

Title:

Survey of human chromosome 21 gene expression effects on early development in

Danio rerio

Short Title:

Chromosome 21 gene effects on embryonic development in *D. rerio*

Authors:

Sarah McGuire Edie^{1#}, Norann A Zaghloul², Carmen C Leitch², Donna K Klinedinst³,
Janette Lebron^{3##}, Joey F Thole¹, Andrew S McCallion¹, Nicholas Katsanis⁴, Roger H.
Reeves^{1,3,*}

Affiliations:

¹ Johns Hopkins University School of Medicine, Institute of Genetic Medicine

² University of Maryland School of Medicine, Department of Medicine, Division of
Endocrinology, Diabetes and Nutrition

³ Johns Hopkins University School of Medicine, Department of Physiology

⁴ Center for Human Disease Modeling, Duke University

Current Address: The Jackson Laboratory, Bar Harbor, ME 04609

Current address: Optogenetics Program, NIDA, Baltimore, MD 21224

*Address correspondence to:

Roger H. Reeves, Ph.D.

Biophysics 201

Johns Hopkins School of Medicine

Baltimore, MD 21205

reeves@jhmi.edu

ABSTRACT:

Trisomy for human chromosome 21 (Hsa21) results in Down syndrome (DS), one of the most genetically complex conditions compatible with human survival. Assessment of the physiological consequences of dosage-driven overexpression of individual Hsa21 genes during early embryogenesis and the resulting contributions to DS pathology in mammals are not tractable in a systematic way. A recent study looked loss-of-function of *C. elegans* orthologues of Hsa21 genes and identified ten candidates with behavioral phenotypes, but the equivalent over-expression experiment has not been done. We turned to zebrafish as a developmental model and, using a number of surrogate phenotypes, we screened Hsa21 genes for dosage sensitive effects on early embryogenesis. We prepared a library of 164 cDNAs of conserved protein coding genes, injected mRNA into early embryos and evaluated up to 5 days post-fertilization (dpf). Twenty-four genes produced a gross morphological phenotype, 11 of which could be reproduced reliably. Seven of these gave a phenotype consistent with down regulation of the sonic hedgehog (Shh) pathway; two showed defects indicative of defective neural crest migration; one resulted consistently in pericardial edema; and one was embryonic lethal. Combinatorial injections of multiple Hsa21 genes revealed both additive and compensatory effects, supporting the notion that complex genetic relationships underlie end phenotypes of trisomy that produce DS. Together, our data suggest that this system is useful in the genetic dissection of dosage-sensitive gene effects on early development and can inform the contribution of both individual loci and their combinatorial effects to phenotypes relevant to the etiopathology of DS.

INTRODUCTION

Down syndrome (DS) occurs in about one of 700 live births due to trisomy for human chromosome 21 (Hsa21) [1]. The consequent ~1.5 fold over expression of most genes on Hsa21 can result in more than 80 clinical phenotypes, many of which originate during prenatal development and vary in both severity and penetrance [2-4]. Among the most consistent features are cognitive impairment, characteristic craniofacial dysmorphism, smaller and hypocellular brain and Alzheimer histopathology [5, 6]. Individuals with DS also have a greatly increased risk of congenital heart disease, Hirschsprung disease and acute megakaryoblastic leukemia in children. However, the incomplete penetrance of many DS phenotypes indicates that trisomy 21 is not sufficient to cause most of these conditions, suggesting an important role for allelic variation of Hsa21 genes and additional modifier genes, as well as potential environmental and stochastic factors [7-9]. DS is one of the few autosomal aneuploidies that is compatible with life, likely related to Hsa21 being the smallest autosome and having very low gene density. Estimates of the gene content on Hsa21 range from ~300-600 genes/transcripts, of which 162 have been identified as well-conserved in other mammals [10]. Understanding how trisomy for these genes affects the presentation of the phenotypes in DS is a major focus for research into this condition.

