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

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

Alicia Blaker-Lee*, Sunny Gupta*, Jasmine M. McCammon*, Gianluca De Rienzo and Hazel Sive#^.

Whitehead Institute for Biomedical Research

Nine Cambridge Center, Cambridge MA 02142

* equal contribution

Massachusetts Institute of Technology

^ corresponding author: sive@wi.mit.edu

Running title: zebrafish 16p11.2 homologs

Keywords: zebrafish, 16p11.2, autism and intellectual disability, copy number variant, aldolase, kinesin

DMM Advance Online Articles. Published 1 May 2012 as doi: 10.1242/dmm.009944
Access the most recent version at http://dmm.biologists.org/lookup/doi/10.1242/dmm.009944

SUMMARY

Deletion or duplication of one copy of the human 16p11.2 interval is tightly associated with impaired brain function, including autism spectrum disorders (ASD), 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. Further, genes whose functional levels change with deletion or duplication (termed "dosage sensors"), which may associate the CNV with pathologies, have not been identified. Using the zebrafish as a tool, a set of 16p11.2 homologs was identified, primarily on chromosomes 3 and 12. Use of eleven phenotypic assays, spanning the first five days of development, demonstrates that this set of genes is highly active, such that 21 out of 22 homologs tested show loss of function phenotypes. Most genes are required for nervous system development – impacting brain morphology, eye development, axonal density or organization, and motor response. In general, human genes can substitute for the fish homolog, demonstrating orthology, and consistent with 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 are reduced by ~50%, suggesting that these genes are deletion dosage sensors. This study leads to two major findings. The first is that the 16p11.2 region comprises a highly active set of genes, which may present a large genetic target, and may 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 amongst the 16p11.2 set, which may link this CNV to brain disorders including ASD and IDD.

INTRODUCTION

Copy number variation (CNV), where intervals of the genome are deleted or duplicated, has 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/developmental delay (often characteristics of intellectual disability disorder (IDD)) (85% of deletion carriers) and autism spectrum disorders (ASD, 19-28% of deletion carriers) (Ciuladaite et al., 2011; The Simons Vip, 2012). In addition to IDD (Bijlsma et al., 2009) and ASD (Kumar et al., 2008; Marshall et al., 2008; Weiss et al., 2008), deletion and/or duplication of 16p11.2 are associated with seizure disorder (Ghebranious et al., 2007), obesity or being underweight (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 hemizygous for the 16p11.2 homologs exhibit a 50% rate of neonatal lethality (Horev et al., 2011).

A key general question is how many 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 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 may 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 were affected by deletion or duplication, as predicted by gene dosage (Horev et al., 2011). Further supporting the notion of dosage sensitivity, in human patients, 16p11.2 duplication or deletion is sometimes correlated with reciprocal phenotypes, such that obesity and macrocephaly are associated with 16p11.2 deficiency, being underweight and microcephaly with duplication (Jacquemont et al., 2011; Shinawi et al., 2010). Similarly, opposite phenotypes exist in the hemizygous 16p11.2 region mice compared to mice containing the duplication, including brain volume and certain behaviors (Horev et al., 2011).

In this study, we use 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 are interested in the high prevalence of associated brain disorders, with particular interest in the CNV'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 mouse, due 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, since they present soon after birth (Konopka et al., 2012; Ploeger et al., 2010). However, maintenance of brain structure, or faulty brain function after birth may play key roles in etiology of these disorders (Okado et al., 2001; Trembath, 1994). The most accessible time window for analysis in the zebrafish is the first five 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 suggesting more targeted mammalian studies.

We show here that most individual zebrafish 16p11.2 homologs are required for normal brain and body development, and that their function is conserved with the human gene. 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 twenty-five protein-coding genes. Twenty-one 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 are 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 (Table S1). These regions are not syntenic with human chromosome 16, since gene order is not conserved. Two sets of syntenic genes were found on chromosome 3; one region comprises *kctd13, sez6l2* and *asphd1,* where the order is not conserved with human; and *mapk3, gdpd3,* and *ypel3*, where the order is conserved. Interestingly, the first set of syntenic genes (*kctd13, sez6l2,* and *asphd1*) is found in a microdeletion associated with ASD (Crepel et al., 2011).

Five homologs were present in two copies (*aldoa, fam57b, gdpd3, ppp4c* and *taok2)* reflecting the partial duplication of the teleost genome (Postlethwait et al., 2000). Genes in each pair had similar but not identical sequences (Table S1), with one found on zebrafish chromosome 12 and the other on chromosome 3. Such duplication may 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 RT-PCR (see Methods, Fig. S1A, Table S2). Most genes are expressed maternally and zygotically, at least until 48 hours post fertilization (hpf). *asphd1, doc2a, prrt2,* and *sez6l2* are expressed only zygotically. Whole mount in situ hybridization (Fig. S1B) showed that all genes, except *tbx24*, are expressed in the brain at 24 hpf (Thisse, 2005).

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 are 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 human, indicate that this gene set is 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 five 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, Table S2), to target zygotic RNA. Where a splice site MO did not give a phenotype, a 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, Table S2) and the resulting effects on 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 (Fig. S1), and both were assayed. Thus, twenty-two 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, as 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 Table 2 and Table 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, as 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, due to species differences or because the MO binding site lies across a splice junction. Fourth, off-target effects of MOs may cause cell death, which can effectively be suppressed by injection of a p53 MO (Robu et al., 2007). Where severe phenotypes were observed, the effect of p53 suppression is tested and if a resulting phenotype was milder, it is the one scored. Finally, comparison of a phenotype obtained with MOs is compared to mutants (or shRNAs) to test similarity of phenotypes. Since 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 Table S3).

Strikingly, LOF for almost all genes (twenty out of twenty-two assayed), with the exception of *prrt2* and *tbx24*, led to changes in brain or eye morphology (Fig. 2A, Fig. 3, 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 (Fig. S2), (Thisse, 2005). Thus, a total of twenty-one out of twenty-two 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, Table S3). Several genes gave very strong phenotypes, suggesting early embryonic defects. In particular, *coro1a, ino80e, 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, 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, while no phenotype was observed after *prrt2* LOF. Since *prrt2* is not highly expressed until 48 hpf, gene function may be required later (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 may result from changes in neuroepithelial specification, morphogenesis, or reduction in cerebrospinal fluid volume (Gato and Desmond, 2009; Lowery and Sive, 2009).

A second group of genes leading to defective brain morphology after LOF showed a straight midbrain (Fig. 2C), where the midbrain hingepoint was essentially absent. This phenotype was seen in *aldoaa*, *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 *aldoaa, fam57ba,* and *maz* LOF embryos, while 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 Figs. 2B,C suggest similar contribution to brain development by all genes in the group, close comparison reveals distinct phenotypes; for example, *aldoaa* and *ypel3* morphants have similar midbrain phenotypes, but MHB and hindbrain phenotypes are 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 and 3, 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 wildtype phenotype. Assays for rescue included observation of the shape of the brain, ventricle size, and eye morphology, movement and axon tracts.

In summary, twenty-one out of twenty-two zebrafish 16p11.2 homologs tested are 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.

Movement deficiency is associated with loss of function 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, kif22, maz* and *ppp4ca*) was associated with spontaneous movement defects. LOF in fourteen genes led to a defect in touch response in at least 70% of embryos for each gene examined (Fig. 3, Table S3). In the most severe cases, *aldoaa* and *fam57ba,* LOF embryos exhibited no response to touch, and these severe phenotypes were rescued by co-injection 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 is 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, Table S3). These data indicate that multiple 16p11.2 homologs were required for normal motor activity, as reflected by spontaneous movement or touch-responsiveness.

A subset of genes is required for normal axon tract development

Motor deficiencies observed in LOF embryos led us to ask whether axon tract formation is 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, kctd13, kif22, mapk3* and *ppp4ca* (Fig. 5, and quantification in legend). LOF in all of these genes led to reduced and disorganized tracts, however the *kctd13* LOF forebrain tract phenotype was mild, and *kif22* LOF hindbrain tracts appeared normal. For each gene, normal phenotypes were observed in at least 75% of embryos examined after co-injection of cognate human mRNA (or fish mRNA for

kctd13) with the antisense MO, further supporting conservation of function and specificity of the axon tract phenotypes caused by MO injection (Fig. 5).

In addition to brain axon tract deficiencies, we demonstrated that pigmentation, indicative of neural crest lineages that include peripheral nerves (Schilling and Kimmel, 1994), was abnormal after LOF in a subset of homologs (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*) where 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, while axon tracts may appear normal, synaptic transmission may be defective. These data show that multiple 16p11.2 homologs are necessary for normal axon tract development – suggesting deficits in formation of neuronal precursors, guidance, or fasciculation – and that axonal deficiencies may 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 may be associated with IDD, ASD and other phenotypes. We define 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 fourteen 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, Table S2). Due 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, *aldoaa* and *kif22*, were reproducibly associated with a phenotype when approximately 45% RNA remained for *aldoaa* and 55% RNA remained for *kif22* (Fig. 6C), 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 *aldoaa* 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, where Western blot analysis showed a 56% and 65% loss of Aldoaa and Kif22 protein expression, respectively, after MO injection (Fig. 6E).

The sensitivity of embryos to *aldoaa* and *kif22* to 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). Since the brains in both *aldoaa* and *kif22* LOF embryos appeared narrow (Fig. 2A), we asked 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 (Obholzer et al., 2008; Ulitsky et al., 2011). After LOF in both *aldoaa* and *kif22,* NeuroD promoter-driven expression of GFP was decreased in the eyes and optic tectum, whereas expression in the cranial neurons and pancreas appeared 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 *aldolase a* (*aldoaa*) and *kinesin family member*

22 (*kif22)* genes each show a robust phenotype at 50% LOF, which may 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 *aldoaa* and *kif22*

In order to address whether the effects of *aldoaa* and *kif22* LOF were due directly to changes in expression within the brain, or secondarily, due to effects on other tissues, shRNAs targeting these genes were expressed in the zebrafish brain. We used the central nervous system-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., in preparation), (Fig. 7A). By testing four hairpins against either *aldoaa* or *kif22*, a targeting construct was identified for each gene. These constructs reduced normal RNA levels to 46% for *aldoaa* 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, 7D). 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, likely reflects defective convergence and extension that can be modulated by *kif22* activity in the spinal cord (De Rienzo et al., 2011).

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

DISCUSSION

This study has uncovered two novel 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, indicating that the set presents a large genetic target, which may explain the penetrant association of 16p11.2 with multiple brain disorders and other phenotypes. The second major finding is that, amongst the 16p11.2 set, there are (at least) two single genes with deletion dosage sensor properties, which may link this CNV to ASD, IDD and other disorders.

The finding that the majority of 16p11.2 homologs is required for normal embryonic nervous system and body development is consistent with the early onset of some of these disorders, which may suggest that key genes underlying the disorder have developmental roles. Alternately, key disorder genes may govern 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 may imply that these are important for maintenance and continued function, however, we did not distinguish whether different mechanisms underlie 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 may be present after extensive knockdown, whereas after partial LOF, later brain function may be altered, with no apparent change to development. By comparison with zebrafish genetic screens, where 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 including the brain, suggesting that the multitude of phenotypes associated with 16p11.2 CNV reflects activity of many genes. While the MOs used for LOF assays may sensitize the embryo due to their unusual nucleic acid backbone, phenotypes we observed were specific, and where tested, similar to that of genetic mutants (as seen in other studies, for example, De Rienzo et al., 2011). Accordingly, the LOF phenotypes for each gene we examined formed a unique phenotypic signature,

where in some cases abnormalities were seen across the entire spectrum of assays, for example, *coro1a* and *fam57ba*, while 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, 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 *Mvp* (Mossink et al., 2002) and *Sez6l2,* which has other copies that are predicted to compensate (Miyazaki et al., 2006). Both *Ppp4c* and *Tbx6* homozygous knockout mice are embryonic lethal; *Ppp4c* heterozygotes showed growth retardation with decreased survival, while *Tbx6* heterozygotes were viable and displayed no obvious phenotypes (Chapman and Papaioannou, 1998; Shui et al., 2007). In contrast, *tbx24* mutant zebrafish are viable (Nikaido et al., 2002), while our *ppp4ca* knockdown fish still had an abnormal phenotype at 5 dpf, implying they will not survive. *Coro1a* mutant mice have lower T cell counts due to defective migration (Foger et al., 2006) and increased apoptosis (Mueller et al., 2011), while *Mapk3* mutant mice have reduced numbers of thymocytes (Pages et al., 1999). We did not evaluate zebrafish immune response, however, *coro1a* LOF zebrafish showed highly abnormal axon tracts, consistent with a migration defect. *Doc2a* knockout mice exhibit defects in excitatory synaptic transmission and long-term potentiation (Sakaguchi et al., 1999), while knockdown of *doc2a* in zebrafish resulted in defective brain morphology, but apparently normal motor responses and axon tracts. Because the mice were studied at later stages, similar phenotypes may develop in older fish. *Mapk3* mutant mice, in combination with *Mapk1* deficiency, exhibit defective neurogenesis (Satoh et al., 2011), and similarly, we observed defective axon tracts in *mapk3* LOF zebrafish. Given the broad set of phenotypes we observed, it seems likely that the extensive postnatal lethality of 16p11.2-region deletion mice is a result of the compound hemizygosity of multiple genes.

Together, the data implicate 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 have been identified in only a handful of CNVs where they are pivotal for association with mental health disorders. 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 *MBD5* in 2q23.1 (Talkowski et al., 2011). We identified two genes in the 16p11.2 CNV, *aldoaa* and *kif22,* where a phenotype is observed after reducing their expression level by ~50% in fish, thus showing characteristics of deletion dosage sensors.

ALDOA is a glycolytic enzyme that catalyzes the conversion of fructose-1,6-bisphosphate to glyceraldehyde-3-phosphate and dihydroxyacetone phosphate. Several other functions and roles have been ascribed to ALDOA, including inhibiting phospholipase D2, binding to the cytoskeleton, and RNAase 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), 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 patients 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 by a protein interactome study as a binding partner for the ASD-linked gene SHANK3 (Sakai et al., 2011) as well as 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, important for chromosome alignment (Santamaria et al., 2008) and compaction during anaphase (Ohsugi et al., 2008). Patients with a point mutation in the motor domain of *KIF22* suffer from the autosomal-dominant skeletal disorder, spondyloepimethaphyseal 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 suggests that this gene is required for formation of neural progenitors. Since mammalian heterozygotes in *Kif22* have not been associated with phenotypes, *Kif22* expression levels may be regulated after loss of one gene copy, or there may be greater redundancy amongst mammalian Kinesins than amongst zebrafish genes. For both *kif22* and *aldoaa*, the stronger phenotypes seen after partial LOF in zebrafish relative to human 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 function of 16p11.2 homologs, without a need to assay for behaviors restricted to humans (Sive, 2011), specifically due to use of the same genetic pathways in mammals and fish. 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. With reference to our specific interest in ASD, we note that several zebrafish phenotypes may be similar to those seen in ASD (but also in IDD) patients, including abnormal brain size and shape, axon tracts and 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 seen in some ASD (but also in IDD) patients (Calhoun et al., 2011; Chen, 1982; Oslejskova 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 genes, in combination with additional 16p11.2 or other genes, may be haploinsufficient for normal brain function. Other dosage sensors in this interval may exist, perhaps as pairs of synergistically functioning genes. Future assays in the zebrafish will augment antisense MO approaches with RNAi and genetic mutants, will 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 associated with disorders arising from abnormal brain development and function.

