show here that most individual zebrafish homologs of genes within 16p11.2 are required for normal brain and body development, and that their function is conserved with the human genes. We also show that zebrafish 16p11.2 homologs include at least two genes with deletion dosage sensor properties, potentially linking hemizygosity of the interval with human phenotypic presentations such as ASD and IDD.

RESULTS

Conservation and expression of zebrafish 16p11.2 homologs

In order to use the zebrafish as an effective tool for functional analysis of the 16p11.2 CNV, we used the strategy shown in Fig. 1A. The 16p11.2 core spans 593 kbp and includes 25 protein-coding genes. In total, 21 of these genes in the human interval were identified in the zebrafish genome (Fig. 1B). Of the remaining genes, SPN, TMEM and C16orf54 are limited to mammals. SPN has a regulatory role in adaptive immunity (Kyoizumi et al., 2004), whereas TMEM and C16orf54 are of unknown function, and all three are of unknown importance in neurodevelopment. Finally, QPRT has a teleost homolog in fugu (52% identity to the human protein) and medaka (54% identity to the human protein), suggesting that a zebrafish homolog exists, but is not yet annotated in the genome.

Zebrafish homologs of 16p11.2 were found to be clustered on either chromosome 3 or 12, previously identified as the genomic counterparts of human chromosome 16 (Taylor et al., 2003), with the exception of ino80e, which is located on chromosome 16 (supplementary material Table S1). These regions are not syntenic with human chromosome 16, because gene order is not conserved. Two sets of syntenic genes were found on chromosome 3: one region comprises kctd13, sez62 and asphd1 genes, which are grouped, but their order on the human chromosome is different. The cluster tbx24, ppdc4b and aldoab genes have conserved order, but the region includes intervening genes. Single fish icon, single homolog; two fish icons, multiple homologs; blue dot, teleost homolog, but no Danio rerio homolog; red box, no teleost homologs identified; black bar, synteny; H. s., Homo sapiens; D. r., Danio rerio; Chr., chromosome.

QPRT has a teleost homolog in fugu (52% identity to the human protein) and medaka (54% identity to the human protein), suggesting that a zebrafish homolog exists, but is not yet annotated in the genome.

Zebrafish homologs of 16p11.2 were found to be clustered on either chromosome 3 or 12, previously identified as the genomic counterparts of human chromosome 16 (Taylor et al., 2003), with the exception of ino80e, which is located on chromosome 16 (supplementary material Table S1). These regions are not syntenic with human chromosome 16, because gene order is not conserved. Two sets of syntenic genes were found on chromosome 3: one region comprises kctd13, sez62 and asphd1, for which the order is not conserved with that in humans; and the other comprises mapk3, gdpd3 and ype13, for which the order is conserved. Interestingly, the first set of syntenic genes (kctd13, sez62 and asphd1) is found in a microdeletion associated with ASD (Crepel et al., 2011).

Five homologs were present in two copies [aldolase a (aldoa), fams57b, gdpd3, ppdc4b and taok2]. reflecting the partial duplication of the teleost genome (Fig. 1B) (Postlethwait et al., 2000). Genes in each pair had similar but not identical sequences (supplementary
material Table S1), with one found on zebrafish chromosome 12 and the other on chromosome 3. Such duplication might result in split or divergent function of the gene (Yamamoto and Vernier, 2011).

In order to determine when zebrafish 16p11.2 homologs are expressed, temporal expression was analyzed by reverse-transcriptase PCR (RT-PCR; see Methods; supplementary material Fig. S1A, Table S2). Most genes were expressed maternally and zygotically, at least until 48 hours post fertilization (hpf). *asphd1*, *doc2a*, *prrt2* and *sez6l2* were expressed only zygotically. Whole-mount in situ hybridization (supplementary material Fig. S1B) showed that almost all genes are expressed in the brain at 24 hpf. Exceptions to these include *tbx24* (Thisse and Thisse, 2005) and one homolog of *aldoa* (*aldoab*) (Rauch et al., 2003; Thisse and Thisse, 2004).

These data indicate that the zebrafish genome includes homologs of 84% of the human 16p11.2 core genes, arranged primarily on two homologous chromosomes. Homologs were all expressed during the first 48 hours of development, as the brain and other organs are forming, with almost all genes showing some expression in the brain. Expression data, chromosomal arrangement of the genes and their sequence conservation with their human counterparts indicated that this gene set was appropriate for further analysis.

### Changes in brain and body morphology accompany loss of function in 16p11.2 homologs

In order to determine which zebrafish 16p11.2 homologs were active as the brain formed and began to function, we screened these for activity during brain development, from 24 hpf through 5 days post-fertilization (dpf), the human developmental equivalent of 5 weeks of gestation to toddlerhood. Loss of function (LOF) was performed by injection of antisense morpholino oligonucleotides (MOs) into one- to two-cell embryos. Where possible, MOs binding to an exon-intron boundary were utilized (Table 1; supplementary material Table S2), to target zygotic RNA. Where a splice site MO did not give a phenotype, an MO directed against the translational start site was tested to determine whether maternal RNA could have prevented observing a phenotype with a splice site MO (Table 1; supplementary material Table S2). The resulting effects of MO action on RNA coding capacity and the predicted protein are included in Table 2. Where two copies of a gene had been characterized, in some cases, only one was highly expressed in the brain and this was assayed for a LOF phenotype. In the case of *taok2*, both genes showed strong brain expression (supplementary material Fig. S1), and both were assayed. Thus, 22 genes were tested for LOF phenotypes.

MO methodology is rapid and allows functional analysis of many genes; however, specificity of phenotypes associated with MOs was carefully tested because off-target effects can sometimes be observed (Bedell et al., 2011; Eisen and Smith, 2008). The criteria employed to test specificity were as follows, and are described more fully in Methods, with results documented in Table 1 and quantified in Tables 2 and 3. First, for MOs targeting a splice donor or acceptor site, a change in RNA splicing and coding capacity should be observed. Second, the corresponding amount of normal RNA should be reduced, and a phenotype should correlate with RNA reduction. Use of splice site MOs allows these assays to be performed quantitatively, because normal RNA can be distinguished from abnormally spliced RNA. Third, a key assay for specificity is the ability of RNA derived from the corresponding human or zebrafish cDNA to prevent the LOF phenotype, when co-injected with the appropriate MO. Such ‘rescue’ RNAs do not contain the MO-binding site, owing to species differences or because the MO-binding site lies across a splice junction. Fourth, off-target effects of MOs might cause cell death, which can effectively be suppressed by injection of a p53 MO (Robu et al., 2007). In cases in which severe phenotypes were observed, the effect of p53 suppression is tested and, if a resulting phenotype was milder, it is the one scored. Finally, a phenotype obtained with MOs was compared with that of mutants (or with that induced by shRNAs) to test similarity of phenotypes. Because mutants are available for a very limited number of genes, we focused on MO-mediated LOF, which is the most feasible way to assay the activity of the large 16p11.2 gene set.