A major challenge in understanding mechanisms of gene action in DS is that trisomy 21 is present from conception and every cell is affected, causing effects throughout development. Trisomic genes may have a primary effect directly on cellular function or may secondarily affect expression and regulation of disomic genes. Trisomy-induced changes in one cell type could alter interactions with neighboring cells, thus

initiating cascades of primary and secondary effects [6, 11]. Use of mouse models trisomic for different segments of Hsa21-orthologous sequences supports to an extent the idea that different genetic segments correlate with some specific phenotypes [12, 13] [14], but even the smallest segmental trisomy still contains many genes. The effort and cost to systematically engineer individual transgenic mouse models of all conserved genes on Hsa21 would be prohibitive, to say nothing of the analysis of the possible combinations of genes. Further, events early in embryogenesis are difficult to access in mammals.

Recently, a screen to examine the effects of down-regulating orthologs of 47 Hsa21 was performed in *Caenorhabditis elegans* [15]. Ten of these conserved genes exhibited neurobehavioral phenotypes: *COL18A1*, *CBS*, *DONSON*, *EVA1C*, *N6AMT*, *NCAM2*, *POFUT2*, *PDXK*, *RUNX1* and *SYNJ1* [15]. Of these ten genes, five were shown to be essential for development based on the lethality phenotype seen in mouse knock-out models. The *C. elegans* screen identified three genes that were previously uncharacterized (*DONSON*, *N6AMT* and *PDXK*) as having a phenotype, providing new information about DS related genes and showing that these types of expression-based screens can provide a valuable resource to the DS research community. The knockdown screen in worms for all of the likely Hsa21 orthologs provided insights into gene function, but an over-expression screen that might be more relevant to over-expression in DS has not been done. Previous studies have shown that the expression and/or suppression in zebrafish embryos of genes that map to disease-associated duplications and deletions in people can distinguish potent drivers of pathology [16-20]. Motivated by such studies, we systematically over-expressed in zebrafish embryos each

1 of 164 Hsa21 cDNAs representing 163 genes and assessed their effects on early
2 development.
3

RESULTS:

Development of the clone set and initial screen:

In a detailed annotation of transcripts from Hsa21, Sturgeon and Gardiner identified 162 genes that are highly conserved with mouse [10]. We assembled a set of Hsa21 cDNA clones consisting of 148 of these conserved genes and 15 human-specific genes [10, 21]. The 15 genes that are not conserved include DSCR genes, long intergenic non-coding RNA, and Hsa21 open reading frames. One gene, *SYNJ1*, is represented by two splice isoforms, for a total of 164 cDNAs (Supplemental Table 1). The majority of clones are from the InVitrogen UltimateORF collection library (now Thermo Fisher Scientific's Ultimate ORF Clones), subcloned into the pCS2+ destination vector (Figure 1) [22]. Thus we selected highly conserved genes from this carefully annotated set for which cDNA could be obtained and which could be cloned into the pCS2+ vector and transcribed. Additional sources and vectors are described in Supplemental Table 1.

Next we used Mouse Mine to interrogate the human gene list and compare it to homology data sets, HomoloGene and Panther, to identify orthologs and ohnologs between zebrafish and mouse (Supplemental Table 1) [23], [24]). In zebrafish, 125 of the 163 Hsa21 genes (77%) are conserved; 35 of those (28%) are represented by more than one zebrafish ortholog (ohnolog). The number of Hsa21 genes conserved with zebrafish is similar to published reports for all human genes (71%), but the proportion of Hsa21 genes with ohnologs is less than the 47% rate for all human genes [25].

Zebrafish Screen

We synthesized mRNAs from the Hsa21 cDNA clones and injected each one into ~100 zebrafish embryos at the 1-2 cell stage at 10, 50 and/or 100 pg (Supplemental Table 2), ranges that have been used in similar studies previously [20]. More than 50,000 embryos were screened for the presence of gross morphological changes present at frequencies greater than controls, typically at five days post fertilization (dpf), but also earlier if mandated by the presence of clear pathologies. We focused primarily on three broad phenotypic classes: a) U-shaped somites and cyclopia, two phenotypes associated frequently but not exclusively with defects in the ciliome and Shh signaling pathway; b) craniofacial abnormalities and pigment differences, which may be related to aberrations affecting neural crest cells; and c) pericardial edema, which can have a number of causes including a structural heart defect (Figure 2, Table 1). We recorded the number of surviving embryos, the phenotype, and the percentage that were affected (Table 2).