METHODS

Identification of zebrafish 16p11.2 homologs

Zebrafish homologs of human16p11.2 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 hours post-fertilization (hpf) at 28°C. The *NeuroD:GFP* line, was previously described (Obholzer et al., 2008; Ulitsky et al., 2011). Additional mutant lines were obtained from ZIRC. The *tbx24* $^{te314a/+}$ line (Nikaido et al., 2002; van Eeden et al., 1996) was incrossed, and homozygotes identified phenotypically at 24 hpf. The *cdipt* $^{hi559Tg/+}$ incrossed line (Amsterdam et al., 2004) was genotyped using the following primers:

hi559_F1: 5'-CTAGCTTGCCAAACCTACAGG-3'

hi559_F2: 5'-ACGCGCCACGCTCATCTACAGTC-3'

hi559_R1: 5'-TGGTTGTAACGTGTAATACTACGC-3'

A 324 bp PCR product was observed in mutants using F1 and R1, but not wildtype embryos. Wildtype embryos show a 524 bp PCR product using F2 and R1 primers.

cDNA constructs

Human or zebrafish cDNAs used for rescue experiments were cloned into pCS2+ (Table S4). Zebrafish cDNAs used for in situ hybridization are also included in Table S4. All human and some zebrafish clones were obtained from Open Biosystems. *asphd1, c16orf53, doc2a, maz, sez6l2, taok2a,* and *taok2b* were cloned by PCR from 24 hpf zebrafish cDNA, using primers listed in Table S2. We thank Dr. Jeremy Green (Kings College, London) for membrane-targeted CAAX-eGFP.

Antisense morpholino-modified oligonucleotide design and use

Antisense morpholino-modified oligonucleotides (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 Table S2. In all cases the top MO listed (in Tables 1 and S2) for a gene was used in phenotypic assays described elsewhere unless otherwise noted. For all experiments, a control MO was injected at the same or greater mass amount.

One nL was injected into a single cell of a 1-2 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'-CCTCTTACCTCAGTTACAATTTATA-3'

Criteria and methodology to assess MO specificity

Specificity of MO-induced LOF phenotypes was determined by the following criteria, as are the standard for the field (Bedell et al., 2011; Eisen and Smith, 2008). First, since initial MOs were designed to target splice junctions (described above), it is predicted that a change in RNA splicing would be observed by RT-PCR. These MOs would target zygotically expressed RNAs. Primers to detect knockdown are included in 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 where 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 the PCR product. Protein coding capacity was determined by 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 methods below).

Third, MO specificity predicts that the LOF phenotype will be prevented ("rescued") after co-injection of the MO with 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 Table S4. RNA synthesis is discussed below. The rescue titration experiments had a minimum of 4 conditions; control MO plus mGFP to serve as a balancer RNA, LOF MO plus mGFP RNA, control MO plus the human/fish RNA, and LOF MO plus the fish RNA.

Fourth, where necrosis, or a severe phenotype was observed, the p53 MO was co-injected, to suppress off-target cell death (Robu et al., 2007), at 1.5-fold greater mass amount than the mass of experimental or control MO, as indicated in Table 2. The p53 MO sequence is:

5'-GCGCCATTGCTTTGCAAGAATTG-3'.

Finally, where the mutant lines were available, for *cdipt* 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 brightfield 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 2 independent experiments, and for the phenotype to be rescued by RNA co-injection.

Morphological assays

LOF embryos were examined by brightfield and fluorescence microscopy, at 24 hpf, for brain morphology, 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 brightfield 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 hpf and 48 hpf. In 24 hpf embryos spontaneous contractions were observed. Spontaneous contractions are previously described (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, as 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 (about 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 brightfield microscopy have been described previously (Gutzman and Sive, 2010). Confocal imaging was performed on a Zeiss LSM710, after fixation in 2% TCA (acetylated α-tubulin) or 4% PFA (phalloidin).

In situ hybridization

In situ hybridization methods are described elsewhere (Wiellette and Sive, 2003). Probes used are described in Table S4 and wildtype 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 (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 and isopropanol precipitation and DNAase treatment or RNeasy kit (Qiagen). cDNA synthesis was performed with Super Script III Reverse Transcriptase (Invitrogen) and oligodT or random hexamers. Primers for RT-PCR and qPCR are shown in Table S2. For detecting changes in splicing, primers were designed around targeted exons.

RT-PCR was performed using Hot Start Taq Plus (Qiagen) and primers in 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 a ABI Prism 7900 (ABI). Fluorescence detection chemistry utilized SYBR green dye master mix (Roche). The relative amount of product was calculated using ΔCT 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 (Ambien), 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 and human anti-ALDOA antibody (Sigma WH0000226M1) used at 1:1000 in 5% milk were detected with anti-mouse HRP secondary antibody (Sigma). Human anti-GAPDH antibody (Abcam ab22555) used at 1: 3000 in 5% milk TBS-T and human anti-eIF4E antibody (Cell Signaling 9742S) used at 1:500 in 5% BSA TBS-T were detected using anti-rabbit HRP secondary antibody (Cell Signaling). Since the Aldoa antibody is not specific for the protein product of *aldoaa* and is expected to cross-react with *aldoab,* the Aldoa Western blot was performed using heads only, as the *aldoab* gene is only expressed in the tail (not shown).

RNAi methods

Hairpins for *aldoaa* and *kif22* were designed using Invitrogen Block-IT RNAi Designer software. A total of four hairpins per gene was designed and analyzed. Hairpin oligonucleotide pairs were purified by SDS-PAGE, annealed by heating to 95^{o}C and slow cooling to 10^{o}C. Annealed oligonucelotides were subcloned into the miR30 backbone (Dong et al., 2009), of the I-SceI-miR124:GFP-miR30-pA plasmid, prepared by Dr. Jennifer Gutzman. 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 addition site, with the expression cassette flanked by I-SceI restriction sites. Transgenesis was effected by the meganuclease (I-SceI) method (Thermes et al., 2002), using fresh I-SceI for each transgenic preparation.

aldoaa hairpin 1318 (binding a site in the 3' UTR):

sense strand, 5'-

GGCTAGCAGTTACTTCCTTATGTGTGAAACACTGGTGCACATGATGGAGTGTTTCACACATGGAAGTAAC

C- 3';

antisense strand, 5'-

TCGTCAATGAAGGAATACACACTTTGTGACCACGTGTACTACCTCACAAAGTGTGTACCTTCATTGGTCG

G- 3';

kif22 hairpin 1616 (binding a site in the 9th exon):

sense strand, 5'-

GGCTAGCAGGCCGTTGTTTTACTCCATTACACTGGTGCACATGATGGAGTGTAATGGAGTAACAACGGCC

C-3';

antisense strand, 5'-

GGCTGGGCCGTTGTTACTCCATTACACTCCATCATGTGCACCAGTGTAATGGAGTAAAACAACGGCCTGC

T- 3'

Acknowledgements

Thanks to Olivier Paugois for expert fish husbandry. We are grateful to Sive lab 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 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, SFARI investigators and members of the Boston Autism Consortium, for discussion. This work was supported by the Simons Foundation Autism Research Initiative, grant 95091.

Financial or competing interests disclosure

The authors declare that they have no financial or competing interests.

Author contributions

1. Alicia Blaker-Lee isolated and characterized 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.

2. Sunny Gupta isolated and characterized expression and function of zebrafish 16p11.2 homologs, standardized the qPCR assay, analyzed the *aldoaa* gene as a putative dosage sensor, prepared figures and contributed to the manuscript.

3. Jasmine McCammon characterized zebrafish 16p11.2 homolog function, including analysis of axon tracts and protein expression and made a major contribution to writing the manuscript.

4. Gianluca De Rienzo performed shRNA analysis of *aldoaa* and *kif22* genes, prepared a figure and contributed to the manuscript.

5. Hazel Sive 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, through grant 95091.

TRANSLATIONAL IMPACT

Clinical issue

Copy number variants (CNVs) are segments of DNA ranging from 1,000 bp to several megabases, where one genomic copy is either duplicated or deleted, changing the number of gene copies in that interval. Because CNVs have been associated with many disorders, from cancer to autoimmune disease to neuropsychiatric disorders, understanding the mechanisms by which CNVs can be deleterious is very important. Carriers of 16p11.2 CNVs present with a wide range of anomalies, including intellectual disability disorder (IDD) and autism spectrum disorders (ASD). Therefore, the genes in this CNV are likely integral to normal brain function. With 25 genes identified in the central core interval, it is hypothesized that dosage changes in one of more of these genes links the CNV to associated pathologies. However, the critical genes in 16p11.2 and the majority of other CNVs are unknown.

Results

This study utilized the zebrafish as a tool to study activity of genes homologous to those in the human 16p11.2 interval, and identify which genes from this CNV may be most important for association with brain disorders. While zebrafish do not possess the behavioral repertoire necessary to recapitulate human behaviors, their gene function is conserved with mammals and they are an ideal system for rapid genetic manipulations. These attributes make zebrafish an appropriate system with which to study the large number of genes in the 16p11.2 CNV. Of twenty-two homologs identified in zebrafish, twenty-one displayed embryonic and larval loss of function phenotypes, demonstrating that this set of genes is very

important during development. Twenty genes are necessary for proper brain size and shape, with subsets also affecting eye development, axon tract organization, movement behaviors, and muscle formation. Because changes in dosage for critical genes in the 16p11.2 are anticipated to be detrimental and lead to onset of associated disorders, the authors examined whether these phenotypes persisted when the gene produced only 50% of its product, equivalent to losing one copy of a gene. The majority of genes analyzed did not have a phenotype at 50% reduction; however, two genes were sensitive to dosage. These encode the glycolytic enzyme Aldolase A (Aldoa), and the microtubule motor Kinesin family member 22 (Kif22). These results suggest that the function of genes encoding Aldoa and Kif22 change with copy number, and could link the 16p11.2 CNV to pathologies.

Implications and future directions

These data show that the 16p11.2 CNV comprises a highly active set of genes, particularly for formation of the nervous system, but likely also for nervous system function. Most importantly, two genes were identified as having functions that were sensitive to dosage, suggesting that these may be critical in connecting the CNV to IDD, ASD and other disorders. Future directions include experiments to understand the molecular pathways by which each works, and whether each gene works together with others in the 16p11.2 region, which human genetic data would predict. Such data will help define targeted assays in mammals, and possible therapeutic directions. This study further shows that zebrafish can be used to identify genes that are dosage sensitive from other CNVs implicated in other disorders.

REFERENCES

Almgren, M., Schalling, M. and Lavebratt, C. (2008). Idiopathic megalencephaly-possible cause and treatment opportunities: from patient to lab. *Eur. J. Paediatr. Neurol*. **12**, 438-445.

Amaral, D. G., Schumann, C. M. and Nordahl, C. W. (2008). Neuroanatomy of autism. *Trends Neurosci.* **31**, 137-145.

Amsterdam, A., Nissen, R. M., Sun, Z., Swindell, E. C., Farrington, S. and Hopkins, N. (2004). Identification of 315 genes essential for early zebrafish development. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 12792-12797.

An, M., Luo, R. and Henion, P. D. (2002). Differentiation and maturation of zebrafish dorsal root and sympathetic ganglion neurons. *J. Comp. Neurol.* **446**, 267-275.

Bardakjian, T. M., Kwok, S., Slavotinek, A. M. and Schneider, A. S. (2010). Clinical report of microphthalmia and optic nerve coloboma associated with a de novo microdeletion of chromosome 16p11.2. *Am. J. Med. Genet. A*. **152A**, 3120-3123.

Beasley, C. L., Pennington, K., Behan, A., Wait, R., Dunn, M. J. and Cotter, D. (2006). Proteomic analysis of the anterior cingulate cortex in the major psychiatric disorders: Evidence for disease-associated changes. *Proteomics*. **6**, 3414-3425.

Bedell, V. M., Westcot, S. E. and Ekker, S. C. (2011). Lessons from morpholino-based screening in zebrafish. *Brief. Funct. Genomics*. **10**, 181-188.

Berger, W., Steiner, E., Grusch, M., Elbling, L. and Micksche, M. (2009). Vaults and the major vault protein: novel roles in signal pathway regulation and immunity. *Cell. Mol. Life Sci.* **66**, 43-61.

Beutler, E., Scott, S., Bishop, A., Margolis, N., Matsumoto, F. and Kuhl, W. (1973). Red cell aldolase deficiency and hemolytic anemia: a new syndrome. *Trans. Assoc. Am. Physicians*. **86**, 154-166.

Bijlsma, E. K., Gijsbers, A. C., Schuurs-Hoeijmakers, J. H., van Haeringen, A., Fransen van de Putte, D. E., Anderlid, B. M., Lundin, J., Lapunzina, P., Perez Jurado, L. A., Delle Chiaie, B. et al. (2009). Extending the phenotype of recurrent rearrangements of 16p11.2: deletions in mentally retarded patients without autism and in normal individuals. *Eur. J. Med. Genet*. **52**, 77-87.

Boyden, E. D., Campos-Xavier, A. B., Kalamajski, S., Cameron, T. L., Suarez, P., Tanackovic, G., Andria, G., Ballhausen, D., Briggs, M. D., Hartley, C. et al. (2011). Recurrent dominant mutations affecting two adjacent residues in the motor domain of the monomeric kinesin KIF22 result in skeletal dysplasia and joint laxity. *Am. J. Hum. Genet*. **89**, 767-772.

Calhoun, M., Longworth, M. and Chester, V. L. (2011). Gait patterns in children with autism. *Clin. Biomech. (Bristol, Avon)*. **26**, 200-206.

Canete-Soler, R., Reddy, K. S., Tolan, D. R. and Zhai, J. (2005). Aldolases a and C are ribonucleolytic components of a neuronal complex that regulates the stability of the light-neurofilament mRNA. *J. Neurosci.* **25**, 4353-4364.

Chapman, D. L. and Papaioannou, V. E. (1998). Three neural tubes in mouse embryos with mutations in the T-box gene Tbx6. *Nature*. **391**, 695-697.

Chen, H. (1982). Skeletal dysplasias and mental retardation. *Prog. Clin. Biol. Res.* **104**, 451-485.

Ciuladaite, Z., Kasnauskiene, J., Cimbalistiene, L., Preiksaitiene, E., Patsalis, P. C. and Kucinskas, V. (2011). Mental retardation and autism associated with recurrent 16p11.2 microdeletion: incomplete penetrance and variable expressivity. *J. Appl. Genet*. **52**, 443-449.

Courchesne, E., Pierce, K., Schumann, C. M., Redcay, E., Buckwalter, J. A., Kennedy, D. P. and Morgan, J. (2007). Mapping early brain development in autism. *Neuron*. **56**, 399-413.

Crepel, A., Steyaert, J., De la Marche, W., De Wolf, V., Fryns, J. P., Noens, I., Devriendt, K. and Peeters, H. (2011). Narrowing the critical deletion region for autism spectrum disorders on 16p11.2. *Am. J. Med. Genet. B Neuropsychiatr. Genet*. **156**, 243-245.

De Rienzo, G., Bishop, J. A., Mao, Y., Pan, L., Ma, T. P., Moens, C. B., Tsai, L. H. and Sive, H. (2011). Disc1 regulates both beta-catenin-mediated and noncanonical Wnt signaling during vertebrate embryogenesis. *FASEB J.* **25**, 4184-4197.

Dong, M., Fu, Y. F., Du, T. T., Jing, C. B., Fu, C. T., Chen, Y., Jin, Y., Deng, M. and Liu, T. X. (2009). Heritable and lineage-specific gene knockdown in zebrafish embryo. *PLoS One*. **4**, e6125.

Durand, C. M., Betancur, C., Boeckers, T. M., Bockmann, J., Chaste, P., Fauchereau, F., Nygren, G., Rastam, M., Gillberg, I. C., Anckarsater, H. et al. (2007). Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nat. Genet.* **39**, 25-27.

Eisen, J. S. and Smith, J. C. (2008). Controlling morpholino experiments: don't stop making antisense. *Development*. **135**, 1735-1743.

Esposito, G., Vitagliano, L., Costanzo, P., Borrelli, L., Barone, R., Pavone, L., Izzo, P., Zagari, A. and Salvatore, F. (2004). Human aldolase A natural mutants: relationship between flexibility of the C-terminal region and enzyme function. *Biochem. J.* **380**, 51-56.

Fanciulli, M., Petretto, E. and Aitman, T. J. (2010). Gene copy number variation and common human disease. *Clin. Genet.* **77**, 201-213.

Foger, N., Jenckel, A., Orinska, Z., Lee, K. H., Chan, A. C. and Bulfone-Paus, S. (2011). Differential regulation of mast cell degranulation versus cytokine secretion by the actin regulatory proteins Coronin1a and Coronin1b. *J. Exp. Med.* **208**, 1777-1787.

Foger, N., Rangell, L., Danilenko, D. M. and Chan, A. C. (2006). Requirement for coronin 1 in T lymphocyte trafficking and cellular homeostasis. *Science*. **313**, 839-842.