LOF embryos were first examined at 24 hpf for brain morphology, after injection of the brain ventricles with Texas Red dextran (Gutzman and Sive, 2009), and scored for brain shape, presence of forebrain, midbrain and hindbrain hingepoints, brain ventricle size, forebrain truncation, and eye morphology. Tail and body phenotypes were assayed as additional indicators. For each gene, each phenotype reported was observed in at least two independent experiments, and observed for at least 70% of embryos examined, using MO amounts that had been titrated (Table 2) and gave a clear phenotype (with quantification presented in supplementary material Table S3). Strikingly, LOF for almost all genes (20 out of the 22 assayed), with the exception of *prrt2* and *tbx24*, led to changes in brain or eye morphology (Fig. 2A, Fig. 3; supplementary material Table S3). These phenotypes are characterized further in Fig. 2B,C. *tbx24* LOF was associated only with a tail phenotype, consistent with a lack of brain expression of this gene, and with the mutant phenotype (supplementary material Fig. S2) (Thisse and Thisse, 2005). Thus, a total of 21 out of the 22 genes examined gave a LOF phenotype. In addition to their brain phenotypes, LOF in all but six genes (*cdipt*, *doc2a*, *fam57ba*, *kctd13*, *prrt2* and *taok2a*) led to tail or body defects, including failure of the yolk cell to extend, a short bent tail and abnormally shaped muscle segments (Fig. 2A, Fig. 3; supplementary material Table S3). Several genes gave very strong phenotypes, suggesting early embryonic defects. In particular, *coro1a*, *inos00e*, *mapk3* and *mvp* ‘morphants’ (defined as LOF embryos caused by MO injection) showed abnormal body length and defective neural tubes. *mapk3* LOF embryos showed a defective forebrain and eyes, and a short body, consistent with a recent study (Krens et al., 2008). *coro1a*, *maz* and *fam57ba* morphants have small eye cups with protruding lenses (Schmitt and Dowling, 1994). LOF in *asphd1* and *sez6l2* gave weak phenotypes, whereas no phenotype was observed after *prrt2* LOF. Because *prrt2* is not highly expressed until 48 hpf, gene function could be required later (supplementary material Fig. S1A).

Two clear groups of brain morphology phenotypes were apparent (Fig. 2B,C). LOF in the first group of genes was associated with reduced brain ventricle size, and the midbrain-hindbrain boundary (MHB) was less sharply defined than in controls. This group includes *c16orf53*, *cdipt*, *doc2a*, *hirip3*, *kctd13* and *taok2b* (Fig. 2B), which have not previously been shown to regulate brain morphology, although *cdipt* is required later in lens and
photoreceptor development in zebrafish (Murphy et al., 2011), and Doc2a is a modulator of synaptic transmission in mice (Yao et al., 2011). These phenotypes could result from changes in neuroepithelial specification, morphogenesis, or a reduction in cerebrospinal fluid volume (Gato and Desmond, 2009; Lowery and Sive, 2009).

### Table 1. Summary of tests for MO specificity

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Type MO</th>
<th>Change in RNA splicing</th>
<th>Resultant change in predicted protein</th>
<th>Phenotype observed</th>
<th>Phenotype correlates with RNA decrease</th>
<th>Rescue by RNA injection</th>
<th>Additional knockdown method</th>
</tr>
</thead>
<tbody>
<tr>
<td>aldoa</td>
<td>SS</td>
<td>Yes</td>
<td>Truncation</td>
<td>Yes</td>
<td>Yes, also protein</td>
<td>Yes</td>
<td>shRNA1</td>
</tr>
<tr>
<td>asphd1</td>
<td>SS</td>
<td>Yes</td>
<td>Truncation</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>c16orf53</td>
<td>SS</td>
<td>Yes</td>
<td>Truncation</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>cdipt</td>
<td>SS</td>
<td>Yes</td>
<td>Truncation</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>coro1a</td>
<td>SS</td>
<td>Yes</td>
<td>Truncation</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SS</td>
<td>Yes</td>
<td>In frame exon deletion</td>
<td>Yes</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>doc2a</td>
<td>SS</td>
<td>Yes</td>
<td>Truncation</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>fam57ba</td>
<td>SS</td>
<td>Yes</td>
<td>Truncation</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>gpd3</td>
<td>TSS</td>
<td>NA</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hirip3</td>
<td>TSS</td>
<td>NA</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SS</td>
<td>Yes</td>
<td>Truncation</td>
<td>No</td>
<td>ND</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>ino80e</td>
<td>TSS</td>
<td>NA</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SS</td>
<td>Yes</td>
<td>Truncation</td>
<td>No</td>
<td>ND</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>kctd13</td>
<td>SS</td>
<td>Yes</td>
<td>Truncation</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>kit22</td>
<td>SS</td>
<td>Yes</td>
<td>Truncation</td>
<td>Yes</td>
<td>Yes, also protein</td>
<td>Yes</td>
<td>shRNA1</td>
</tr>
<tr>
<td>mapk3</td>
<td>TSS</td>
<td>NA</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>maz</td>
<td>SS</td>
<td>Yes</td>
<td>Truncation</td>
<td>No</td>
<td>ND</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>mvp</td>
<td>SS1</td>
<td>Yes</td>
<td>Truncation</td>
<td>Yes</td>
<td>ND</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SS2</td>
<td>NA</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SS3</td>
<td>ND</td>
<td>NA</td>
<td>Yes</td>
<td>ND</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>pgrp4ca</td>
<td>SS</td>
<td>Yes</td>
<td>Truncation</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>pgrp4cb</td>
<td>SS</td>
<td>Yes</td>
<td>Truncation</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>prrt2</td>
<td>TSS</td>
<td>NA</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sez6l2</td>
<td>TSS</td>
<td>NA</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>took2a</td>
<td>SS</td>
<td>Yes</td>
<td>Truncation</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>took2b</td>
<td>SS</td>
<td>Yes</td>
<td>Truncation</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>tbx24</td>
<td>TSS</td>
<td>NA</td>
<td>Yes</td>
<td></td>
<td>Published</td>
<td>Mutant</td>
<td></td>
</tr>
<tr>
<td>ypef3</td>
<td>SS</td>
<td>Yes</td>
<td>Truncation</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>

For gene assignments, see Fig. 1 and supplementary material Table S1.

MOs were designed as indicated in Methods. See supplementary material Table S2 for MO sequences. SS, splice site; TSS, translational start site.

For MOs targeting a splice site, RT-PCR (detailed in Methods) was used to determine whether a change in splicing occurred. Primers are included in supplementary material Table S2.

Changes in predicted protein resulting from the RT-PCR-detected change in RNA splicing. Changes are described in Table 2.

Phenotypes were scored as indicated in Methods (see Figs 2, 3; supplementary material Table S3).

MOs were administered in titrations. For splice site MOs where the mass correlated with an increase in phenotype severity, qPCR was performed to measure the normal RNA (Fig. 6).

RNA was co-injected with MOs to confirm phenotypes were specific effects of the MO (Fig. 3, Table 2).

Similar phenotypes were observed for additional knockdown methods (Fig. 7; supplementary material Fig. S2).

A decrease in protein was also observed by western blot analysis of 24 hpf embryos (Fig. 6 and Methods).

shRNA was used to target deletion dosage sensors. Similar phenotypes were observed (Fig. 7).

Similar phenotypes were observed (supplementary material Fig. S2 and Methods).

No splicing change was observed in RNA.

The rescue for the tbx24 MO was previously published in Nikaido et al. (Nikaido et al., 2002).

The rescue for the mvp MOs 1, 2 and 3, correspond to those included in Table 2 and supplementary material Tables S2 and S3, respectively, although no mvp LOF conditions were rescued. ND, not determined; NA, not applicable.
A second group of genes leading to defective brain morphology after LOF showed a straight midbrain (Fig. 2C), where the midbrain hindgpoint was essentially absent. This phenotype was seen in aldooa, fam57ba, gdpd3, maz, ppp4ca and ypel3 morphants, a group of genes with previously undefined contributions to brain development. gdpd3, ppp4ca and ypel3 LOF embryos showed a wider opening at the MHB, relative to the narrowing seen in control embryos. A subset of embryos from both groups showed a narrowed forebrain, such as those seen in aldooa, fam57ba and maz LOF embryos, whereas the area rostral of the eyes was reduced after fam57ba, gdpd3, hirip3 and maz LOF. Expression of pax2a at the MHB was normal, indicating correct specification of this region, and suggesting that later steps resulted in the phenotypes observed (not shown). Although the groupings shown in Fig. 2B,C suggest a similar contribution to brain development by all genes in the group, close comparison reveals distinct phenotypes; for example, aldooa and ypel3 morphants have similar midbrain phenotypes, but their MHB and hindbrain phenotypes were unique.