Of the 164 RNAs, 24 showed a phenotype after the initial screen (Supplemental Tables 2 and 3); the remaining 140 did not yield a significant number of affected embryos. Thus, again consistent with previous studies [20], our approach is not generally toxic to zebrafish embryos. Expression of 14 of these 24 genes gave a phenotype consistent with perturbation of the ciliome/Shh- pathway (eight with U-shaped somites and a partially overlapping set of six with cyclopia). An additional seven genes resulted in phenotypes that may be due to neural crest defects (four with craniofacial abnormalities and three with pigment differences). Finally, one gene induced pericardial edema, one resulted in dysmorphic fins, and one exhibited elevated lethality (Table 1 and

Supplemental Table 2). Of the 24 first-round candidates, 19 are conserved in both mouse and zebrafish; four are conserved in between human and mouse but not zebrafish (*BACH1*, *Clic6*, *MAP3K7CL*, *SPATC1L*); and one gene (*LINC00313*) is human specific. Six of the 19 conserved Hsa21 genes have ohnologs in zebrafish (Supplemental Table 1).

Fresh mRNA was prepared from these first-round candidates and injections were repeated. Eleven candidates recapitulated the original phenotypes robustly, seven from the Shh group, two with neural crest related phenotypes, and the single genes resulting in pericardial edema and embryonic lethality, respectively (Figure 2, Table 2, Supplemental Tables 2 and 3). Of the 11 second-round candidates conserved in zebrafish, two have ohnologs in zebrafish (*JAM2* and *CBR3*).

Phenotypic rescue via Morpholinos

The first indication that the phenotypes observed in repeated injections are due to the expression of the specific RNA and not to a general RNA effect is that there are at least 140 clones that produced no phenotype in this screen. Next, we selected randomly four of the candidate genes and performed a rescue experiment using morpholino (MO)-based knockdown. We designed translation-blocking MOs against the human copies of *SOD1*, *RWDD2B* or *CCT8* to target the ATG start site of the gene to suppress specifically the introduced human mRNA. We also designed MOs to knock down the endogenous zebrafish orthologs of *JAM2*, using a previously validated MO [26]. One hundred embryos were injected with 2ng MO and 100pg of RNA. For all four genes,

injection of the RNA alone produced significantly higher penetrance than the uninjected controls, the MO alone, or the MO+RNA injected embryos ($p < 0.05$ [Figure 3](#)). For *SOD1*, *RWDD2B*, and *CCT8* the MO+RNA was not significantly different from the controls. *JAM2* MO+RNA showed significantly lower penetrance of heart edema relative to RNA-injection alone [27]. The *JAM2* experiment suggests that this phenotype is a product of additive effects of the human cDNA with its fish orthologs/ohnologs.

Dosage and Gene expression patterns

This screen was designed to identify Hsa21 genes with an effect on broad aspects of early embryonic development and not as a study of human dosage effects on fish (see Discussion). For those cDNAs producing phenotypes, penetrance was generally correlated with but not directly proportional to penetrance ([Supplemental Table 2](#)). In the case of *ERG*, however, the frequency of embryonic lethality was proportional to RNA dose. Approximately 1/3 of embryos (24/67) survived after injection with 10 pg *ERG* RNA, while 14% survived injection with 50 pg and 6.5% (24/358) were alive two days after injection of 100 pg. Retrospective examination also showed somewhat elevated mortality following injection of either of the additional two ETS family transcription factors in the Hsa21 clone set, *ETS2* and *GABPA* ([Supplemental Table 2](#)). *HMGN1* also showed a trend toward higher lethality.

Among the final candidates with repeatable phenotypes, only *ERG* showed increased lethality with increasing RNA concentrations. The 10 remaining candidates were injected at three or more concentrations ranging from 10pg to 200pg but no correlation between dosage and penetrance was observed ([Supplemental Figure 1](#)). The mRNA concentrations that we used are consistent with those commonly reported

[20, 28]. We also chose two genes, *SOD1* and *RRP1*, to look for effects of low RNA doses. These genes were injected at 2pg and 5pg, and the penetrance of U-shaped somites was examined. Both genes showed low penetrance of the phenotype at these low doses, although each reproducibly produced a phenotype after injection of 50-100pg (Supplemental Figure 2).