Gato, A. and Desmond, M. E. (2009). Why the embryo still matters: CSF and the neuroepithelium as interdependent regulators of embryonic brain growth, morphogenesis and histiogenesis. *Dev. Biol.* **327**, 263-272.

Ghebranious, N., Giampietro, P. F., Wesbrook, F. P. and Rezkalla, S. H. (2007). A novel microdeletion at 16p11.2 harbors candidate genes for aortic valve development, seizure disorder, and mild mental retardation. *Am. J. Med. Genet. A*. **143A**, 1462-1471.

Giulivi, C., Zhang, Y. F., Omanska-Klusek, A., Ross-Inta, C., Wong, S., Hertz-Picciotto, I., Tassone, F. and Pessah, I. N. (2010). Mitochondrial dysfunction in autism. *JAMA*. **304**, 2389-2396.

Gutzman, J. H., Graeden, E. G., Lowery, L. A., Holley, H. S. and Sive, H. (2008). Formation of the zebrafish midbrain-hindbrain boundary constriction requires laminin-dependent basal constriction. *Mech. Dev.* **125**, 974-983.

Gutzman, J. H. and Sive, H. (2009). Zebrafish brain ventricle injection. *J. Vis. Exp*.

Gutzman, J. H. and Sive, H. (2010). Epithelial relaxation mediated by the myosin phosphatase regulator Mypt1 is required for brain ventricle lumen expansion and hindbrain morphogenesis. *Development*. **137**, 795-804.

Haffter, P., Granato, M., Brand, M., Mullins, M. C., Hammerschmidt, M., Kane, D. A., Odenthal, J., van Eeden, F. J., Jiang, Y. J., Heisenberg, C. P. et al. (1996). The identification of genes with unique and essential functions in the development of the zebrafish, Danio rerio. *Development*. **123**, 1-36.

Hashimoto, T., Tayama, M., Miyazaki, M., Murakawa, K., Sakurama, N., Yoshimoto, T. and Kuroda, Y. (1991). Reduced midbrain and pons size in children with autism. *Tokushima J. Exp. Med.* **38**, 15-18.

Horev, G., Ellegood, J., Lerch, J. P., Son, Y. E., Muthuswamy, L., Vogel, H., Krieger, A. M., Buja, A., Henkelman, R. M., Wigler, M. et al. (2011). Dosage-dependent phenotypes in models of 16p11.2 lesions found in autism. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 17076-17081.

Issa, F. A., O'Brien, G., Kettunen, P., Sagasti, A., Glanzman, D. L. and Papazian, D. M. (2011). Neural circuit activity in freely behaving zebrafish (Danio rerio). *J. Exp. Biol.* **214**, 1028-1038.

Jacquemont, S. Reymond, A. Zufferey, F. Harewood, L. Walters, R. G. Kutalik, Z. Martinet, D. Shen, Y. Valsesia, A. Beckmann, N. D. et al. (2011). Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. *Nature*. **478**, 97-102.

Jiang, Y. J., Brand, M., Heisenberg, C. P., Beuchle, D., Furutani-Seiki, M., Kelsh, R. N., Warga, R. M., Granato, M., Haffter, P., Hammerschmidt, M. et al. (1996). Mutations affecting neurogenesis and brain morphology in the zebrafish, Danio rerio. *Development*. **123**, 205-216.

Kim, J. H., Lee, S., Kim, J. H., Lee, T. G., Hirata, M., Suh, P. G. and Ryu, S. H. (2002). Phospholipase D2 directly interacts with aldolase via Its PH domain. *Biochemistry (Mosc)*. **41**, 3414-3421.

Kirov, G., Pocklington, A. J., Holmans, P., Ivanov, D., Ikeda, M., Ruderfer, D., Moran, J., Chambert, K., Toncheva, D., Georgieva, L. et al. (2012). De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. *Mol. Psychiatry*. **17**, 142-153.

Kishi, H., Mukai, T., Hirono, A., Fujii, H., Miwa, S. and Hori, K. (1987). Human aldolase A deficiency associated with a hemolytic anemia: thermolabile aldolase due to a single base mutation. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 8623-8627.

Konopka, G., Wexler, E., Rosen, E., Mukamel, Z., Osborn, G. E., Chen, L., Lu, D., Gao, F., Gao, K., Lowe, J. K. et al. (2012). Modeling the functional genomics of autism using human neurons. *Mol. Psychiatry*. **17**, 202-214.

Krens, S. F., He, S., Lamers, G. E., Meijer, A. H., Bakkers, J., Schmidt, T., Spaink, H. P. and Snaar-Jagalska, B. E. (2008). Distinct functions for ERK1 and ERK2 in cell migration processes during zebrafish gastrulation. *Dev. Biol.* **319**, 370-383.

Kreuder, J., Borkhardt, A., Repp, R., Pekrun, A., Gottsche, B., Gottschalk, U., Reichmann, H., Schachenmayr, W., Schlegel, K. and Lampert, F. (1996). Brief report: inherited metabolic myopathy and hemolysis due to a mutation in aldolase A. *N. Engl. J. Med.* **334**, 1100-1104.

Kumar, R. A., KaraMohamed, S., Sudi, J., Conrad, D. F., Brune, C., Badner, J. A., Gilliam, T. C., Nowak, N. J., Cook, E. H., Jr., Dobyns, W. B. et al. (2008). Recurrent 16p11.2 microdeletions in autism. *Hum. Mol. Genet.* **17**, 628-638.

Kusakabe, T., Motoki, K. and Hori, K. (1997). Mode of interactions of human aldolase isozymes with cytoskeletons. *Arch. Biochem. Biophys.* **344**, 184-193.

Kyoizumi, S., Ohara, T., Kusunoki, Y., Hayashi, T., Koyama, K. and Tsuyama, N. (2004). Expression characteristics and stimulatory functions of CD43 in human CD4+ memory T cells: analysis using a monoclonal antibody to CD43 that has a novel lineage specificity. *J. Immunol.* **172**, 7246-7253.

Levy, D., Ronemus, M., Yamrom, B., Lee, Y. H., Leotta, A., Kendall, J., Marks, S., Lakshmi, B., Pai, D., Ye, K. et al. (2011). Rare de novo and transmitted copy-number variation in autistic spectrum disorders. *Neuron*. **70**, 886-897.

Lionel, A. C., Crosbie, J., Barbosa, N., Goodale, T., Thiruvahindrapuram, B., Rickaby, J., Gazzellone, M., Carson, A. R., Howe, J. L., Wang, Z. et al. (2011). Rare copy number variation discovery and cross-disorder comparisons identify risk genes for ADHD. *Sci. Transl. Med*. **3**, 95ra75.

Liu, Y. C., Bailey, I. and Hale, M. E. (2012). Alternative startle motor patterns and behaviors in the larval zebrafish (Danio rerio). *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol*. **198**, 11-24.

Lowery, L. A. and Sive, H. (2009). Totally tubular: the mystery behind function and origin of the brain ventricular system. *Bioessays*. **31**, 446-458.

Marin-Padilla, M. (1975). Abnormal neuronal differentiation (functional maturation) in mental retardation. *Birth Defects Orig. Artic. Ser.* **11**, 133-153.

Marmigere, F. and Ernfors, P. (2007). Specification and connectivity of neuronal subtypes in the sensory lineage. *Nat. Rev. Neurosci*. **8**, 114-127.

Marshall, C. R., Noor, A., Vincent, J. B., Lionel, A. C., Feuk, L., Skaug, J., Shago, M., Moessner, R., Pinto, D., Ren, Y. et al. (2008). Structural variation of chromosomes in autism spectrum disorder. *Am. J. Hum. Genet*. **82**, 477-488.

Matson, M. L., Matson, J. L. and Beighley, J. S. (2011). Comorbidity of physical and motor problems in children with autism. *Res. Dev. Disabil.* **32**, 2304-2308.

McCarthy, S. E., Makarov, V., Kirov, G., Addington, A. M., McClellan, J., Yoon, S., Perkins, D. O., Dickel, D. E., Kusenda, M., Krastoshevsky, O. et al. (2009). Microduplications of 16p11.2 are associated with schizophrenia. *Nat. Genet.* **41**, 1223-1227.

Min, B. J., Kim, N., Chung, T., Kim, O. H., Nishimura, G., Chung, C. Y., Song, H. R., Kim, H. W., Lee, H. R., Kim, J. et al. (2011). Whole-exome sequencing identifies mutations of KIF22 in spondyloepimetaphyseal dysplasia with joint laxity, leptodactylic type. *Am. J. Hum. Genet*. **89**, 760-766.

Miwa, S., Fujii, H., Tani, K., Takahashi, K., Takegawa, S., Fujinami, N., Sakurai, M., Kubo, M., Tanimoto, Y., Kato, T. et al. (1981). Two cases of red cell aldolase deficiency associated with hereditary hemolytic anemia in a Japanese family. *Am. J. Hematol.* **11**, 425-437.

Miyazaki, T., Hashimoto, K., Uda, A., Sakagami, H., Nakamura, Y., Saito, S. Y., Nishi, M., Kume, H., Tohgo, A., Kaneko, I. et al. (2006). Disturbance of cerebellar synaptic maturation in mutant mice lacking BSRPs, a novel brain-specific receptor-like protein family. *FEBS Lett.* **580**, 4057-4064.

Mossink, M. H., van Zon, A., Franzel-Luiten, E., Schoester, M., Kickhoefer, V. A., Scheffer, G. L., Scheper, R. J., Sonneveld, P. and Wiemer, E. A. (2002). Disruption of the murine major vault protein (MVP/LRP) gene does not induce hypersensitivity to cytostatics. *Cancer Res.* **62**, 7298-7304.

Mueller, P., Liu, X. and Pieters, J. (2011). Migration and homeostasis of naive T cells depends on coronin 1-mediated prosurvival signals and not on coronin 1-dependent filamentous actin modulation. *J. Immunol.* **186**, 4039-4050.

Murphy, T. R., Vihtelic, T. S., Ile, K. E., Watson, C. T., Willer, G. B., Gregg, R. G., Bankaitis, V. A. and Hyde, D. R. (2011). Phosphatidylinositol synthase is required for lens structural integrity and photoreceptor cell survival in the zebrafish eye. *Exp. Eye Res.* **93**, 460-474.

Naganawa, Y. and Hirata, H. (2011). Developmental transition of touch response from slow muscle-mediated coilings to fast muscle-mediated burst swimming in zebrafish. *Dev. Biol.* **355**, 194-204.

Nikaido, M., Kawakami, A., Sawada, A., Furutani-Seiki, M., Takeda, H. and Araki, K. (2002). Tbx24, encoding a T-box protein, is mutated in the zebrafish somite-segmentation mutant fused somites. *Nat. Genet.* **31**, 195-199.

Nord, A. S., Lee, M., King, M. C. and Walsh, T. (2011). Accurate and exact CNV identification from targeted high-throughput sequence data. *BMC Genomics*. **12**, 184.

Obholzer, N., Wolfson, S., Trapani, J. G., Mo, W., Nechiporuk, A., Busch-Nentwich, E., Seiler, C., Sidi, S., Sollner, C., Duncan, R. N. et al. (2008). Vesicular glutamate transporter 3 is required for synaptic transmission in zebrafish hair cells. *J. Neurosci.* **28**, 2110-2118.

Ohsugi, M., Adachi, K., Horai, R., Kakuta, S., Sudo, K., Kotaki, H., Tokai-Nishizumi, N., Sagara, H., Iwakura, Y. and Yamamoto, T. (2008). Kid-mediated chromosome compaction ensures proper nuclear envelope formation. *Cell*. **132**, 771-782.

Okado, N., Narita, M. and Narita, N. (2001). A biogenic amine-synapse mechanism for mental retardation and developmental disabilities. *Brain Dev.* **23 Suppl 1**, S11-15.

Oslejskova, H., Dusek, L., Makovska, Z. and Rektor, I. (2007). Epilepsia, epileptiform abnormalities, non-right-handedness, hypotonia and severe decreased IQ are associated with language impairment in autism. *Epileptic Disord*. **9 Suppl 1**, S9-18.

Outwin, E., Carpenter, G., Bi, W., Withers, M. A., Lupski, J. R. and O'Driscoll, M. (2011). Increased RPA1 gene dosage affects genomic stability potentially contributing to 17p13.3 duplication syndrome. *PLoS Genet*. **7**, e1002247.

Pages, G., Guerin, S., Grall, D., Bonino, F., Smith, A., Anjuere, F., Auberger, P. and Pouyssegur, J. (1999). Defective thymocyte maturation in p44 MAP kinase (Erk 1) knockout mice. *Science*. **286**, 1374-1377.

Ploeger, A., Raijmakers, M. E., van der Maas, H. L. and Galis, F. (2010). The association between autism and errors in early embryogenesis: what is the causal mechanism? *Biol. Psychiatry*. **67**, 602-607.

Postlethwait, J. H., Woods, I. G., Ngo-Hazelett, P., Yan, Y. L., Kelly, P. D., Chu, F., Huang, H., Hill-Force, A. and Talbot, W. S. (2000). Zebrafish comparative genomics and the origins of vertebrate chromosomes. *Genome Res.* **10**, 1890-1902.

Puvabanditsin, S., Nagar, M. S., Joshi, M., Lambert, G., Garrow, E. and Brandsma, E. (2010). Microdeletion of 16p11.2 associated with endocardial fibroelastosis. *Am. J. Med. Genet. A*. **152A**, 2383-2386.

Ricard, G., Molina, J., Chrast, J., Gu, W., Gheldof, N., Pradervand, S., Schutz, F., Young, J. I., Lupski, J. R., Reymond, A. et al. (2010). Phenotypic consequences of copy number variation: insights from Smith-Magenis and Potocki-Lupski syndrome mouse models. *PLoS Biol*. **8**, e1000543.

Ritvo, E. R., Creel, D., Crandall, A. S., Freeman, B. J., Pingree, C., Barr, R. and Realmuto, G. (1986). Retinal pathology in autistic children--a possible biological marker for a subtype? *J. Am. Acad. Child Psychiatry*. **25**, 137.

Robu, M. E., Larson, J. D., Nasevicius, A., Beiraghi, S., Brenner, C., Farber, S. A. and Ekker, S. C. (2007). p53 activation by knockdown technologies. *PLoS Genet*. **3**, e78.

Saint-Amant, L. and Drapeau, P. (1998). Time course of the development of motor behaviors in the zebrafish embryo. *J. Neurobiol.* **37**, 622-632.

Sakaguchi, G., Manabe, T., Kobayashi, K., Orita, S., Sasaki, T., Naito, A., Maeda, M., Igarashi, H., Katsuura, G., Nishioka, H. et al. (1999). Doc2alpha is an activity-dependent modulator of excitatory synaptic transmission. *Eur. J. Neurosci.* **11**, 4262-4268.

Sakai, Y., Shaw, C. A., Dawson, B. C., Dugas, D. V., Al-Mohtaseb, Z., Hill, D. E. and Zoghbi, H. Y. (2011). Protein interactome reveals converging molecular pathways among autism disorders. *Sci. Transl. Med*. **3**, 86ra49.

Sanders, S. J., Ercan-Sencicek, A. G., Hus, V., Luo, R., Murtha, M. T., Moreno-De-Luca, D., Chu, S. H., Moreau, M. P., Gupta, A. R., Thomson, S. A. et al. (2011). Multiple recurrent de novo CNVs,

including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron*. **70**, 863-885.

Santamaria, A., Nagel, S., Sillje, H. H. and Nigg, E. A. (2008). The spindle protein CHICA mediates localization of the chromokinesin Kid to the mitotic spindle. *Curr. Biol.* **18**, 723-729.

Satoh, Y., Kobayashi, Y., Takeuchi, A., Pages, G., Pouyssegur, J. and Kazama, T. (2011). Deletion of ERK1 and ERK2 in the CNS causes cortical abnormalities and neonatal lethality: Erk1 deficiency enhances the impairment of neurogenesis in Erk2-deficient mice. *J. Neurosci.* **31**, 1149-1155.