For all but three genes, co-injection of RNA derived from the human cDNA together with the MO generally restored the phenotype, indicating fish-human orthology, and confirming the specificity of the phenotype (Tables 1, 3 and Fig. 4). For kctd13 and maz, the zebrafish, but not the human, gene rescued the LOF phenotype. mvp LOF gave a similar phenotype with three tested MOs, but was not reproducibly rescued by fish or human cDNAs, perhaps reflecting a stoichiometric requirement for other components with which mvp complexes (Berger et al., 2009).

Embryos were scored as rescued if morphological and behavioral phenotypes were ameliorated in ~50% or more embryos (see Fig. 4 and Table 3). For strong phenotypes, such as gdpd3 and mapk3 LOF, rescues vastly improved brain and body morphology, but did not fully restore a wild-type phenotype. Assays for rescue included

### Table 2. MO amounts and effects on gene expression

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>MO mass (ng/embryo)</th>
<th>Results of MO action on RNA coding capacity</th>
<th>Predicted protein (truncated/normal)</th>
<th>p53 MO included</th>
</tr>
</thead>
<tbody>
<tr>
<td>aldooa</td>
<td>1</td>
<td>Exon deletion causing frame shift and early stop codon</td>
<td>202/364</td>
<td>No</td>
</tr>
<tr>
<td>asphd1</td>
<td>3.5</td>
<td>Intron inclusion containing stop codon</td>
<td>144/354</td>
<td>Yes</td>
</tr>
<tr>
<td>c16orf53</td>
<td>7.5</td>
<td>Intron inclusion containing stop codon</td>
<td>52/273</td>
<td>No</td>
</tr>
<tr>
<td>cdip7</td>
<td>7.5</td>
<td>Intron inclusion causing frame shift and early stop codon</td>
<td>144/214</td>
<td>No</td>
</tr>
<tr>
<td>coro1a</td>
<td>5</td>
<td>Intron inclusion causing frame shift and early stop codon</td>
<td>61/455</td>
<td>Yes</td>
</tr>
<tr>
<td>doc2a</td>
<td>7.5</td>
<td>Intron inclusion causing frame shift and early stop codon</td>
<td>184/401</td>
<td>No</td>
</tr>
<tr>
<td>fam57ba</td>
<td>1.2</td>
<td>Exon deletion causing frame shift and early stop codon</td>
<td>33/202</td>
<td>No</td>
</tr>
<tr>
<td>gdpd3</td>
<td>7</td>
<td>Translation blocking</td>
<td>0/315</td>
<td>Yes</td>
</tr>
<tr>
<td>hirip3</td>
<td>6.5</td>
<td>Translation blocking</td>
<td>0/510</td>
<td>Yes</td>
</tr>
<tr>
<td>ino80e</td>
<td>5</td>
<td>Translation blocking</td>
<td>0/231</td>
<td>Yes</td>
</tr>
<tr>
<td>kctd13</td>
<td>7.5</td>
<td>Intron inclusion causing frame shift and early stop codon</td>
<td>114/330</td>
<td>No</td>
</tr>
<tr>
<td>kf22</td>
<td>5</td>
<td>Exon deletion causing frame shift and early stop codon</td>
<td>123/634</td>
<td>Yes</td>
</tr>
<tr>
<td>mapk3</td>
<td>5</td>
<td>Translation blocking</td>
<td>0/392</td>
<td>Yes</td>
</tr>
<tr>
<td>maz</td>
<td>5.5</td>
<td>Intron inclusion causing frame shift and early stop codon</td>
<td>97/493</td>
<td>No</td>
</tr>
<tr>
<td>mvp</td>
<td>5</td>
<td>Intron inclusion causing frame shift and early stop codon</td>
<td>45/863</td>
<td>Yes</td>
</tr>
<tr>
<td>mvp</td>
<td>6</td>
<td>Translation blocking</td>
<td>0/863</td>
<td>Yes</td>
</tr>
<tr>
<td>mvp</td>
<td>5</td>
<td>Translation blocking</td>
<td>0/231</td>
<td>Yes</td>
</tr>
<tr>
<td>ppp4ca</td>
<td>3.5</td>
<td>Intron inclusion causing frame shift and early stop codon</td>
<td>37/311</td>
<td>Yes</td>
</tr>
<tr>
<td>prtt2</td>
<td>10</td>
<td>Intron inclusion causing frame shift and early stop codon</td>
<td>79/226</td>
<td>No</td>
</tr>
<tr>
<td>seqb2</td>
<td>5</td>
<td>Translation blocking</td>
<td>0/892</td>
<td>Yes</td>
</tr>
<tr>
<td>took2a</td>
<td>5</td>
<td>Intron inclusion causing frame shift and early stop codon</td>
<td>65/1138</td>
<td>No</td>
</tr>
<tr>
<td>took2b</td>
<td>5</td>
<td>Intron inclusion containing stop codon</td>
<td>188/1182</td>
<td>No</td>
</tr>
<tr>
<td>tbx24</td>
<td>4</td>
<td>Translation blocking</td>
<td>0/874</td>
<td>Yes</td>
</tr>
<tr>
<td>ypel3</td>
<td>2</td>
<td>Exon deletion causing frame shift and early stop codon</td>
<td>60/119</td>
<td>Yes</td>
</tr>
</tbody>
</table>

*Genes indicated in supplementary material Table S1 were targeted in LOF experiments and quantification is included in Table 3 and supplementary material Table S3.

*The MOs (supplementary material Table S2) were injected at the one- to two-cell stage in titration experiments. The mass of MO that was used for phenotype characterization, scoring (Figs 2, 3; supplementary material Table S3) and rescue experiments (Figs 4, 5 and Table 3) is indicated.

*Splicing changes were monitored by RT-PCR (Methods) using primers indicated in supplementary material Table S2 and detected changes were sequenced to identify changes in mRNA, including frame shifts.

*The predicted changes in protein, based on the sequencing data obtained, are included as the number of amino acids in truncated vs normal protein. Truncated proteins listed might contain abnormal amino acids, because amino acids encoded by an intron inclusion or a frame shift are predicted (in some cases) before an early stop codon.

*The p53 MO was used to eliminate off-target effects where indicated.

*The mvp MOs are included in the same order as in Table 1, and supplementary material Tables S2 and S3.

*Indicates the mvp MO that was used for the LOF collation in Fig. 2A and Fig. 3.

ND, not determined.
A subset of genes is required for normal axon tract development

Motor deficiencies observed in LOF embryos led us to investigate whether axon tract formation was affected. Both forebrain and hindbrain axon tracts were analyzed by immunostaining for acetylated α-tubulin and confocal imaging at 36 hpf, when initial scaffolding has formed (Fig. 5). Embryos with deficient axon tracts were seen in more than 80% of embryos, after LOF in each of six genes: coro1a, fam57ba, hirip3, kinesin family member 22 (kif22), maz, and ppp4ca. This was associated with spontaneous movement defects. LOF in these genes led to a defect in touch response in at least 70% of embryos for each gene examined (Fig. 3; supplementary material Table S3). These data indicate that multiple 16p11.2 homologs are required for normal motor activity, as reflected by spontaneous movement or touch-responsiveness.

Movement deficiency is associated with LOF in some 16p11.2 homologs

We next assessed motor function as a read-out of neural circuitry, by assaying two early behaviors: spontaneous movement at 24 hpf and the touch response at 48 hpf as described in Methods (Liu et al., 2012; Naganawa and Hirata, 2011). LOF in seven genes [coro1a, fam57ba, gdpd3, hirip3, kinesin family member 22 (kif22), maz, and ppp4ca] was associated with spontaneous movement defects. LOF in 14 genes led to a defect in touch response in at least 70% of embryos for each gene examined (Fig. 3; supplementary material Table S3). In the most severe cases, aldoaa and fam57ba, LOF embryos exhibited no response to touch, and these severe phenotypes were rescued by coinjection with human RNA. A sluggish response to touch, rather than the normal rapid C-bend and brief swim, was observed in coro1a, gdpd3, hirip3, kctd13, kif22, maz, and ppp4ca morphants. The response to touch was improved by the addition of rescue RNA. Abnormal, U-shaped muscle segments and/or a short bent tail were seen in the spontaneous-movement- and touch-response-defective aldoaa, coro1a, fam57ba, gdpd3, kif22 and hirip3 LOF embryos, perhaps explaining the movement defects (Fig. 2A, Fig. 3; supplementary material Table S3). These data indicate that multiple 16p11.2 homologs are required for normal motor activity, as reflected by spontaneous movement or touch-responsiveness.