Nine of the eleven candidates have zebrafish homologs, the exceptions being *C21ORF84* and *CBR3*. For eight of these genes, representative *in situ* hybridization data are available in ZFIN [29]. In most cases, the structure(s) affected by RNA injection is consistent with the *in situ* expression data. For example, *sod1*, *rrp1* and *ybey* are expressed in somites, *cct8* is expressed ubiquitously, *jam2* shows expression in the region of the developing heart and *erg* is expressed in the vasculature (Table 2). *pofut2* is expressed in the brain and eye [30]; knockdown of a *C. elegans* homolog of this gene produced a neuromuscular phenotype [15]. *pcb3* is expressed in the retina and telencephalon beginning at the Prim 15 stage [29]. No information about *rwdd2b* expression was available in ZFIN. Ten of the eleven candidates have mouse orthologs whose expression has been examined at mid-gestation showing that the genes are expressed in the corresponding tissues during embryonic development in mouse (Table 2, [31, 32]). *C21ORF84* is a human specific lncRNA [21].

Combinatorial injections

DS is a contiguous gene defect with effects on development that exceed those of individual genes. Examination of all pairwise interactions of 163 genes would require more than 13,000 pairs, posing a logistical challenge even in this system. We took an

initial step toward understanding this process with respect to a small set of candidate genes. We first examined the subset of genes that produced possible Shh-related phenotypes in our screen. *SOD1*, *RWDD2B*, *YBEY*, *PCBP3*, and *RRP1* all produced U-shaped somites, *POFUT2* resulted in cyclopia and *C21ORF84*-injected embryos showed both phenotypes. We used *C21ORF84* as the reference gene to pair with the other six genes. For each set of pairwise injections, *C21ORF84* was injected alone at 100pg, the second gene was injected alone at 100pg, and both genes were injected together at 100pg each.

Injection of *C21ORF84* plus *YBEY*, *PCBP3* or *POFUT2* showed a significant increase in the frequency of embryos with U-shaped somites (Fig. 4, $p < 0.05$ for each combination). For two genes, *SOD1* and *RWDD2B*, there was not a significant difference between the individual injections and the combinatorial injection. *RRP1* alone had a penetrance of 42%, the highest of all genes tested. This frequency was reduced significantly in embryos injected with *RRP1* and *C21ORF84* together.

We repeated the pairwise injections using a different reference gene from this group, *SOD1*. As in the previous experiment, *SOD1* was injected individually and in combination with one of the other six genes, and the embryos were examined for the presence of U-shaped somites. In contrast to *ORF84*, *SOD1* did not show an interaction with any of the other six genes (Supplemental Figure 3). Finally, combinatorial injections using *C21ORF84* were repeated with freshly prepared RNAs at a separate institution (by NAZ and CCL) as a means of independent replication. We observed a similar increase in affected embryos upon injection of *C21ORF84* with *YBEY* and *PCBP3* compared to each gene alone.

We also selected candidate gene sets for combinatorial injection based on reported roles in other systems and assessed them for possible combinatorial effects on developmental phenotypes. Several genes associated previously with heart anomalies in Down syndrome, *SH3BGR*, *DCSR6* and *ADAMTS1*, were co-injected (30pg each). This resulted in cyclopia in 3.6% and pericardial edema in a non-overlapping 3.6% of embryos, whereas no controls were observed to have either edema or cyclopia, a trend though not formally significant (Fisher's exact test, $p=0.056$ for either cyclopia or edema). Several Hsa21 gene combinations have been implicated in the high frequency of congenital heart disease in DS [33]. However, neither injection of all three collagens together (*COL6A1*, *COL6A2*, *COL18A1*) nor coinjection of *DSCAM* and *SH3BGR* produced a significant frequency of heart or other gross anatomical defects.