Schier, A. F., Neuhauss, S. C., Harvey, M., Malicki, J., Solnica-Krezel, L., Stainier, D. Y., Zwartkruis, F., Abdelilah, S., Stemple, D. L., Rangini, Z. et al. (1996). Mutations affecting the development of the embryonic zebrafish brain. *Development*. **123**, 165-178.

Schilling, T. F. and Kimmel, C. B. (1994). Segment and cell type lineage restrictions during pharyngeal arch development in the zebrafish embryo. *Development*. **120**, 483-494.

Schmitt, E. A. and Dowling, J. E. (1994). Early eye morphogenesis in the zebrafish, Brachydanio rerio. *J. Comp. Neurol.* **344**, 532-542.

Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., Yamrom, B., Yoon, S., Krasnitz, A., Kendall, J. et al. (2007). Strong association of de novo copy number mutations with autism. *Science*. **316**, 445-449.

Shimojima, K., Inoue, T., Fujii, Y., Ohno, K. and Yamamoto, T. (2009). A familial 593-kb microdeletion of 16p11.2 associated with mental retardation and hemivertebrae. *Eur. J. Med. Genet*. **52**, 433-435.

Shinawi, M., Liu, P., Kang, S. H., Shen, J., Belmont, J. W., Scott, D. A., Probst, F. J., Craigen, W. J., Graham, B. H., Pursley, A. et al. (2010). Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. *J. Med. Genet.* **47**, 332-341.

Shiow, L. R., Paris, K., Akana, M. C., Cyster, J. G., Sorensen, R. U. and Puck, J. M. (2009). Severe combined immunodeficiency (SCID) and attention deficit hyperactivity disorder (ADHD) associated with a Coronin-1A mutation and a chromosome 16p11.2 deletion. *Clin. Immunol.* **131**, 24-30.

Shkumatava, A., Stark, A., Sive, H. and Bartel, D. P. (2009). Coherent but overlapping expression of microRNAs and their targets during vertebrate development. *Genes Dev.* **23**, 466-481.

Shui, J. W., Hu, M. C. and Tan, T. H. (2007). Conditional knockout mice reveal an essential role of protein phosphatase 4 in thymocyte development and pre-T-cell receptor signaling. *Mol. Cell. Biol.* **27**, 79-91.

Sive, H. (2011). 'Model' or 'tool'? New definitions for translational research. *Dis. Model Mech*. **4**, 137-138.

Skarnes, W. C., Rosen, B., West, A. P., Koutsourakis, M., Bushell, W., Iyer, V., Mujica, A. O., Thomas, M., Harrow, J., Cox, T. et al. (2011). A conditional knockout resource for the genome-wide study of mouse gene function. *Nature*. **474**, 337-342.

Talkowski, M. E., Mullegama, S. V., Rosenfeld, J. A., van Bon, B. W., Shen, Y., Repnikova, E. A., Gastier-Foster, J., Thrush, D. L., Kathiresan, S., Ruderfer, D. M. et al. (2011). Assessment of 2q23.1 microdeletion syndrome implicates MBD5 as a single causal locus of intellectual disability, epilepsy, and autism spectrum disorder. *Am. J. Hum. Genet*. **89**, 551-563.

Tannour-Louet, M., Han, S., Corbett, S. T., Louet, J. F., Yatsenko, S., Meyers, L., Shaw, C. A., Kang, S. H., Cheung, S. W. and Lamb, D. J. (2010). Identification of de novo copy number variants associated with human disorders of sexual development. *PLoS One*. **5**, e15392.

Taylor, J. S., Braasch, I., Frickey, T., Meyer, A. and Van de Peer, Y. (2003). Genome duplication, a trait shared by 22000 species of ray-finned fish. *Genome Res.* **13**, 382-390.

The Simons Vip, C. (2012). Simons Variation in Individuals Project (Simons VIP): A Genetics-First Approach to Studying Autism Spectrum and Related Neurodevelopmental Disorders. *Neuron*. **73**, 1063-1067.

Thermes, V., Grabher, C., Ristoratore, F., Bourrat, F., Choulika, A., Wittbrodt, J. and Joly, J. S. (2002). I-SceI meganuclease mediates highly efficient transgenesis in fish. *Mech. Dev.* **118**, 91-98.

Thisse, C. and Thisse, B. (2005). High Throughput Expression Analysis of ZF-Models Consortium Clones. *ZFIN Direct Data Submission (*http://zfin.org).

Trembath, R. C. (1994). Genetic mechanisms and mental retardation. *J. R. Coll. Physicians Lond.* **28**, 121-125.

Ulitsky, I., Shkumatava, A., Jan, C. H., Sive, H. and Bartel, D. P. (2011). Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell*. **147**, 1537-1550.

Vacic, V., McCarthy, S., Malhotra, D., Murray, F., Chou, H. H., Peoples, A., Makarov, V., Yoon, S., Bhandari, A., Corominas, R. et al. (2011). Duplications of the neuropeptide receptor gene VIPR2 confer significant risk for schizophrenia. *Nature*. **471**, 499-503.

van Eeden, F. J., Granato, M., Schach, U., Brand, M., Furutani-Seiki, M., Haffter, P., Hammerschmidt, M., Heisenberg, C. P., Jiang, Y. J., Kane, D. A. et al. (1996). Mutations affecting somite formation and patterning in the zebrafish, Danio rerio. *Development*. **123**, 153-164.

Weiss, L. A., Shen, Y., Korn, J. M., Arking, D. E., Miller, D. T., Fossdal, R., Saemundsen, E., Stefansson, H., Ferreira, M. A., Green, T. et al. (2008). Association between microdeletion and microduplication at 16p11.2 and autism. *N. Engl. J. Med.* **358**, 667-675.

Wiellette, E. L. and Sive, H. (2003). vhnf1 and Fgf signals synergize to specify rhombomere identity in the zebrafish hindbrain. *Development*. **130**, 3821-3829.

Yamamoto, K. and Vernier, P. (2011). The evolution of dopamine systems in chordates. *Front Neuroanat*. **5**, 21.

Yao, D. C., Tolan, D. R., Murray, M. F., Harris, D. J., Darras, B. T., Geva, A. and Neufeld, E. J. (2004). Hemolytic anemia and severe rhabdomyolysis caused by compound heterozygous mutations of

the gene for erythrocyte/muscle isozyme of aldolase, ALDOA(Arg303X/Cys338Tyr). *Blood*. **103**, 2401-2403.

Yao, J., Gaffaney, J. D., Kwon, S. E. and Chapman, E. R. (2011). Doc2 is a Ca2+ sensor required for asynchronous neurotransmitter release. *Cell*. **147**, 666-677.

FIGURE LEGENDS

Figure 1. Strategy and isolation of zebrafish (*Danio rerio*) 16p11.2 homologs. (A) Strategy to use zebrafish as a tool to analyze 16p11.2 gene activity. (B) Homologous human and zebrafish genes. Each homologous pair is connected by a red line. Genes are shown in relative chromosomal positions (Table S1). The *mapk3, gdpd3,* and *ypel3* loci are syntenic, whereas *kctd13, sez6l2,* and *asphd1* genes are grouped, but order on the human chromosome is different. The cluster *tbx24, ppp4cb,* 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.

Figure 2. Loss of function (LOF) embryos have abnormal brain and tail morphology. (A) Embryonic phenotypes were observed at 24 hpf, after LOF effected by injection of antisense morpholino oligonucleotides (MO) (Table S2) at the one to two cell stage (Table S3). Genes assayed (Table S1) are indicated above each set of images. "Control" embryos were injected with control MO (Methods). Brain ventricles are injected with Texas Red dextran, and brightfield and fluorescence images superimposed. Images are representative of phenotypes observed in at least 70% of embryos, over 2-7 independent experiments, with 50-350 embryos total assayed per gene (Table S3). (a-w): dorsal views; (a'-w') lateral close-up; (a"-w") full embryo lateral view. (a-a") schematics of embryo landmarks. F, forebrain ventricle; M, midbrain ventricle; H, hindbrain ventricle. (b-w) and (b'-w') anterior to the left, and images are shown at equivalent magnification. (B) Phenotypic group where LOF embryos have narrow midbrain and hindbrain ventricles. 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 where LOF embryos have a straight midbrain. Gene assayed is indicated above each panel. Dorsal views, anterior to the left. FB, forebrain; MHB, midbrain-hindbrain boundary; asterisk,

midbrain hingepoint; scale bar = 150 μm. Embryo images in (B) and (C) are the dorsal view panels of (A).

Figure 3. Phenotypes obtained after LOF of zebrafish 16p11.2 homologs. Embryonic phenotypes observed after LOF (Fig. 2) are cataloged and structures or assays are indicated at the top of each column. All assays were done in comparison to embryos injected with control MO. Quantification of phenotypes is given in Table S3 and rescue conditions are shown in Fig. 4 and Table 3.

[a] Genes assayed are indicated in the left column and gene identifiers are included in Table S1. MOs targeting these genes are included in Table S2.

[b] Morphological analyses addressed head morphology (brain ventricle shape and eye formation), tail shape and length, and muscle segment shape (chevron vs. U-shape).

[c] Two types of movement were tested: spontaneous movement at 24 hpf, and touch response at 48 hpf. The *ino80e* LOF embryos respond with one flip of the tail or not at all. *mapk3* LOF embryos have a jerky response and the *ypel3* LOF embryos move in small, jerky circles. *mvp* LOF embryos range from not responding to spinning in response to touch. Otherwise, touch response was weak, sluggish, or absent.

[d] Initial axon tracts form by 36 hpf and were assayed by immunostaining for acetylated tubulin. Axon tracts were not assayed in *mvp* LOF due to lack of rescue, in *prrt2* LOF due to lack of observable phenotype at these time points, and in *tbx24* LOF due to lack of head expression.

[e] Pigmentation was observed at 48 hpf in LOF embryos and images are included in Fig. S3.

[f] Persistence of early phenotypic abnormalities was monitored up to 5 dpf.

[1] The *mvp* MO used for the phenotype reported is indicated with matching superscript and listed first in Tables 1 and S2.

Red boxes: abnormal phenotype in >70% embryos; speckled red boxes: abnormal phenotype in >70% of embryos, but the phenotype was mild. ND, not determined.

Figure 4. Human orthologs rescue zebrafish LOF embryos. Single cell embryos were injected with MO, either alone or together with human or fish mRNA, and imaged at 24 hpf after injecting Texas Red dextran into the brain ventricles. The human homolog co-injection with a zebrafish LOF whose phenotype is thusly significantly restored to normal, indicates functional equivalence with the zebrafish gene (orthology). *kctd13* ((L^0-L"') and *maz* LOF (O^0-O"') were rescued by zebrafish but not human RNA. (B^0-T^0) dorsal views and (B"-T") lateral views of LOF embryos. (B'-T') dorsal views and (B"'-T"') lateral views of LOF embryos plus rescue mRNA. Human RNAs co-injected for rescue are indicated by upper case letters, except for *z.kctd13* and *z.maz*, which refer to RNA from the zebrafish genes. The rescue experiments shown in D, E, and G are from the experiment shown in Fig. 2, in other cases, the images are taken from different experiments, with data consistent to that shown in Fig. 2. Images are representative of 2-4 independent experiments per gene, with 40-194 embryos total assayed per gene. Rescue of the abnormal LOF phenotypes was achieved in ~50% or more of the embryos. Representative images are shown here and quantification is included in Table 3.

Figure 5. Axon tracts are abnormal in LOF embryos. Forebrain and hindbrain axon tracts after LOF. Axons were labeled with acetylated α-tubulin antibody, and imaged by scanning confocal microscopy of fixed, flatmounted 36 hpf LOF embryos. (A-M) lateral view, showing forebrain axons. (A'-M') dorsal view, showing hindbrain axons. Genes targeted for LOF by MO injection are indicated above each set of panels in lowercase. Over two independent experiments, an average of 8 embryos per gene was imaged for effects of LOF and rescue. The percentage of affected embryos was 80% or greater. Hindbrain and forebrain tracts were affected in all LOF conditions, except *kctd13,* which only showed defects in the hindbrain and *kif22,* which only showed defects in the forebrain. The rescues with cognate RNA led to rescue in 75-100% of embryos assayed. Human RNAs co-injected for rescue are indicated by upper case letters, except for *z.kctd13*, which refers to RNA from the zebrafish gene. "Control" embryos were injected with control MO (Methods). ac, anterior commissures; sot, supra optic tract; poc, post optic

commissure; tpoc, tract of the post optic commissure; tpc, tract of the posterior commissure; r2, r4, r6, rhombomeres 2, 4, and 6; asterisk, reduced or disorganized ac; arrowhead, reduced or disorganized sot; arrow, reduced tpc; open arrowhead, reduced or disorganized tpoc; dotted arrow, reduced poc.

Figure 6. Identification of deletion dosage sensor genes. (A) Strategy for identification of deletion dosage sensor genes. MOs designed against splice sites are titrated to find the lowest amount resulting in a phenotype, with at least 70% penetrance, and normal RNA remaining at this MO concentration is determined. A "deletion dosage sensor" is defined as a gene where a phenotype is observed when ~50% of the normal mRNA remains. (B) Strategy to quantify normal mRNA remaining in LOF embryos. An antisense MO is designed to an intron/exon boundary, and typically results in exon exclusion or intron inclusion (Table 2). qPCR primers are designed to detect the normally processed mRNA, where one primer in each set hybridizes to the normal but not the abnormally processed transcript. For, forward primer; Rev, reverse primer; Ex, exon; Intr, Intron. (C) Percentage of normal mRNA remaining in 24 hpf LOF embryos. RNA levels were quantified by qPCR, normalized to *ef1α* and expressed relative to levels of experimental RNA in control MO-injected embryos. LOF was performed at two MO concentrations, one that did not give a phenotype ("Low MO") and one that did ("High MO"). Genes assayed are indicated below relevant histograms. (D) Quantification of normal *aldoaa* and *kif22* mRNA after LOF in 18, 24 and 36 hpf embryos, at the same MO concentration used in (C). qPCR was performed and RNA levels normalized to *ef1a* and expressed relative to levels of experimental RNA in control MO-injected embryos. (E) Western blots of 24 hpf LOF embryos, same MO concentration used in (C), representative image of 3 experiments. Protein was extracted from embryos injected with *aldoaa* or *kif22* MOs. After LOF, 56% of Aldoa (from head-dissected protein, thus Aldoaa-enriched, see Methods) and 65% of Kif22 (whole embryo) protein remains when normalized to control proteins (Eif4e and Gapdh, respectively) compared to control MO-injected embryos. (F) Muscle segments in 24 hpf LOF embryos. Actin is stained with phalloidin, and muscle shape is indicated by white dotted lines. Over two experiments, chevron

shape was abnormal in 0% of control MO-injected embryos, 100% *aldoaa* LOF embryos, and 100% *kif22* LOF embryos (n=10 for each condition). (G) GFP expression in the *NeuroD:GFP* line. 0% (n=106) control embryos (injected with control MO); 94% (n=97) *aldoaa* LOF embryos; 100% (n=100) *kif22* LOF embryos affected as observed over four independent experiments. Dotted arrow, retina; arrow, tectum; oval, pancreas. (a-a', c-c', e-e') lateral view, anterior to the left, (b-b', d-d', f-f') ventral view.

Figure 7. *aldoaa* 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 I-SceI meganuclease. (B) Relative expression of *aldoaa* mRNA after inhibition by shRNA. Expression was quantified by qPCR in 24 hpf embryos injected with an *aldoaa* hairpin (*aldoah*) shown relative to embryos injected with a control hairpin (*YFPh*). Data was normalized to either (a) actin or (b) GFP expression. (c) Touch response in 24 hpf *aldoah* embryos versus control *YFPh* embryos (81% abnormal, n=26). (C) Phenotype of shRNA-injected embryos. (a-a") 24 hpf control YFPh embryos (16% abnormal, n=51, in 2 independent experiments). (b-b") 24 hpf aldoah embryos (89% abnormal, n=53, in 2 independent experiments). (a, b), lateral view of the head; (a', b'), dorsal view of the head after brain ventricle injection; (a", b"), 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 (a) actin or (b) GFP expression. (E) Phenotype of shRNA-injected embryos. (a-a") 24 hpf *YFPh* embryos (8% abnormal, n=27, in 2 independent experiments). (b-b") 24 hpf *kif22h* embryos (81% abnormal, n=26, in 2 independent experiments). (a, b), lateral view of the head; (a', b'), dorsal view of the head after brain ventricle injection; (a", b"), lateral view of whole embryo.