In summary, 21 out of the 22 zebrafish 16p11.2 homologs tested were required for early brain and/or body development, with the majority showing conserved function with the cognate human gene. The data therefore show that this set of genes is highly active during early development.

A sluggish response to touch, rather than the normal rapid C-bend and brief swim, was observed in embryos after LOF for each of the tested genes, with 50-350 embryos assayed in total per gene (supplementary material Table S3). (Ab-Aw) Dorsal views; (Ab'-Aw') lateral close-up; (Ab'-Aw') full-embryo lateral view. (Ab-Aa) Schematics of embryo landmarks. F, forebrain ventricle; M, midbrain ventricle; H, hindbrain ventricle. (Ab-hu, Ab'-hu) Anterior to the left, and images are shown at equivalent magnification. (B) Phenotypic group for which LOF embryos have narrow midbrain and hindbrain ventricles. The gene assayed is indicated above each panel. Dorsal views, anterior to the left. F, forebrain ventricle; M, midbrain ventricle; H, hindbrain ventricle; MHB, midbrain-hindbrain boundary. (C) Phenotypic group for which LOF embryos have a straight midbrain. The gene assayed is indicated above each panel. Dorsal views, anterior to the left. FB, forebrain; MHB, midbrain-hindbrain boundary; asterisk, midbrain hingepoint. Scale bars: 150 μm. Embryo images in B and C are the dorsal view of panels A except for the control embryo and the gdpd3 LOF embryo in C.
peripheral nerves (Schilling and Kimmel, 1994), was abnormal after LOF in a subset of homologs (supplementary material Fig. S3). Some of these also presented with axon tract abnormalities (coro1a, fam57ba, mapk3 and ppp4ca). Interestingly, for six genes (coro1a, fam57ba, kctd13, kif22, mapk3 and ppp4ca) for which LOF gave a movement or touch response phenotype, an axon tract and/or pigmentation defect was also apparent (Fig. 3), perhaps connecting these phenotypes to the abnormal behavior (An et al., 2002; Haffter et al., 1996; Marmigere and Ernfors, 2007). For four genes (maz, mvp, tbx24 and ypel3), a touch response phenotype seen after LOF was not accompanied by axon tract or pigmentation aberrations, implicating abnormal muscle activity in the phenotype. However, although axon tracts might appear normal, synaptic transmission could be defective. These data show that multiple 16p11.2 homologs are necessary for normal axon tract development – suggesting deficits in the formation of neuronal precursors, guidance or fasciculation – and that axonal deficiencies could be linked to motor phenotypes.

Identification of aldoaa and kif22 as deletion dosage sensor genes

A major goal of this study was to determine whether any 16p11.2 homologs had the properties of deletion dosage sensors, which could be associated with IDD, ASD and other phenotypes. We defined a deletion dosage sensor as a gene that gives a phenotype after a 50% decrease in expression, in accord with the simplest outcome of loss of one gene copy. Our initial assays for function of 16p11.2 homologs (Figs 2, 3) used MO concentrations that led to a clear phenotype; however, the associated decreases in RNA expression were not determined. In order to identify whether any of these genes are deletion dosage sensors, we assessed the lowest dose of MO that led to a phenotype, and quantified the amount of normal RNA remaining at this concentration using qPCR (Fig. 6). This approach was used for the 14 genes whose function was inhibited by splice site MOs, where abnormal splicing had been detected and the normal and abnormal transcript could be distinguished, using appropriate primers (Fig. 6B, Table 1; supplementary material Table S2). Owing to standard error intrinsic to the qPCR process, genes were designated putative deletion dosage sensors if a phenotype was observed when between 40% and 60% normal RNA remained, although a deletion dosage sensor could also be sensitive to smaller decreases in expression (>60% of RNA remaining).

Most genes tested showed a LOF phenotype only when 25% or less normally spliced RNA remained (Fig. 6C). However, two genes,
**Table 3. Rescue assays with human or fish genes after MO LOF**

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Rescue mRNA mass (pg/embryo)</th>
<th>Total embryos</th>
<th>Rescued embryos</th>
<th>% rescued</th>
<th>Number of independent experiments</th>
</tr>
</thead>
<tbody>
<tr>
<td>aldoa</td>
<td>200</td>
<td>52</td>
<td>37</td>
<td>71</td>
<td>3</td>
</tr>
<tr>
<td>asphd1</td>
<td>100</td>
<td>110</td>
<td>58</td>
<td>53</td>
<td>2</td>
</tr>
<tr>
<td>c16orf53</td>
<td>80</td>
<td>48</td>
<td>29</td>
<td>60</td>
<td>2</td>
</tr>
<tr>
<td>cdpt</td>
<td>50</td>
<td>107</td>
<td>93</td>
<td>87</td>
<td>3</td>
</tr>
<tr>
<td>coro1a</td>
<td>10</td>
<td>74</td>
<td>46</td>
<td>62</td>
<td>2</td>
</tr>
<tr>
<td>doc2a</td>
<td>100</td>
<td>90</td>
<td>84</td>
<td>93</td>
<td>2</td>
</tr>
<tr>
<td>fam57ba</td>
<td>320</td>
<td>82</td>
<td>69</td>
<td>84</td>
<td>3</td>
</tr>
<tr>
<td>gpd3</td>
<td>20</td>
<td>70</td>
<td>34</td>
<td>49</td>
<td>2</td>
</tr>
<tr>
<td>hirip3</td>
<td>100</td>
<td>95</td>
<td>81</td>
<td>85</td>
<td>2</td>
</tr>
<tr>
<td>ino80e</td>
<td>200</td>
<td>56</td>
<td>45</td>
<td>80</td>
<td>2</td>
</tr>
<tr>
<td>kctd13</td>
<td>200</td>
<td>80</td>
<td>62</td>
<td>78</td>
<td>2</td>
</tr>
<tr>
<td>kif22</td>
<td>15</td>
<td>194</td>
<td>140</td>
<td>72</td>
<td>4</td>
</tr>
<tr>
<td>mapk3</td>
<td>5</td>
<td>96</td>
<td>85</td>
<td>89</td>
<td>4</td>
</tr>
<tr>
<td>maz</td>
<td>50</td>
<td>102</td>
<td>80</td>
<td>78</td>
<td>2</td>
</tr>
<tr>
<td>ppp4ca</td>
<td>30</td>
<td>116</td>
<td>116</td>
<td>100</td>
<td>3</td>
</tr>
<tr>
<td>sez6l2</td>
<td>100</td>
<td>43</td>
<td>24</td>
<td>56</td>
<td>2</td>
</tr>
<tr>
<td>tak2a</td>
<td>200</td>
<td>85</td>
<td>49</td>
<td>58</td>
<td>2</td>
</tr>
<tr>
<td>tak2b</td>
<td>200</td>
<td>77</td>
<td>51</td>
<td>66</td>
<td>2</td>
</tr>
<tr>
<td>ypel3</td>
<td>75</td>
<td>40</td>
<td>33</td>
<td>83</td>
<td>2</td>
</tr>
</tbody>
</table>

Representative images of embryos from rescue experiments are shown in Fig. 4.

aGenes indicated in supplementary material Table S1 were targeted in LOF experiments by MOs included in supplementary material Table S2, using MO masses described in Table 2.
bMass of mRNA for rescue was determined based on a titration used for single cell embryo injections. In control and LOF conditions a balancing amount of membrane targeted GFP mRNA was co-injected to achieve the equivalent load in the rescue condition (Methods). Where more than one MO was tested per gene, the rescue of LOF experiment was performed using the first MO listed for a gene in supplementary material Table S2. cDNA constructs included in supplementary material Table S4 and RNA preparation described in Methods.
cTotal embryos scored over two to four experiments.
dNumber of embryos whose phenotype was significantly restored to normal.
ePercentage of rescued embryos relative to total embryos injected.
fTwo to four independent experiments were performed per rescue condition after the preferred titrated dosage was determined.
gktld13 and maz LOF phenotypes were rescued using zebrafish mRNA using the indicated masses. All other rescue mRNA masses refer to human mRNA.