DISCUSSION

We have developed a study set of Hsa21 gene expression clones (available from AddGene, see Methods) and used it to conduct the first large-scale study of the effects of Hsa21 gene expression on early vertebrate embryogenesis. Previous analyses of Hsa21 gene expression in early development used *in situ* hybridization [31, 34] or microarrays to examine the localization, timing, and/or levels of Hsa21 gene up-regulation [2]. Here, we used a functional assay in zebrafish embryos to find candidate genes with effects on early development in a systematic, unbiased approach. This stage of development is difficult to study in mammalian models. Of the 163 genes (164 cDNAs) assayed, only eleven genes consistently showed an effect. These genes are implicated in ciliome/Shh signaling, neural crest cell (NCC) generation and differentiation, heart development, and/or embryonic lethality. *ERG* has recently been shown by the Knock-Out Mouse Phenotyping Project (KOMP2) to be required during mammalian embryogenesis; embryos homozygous for a null allele of *Erg* are lethal prior to E15.5 (www.mousephenotype.org accessed March 27, 2017 [35]). Several of the genes have little or no known function and none have been implicated previously in either Shh signaling or NCC generation and development. Notably, the *POFUT2* gene, was also found to have a neurobehavioral phenotype in a reciprocal loss-of-function screen in *C. elegans*, indicating that it may be a good candidate for further study [15].

There are several caveats to interpretation of this large scale screen. A negative result in the screen does not rule out a contribution of that gene to DS. First, some human genes could fail to produce an effect due to inefficient translation related to the fact that these are mammalian mRNAs that may not be properly processed in fish.

Some human proteins may simply be hypo- or non-functional in zebrafish. Except for *SYNJ1*, we selected a single isoform for the cDNA clone set. Next, DS is a contiguous gene syndrome, with many interactions among the over-expressed genes and with disomic genes at different stages of development [11, 36]. Finally, it should be clear that *this is not a study of the effects of dosage imbalance* that occur in DS *per se*. Indeed, it would be difficult to define what dose of a human mRNA might result in human protein levels that, together with the orthologous fish proteins, replicate the *functional* stoichiometry of DS in a fish embryo. Genes with ohnologs that function in a complementary manner in fish represent an even more complicated situation when additional expression of (possibly) related but non-identical function is introduced via the human gene product. We did not observe a lower frequency of phenotypes among genes with ohnologs than in the Hsa21 gene set as a whole, but the numbers are very small. What we present is instead a screen for genes whose effects in this assay indicate candidates to pursue in far more labor intensive studies of early mammalian development. We have succeeded in one such application with *Jam2* [27].

The consistent occurrence of heart edema in zebrafish after injection of JAM2 RNA led us to examine this candidate in conjunction with increased penetrance of heart defects in trisomic mouse models. Ts65Dn “Down syndrome” mice are trisomic for about 104 genes orthologous to Hsa21 [37]. Breeding a null allele of the disomic gene, *Creld1*, onto Ts65Dn significantly increased penetrance of septal defects in the heart from 4% to 33% [38]. However, putting the same null allele on a related trisomy (Ts1Cje) that was trisomic for about 81 of the genes triplicated in Ts65Dn, had no impact on penetrance. *Jam2* was one of the 23 genes that are trisomic in Ts65Dn but

not Ts1Cje, and one of 14 of these 23 that are expressed prenatally and/or in heart. Based on its effect on heart development in the zebrafish screen, we tested its role in mice by introducing a null allele of *Jam2* to produce Ts65Dn;*Creld1*^{+/-};*Jam2*^{+/-} mice (returning *Jam2* to the normal two copies). Instead of the expected increase in penetrance in Ts65Dn;*Creld1*^{+/-}, trisomy for all of the Ts65Dn genes except *Jam2* resulted in penetrance in these mice that was the same as Ts65Dn alone (i.e., 4%) [27]. Thus *Jam2* is a trisomic potentiator of the disomic modifier of heart disease penetrance, *Creld1*. This was the first demonstration of this type of genetic relationship in Down syndrome. It would not have been possible to pursue this relationship for all 14 candidate genes trisomic in Ts65Dn but not in Ts1Cje in mice.

In this screen, it was somewhat surprising that several Hsa21 genes that have been associated with robust phenotypes in mouse models of DS did not produce a phenotype. For example, *DSCAM*, a cell adhesion molecule that is involved in cell recognition [39], has been implicated in both heart and neurogenesis defects based on work done in mice as well as in *Drosophila* [40, 41] but did not produce a phenotype in our zebrafish screen. *DYRK1A*, a dual specificity kinase expressed during early neurogenesis that has been a target in pilot studies for treatment of cognitive deficits in DS [42, 43], also produced no phenotype. Many types of refined screens with greater sensitivity and specificity are possible, taking advantage of transgenically marked zebrafish lines to ask specific questions about development of specific structures. Furthermore, our identification of phenotypes associated with pathways previously identified as central to multiple manifestations of trisomy, such as Shh signaling [44-46], support the use of *D. rerio* for this type of large-scale systematic screen. This system

represents a useful screening tool to identify individual candidate genes that may be significant drivers of DS phenotypes.