Figure 1

A

B

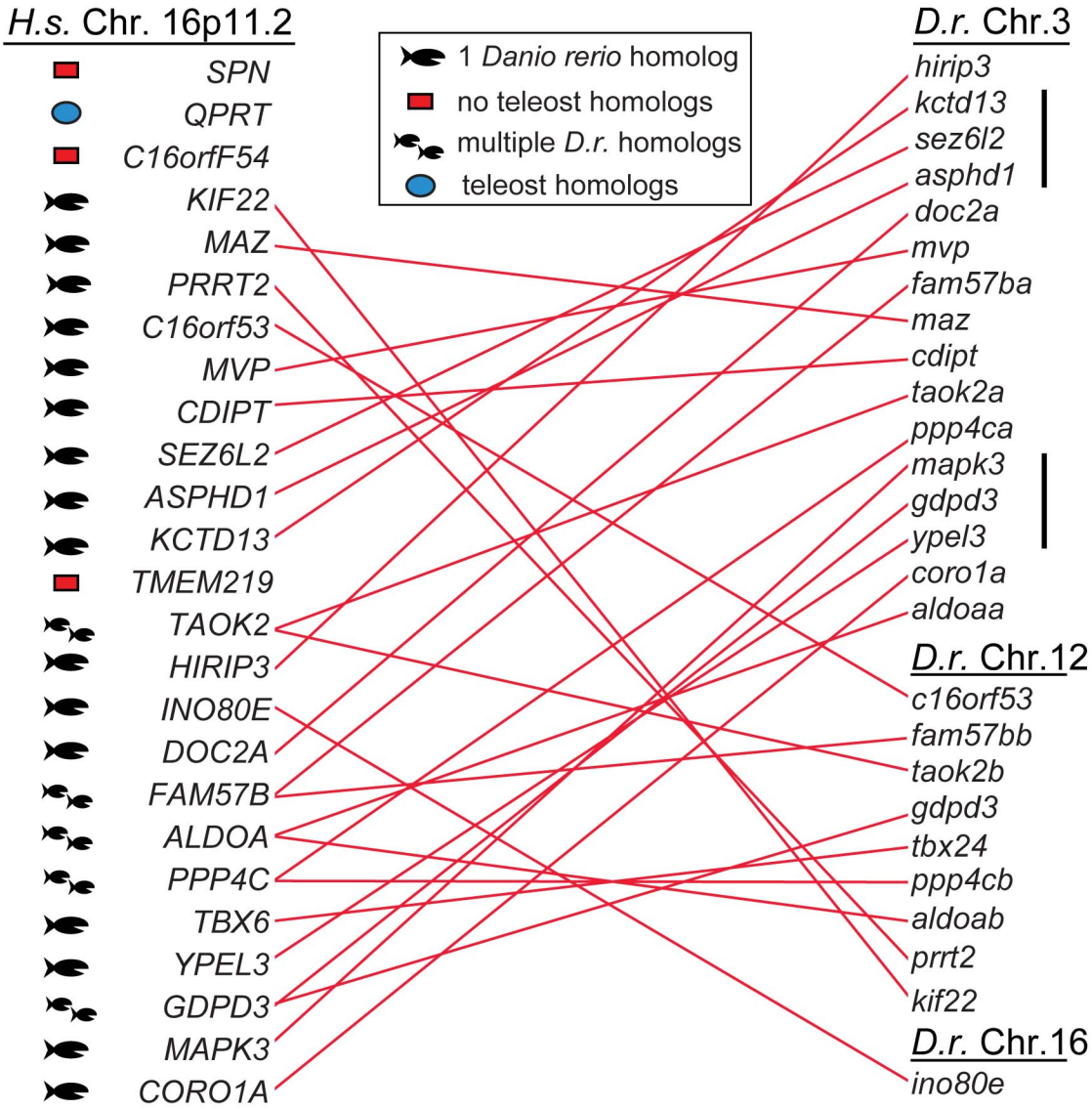

Figure 2

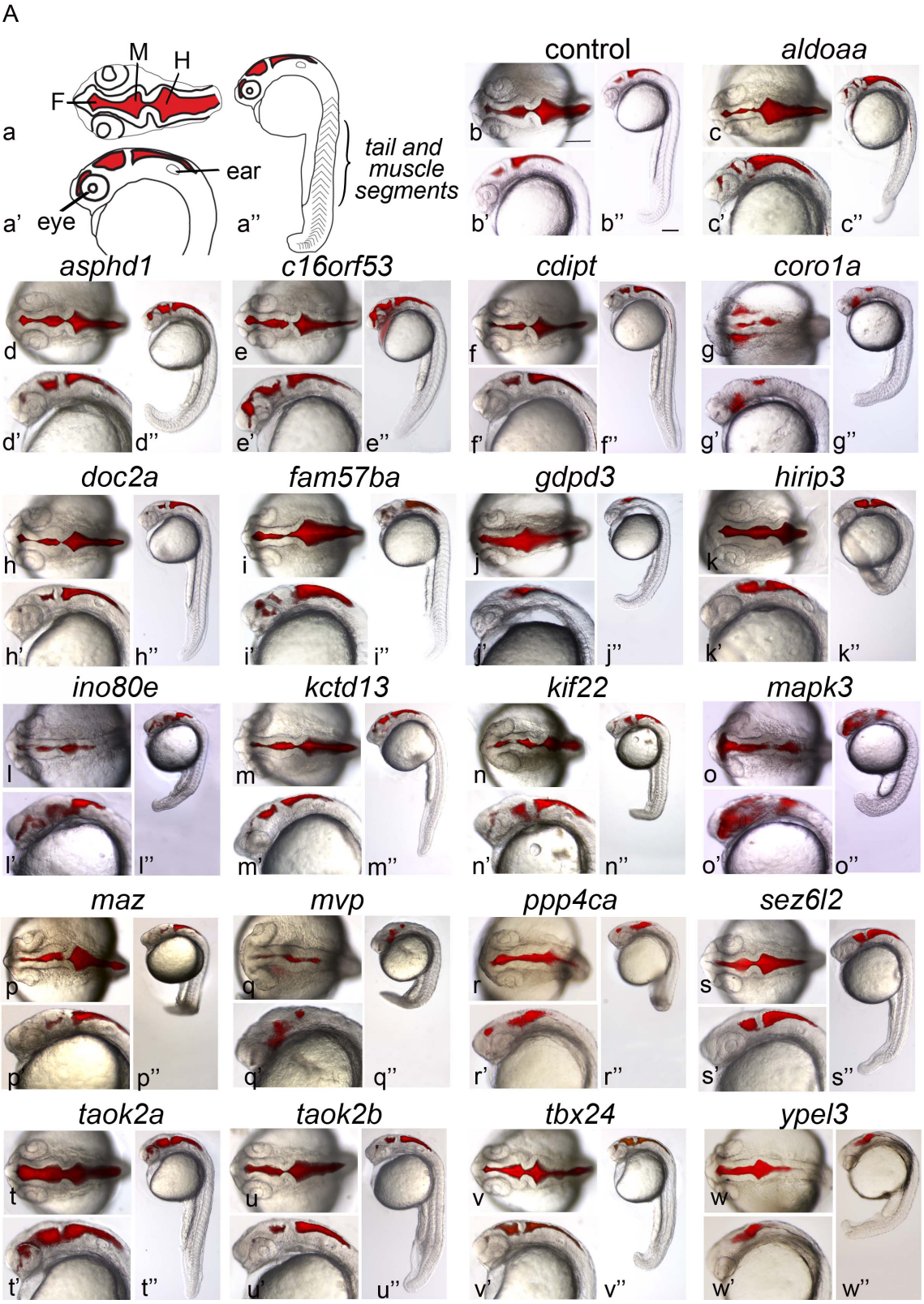

Figure 2 continued

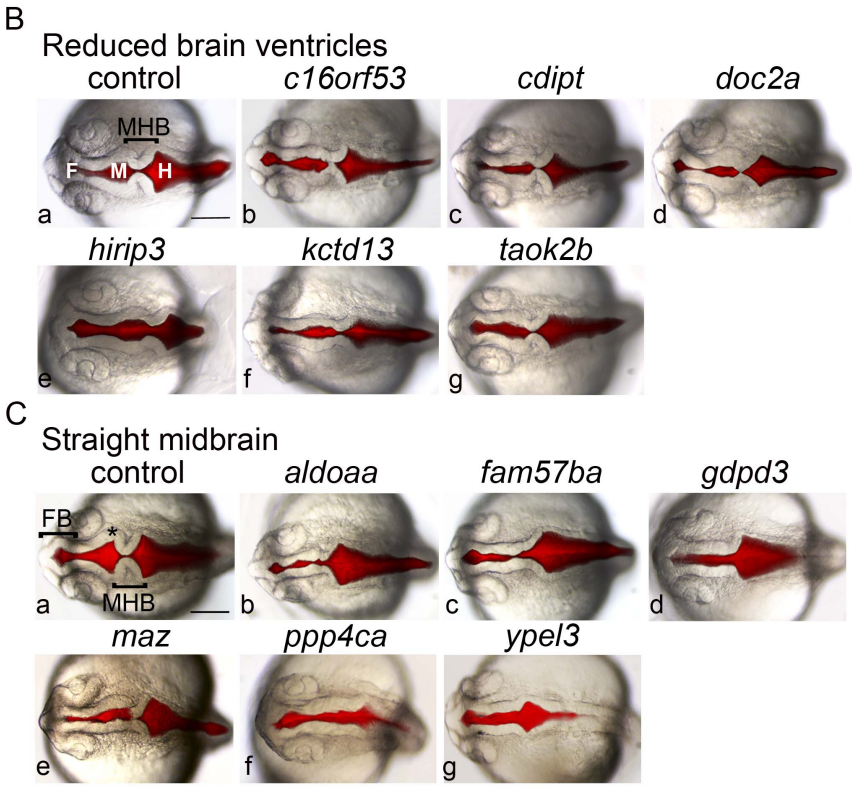

Figure 3

Gene Name[a]	Phenotypes										
	Morphology[b]						Spontaneous Movement[c] (24 hpf)	Axon Tracts[d] (36 hpf)	Touch Response[c] (48 hpf)	Pigmentation[e] (48 hpf)	Abnormal 5 dpf Phenotype[f]
	Forebrain (24 hpf)	Midbrain (24 hpf)	Hindbrain (24 hpf)	Tail (24 hpf)	Muscle Segments (24 hpf)	Eyes (24 hpf)					
aldoaa	■	■	■	■	■				■	■	■
asphd1	▒	▒	▒	■	■						■
c16orf53		■	■	■	■						
cdipt		■	■								■
coro1a	■	■	■	■	■	■	■	■	■	■	■
doc2a		■	▒								
fam57ba	■	■	■		■	■	■	■	■	■	■
gdpd3	■	■	■	■	■		■		■	■	■
hirip3	■	■	■	■	■		■		■	■	■
ino80e	■	■	■	■	■				■	■	■
kctd13	■	■	■		■			■	■		■
kif22	■	■	■	■	■			■	■		■
mapk3	■	■	■	■	■	■		■	■	■	■
maz	■	■	■	■	■	■			■		■
mvp[1]	■	■	■	■	■	■		ND	■		■
ppp4ca	■	■	■	■	■			■	■	■	■
prrt2								ND			
sez6l2	▒	▒	▒	■	■						■
taok2a	■	■	■		■						■
taok2b		■		■	■						■
tbx24				■	■			ND	■		■
ypel3	■	■	■	■	■				■		■