*aldoa* and *kif22*, were reproducibly associated with a phenotype when approximately 45% and 55% RNA remained (Fig. 6C), respectively, suggesting that these genes are deletion dosage sensors. Although RNA levels were initially quantified at 24 hpf, approximately 50% of normal levels were present also at 18 and 36 hpf (Fig. 6D) (for *aldoa* and *kif22*, respectively, 64% and 52% of normal levels were present at 18 hpf, and 67% and 41% of normal levels were present at 36 hpf). Importantly, the decrease in RNA expression was mirrored at the protein level: western blot analysis showed a 56% and 65% loss of Aldoa and Kif22 protein expression, respectively, when normalized to *β-actin* (Fig. 6E).

The sensitivity of embryos to *aldoa* and *kif22* 50% LOF led us to examine the phenotypes obtained in more detail. Thus, abnormal muscle segment formation (Fig. 2) was further characterized after phalloidin staining, which showed U-shaped muscle segments, as well as muscle fibers that were wavy and poorly aligned (Fig. 6F). Because the brains in both *aldoa* and *kif22* LOF embryos appeared narrow (Fig. 2A), we investigated whether formation of neural progenitors was affected at 48 hpf, using a transgenic line with GFP driven by the promoter of NeuroD, a pan neuronal transcription factor (Gibson et al., 2008; Ulitsky et al., 2011). After LOF in both *aldoa* and *kif22*, NeuroD-promoter-driven expression of GFP was decreased in the eyes and optic tectum, whereas expression in the cranial neurons and pancreas seemed unaffected (Fig. 6G).

These data indicate that most zebrafish homologs in the 16p11.2 cohort do not have the characteristics of deletion dosage sensors. However, the *aldoa* and *kif22* genes each showed a robust phenotype at 50% LOF, which might indicate the effects of genetic hemizygosity at these loci, and designate these genes as putative deletion dosage sensors.

*Tissue-specific shRNA expression indicates nervous system function for aldoa and kif22*

In order to address whether the effects of *aldoa* and *kif22* LOF were due directly to changes in expression within the brain, or secondarily, owing to effects on other tissues, shRNAs targeting these genes were expressed in the zebrafish brain. We used the central nervous system (CNS)-specific miR124 promoter (De Rienzo et al., 2011; Shkumatava et al., 2009) and expressed shRNAs from the miR30 backbone (see Methods) (Dong et al., 2009; De Rienzo et al., 2011; De Rienzo et al., 2012) (Fig. 7A). By testing four hairpins against either *aldoa* or *kif22*, a targeting construct was identified for each gene. These constructs reduced normal RNA levels to 46% for *aldoa* and 42% for *kif22* when normalized to GFP in the whole embryo (reflecting total number of cells
expressing the shRNA) or 64% for aldoaa and 57% for kif22 when normalized to β-actin in microdissected brain (reflecting total brain RNA) (Fig. 7B,D). For aldoaa, a phenotype very similar to that of the antisense MO was observed, such that touch response was highly defective (Fig. 7B) and the forebrain was narrow (Fig. 7C). However, the tail and muscle segment phenotype was not observed,
in accord with nervous-system-specific expression of this shRNA. The kif22 RNAi phenotype was very similar to that seen with the antisense MO, including abnormal brain morphology and a bent tail (Fig. 7E). The persistent bent tail phenotype, even after nervous-system-specific expression of the kif22 shRNA, probably reflects defective convergence and extension that could be modulated by kif22 activity in the spinal cord (De Rienzo et al., 2011).

These shRNA data further confirm specificity of the aldoa and kif22 phenotypes observed using MOs. Similar confirmation of phenotypes obtained from MO injection or genetic mutants was seen for the cdip1 and tbx24 genes (supplementary material Fig. S2). In summary, the data demonstrate that the aldoaa and kif22 phenotypes were consistently seen after 50% LOF, and are due to activity of these genes in the brain.

**DISCUSSION**

This study has uncovered two previously unknown and fundamental aspects of genes within the 16p11.2 CNV. The first major finding is that the 16p11.2 set of genes is highly active and necessary for early development. This indicates that the set presents a large genetic target, helping to explain the penetrant association of 16p11.2 with multiple brain disorders and other phenotypes. The second major finding is that, among the 16p11.2 set, there are (at least) two single genes with deletion dosage sensor properties, which could link this CNV to ASD, IDD and other disorders.

The finding that the majority of 16p11.2 homologs are required for normal embryonic nervous system and body development is consistent with the early onset of some of these disorders, which suggests that key genes underlying the disorder have developmental roles. Alternatively, key disorder genes might govern the maintenance of aspects of the brain, or lead to a postnatal change in brain function. The finding that LOF for a majority of these genes results in persistent phenotypes through 5 dpf could imply that these are important for maintenance and continued function; however, we did not distinguish whether different mechanisms underlie the early and later phenotypes. We further note that LOF phenotypes observed after extensive RNA knockdown will generally be more severe than those seen in hemizygous mutant fish or human patients. Thus, developmental phenotypes might be present after extensive knockdown, whereas, after partial LOF, later brain function might be altered, with no apparent change to development. By comparison with zebrafish genetic screens, in which approximately 10% of genes give an embryonic phenotype after mutation (Jiang et al., 1996; Schier et al., 1996), 95% of the genes tested here gave a LOF phenotype, almost all involving the brain, suggesting that the multitude of phenotypes associated with the 16p11.2 CNV reflects the activity of many genes. Although the MOs used for LOF assays might sensitize the embryo owing to their unusual nucleic acid backbone, the observed phenotypes were specific and, where tested, were similar to that of genetic mutants [as seen in other studies (e.g. De Rienzo et al., 2011)]. Accordingly, the LOF phenotypes for each gene that we examined formed a unique phenotypic signature: in some cases, abnormalities were seen across the entire spectrum of assays (for example, coro1a and fam57ba), whereas, for other genes, only a subset of phenotypes was observed.

Overall, it is not clear whether the mammalian set of 16p11.2 genes is as active as the fish gene set; however, for 16p11.2 homolog activity that has been reported, mouse and fish LOF phenotypes are often similar. Mouse null knockouts have been reported for only eight 16p11.2 homologs (Coro1a, Doc2a, Kif22, Mapk3, Mvp, Ppp4c, Sez6l2 and Tbx6) (Chapman and Papaioannou, 1998; Foger et al., 2011; Miyazaki et al., 2006; Mossink et al., 2002; Ohsugi et al., 2008; Pages et al., 1999; Sakaguchi et al., 1999; Shui et al., 2007), with one conditional knock out (Prrt2) (Skarnes et al., 2011). All of the homozygous knockouts are associated with phenotypes, except for Sez6l2, which has other copies that are predicted to compensate (Miyazaki et al., 2006), and Mvp (Mossink et al., 2002).
Both Ppp4c and Tbx6 homozygous knockout mice are embryonic lethal; Ppp4c heterozygotes showed growth retardation with decreased survival, whereas Tbx6 heterozygotes were viable and displayed no obvious phenotypes (Chapman and Papaioannou, 1998; Shui et al., 2007). By contrast, tbx24 mutant zebrafish are viable (Nikaido et al., 2002), whereas our ppp4ca knockout mice still had an abnormal phenotype at 5 dpf, implying that they would not survive. Coro1a mutant mice have lower T cell counts, owing to defective migration (Foger et al., 2006) and increased apoptosis. By contrast, mouse knock-in embryos, which a phenotype is observed when ~50% of the normal mRNA remains. We identified two genes in the 16p11.2 CNV, which include Tbx6 and Mapk3, and expressed relative to levels of experimental RNA in control-MO-injected embryos. The broad set of phenotypes that we observed, it is clear that the extensive postnatal lethality of 16p11.2-region deletion mice is a result of the compound hemizygosity of multiple genes. Together, the data imitate the activity of many 16p11.2 genes in development and/or function of the brain and body.