We observed that some Hsa21 genes that produce phenotypes on their own can interact in an additive manner, some have no apparent interaction and one pair had a compensatory interaction. Compensatory interaction implies that in some cases overexpression of one gene can balance the increased expression of another. To date, several genes of major affect have been associated with manifestations of Down syndrome in the mouse models. However, no single gene has been identified that is sufficient to produce completely a complex developmental phenotype of DS, consistent with the understanding of DS as a product of complex multi-gene interactions. Given the large number of possible gene-gene interactions on chr21 alone, the system described here provides a useful way to interrogate more complex interactions of non-contiguous genes from the earliest stages of development.

METHODS:

Hsa21 Gene Expression Library preparation

cDNAs were selected from lists of conserved genes on Hsa21 [2, 21]. For 84 genes, plasmids containing the gene in the pENTR221 entry vector were obtained through the Invitrogen UltimateORF collection. Of the remaining genes, 49 were subcloned from a variety of vectors into one of the Invitrogen Gateway entry vectors (for complete list of original vectors and sources see Supplemental Table 1) and 40 genes were commercially cloned into the pCS2DEST vector. Entry vector clones were selected

using kanamycin and then sequenced to confirm correct insertion of the gene. All genes in the entry vector were subcloned into the pCS2DEST vector (Addgene) using LR clonase as previously described by Katzen *et al.* 2007 [47]. Genes in the pCSDEST vector were selected by ampicillin. The entire pCS2+ vector clone set, named the Hsa21 Gene Expression Set, is available through Addgene (https://www.addgene.org/Roger_Reeves/). A limited set of HH pathway-related genes was used as a training set for the system including recognition of the U-shaped somite phenotype.

Bioinformatics

Comparisons of human, mouse, and zebrafish orthologs and ohnologs was performed using MouseMine (www.mousemine.org, accessed October 1, 2017 [23]. Briefly, Hsa21 gene symbols for the 163 genes in this screen were uploaded as a list in MouseMine and interrogated against the mouse and zebrafish using the HomoloGene data set from NCBI and the PANTHER data set from MGI [24]. These lists were used to compile the ortholog and ohnolog lists in Supplemental Table 1.

In Vitro Transcription of mRNA

Plasmids were transcribed *in vitro* using the mMessage mMachine SP6 kit (Ambion, Austin, TX). Plasmids were linearized and purified by precipitation. Transcribed sequence reactions were treated with DNase1 and mRNA was purified with lithium chloride. mRNA quality and quantity were confirmed with a formaldehyde agarose gel and the Nanodrop8000, respectively.

Zebrafish maintenance and injections

All procedures were approved by the Johns Hopkins University Animal Care and Use Committee, protocol no. FI15M197. Zebrafish were raised in the FINZ center at the Institute for Genetic Medicine (Johns Hopkins University) as described previously [48]. Zebrafish were maintained at 28°C. Male and female Tubingen zebrafish were placed in the same breeding tank in the morning and embryos were collected 30 minutes later. One hundred embryos were then injected at the 1-4 cell blastula stage using a Zeiss Stemi 2000 microscope and PV820 Pneumatic picopump injector. All genes were injected at 50pg or 100pg and most were injected a second time at a different dose (Supplemental. Table 2). All of those producing a phenotype were re-injected at 100pg. Embryos were raised to 5 days post fertilization and then phenotyped using a Nikon SMZ1500 microscope and imaged with NIS Elements Imaging Software. After imaging, embryos were fixed in 4% PFA overnight then transferred to 100% Methanol for storage at -20°C. For low dosage experiments, *SOD1* and *RRP1* were injected at 2pg, 5pg and 10pg and examined at 5 dpf.