Figure 4

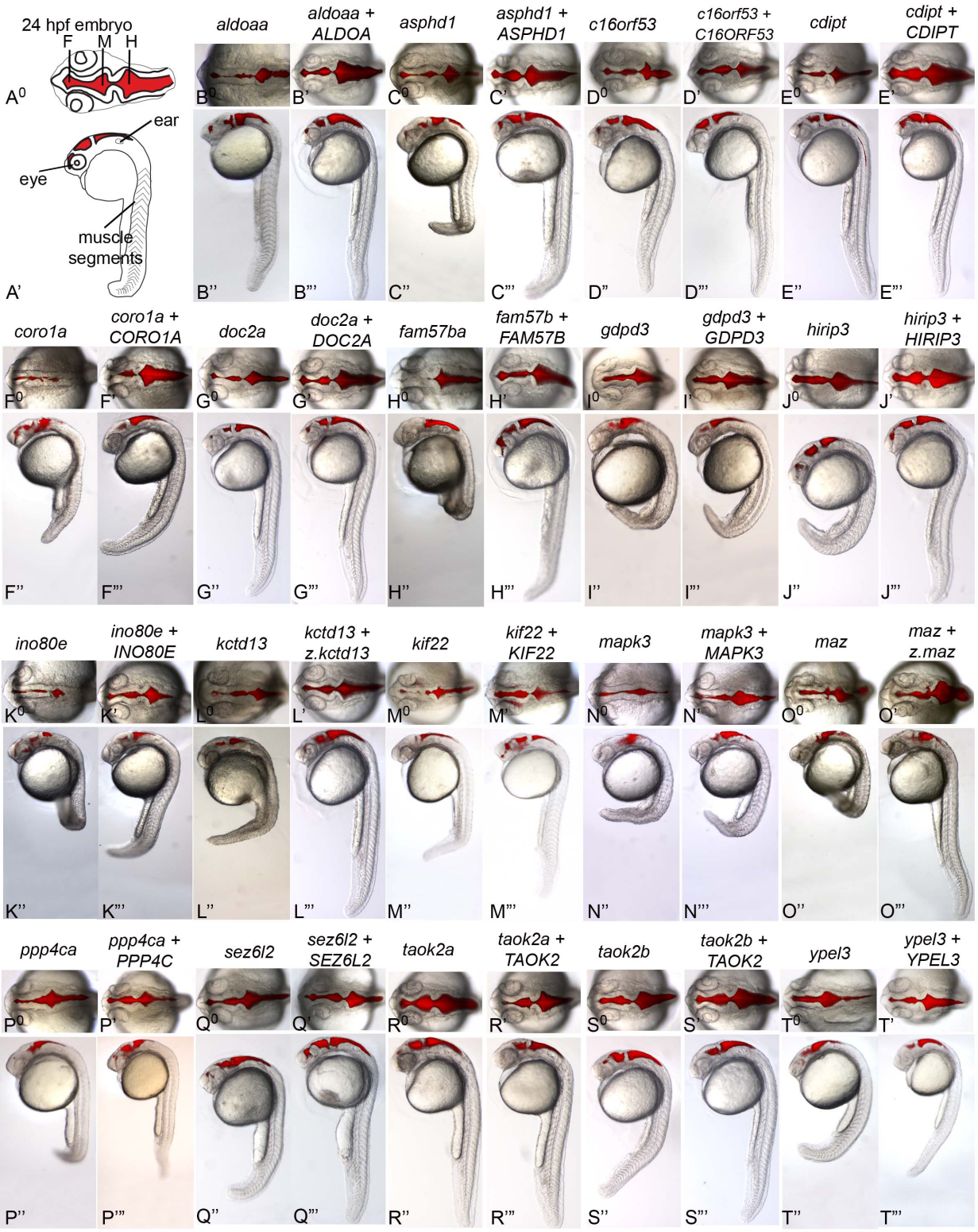

Figure 5

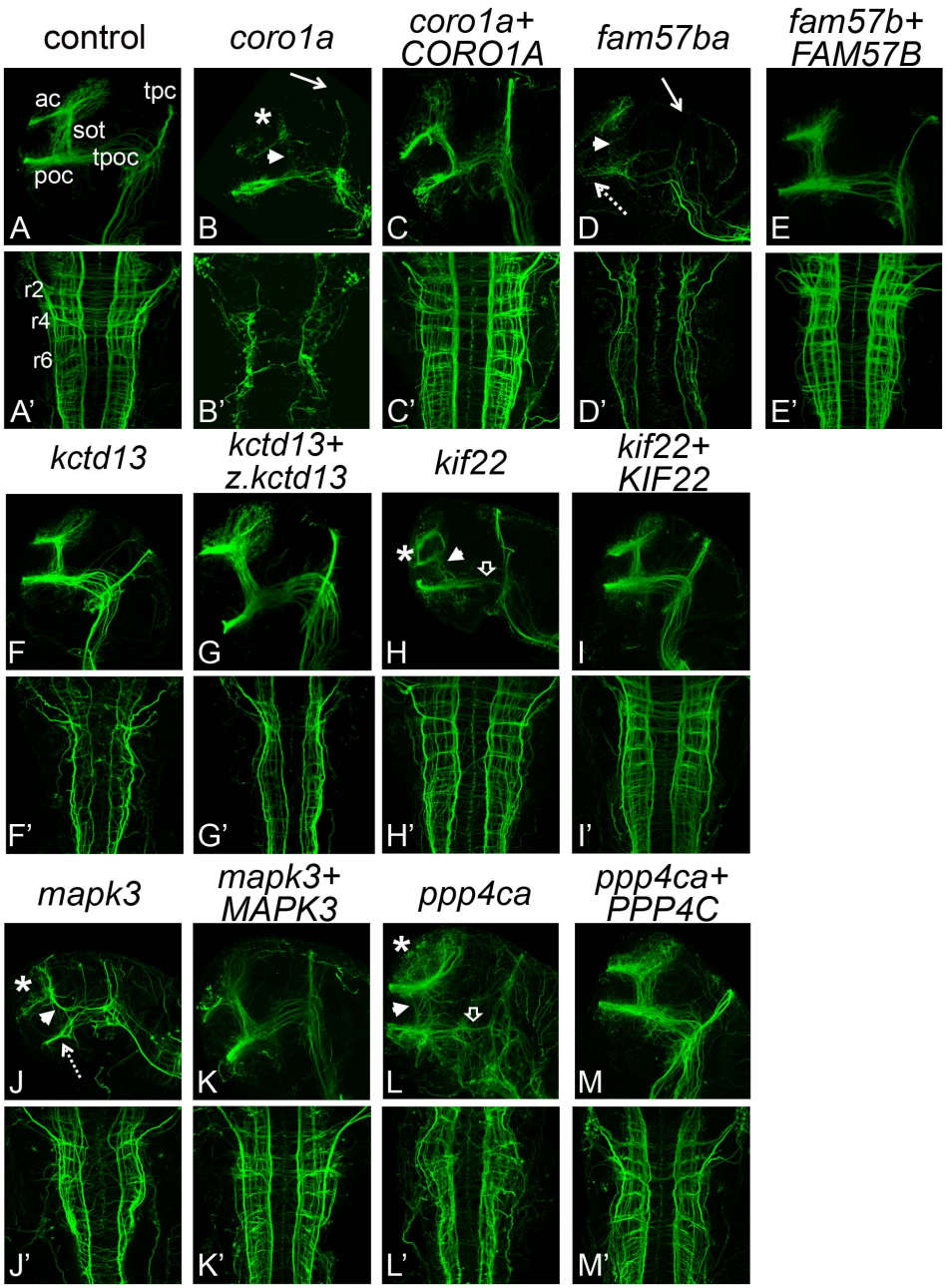

Figure 6

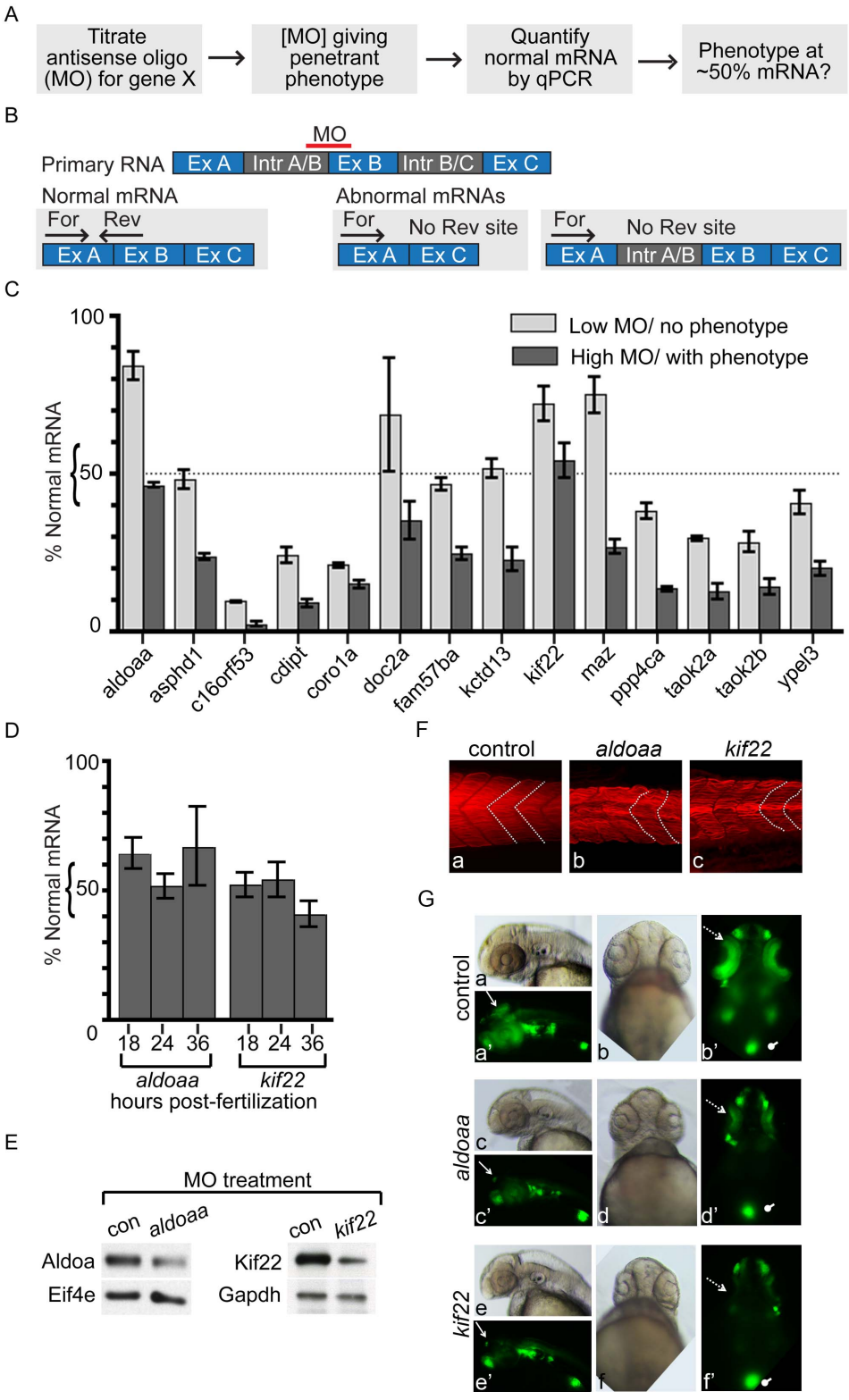

Figure 7

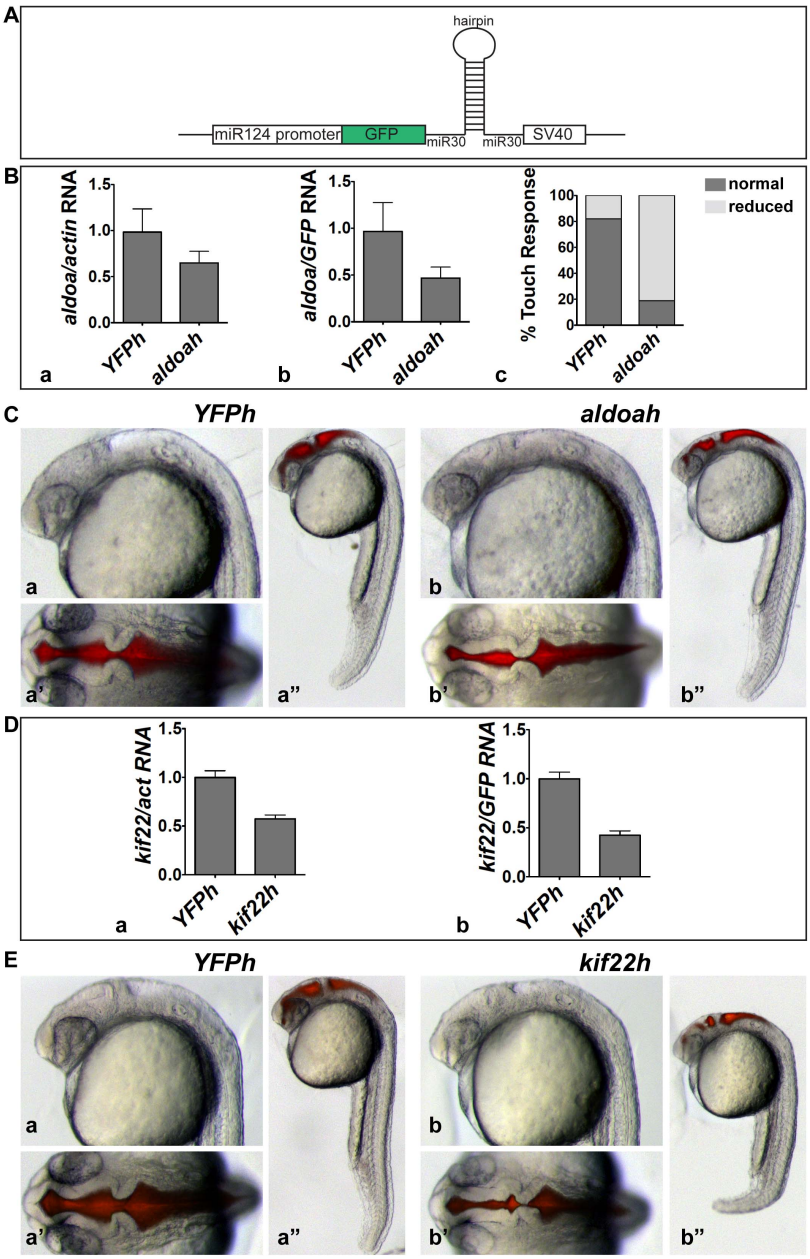

Table 1

Gene Name[a]	Translation Start Site MO[b]	Splice Site MO[c]	Change in RNA Splicing[d]	Resultant Change in Predicted Protein[e]	Phenotype Observed[f]	Phenotype Correlates with RNA Decrease[g]	Rescue by RNA Injection[h]	Additional Knockdown Method[i]
aldoaa		yes	yes	truncation	yes	yes, also protein[j]	yes	shRNA[k]
asphd1		yes	yes	truncation	yes	yes	yes	
c16orf53		yes	yes	truncation	yes	yes	yes	
cdipt		yes	yes	truncation	yes	yes	yes	mutant[l]
coro1a		yes	yes	truncation	yes	yes	yes	
		yes	yes	in frame exon deletion	yes	ND	ND	
doc2a		yes	yes	truncation	yes	yes	yes	
fam57ba		yes	yes	truncation	yes	yes	yes	
gdpd3	yes		NA		yes		yes	
hirip3	yes		NA		yes		yes	
		yes	yes	truncation	no	ND	NA	
ino80e	yes		NA		yes		yes	
		yes	yes	truncation	no	ND	no	
kctd13		yes	yes	truncation	yes	yes	yes	
kif22		yes	yes	truncation	yes	yes, also protein[j]	yes	shRNA[k]
mapk3	yes		NA		yes		yes	
		yes	yes	truncation	no	ND	NA	
maz		yes	yes	truncation	yes	yes	yes	
mvp		yes[1]	yes	truncation	yes	ND	no	
	yes[2]		NA		yes		no	
		yes[3]	ND[m]	NA	yes	ND	no	
ppp4ca		yes	yes	truncation	yes	yes	yes	
		yes	yes	in frame exon deletion	no	NA	NA	
prrt2	yes		NA		no		NA	
		yes	yes	truncation	no	NA	NA	
sez6l2	yes		NA		yes		yes	
		yes	yes	truncation	no	NA	NA	
taok2a		yes	yes	truncation	yes	yes	yes	
taok2b		yes	yes	truncation	yes	yes	yes	
tbx24	yes		NA		yes		published[n]	mutant[l]
ypel3		yes	yes	truncation	yes	yes	yes	

Table 1. Summary of tests for antisense morpholino oligonucleotide specificity.

[a] For gene assignments, see Fig. 1 and Table S1.

[b,c] MOs were designed as indicated in Methods. See Table S2 for MO sequences.

[d] 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 Table S2.

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

[f] Phenotypes were scored as indicated in Methods (See Fig. 2, 3, and Table S3).

[g] 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).

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

[i] Similar phenotypes were observed for additional knockdown methods (Fig. 7, S2).

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

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

[l] Mutant lines were available for *cdipt* and *tbx24* and similar phenotypes were observed (Fig. S2 and Methods).

[m] No splicing change was observed in RNA.

[n] The rescue for the *tbx24* MO was previously published in Nikaido et al., 2002.

[1, 2, 3] *mvp* MOs 1, 2, and 3, correspond to those included in Tables 2, S2, and S3, although no *mvp* LOF conditions were rescued.

ND, not determined; NA, not applicable.

Table 2

Gene Name[a]	MO Mass[b] (ng/embryo)	Results of MO Action on RNA Coding Capacity[c]	Predicted Protein[d] (Truncated/Normal) (Amino Acids)	p53 MO Included[e]
aldoaa	1	exon deletion causing frame shift and early stop codon	202/364	no
asphd1	3.5	intron inclusion containing stop codon	144/354	yes
c16orf53	7.5	intron inclusion containing stop codon	52/273	no
cdipt	7.5	intron inclusion causing frame shift and early stop codon	144/214	no
coro1a	5	intron inclusion causing frame shift and early stop codon	61/455	yes
doc2a	7.5	intron inclusion causing frame shift and early stop codon	184/401	no
fam57ba	1.2	exon deletion causing frame shift and early stop codon	33/202	no
gdpd3	7	translation blocking	0/315	yes
hirip3	6.5	translation blocking	0/510	yes
ino80e	5	translation blocking	0/231	yes
kctd13	7.5	intron inclusion causing frame shift and early stop codon	114/330	no
kif22	5	exon deletion causing frame shift and early stop codon	123/634	yes
mapk3	5	translation blocking	0/392	yes
maz	1.5	intron inclusion causing frame shift and early stop codon	97/493	no
mvp[1]	5	intron inclusion causing frame shift and early stop codon	45/863	yes
mvp[2]	6	translation blocking	0/863	yes
mvp[3]	5	ND	ND	yes
ppp4ca	3.5	intron inclusion causing frame shift and early stop codon	37/311	yes
prrt2	10	intron inclusion causing frame shift and early stop codon	79/226	no
sez6l2	5	translation blocking	0/892	yes
taok2a	5	intron inclusion causing frame shift and early stop codon	65/1138	no
taok2b	5	intron inclusion containing stop codon	188/1182	no
tbx24	4	translation blocking	0/874	yes
ypel3	2	exon deletion causing frame shift and early stop codon	60/119	yes

Table 2. MO Amounts and Effects on Gene Expression.

[a] Genes indicated in Table S1 were targeted in LOF experiments and quantification is included in Tables 3 and S3.

[b] The MOs (Table S2) were injected at the 1-2 cell stage in titration experiments. The mass of MO that was used for phenotype characterization, scoring (Figs. 2, 3 and Table S3) and rescue experiments (Figs. 4, 5 and Table 3) is indicated.

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

[d] 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 may contain abnormal amino acids, as amino acids encoded by an intron inclusion or a frame shift are predicted (in some cases) before an early stop codon.

[e] The p53 MO was used to eliminate off-target effects where indicated.

[1, 2, 3] The *mvp* MOs are included in the same order as Tables 1, S2, and S3.

[1] Indicates the *mvp* MO that was used for the LOF collation in Figs. 2A and 3.

ND, not determined.