The second major finding from this zebrafish study, definition of the first dosage sensor genes in the 16p11.2 CNV, is groundbreaking because dosage sensor genes that are pivotal for association with mental health disorders have been identified in only a handful of CNVs. These include SHANK3 in 22q13 (Durand et al., 2007), RPA1 in 17q13.3 (Outwin et al., 2011), VIPR2 in 7q36.3 (Vacic et al., 2011) and MBDS in 2q23.1 (Talkowski et al., 2011). We identified two genes in the 16p11.2 CNV, aldoaa and kif22, for which a phenotype is observed after reducing their expression level by ~50% in fish, thus showing characteristics of dosage sensors.

ALDOA is a glycolytic enzyme that catalyzes the conversion of fructose-1,6-bisphosphate to glyceraldehyde-3-phosphate and...
Fig. 7. aldooa and kif22 function are required in the brain. (A) shRNA expression strategy. shRNAs were expressed from a mir30 backbone, under the CNS-specific mir124 promoter. Transient transgenesis was induced using L- Scel meganuclease. (B) Relative expression of aldooa mRNA after inhibition by shRNA. Expression was quantified by qPCR in 24 hpf embryos injected with an aldooa hairpin (aldooa), shown relative to embryos injected with a control hairpin (YFPh). Data was normalized to either (Ba) actin or (Bb) GFP expression. (Bc) Touch response in 24 hpf aldooa embryos versus control YFPh embryos (81% abnormal, n=26). (C) Phenotype of shRNA-injected embryos. (Ca-Ca*) 24 hpf control YFPh embryos (16% abnormal, n=51, in two independent experiments). (Cb-Cb*) 24 hpf aldooa embryos (89% abnormal, n=53, in two independent experiments). (Ca,Cb) Lateral view of the head; (Ca,Cb') dorsal view of the head after brain ventricle injection; (Ca*,Cb*) lateral view of whole embryo. (D) Relative expression of kif22 mRNA after inhibition by shRNA. kif22 mRNA relative expression was quantified by qPCR in 24 hpf embryos injected with a kif22 hairpin (kif22h) relative to those injected with the YFPh control hairpin. Data was normalized to (Da) actin or (Db) GFP expression. (E) Phenotype of shRNA-injected embryos. (Ea-Ea*) 24 hpf YFPh embryos (8% abnormal, n=27, in two independent experiments). (Eb-Eb*) 24 hpf kif22h embryos (81% abnormal, n=26, in two independent experiments). (Ea,Eb) Lateral view of the head; (Ea*,Eb*) dorsal view of the head after brain ventricle injection; (Ea*,Eb*) lateral view of whole embryo.

dihydroxyacetone phosphate. Several other functions and roles have been ascribed to ALDOA, including inhibiting phospholipase D2, binding to the cytoskeleton and RNase activity (Canete-Soler et al., 2005; Kim et al., 2002; Kusakabe et al., 1997). No homozygous null mutations have been identified in humans, indicating that ALDOA is essential (Esposito et al., 2004). Six cases of hemolytic anemia and myopathy have been associated with point mutations and reduced ALDOA activity (Beutler et al., 1973; Esposito et al., 2004; Kishi et al., 1987; Kreuder et al., 1996; Miwa et al., 1981; Yao et al., 2004). One case presented with mental retardation (Beutler et al., 1973), and another with microcephaly and language delay (Kreuder et al., 1996). Further connecting this gene with mental health disorders, expression of ALDOA is upregulated in the cortex of individuals with schizophrenia and depression (Beasley et al., 2006). The mitochondrial citric acid cycle, into which glycolytic end products feed, has been shown to be dysregulated in children with autism (Giulivi et al., 2010), pointing to glycolysis and energy production as possible ALDOA targets. ALDOA was identified as a binding partner for the ASD-linked protein SHANK3 by a protein interactome study (Sakai et al., 2011), as well as being identified in a study implicating postsynaptic signaling complexes in ASD (Kirov et al., 2012). The association of partial LOF in ALDOA with patient phenotypes, as well as these other considerations, suggest that ALDOA is a player in 16p11.2 pathologies.

KIF22 is a microtubule- and DNA-binding molecular motor that is important for chromosome alignment (Santamaria et al., 2008) and compaction during anaphase (Ohsugi et al., 2008). Individuals with a point mutation in the motor domain of KIF22 suffer from the autosomal-dominant skeletal disorder spondyloepimetaphyseal dysplasia with joint laxity (Boyden et al., 2011; Min et al., 2011). No phenotype has been reported in Kif22+/- mice; however, ~50% of Kif22-/- mouse embryos do not survive past the morula stage (Ohsugi et al., 2008). KIF22 has not previously been implicated in brain function disorders, but our data suggest that this gene is required for the formation of neural progenitors. The fact that mammalian heterozygotes in Kif22 have not been associated with phenotypes suggests that Kif22 expression levels might be regulated after loss of one gene copy, or there might be greater redundancy among mammalian kinesins than among the zebrafish genes. For both kif22 and aldooa, the stronger phenotypes seen after partial LOF in zebrafish relative to humans suggest that an additional gene(s) must synergize with ALDOA or KIF22 to convey ASD, IDD or other phenotypic risk in humans.

We suggested that the zebrafish could be a useful tool to address the function of 16p11.2 homologs, without a need to assay for behaviors that are restricted to humans (Sive, 2011). This suggestion is made because the same genetic pathways are active in mammals and fish, and other indicators of pathway activity can be employed. Consistently, almost all zebrafish LOF phenotypes could be prevented by expression of the homologous human gene, supporting gene orthology and shared gene function at the molecular or cellular level. We note that several zebrafish phenotypes could be similar to those seen in individuals with ASD and/or IDD; such phenotypes include abnormal brain size, brain shape, axon tracts, motor readouts, and specification of retinal and tectal neural progenitors (Almgren et al., 2008; Amaral et al., 2008; Courchesne et al., 2007; Hashimoto et al., 1991; Marin-Padilla, 1975; Matson et al., 2011; Ritvo et al., 1986), as well as musculoskeletal defects that are seen in some individuals with ASD and/or IDD (Calhoun et al., 2011; Chen, 1982; Oleskova et al., 2007; Shimojima et al., 2009).
This work identifies the 16p11.2 CNV as an active genomic region and delineates two putative deletion dosage sensor genes in the region, with predictable connection to functional brain syndromes associated with the CNV. These two genes, in combination with additional 16p11.2 or other genes, might be haploinsufficient in ASD and related disorders, leading to abnormal brain function. Other dosage sensors in this interval might exist, perhaps as pairs of synergistically functioning genes. Future assays in the zebrafish will augment antisense MO approaches with RNAi and genetic mutants, determine whether duplication and deletion sensor genes are the same, and screen for synergistic deletion and duplication dosage sensor genes in the 16p11.2 gene set. These unbiased screening approaches are a powerful step in translational research focusing on CNVs that are associated with disorders arising from abnormal brain development and function.

METHODS
Identification of zebrafish 16p11.2 homologs
Zebrafish homologs of human 16p11.2 genes were identified using UniGene, Ensembl and UCSC Genome browsers with alignment and family tree comparisons.