Table 3

Gene Name[a]	Rescue mRNA Mass[b] (pg/embryo)	Total Embryos[c]	Rescued Embryos[d]	% Rescued[e]	Number of Independent Experiments[f]
aldoaa	200	52	37	71	3
asphd1	100	110	58	53	2
c16orf53	80	48	29	60	2
cdipt	50	107	93	87	3
coro1a	10	74	46	62	2
doc2a	100	90	84	93	2
fam57ba	320	82	69	84	3
gdpd3	20	70	34	49	2
hirip3	100	95	81	85	2
ino80e	200	56	45	80	2
kctd13[g]	200	80	62	78	2
kif22	15	194	140	72	4
mapk3	5	96	85	89	4
maz[g]	50	102	80	78	2
ppp4ca	30	116	116	100	3
sez6l2	100	43	24	56	2
taok2a	200	85	49	58	2
taok2b	200	77	51	66	2
ypel3	75	40	33	83	2

Table 3. Rescue Assays with Human or Fish Genes after MO LOF. Representative images of embryos from rescue experiments are shown in Fig. 4.

[a] Genes indicated in Table S1 were targeted in LOF experiments by MOs included in Table S2, using MO masses described in Table 2.

[b] Mass 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 Table S2. cDNA constructs included in Table S4 and RNA preparation described in Methods.

[c] Total embryos scored over 2-4 experiments.

[d] Number of embryos whose phenotype was significantly restored to normal.

[e] Percentage of rescued embryos relative to total embryos injected.

[f] 2-4 independent experiments were performed per rescue condition after the preferred titrated dosage was determined.

[g] 2-4 *kctd13* and *maz* LOF phenotypes were rescued using zebrafish mRNA using the indicated masses. All other rescue mRNA masses refer to human mRNA.

Figure S1

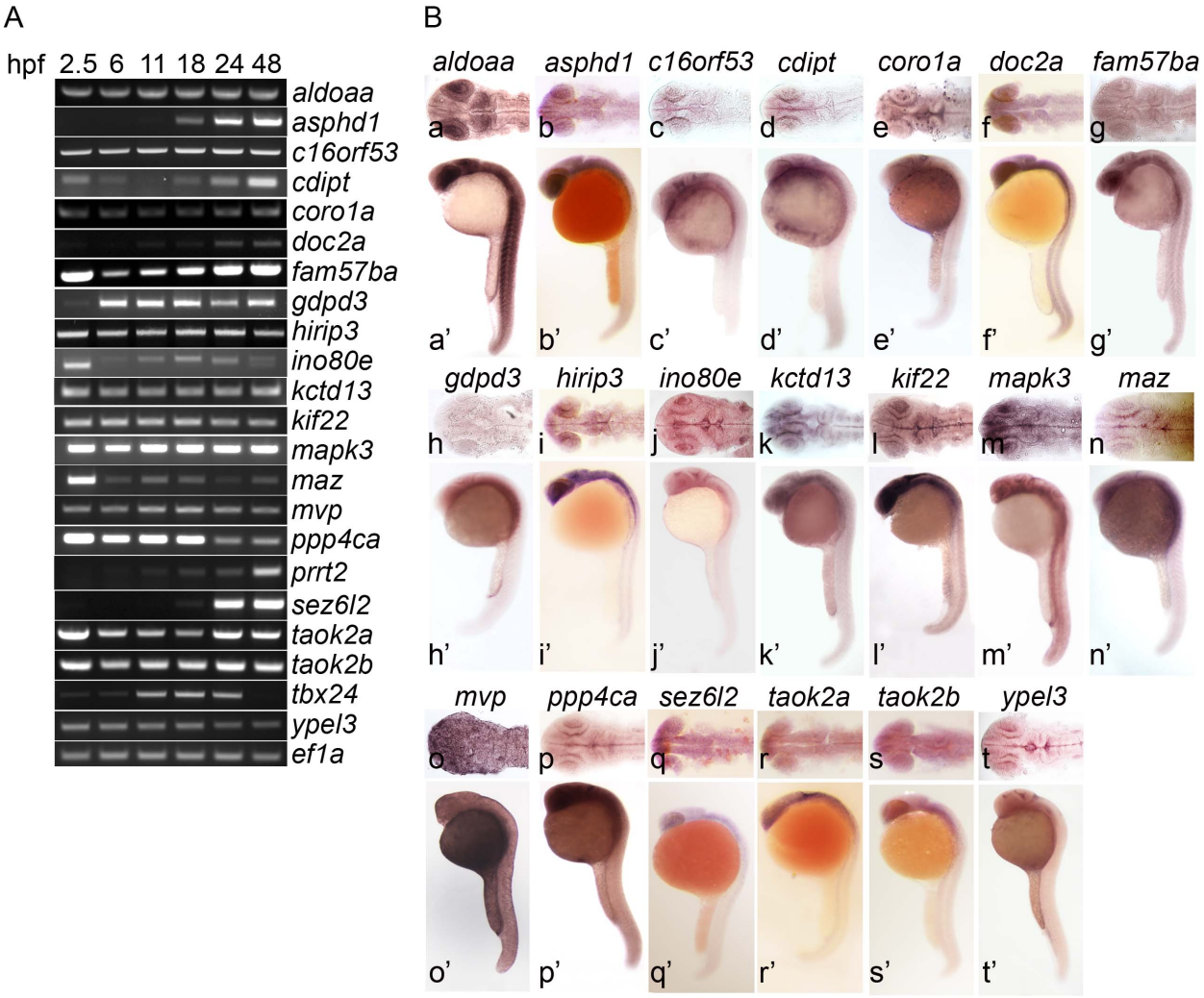

Figure S1. 16p11.2 homologs are expressed during development and in the brain. (A) RNA expression analysis by RT-PCR. Total embryo RNA was analyzed for expression of 16p11.2 homologs at 2.5 hpf (blastula), 6 hpf (early gastrula), 11 hpf (late gastrula), 18 hpf (mid-somitogenesis), 24 hpf and 48 hpf. Most genes are expressed maternally, at 2.5 hpf, and expression persists zygotically (6 hpf and later). *asphd1, doc2a, prrt2* and *sez6L2* show zygotic expression only, suggesting that these genes act later in development. *tbx24* expression declines sharply by 48 hpf, suggesting a narrow temporal activity window. Further, this gene is not expressed in brain**.** (B) Spatial expression analysis by whole mount in situ hybridization. Gene expression was analyzed at 24 hpf (see Methods and Table S4). (a-t) dorsal view, (a'-t') lateral view. All genes except *tbx24* were expressed in the brain.

Figure S2

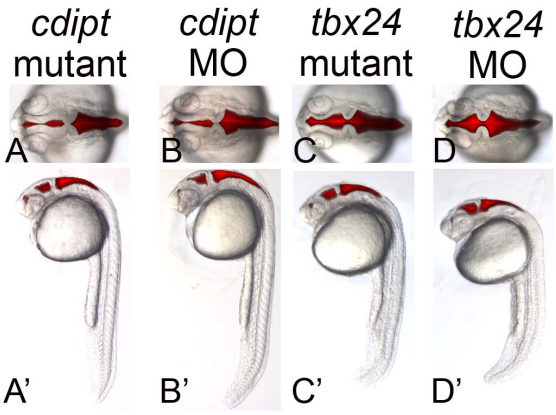

Figure S2. Similar phenotypes after LOF by MO injection and in genetic mutants. (A, A') *cdipt* $^{hi559Tg/hi559Tg}$ were scored by genotyping (see Methods for details). (B, B') *cdipt* antisense morpholino LOF embryo. In both cases, embryos have a narrow forebrain and midbrain ventricle. (C-C') *tbx24* $^{te314/te314}$ were scored by morphology. (D-D') *tbx24* antisense morpholino LOF embryo. In both cases, fused muscle segment borders and shorter tails are apparent. (A-D) dorsal view, (A'-D') lateral view.

Figure S3

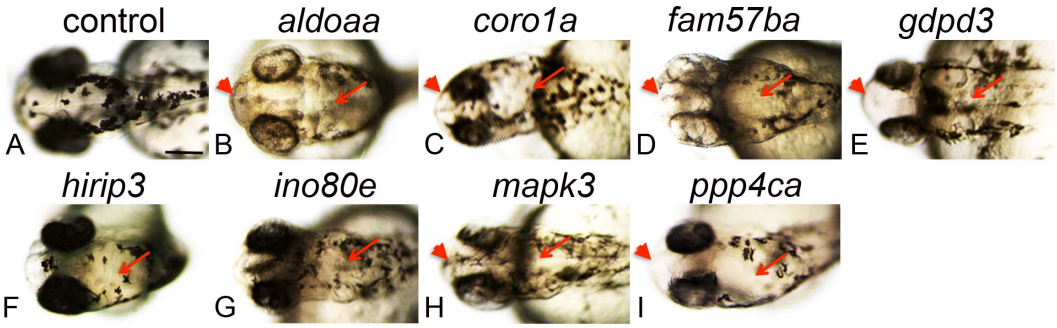

Figure S3. Pigmentation is abnormal in LOF embryos. Pigment cells were observed after LOF, at 48 hpf, by brightfield imaging. Rostral and midline cells were specifically scored as a measure of normal melanocyte migration. (A-I) dorsal views. Genes targeted for LOF by MO injection are indicated above each set of panels. Red arrowhead, no rostral pigment cells; red arrow, no midline pigment cells; scale bar, 150 μM. Average of 35 embryos assayed per gene, over 2-4 independent experiments, 70-100% affected compared to control MO-injected embryos.

Table S1

			Zebrafish Gene Percent Identity with Homolog in[d]:			
D.rerio Gene Name[a]	Ensembl Identifier[b]	Chromosomal Location[c]	Human	Fugu	Medaka	Mouse
aldoaa	ENSDARG00000011665	Chr. 3: 39.56	84	92	93	85
aldoab	ENSDARG00000034470	Chr. 12: 4.75Rev	84	ND	ND	82
asphd1	ENSDARG00000075813	Chr. 3: 14.87	52	ND	72	50
c16orf53	ENSDARG00000076966	Chr. 12: 4.31	50	54	60	48
cdipt	ENSDARG00000070686	Chr. 3: 21.35	69	70	76	60
coro1a	ENSDARG00000054610	Chr. 3: 31.09Rev	69	85	84	69
doc2a	ENSDARG00000078736	Chr. 3: 15.15	63	81	48	63
fam57ba	ENSDARG00000026875	Chr. 3: 21.15	58	ND	73	58
fam57bb	ENSDARG00000074564	Chr. 12: 4.34Rev	60	ND	74	55
gdpd3	ENSDARG00000074466	Chr. 3: 23.47Rev	46	67	66	44
gdpd3 (incomplete)[e]	ENSDARG00000006944	Chr. 12: 4.58Rev	48	61	60	47
hirip3	ENSDARG00000027749	Chr. 3: 14.81Rev	32	ND	37	30
ino80e	ENSDARG00000022939	Chr. 16: 12.02	48	70	74	49
kctd13	ENSDARG00000044769	Chr. 3: 14.83Rev	62	80	80	62
kif22	ENSDARG00000077375	Chr. 12: 5.80Rev	52	70	69	51
mapk3	ENSDARG00000070573	Chr. 3: 26.81	81	83	83	81
maz	ENSDARG00000087330	Chr. 3: 21.33Rev	45	ND	44	36
mvp	ENSDARG00000021242	Chr. 3: 15.33Rev	68	83	46	68
ppp4ca	ENSDARG00000070570	Chr. 3: 26.79	93	95	94	93
ppp4cb	ENSDARG00000076439	Chr. 12: 4.65Rev	98	98	98	98
prrt2	ENSDARG00000089367	Chr. 12: 5.77Rev	36	15	9	24
sez6l2	ENSDARG00000076052	Chr. 3: 14.86Rev	56	80	78	56
taok2a	ENSDARG00000074899	Chr. 3: 21.37Rev	70	89	87	69
taok2b	ENSDARG00000079261	Chr. 12: 4.50Rev	62	70	85	62
tbx24	ENSDARG00000011785	Chr.12: 4.59	19	18	19	20
ypel3	ENSDARG00000055510	Chr. 3: 26.84	89	ND	98	89

Table S1. Gene Assignment and Conservation.

[a] *D. rerio* gene names, which correspond to homologs indicated in Fig. 1.

[b] Ensembl identifiers (Release 62, April 2011).

[c] Chromosomal location (Release 62, April 2011). Chromosomal orientation of the zebrafish gene is on the plus strand unless the location number is followed by "Rev".

[d] Zebrafish percent identity with human, fugu, medaka, and mouse homologs of human 16p11.2 genes is shown.

[e] As the *gdpd3* gene on chromosome 12 is incomplete, the *gdpd3* gene on chromosome 3 was used for all phenotypic assays performed.

D. rerio, Danio rerio; Chr., chromosome; ND, not determined.

Table S2

	Antisense Morpholino Oligonucleutide[b]		Primer Sequences (5' to 3')[c]	
Gene Name[a]	MO Sequence (5' to 3')	Splice or Translation Blocking	Forward	Reverse
aldoaa	CCATGCTGAGGAGGAATGCAAGATT	Splice Site	GACACTGGTTCATGGCATTG	GCATCAAAGTGGACAAGGGT
			AGATGCATGGCATCGTTCCAATCG#	GGGCCTTGTACACAGCAGCCA#
asphd1	GAGAGCTTCCCATGTCTGACCTGAT	Splice Site	AGCAGCTCAGGCCGTTCCTCAT	GAGTGGGAACACAGTCTGGCCCTTT
			GCAATCAGGCCAGGATCAAT#	ATAGCTGGCCTCCAGAACCT#
c16orf53	GGAGTGTTGTAAAGCTCATACTGTT	Splice Site	CGCTGGGAGTCGGTGAACGG	CTCCTGCGTGTGGTCCTGCG
			CGACGGAGGACAACACTGCAGAAG#	CTCCTCATCACTGTAAGGGATCATCC#
cdipt	TTGTACCTAAGGACGAGGCACGAAT	Splice Site	GCCAGACTTCAGCAGACACA	TTTAATCGGATATGCCCGTATCGTAC
			GTGCGCTCAATCAAGGTACAAAATTTGGTGC#	ACGGGTACAGCAGTGCCAGGT#
coro1a	TGGTTTCTGATGTGGTTTACCTGCA	Splice Site	GAGGAGCCTTCATTGTCCTG	ACGTTCCAGAGGATCACCAC
			TGTAGTGTCATGATCTGGGAGA#	TCACAGCCTGCACTCATTAG#
	ATGAAGTCCTTGTCACTCACCATGA	Splice Site	ATCTGGGAGATTCCAGAGGG^	ATTATCACAGCCTGCACTCAT^
doc2a	CGCAGCTACACAATCACAGGAATGA	Splice Site	ACTGCGCTGGGGACGCTAGA	AATCCCCCTCGCCGGGGACTC
			GCACCATCATCAGAGCAAAGGGTTTG#	TTGGCCTTGCATGCTCCTGGTAGA#
fam57ba	ATGCCTGGAAGGAAAATGGAAGAGA	Splice Site	ATCTGACGCTGCTCAACCTT	ATAATCCCCTTTGCCCTGTC
			CTTGTAAAGACATCATCGAGGACCAGC#	GTACCAGTAGCACAAGAACATGGCA#
gdpd3	GCTCGCCATATTTCCCAAAACAATT	Translation Start Site	NA	NA
			TGTTTTCCGCAAATTTCCTC*	CAGCGCTTTTCTCATTGTCA*
hirip3	TTATCGCGTCCTTTTCTTTGGCCAT	Translation Start Site	NA	NA
	CTTTTCAGTCACTTGTTTACCTGCA	Splice Site	TGGGAAGAGAGTCGCTCAGT*^	TTCCTCGCTGTCTGAATCCT*^
ino80e	CCTCTACTTCTGCTTGCCCGTTCAT	Translation Start Site	NA	NA
			AACGGGCAAGCAGAAGTAGA*^	GGAAGAGAGTGGACCAACCA*^
	AGATGGTAACGTACTTACATACACC	Splice Site		
kctd13	CCAGCCTTCAATACAAATGGACACA	Splice Site	GTACTCCCCTGATGTCTGCC	CTCTGCGTTTCTGTAATGTGGT
			AGCATCGACTCTGAAGGCTGGGTG#	CTCCTCCAGCTCCCTCGTGCT#
kif22	CTTTCCTGAGGTGAAGAACAAGAAT	Splice Site	GTGTTGCTGTGAACGATGGA	GGAGCGCTGGTTGAGTTTAG
			CAGGCCGTTGTTACTCCATT*	CAAAGTGGTAGATGCGCTGA*
			GTTTCTGACCTCCGTCAAGC#	TCCAACCAGATTGAAGACCTC#
mapk3	GCTACTGCCCGATTCCGCCATCGTT	Translation Start Site	NA	NA
			CCTGAACATGACCACACTGG*	GGAGGTAGTTCCTGGCCTTC*^
	GTACGATATAGCTGAAACGCGGTTA	Splice Site	CATCATCGGGATCAATGACA^	
maz	ACTGGTTTTACTGGGAAAGAGAAGA	Splice Site	ATGCAGCTTGGAGCAATTTT	CCTGCTTCCCAAAAGTATGC
			ACCCTCCAATCAGATCTTCTC*#	GATTCTTCCTGACCGGATTG*
				GTTTTACCGGCGGCATGG#
mvp	TGCAGTCTTTCAGATTACCTCTCAT[1]	Splice Site	CCCACCTCACCACTACATCC^	AATGTATCCCGGCAACAAAG
			GAGGGACCAGGGACCTACAT*	TCTGCAACTGAGGGAATGTG*^
	TCCATGTTTCTTCCTACGGGCATTC[2]	Translation Start Site	NA	NA
	GGTCCCTGCCGGATTTAAAACACAG[3]	Splice Site		
ppp4ca	GTGGAGTAAAAGGAGTTACCTTGCT	Splice Site	GCGACTTCACTGACCTGGAC	CTGGCCATGTATGTCACCAC
			CAAAGCAAGGGAGATTCTGG*#^	CTGTCAGGGTATCGCACCTT*
				GTTTGTTTCCGGGACTTCAC#
	CCGACCTAAAGACAATCATCAGAGA	Splice Site		GCAGAGCCGTATTTTCGATG^
prrt2	ACGCCTGATGGTACTCTGTTTCCAT	Translation Start Site	NA	NA
	AGAATTTGAAAGTACCGCTCTCTGC	Splice Site	GGAGAGGTCATCGTGATGGT*^	GCTCCAGACTGTTCCTCGAC*^
sez6l2	TCACAGCAAAAACTGTGGACACCAT	Translation Start Site	NA	NA
	GCACTGAACTGAAGAAACAGTTTGT	Splice Site	TTTCTAGTCCCCGCCCCCTGG*^	AGGTGCACTCCAGTGGGGAAA*^
taok2a	TAATGGTGTGGCTTGCTCACAAAGT	Splice Site	CCGTGTGTTGATTTGAAACG	GCCCTGAAGTGCACCGTGGG
			GGAGCAGTGTACTTTGCTCACGAT#	GGGGTGTCTGAGCTTCTGCAAGA#
taok2b	CCATCCTGAAAACACAGCACAACTC	Splice Site	AAAGCTGCGGCACCCCAACA	GCAGGTGATGCCGAGTGACCA
			GTGGGCACACCTTACTGGATGGC#	TGCCGAGTGACCACACGTCA#
tbx24	CATTTCCACACCCAGCATGTCTCGG	Translation Start Site	ACATGTGAGATGGCAGCTCGCG*	TGAGTTTGGCCCGGTGGAAGGA*
ypel3	CCTACGTTCACCCTAAAAACACAAA	Splice Site	AGCAGACCAAGGCCAAAACT	TCAGTCCCAGCCATTATCCT
			GCAGTCAAGGACGAGCCTAC*	
			TCAAGGACGAGCCTACCTGT#	CAGTGTGGTGTGGCAATTCT#

Table S2. Primer and MO sequences.

[a] Gene names corresponding with zebrafish genes included in Table S1 that were targeted for loss of function (LOF) by antisense morpholino oligonucleotides (MOs).

[b] MO sequences and primer sequences are listed 5'-3'. Splice site MOs are designed to either the donor or acceptor side of the intron - exon boundary**.** Where multiple MOs are shown for one gene, the top-most MO was used for phenotypes reported.

[c] The first set of primers listed for each gene was used to monitor changes in splicing after MO injection, except for translation-blocking MOs (NA).

[1, 2, 3] *mvp* MOs are included in this order in Tables 1, 2, and S3. *mvp* MO[1] is included in Figs. 2A and 3.

***** Indicates primers used for the developmental expression timing RT-PCR if different from those used to monitor changes in splicing.

Represents primers used for qPCR.

^ Indicates primers used to monitor changes in splicing for second or third MOs (see Table 1), whose phenotypes are otherwise not reported (aside from reporting for *mvp* MOs in Tables 2 and S3).

NA, not applicable.

Table S3

Gene Name[a]	Brain Phenotypes[b]			Tail & Muscle Phenotypes[c]			Movement Phenotypes[d]		
	Total Embryos	Affected Embryos	Abnormal Brain %	Total Embryos	Affected Embryos	Abnormal Tail %	Total Embryos	Affected Embryos	Abnormal Movement %
aldoaa	152	141	93	152	140	92	104	88	85
asphd1	198	193	97	115	115	100	115	0	0
c16orf53	90	66	73	90	66	73	90	0	0
cdipt	117	104	89	117	0	0	117	0	0
coro1a	106	94	89	106	92	87	91	87	96
doc2a	140	83	59	140	0	0	140	0	0
fam57ba	169	162	96	169	162	96	71	65	92
gdpd3	88	65	74	100	83	83	99	89	90
hirip3	102	102	100	150	131	87	150	123	82
ino80e	103	100	97	103	97	94	68	68	100
kctd13	165	149	90	115	107	93	75	65	87
kif22	351	308	88	351	313	89	242	238	98
mapk3	158	154	97	158	156	99	124	123	99
maz	202	170	84	201	149	74	46	44	96
mvp[1]	167	145	87	167	145	87	67	65	97
mvp[2]	28	18	64	28	18	64	ND	ND	ND
mvp[3]	50	50	100	50	50	100	40	40	100
ppp4ca	101	96	95	101	83	82	97	71	73
prrt2	64	5	8	64	5	8	64	1	2
sez6l2	47	42	89	109	99	91	109	0	0
taok2a	71	62	87	71	0	0	21	10	48
taok2b	74	67	91	78	65	83	78	0	0
tbx24	95	0	0	95	95	100	95	95	100
ypel3	49	47	96	49	38	78	47	42	89

Table S3. Quantification of brain, tail, and movement phenotypes. Data was gathered in 2-7 experiments for each LOF condition and includes 24 - 48 hpf experimental timepoints. Data is summarized in Fig. 3 and representative images are shown in Fig. 2.

[a] Genes indicated in Table S1 were targeted in LOF experiments by MOs included in Table S2, using MO mass tabulated in Table 2. Where multiple MOs were designed, this table represents data obtained using the first MO listed in Table S2, aside from *mvp*, where three genes are listed.

[b] Brain phenotypes, including abnormally shaped ventricle neuroepithelium were observed in 24 hpf LOF embryos using criteria described in Methods and shown in Figs. 2 and 3. Phenotypes observed for forebrain, midbrain, and hindbrain were included as a group in the number of abnormally affected embryos that is indicated.

[c] Summation of major categories of tail and muscle phenotypes (including tail length, extension, and shape and muscle segment chevron abnormalities) were observed in 24 hpf LOF embryos using criteria described in Methods and data shown in Figs. 2 and 3. The number of abnormally affected embryos is indicated.

[d] Movement abnormalities (which may arise from motor neuron or muscular defects) include grouping of spontaneous movement and touch response defects, which are assays described in Methods. The number of abnormally affected embryos are indicated.

[1, 2, 3] The *mvp* MOs are included in the same order as Tables 2 and S2.

[1] Indicates the *mvp* MO that is included in Figs. 2 and 3.

ND, not determined.

Table S4

Gene Name[a]	Human or Fish Gene[b]	GenBank Accession No.[c]	Parent Vector[d]	Cloning into pCS2+[e] (primers 5' to 3' or restriction enzymes)	in situ sense probe preparation[f]	in situ antisense probe preparation[f]	mMessage rescue mRNA[g]
ALDOA	human	NM_184043.1	pINCY	EcoRI/SalI			SP6 Pol, NotI RE
aldoaa	zebrafish	BC065320	pCMV-SPORT6.1	NA[h]	SP6 Pol, NotI RE	T7 Pol, KpnI RE	
ASPHD1	human	BC126319	pCR4-TOPO	EcoRI/XbaI			SP6 Pol, NotI RE
asphd1	zebrafish	NA	NA	GATCGGATCCTCTGATCCTCCATCCCACTC	SP6 Pol, NotI RE	T7 Pol, HindIII RE	
				GATCCTCGAGAGGTCCACGCTGAAGATGAC			
C16ORF53	human	BC003640	pOTB7	AAGAATTCCTGCCCTAGTGGCCTATGTCCC			SP6 Pol, NotI RE
				AACTCGAGTCAGTATTTCCGCTGCCGAGGG			
c16orf53	zebrafish	NA	NA	AAGAATTCTTAATCACCCCGCGCGGCTC	SP6 Pol, XhoI RE	T7 Pol, EcoRI RE	
				AACTCGAGCCCCTCTGAAAGACTGAGCTGTCC			
CDIPT	human	BC001444.2	pOTB7	EcoRI/BglII (blunt) to StuI site in pcs2+			SP6 Pol, NotI RE
cdipt	zebrafish	BC066511	pCMV-SPORT6.1	NA[h]	SP6 Pol, NotI RE	T7 Pol, SpeI RE	
CORO1A	human	BC110374	pCMV-SPORT6	ATTAGGATCCCAGAATGAGCCGGCAGGTGGT			SP6 Pol, NotI RE
				GCATCTCGAGGGCTCTACTTGGCCTGGACTG			
coro1a	zebrafish	BC055237.1	pME18SFL3	EcoRI/XbaI	SP6 Pol, XbaI RE	T7 Pol, EcoRI RE	
DOC2A	human	BC041769	pBluescriptR	AAGAATTCGCCAGGGGTGCTGCATGAGG			SP6 Pol, NotI RE
				AATCTAGATGTGCCGGGACACTGCTGTC			
doc2a	zebrafish	NA	NA	GATCGGATCCCCCTGCTTCTCCACTGTCTC	SP6 Pol, NotI RE	T7 Pol, HindIII RE	
				GATCCTCGAGACTGTGACCTCCAGCGTCTT			
FAM57B	human	BC007892	pOTB7	EcoRI/XhoI			SP6 Pol, NotI RE
fam57b	zebrafish	BC059567	pBluecript SK(-)	NA[h]	T3 Pol, XhoI RE	T7 Pol, EcoRI RE	na
GDPD3	human	BC002714	pOTB7	EcoRI/XhoI (Removes 3'UTR)			SP6 Pol, NotI RE
gdpd3	zebrafish	BC064306	pExpress-1	NA[h]	SP6 Pol, NotI RE	T7 Pol, EcoRV RE	
HIRIP3	human	NM_003609	pCMV6-AC	AAGAATTCGAGCCGTCAATCCCGGGTTG			SP6 Pol, NotI RE
				AACTCGAGTGGGGGTGGCAGAGCTCAGT			
hirip3	zebrafish	BC055674	pME18SFL3	XhoI	SP6 Pol, NotI RE	T7 Pol, HindIII RE	
INO80E	human	BC047712.2	pCMV-SPORT6	ATTAGGATCCGGTCATGAACGGGCCGGCGGACG			SP6 Pol, NotI RE
				GCATCTCGAGCACGGTCACTCCGGGATGTCGAT			
ino80e	zebrafish	BC065338	pCMV-SPORT6	NA[h]	SP6 Pol, XhoI RE	T7 Pol, StuI RE	
KCTD13	human	BI548684.1	pBluescriptR				
kctd13	zebrafish	BC122284	pME18SFL3	XhoI	SP6 Pol, BglII RE	T7 Pol, EcoRI RE	SP6 Pol, NotI RE
KIF22	human	BC028155	pCMV-SPORT6	EcoRI/XhoI (114 n.t. of 3' UTR remains)			SP6 Pol, NotI RE
kif22	zebrafish	BC154464	pME18SFL3	XhoI	SP6 Pol, XbaI RE	T7 Pol, EcoRI RE	
MAPK3	human	BC013992	pOTB7	TGTAAAACGACGGCCAGTAA			SP6 Pol, NotI RE
				GGTGCAGAGATGTCTGTCTGG			
mapk3	zebrafish	BC097073	pCMV-SPORT6	NA[h]	SP6 Pol, XhoI RE	T7 Pol, EcoRV RE	
maz	zebrafish	NA	NA	ACGTCATGATGGATGCAGCTTGG	SP6 Pol, NotI RE	T7 Pol, EcoRI RE	SP6 Pol, NotI RE
				CCTCTCTCCTGTTGGTTAGCTG			
MVP	human	BC015623	pOTB7	EcoRI/XhoI (93 n.t. of 3' UTR remains)			SP6 Pol, NotI RE
mvp	zebrafish	BC063949.1	pExpress-1	NA[h]	SP6 Pol, NotI RE	T7 Pol, EcoRI RE	
mvp	zebrafish	BC063949.2	pExpress-1	EcoRI/NsiI			SP6 Pol, NotI RE
PPP4C	human	BC001416	pOTB7	EcoRI/XhoI (301 n.t. of 3' UTR remains)			SP6 Pol, NotI RE
ppp4ca	zebrafish	BC049430	pME18SFL3	EcoRI/XbaI	SP6 Pol, XbaI RE	T7 Pol, EcoRI RE	
ppp4cb	zebrafish	BC155609	pExpress-1	NA[h]	SP6 Pol, NotI RE	T7 Pol, EcoRV RE	
prrt2	zebrafish	BC053594	pCMV-SPORT6	NA	NA	NA	NA
SEZ6L2	human	BC000567	pOTB7	AAGAATTCAGAGATCGGGGTGAGTCGCCA			SP6 Pol, NotI RE
				AAACTAGTTGCAGCTGTAGTCTTGGGGTTCA			
sez6l2	zebrafish	NA	NA	GATCGGATCCCATACGCCACAATGACGTTC	SP6 Pol, NotI RE	T7 Pol, HindIII RE	
				GATCTCTAGACTCCACTGTAATGGGGCTGT			
TAOK2	human	BC142663	pCMV-SPORT6	AAGAATTCTACCAGGCCAGGCCCCACTC			SP6 Pol, NotI RE
				AAACTAGTCAGCTaCCTCCAGGGGGGCAG			
taok2a	zebrafish		pGEM-T Easy	AGCGGGCGAGCAGGTACAC	T7 Pol, SpeI RE	SP6 Pol, SacII RE	
				ATGGGGGCGCTCTCGGCATAA			
taok2b	zebrafish		pGEM-T Easy	CCTTCAAAAGCTGCGGCACCC	T7 Pol, SpeI RE	SP6 Pol, SacII RE	
				TGATTCAGTGCTTTGTGCAGCACG			
TBX6	human	BC026031	pOTB7				
YPEL3	human	BC050664	pCMV-SPORT6	EcoRI/XhoI (48 n.t. of 3' UTR remains)			SP6 Pol, SacII RE
ypel3	zebrafish	BC067578	pCMV-SPORT6	NA[h]	SP6 Pol, XhoI RE	T7 Pol, KpnI RE	

Table S4. Human and zebrafish cDNA clones.

[a] Genes indicated in Fig. 1 for zebrafish and human.

[b] Human genes were used for rescue experiments, except for *kctd13* and *maz*. Zebrafish genes were used for in situ probe preparation, and for rescue of *kctd13* and *maz*.

[c] Clones were ordered for human genes, and where available, for zebrafish genes.

[d] Parent vector of clones.

[e] cDNAs were subcloned into pCS2+ for in vitro mRNA transcription. Primers or restriction sites used for subcloning are listed and were designed to remove the majority of untranslated regions of the cDNA, as these can promote degradation in the fish embryo. Primers were also used to amplify zebrafish cDNAs where clones were not available.

[f] Linearizing enzyme and polymerase used to generate sense and antisense probes.

[g] Linearizing enzyme and polymerase used to generate rescue mRNA.

[h] The cDNA used for in situ probes was left in the parent vector if appropriate polymerase and restriction binding sites were present.

Pol, polymerase; RE, restriction enzyme; n.t., nucleotide; UTR, untranslated region; NA, not applicable.