Fish lines and maintenance
Embryos were obtained from natural spawnings. Developmental stages are reported as hpf at 28°C. The NeuroD:GFP line was previously described (Ohbozler et al., 2008; Ulitsky et al., 2011). Additional mutant lines were obtained from the Zebrafish International Resource Center (ZIRC). The tbx24<sup>gln114a/+</sup> line (Nikaido et al., 2002; van Eeden et al., 1996) was increased and homozygotes identified phenotypically at 24 hpf. The cdip<sup>h11032/+</sup> line (Amsterdam et al., 2004) was genotyped using the following primers: hi559_F1: 5’-CTAG-GCGCCATTGCTTTGCAAG-3’; hi559_F2: 5’-ACGCGCCAC-GCTCACTACAGTC-3’; hi559_R1: 5’-TGTTTTGAACGTGTGAATACTACGGC-3’. A 324 bp PCR product was observed in mutants using F1 and R1, but not in wild-type embryos. Wild-type embryos show a 524 bp PCR product using the F2 and R1 primers.

cDNA constructs
Human or zebrafish cDNAs that were used for rescue experiments were cloned into pCS2+ (supplementary material Table S4). Zebrafish cDNAs used for in situ hybridization are also included in supplementary material Table S4. All human and some zebrafish clones were obtained from Open Biosystems. asphd1, c16orf53, doc2a, maz, zebf2, taok2a and taok2b were cloned by PCR from 24-hpf zebrafish cDNA, using primers listed in supplementary material Table S2. We thank Dr Jeremy Green (Kings College, London) for membrane-targeted CAAX-eGFP.

MO design and use
MOs were designed by Gene Tools, LLC, to a splice donor or acceptor site, as close to the 5′ end of the predicted primary RNA as possible. Where a splice site MO gave no phenotype, a translational start site MO was designed. The designed MO sequences are shown in supplementary material Table S2. In all cases, the top MO listed (in Table 1; supplementary material Table S2) for a gene was used in the phenotypic assays described elsewhere, unless otherwise noted. For all experiments, a control MO was injected at the same or greater mass amount. MO (1 nl) was injected into a single cell of a one- to two-cell embryo, using a range of concentrations to determine the lowest concentration at which a phenotype was observed. No more than 7.5 ng of a MO was injected. Unless otherwise stated, the ‘control’ condition refers to control-MO-injected embryos. The control MO sequence is 5′-CCTCTTACCTCAGTTCAATTATA-3′.

Criteria and methodology to assess MO specificity
Specificity of MO-induced LOF phenotypes was determined by the following criteria, as is standard for the field (Bedell et al., 2011; Eisen and Smith, 2008). First, because initial MOs were designed to target splice junctions (described above), it is predicted that a change in RNA splicing will be observed by RT-PCR. These MOs would target zygotically expressed RNAs. Primers to detect knockdown are included in supplementary material Table S2, and RT-PCR methods are discussed below. Where a change in splicing was not detected, an additional splice site MO was designed. For genes for which splice site MOs changed splicing but did not result in an observable phenotype, a translation blocking MO was designed to target maternal transcripts, as well as zygotic. It is further predicted that protein-coding capacity would be altered. This was determined by gel purification of the RT-PCR products after control or test-gene MO injection and sequencing of the PCR product. Protein-coding capacity was determined using Sequencher and MacVector software. In the cases of putative deletion dosage sensor genes, change in expression resulting from MO injection was monitored at the protein level by western blot analysis (described below).

Second, for phenotypes observed after injection of splice blocking MOs, it is predicted that normal RNA levels will decrease, and this was monitored by qPCR, as discussed in the specific Methods section. Correlation between phenotype and MO mass is also predicted, and was assayed in MO titration experiments (Fig. 6, and below).

Third, MO specificity predicts that the LOF phenotype will be prevented (‘rescued’) after co-injection of the MO with the cognate human or zebrafish mRNA that lacks the MO-binding site but preserves protein coding capacity. The appropriate mass of RNA used in rescue experiments was based on both rescue of the LOF phenotype as well as the lack of an overexpression phenotype when the same RNA mass was co-injected with the control MO. Mass of RNA injected in rescue assays and the success of rescue is listed in Table 3. GenBank Accession numbers and cDNA constructs used to synthesize RNA are shown in supplementary material Table S4. RNA synthesis is discussed below. The rescue titration experiments had a minimum of four conditions: control MO plus mGFP to serve as a balancer RNA; LOF MO plus the human/fish RNA; LOF MO plus the human/fish RNA; and LOF MO plus mGFP RNA. Control MO plus mGFP RNA; control MO plus the human/fish RNA; and LOF MO plus the human/fish RNA.

Fourth, for cases in which necrosis or a severe phenotype was observed, the p53 MO was co-injected to suppress off-target cell death (Robu et al., 2007) at a 1.5-fold greater mass amount than elsewhere, unless otherwise noted. For all experiments, a control MO was injected at the same or greater mass amount. MO (1 nl) was injected into a single cell of a one- to two-cell embryo, using a range of concentrations to determine the lowest concentration at which a phenotype was observed. No more than 7.5 ng of a MO was injected. Unless otherwise stated, the ‘control’ condition refers to control-MO-injected embryos. The control MO sequence is 5′-CCTCTTACCTCAGTTCAATTATA-3′.

Finally, where the mutant lines were available, for cdip and tbx24 the MO-induced LOF phenotypes were compared, or, in the cases of aldoaa and kif22, the effects of shRNAs were examined.
Phenotypic scoring procedures
Embryos were scored live using bright-field imaging, or after fixation for axon tracts using scanning confocal imaging. Where there was any ambiguity, or a question of whether phenotypic rescue had been achieved, another lab member scored the embryos. For most genes, more than one of the authors independently assayed MO effects or ability to be rescued by RNA injection. Results were almost always concordant. We required ~70% of embryos in a condition to have an aberrant phenotype in at least two independent experiments, and for the phenotype to be rescued by RNA co-injection, for inclusion in further experiments.

Morphological assays
Brain morphology of LOF embryos was examined at 24 hpf by bright-field and fluorescence microscopy, after injection of the brain ventricles with Texas Red dextran (Gutzman and Sive, 2009). Embryos were scored for: presence of forebrain, midbrain and hindbrain hingepoints; brain ventricle size and volume; forebrain truncation; and eye morphology. Trunk and tail morphology, and the shape of muscle segments, were scored by bright-field microscopy or by staining actin filaments with phalloidin (described below). Phenotypes existing in greater or equal to 70% of embryos and also meeting specificity criteria were included in results unless otherwise stated.

Movement assays
Movement in LOF embryos was monitored at 24 and 48 hpf. In 24 hpf embryos, spontaneous contractions were observed. Spontaneous contractions have been described previously (Saint-Amant and Drapeau, 1998). At 24 hpf, the typical movement consists of side-to-side contractions that result in slow coils. Embryos were observed for several minutes because, by 24 hpf, the contractions are sporadic. Touch response assays were administered at 48 hpf. For this, a loop of thread was used to gently touch the embryos on both the head and the tail. The normal response of a tail stimulus involves the embryo briefly swimming (approximately the length of its body) and landing again on the bottom of the dish, whereas a stimulus to the head begins with full coiling of the embryo resulting in repositioning (C-start) (Issa et al., 2011; Saint-Amant and Drapeau, 1998).

Microscopy
Methods for bright-field microscopy have been described previously (Gutzman and Sive, 2010). Confocal imaging was performed on a Zeiss LSM710, after fixation in 2% TCA (acetylated β-tubulin) or 4% PFA (phalloidin).

In situ hybridization
In situ hybridization methods are described elsewhere (Wiellette and Sive, 2003). Probes used are described in supplementary material Table S4 and wild-type 24 hpf embryos that were fixed in 4% PFA were used to assay spatial expression.

Immunohistochemistry
Whole-mount immunostaining used mouse anti-acetylated α-tubulin (Sigma, 1:1000). Goat anti-mouse Alexa Fluor 488 (Molecular Probes, 1:500) was used as a secondary antibody. Staining by phalloidin Texas Red was performed using a previously described method (De Rienzo et al, 2011).

RT-PCR and qPCR
RT-PCR was performed to monitor expression at developmental time points (supplementary material Fig. S1) and changes in splicing that resulted from MO targeting (Table 2). Total embryo RNA was extracted using Trizol (Invitrogen) followed by chloroform extraction, isopropanol precipitation and DNase treatment or use of the RNaseasy kit (Qiagen). cDNA synthesis was performed with Super Script III Reverse Transcriptase (Invitrogen) and oligo dT or random hexamers. Primers for RT-PCR and qPCR are shown in supplementary material Table S2. To detect changes in splicing, primers were designed around targeted exons.

RT-PCR was performed using Hot Start Taq Plus (Qiagen) and primers shown in supplementary material Table S2. Primers were designed to only recognize normal transcript (see Fig. 5 for primer design strategy). Knockdown was confirmed by sequencing PCR products, and MO effects on expression of predicted protein are included in Table 2. qPCR was performed using an ABI Prism 7900 (ABI). Fluorescence detection chemistry utilized SYBR Green dye master mix (Roche). The relative amount of product was calculated using ACT and normalized to Ef1α. Values are reported with standard deviation.

Each assay was performed in at least two independent experiments. Each experiment contained at least 90 embryos per condition divided into three separate RNA preparations (biological replicates). Each RNA preparation was used for one reverse transcription reaction, which was then used in triplicate for each qPCR reaction (technical replicates).

RNA injections
RNA was synthesized using the Message Machine kit (Ambion) and injected as described previously (Gutzman et al., 2008).

Western blot analysis
Methods for western blot analysis have been described previously (Gutzman and Sive, 2010). Human anti-KIF22 antibody (Sigma K1390) used at 1:1000 in 3% BSA TBST and human anti-ALDOA antibody (Sigma WHH000226M1) used at 1:1000 in 5% milk TBST were detected with anti-mouse HRP secondary antibody (Sigma). Human anti-GAPDH antibody (Abcam ab22555) used at 1:3000 in 5% milk TBST and human anti-eIF4E antibody (Cell Signaling 9742S) used at 1:500 in 5% BSA TBST were detected using anti-rabbit HRP secondary antibody (Cell Signaling). Because the anti-ALDOA antibody is not specific for the protein product of aldoa and is expected to cross-react with that of aldoab, the Aldoa western blot was performed using heads only, because the aldoab gene is only expressed in the tail (not shown).

RNAi methods
Hairpins for aldoa and kif22 were designed using Invitrogen Block-IT RNAi Designer software. A total of four hairpins per gene were designed and analyzed. Hairpin oligonucleotide pairs were purified by SDS-PAGE, and annealed by heating to 95°C and slow cooling to 10°C. Annealed oligonucleotides were subcloned into the mir30 backbone (Dong et al., 2009) of the I-SceI-miR124:GFP-mir30-pA plasmid, prepared by Dr Jennifer Gutzman, University of Wisconsin-Milwaukee, WI. This plasmid consists of the CNS-specific promoter miR124 (Shkumatava et al., 2009) driving a GFP reporter upstream of the mir30 backbone and an SV40 polyA
Clinical issue
Copy number variant regions (CNVs) are intervals of DNA ranging from 1000 bp to several megabases, in which one genomic copy is either duplicated or deleted, changing the number of gene copies in that interval. Because CNVs have been associated with many diseases, from cancer to autoimmune disease to neuropsychiatric disorders, understanding how they cause deleterious effects is important. Carriers of 16p11.2 CNVs present with a wide range of disorders, including intellectual disability disorder (IDD) and autism spectrum disorders (ASD). The genes in the 16p11.2 CNV are probably integral to normal brain function. There are 25 genes in the central core interval, and it is hypothesized that dosage changes in one or more of these genes underlie the pathologies associated with the 16p11.2 CNV. However, the crucial genes in the 16p11.2 interval – and in many CNVs associated with other disorders – are unknown.

Results
This study used the zebrafish as a tool to study the activity of genes that are homologous to those in the human 16p11.2 interval, and to identify which might be most important for the association of the 16p11.2 CNV with human brain disorders. Of the 25 human genes in this interval, 22 homologs were identified in zebrafish, and 21 displayed embryonic and larval loss-of-function phenotypes, demonstrating that this set of genes is very important during development. In total, 20 genes were found to be necessary for proper brain size and shape, with subsets also affecting eye development, axon tract organization, movement behaviors and muscle formation. The authors also examined whether these phenotypes persisted when each gene produced only 50% of its product (equivalent to losing one copy of a gene). Two genes were sensitive to dosage: one encoding glycolytic enzyme aldolase A (Aldoa) and the other microtubule motor kinesin family member 22 (kif22). Thus, the function of these genes changes with copy number, which might explain the link between 16p11.2 CNVs and associated disorders.

Implications and future directions
These data show that the 16p11.2 CNV comprises a highly active set of genes that are important for the formation of the nervous system and probably also for its function. Most importantly, two genes were identified as having functions that were sensitive to dosage, indicating that these might be crucial in connecting this CNV to IDD, ASDs and other disorders. Future directions include experiments to understand the molecular pathways by which each gene works, and whether each works together with other genes in the 16p11.2 interval, as predicted by human genetic data. This information will help to define targeted assays in mammals, and possibly guide therapeutic directions. This study further shows that zebrafish could be used to identify conserved, dosage-sensitive genes in other CNVs that are implicated in other directions. This study further shows that zebrafish could be used to identify conserved, dosage-sensitive genes in other CNVs that are implicated in other human disorders.

ACKNOWLEDGEMENTS
Thanks to Olivier Pauget for expert fish husbandry. We are grateful to Sive laboratory members for support, and comments on the manuscript, especially Isabel Brachmann and Laura Jacox. Special thanks to Michael Lee for comments, Jennifer Gutzman for the I-SceI-miR124/GFP-miR30-pA plasmid, and Jacob Austin-Breneman for shRNA design. Thanks to George Bell and Prathapan Thiru of BARC for identification of homologs, Jeong-Ah Kwon in the Genome Technology Core for help with qPCR, and Nicki Watson and Wendy Salmon in the Keck Imaging Facility. Thanks to Mark Daly and Stephen Haggarty for useful discussions, and to Ed Skolnick for encouragement. Thanks to colleagues in the Simons Center for the Social Brain at MIT, SPARR investigators and members of the Boston Autism Consortium, for discussion.

COMPETING INTERESTS
The authors declare that they do not have any competing or financial interests.

AUTHOR CONTRIBUTIONS
A.B.-L. isolated, and characterized the expression and function of, zebrafish 16p11.2 homologs, analyzed the kif22 gene as a putative dosage sensor, prepared figures and made a major contribution to writing the manuscript. S.G. isolated, and characterized the expression and function of, zebrafish 16p11.2 homologs, standardized the qPCR assay, analyzed the aldooa gene as a putative dosage sensor, prepared figures and contributed to the manuscript. J.M.M. characterized zebrafish 16p11.2 homolog function, including analysis of axon tracts and protein expression, and made a major contribution to writing the manuscript. G.D.R. performed shRNA analysis of aldooa and kif22 genes, prepared a figure and contributed to the manuscript. H.S. initiated and directed the study, and made a major contribution to writing the manuscript.

FUNDING
This work was supported by the Simons Foundation Autism Research Initiative (grant 95091).

SUPPLEMENTARY MATERIAL
Supplementary material for this article is available at http://dmm.biologists.org/lookup/suppl/doi:10.1242/dmm.009944/-/DC1

REFERENCES
Aldolases a and C are deficiencies associated with hereditary hemolytic anemia in a Japanese family.


Vaccarino, M. A., Malhotra, D., Murray, F., Chou, H. H., Peoples, A., Makarova, V., Yoon, S., Bhandari, A., Corominas, R. et al. (2011). Duplication of the...