Morpholino Rescue

Translation-blocking antisense morpholinos (MO) were designed against the human sequence for the genes *SOD1*, *RWDD2B*, and *CCT8*, designed to bind to the ATG start codon of the mRNA using Gene Tools (Philomath, OR): Hs-*SOD1* 5'-GCACGCACACGGCCTTCGTCGCCAT-3'; Hs-*RWDD2B* 5'-GCTGCATGGACAGCTCAATTTTCAT-3'; and Hs-*CCT8* 5'-

GAGCCTTGGGAACGTGAAGCGCCAT-3'. The MOs were checked using BLAST for sequence specificity to the human homolog and insure that they were unique in the either the human or zebrafish genomes. For each gene, 100 embryos were injected with 2ng MO, 100 embryos were injected with 100pg of mRNA and 100 embryos were injected with both 2ng MO and 100pg mRNA; 100 uninjected embryos were used as a control. Embryos were examined at 5 dpf for Hs-*SOD1* and Hs-*RWDD2B*, 4 dpf for Hs-*CCT8*, and 24 hrs. post-fertilization (hpf) for Dr-*JAM2*.

Combinatorial Injections:

RNA from *C21ORF84* was coinjected with RNA from the following genes: *SOD1*, *RWDD2B*, *RRP1*, *PCBP3*, *POFUT2*, and *YBEY*. Each gene was injected individually at 100pg into 100 embryos, and then coinjected at 100pg of each RNA (200pg RNA). Embryos were phenotyped at 24 hpf for the presence of U shaped somites and cyclopia. Each coinjection was performed twice. The entire experiment was carried out independently at a different institution (NAZ and CCL) injecting 200 pg mRNA into 50-100 embryos of the Tubingen line using the pairs listed and phenotyping at 24 hpf.

The combinatorial strategy was repeated using *SOD1* as the reference gene and coinjected with the same genes listed above. In this case, 50pg of each RNA was injected individually into 100 embryos each, and then 50pg each of both RNAs were coinjected into 100 embryos, with 100 control embryos. Embryos were examined at 24hpf for the presence of U-shaped somites.

1 Statistical Tests

2 For all injections, penetrance differences were examined using a Fisher's Exact test
3 with $p < 0.05$ required for significance.

4

5

Works cited

1. Parker SE, Mai CT, Canfield MA, Rickard R, Wang Y, Meyer RE, et al. Updated National Birth Prevalence estimates for selected birth defects in the United States, 2004-2006. *Birth Defects Res A Clin Mol Teratol.* 2010;88(12):1008-16. Epub 2010/09/30. doi: 10.1002/bdra.20735. PubMed PMID: 20878909.
2. Kahlem P, Sultan M, Herwig R, Steinfath M, Balzereit D, Eppens B, et al. Transcript level alterations reflect gene dosage effects across multiple tissues in a mouse model of down syndrome. *Genome research.* 2004;14(7):1258-67. PubMed PMID: 15231742.
3. Deutsch S, Lyle R, Dermitzakis ET, Attar H, Subrahmanyam L, Gehrig C, et al. Gene expression variation and expression quantitative trait mapping of human chromosome 21 genes. *Hum Mol Genet.* 2005;14(23):3741-9. PubMed PMID: 16251198.
4. Epstein CJ, Korenberg JR, Anneren G, Antonarakis SE, Ayme S, Courchesne E, et al. Protocols to establish genotype-phenotype correlations in Down syndrome. *Am J Hum Genet.* 1991;49(1):207-35. Epub 1991/07/01. PubMed PMID: 1829580; PubMed Central PMCID: PMC1683195.
5. Ait Yahya-Graison E, Aubert J, Dauphinot L, Rivals I, Prieur M, Golfier G, et al. Classification of human chromosome 21 gene-expression variations in Down syndrome: impact on disease phenotypes. *Am J Hum Genet.* 2007;81(3):475-91. Epub 2007/08/19. doi: S0002-9297(07)61345-7 [pii] 10.1086/520000. PubMed PMID: 17701894.
6. Roper RJ, Reeves RH. Understanding the basis for Down syndrome phenotypes. *PLoS Genet.* 2006;2(3):e50. Epub 2006/04/06. doi: 10.1371/journal.pgen.0020050. PubMed PMID: 16596169; PubMed Central PMCID: PMC1420680.
7. Yang Q, Sherman SL, Hassold TJ, Allran K, Taft L, Pettay D, et al. Risk factors for trisomy 21: maternal cigarette smoking and oral contraceptive use in a population-based case-control study. *Genetics in Medicine.* 1999;1(3):80-8. PubMed PMID: 11336457.
8. Li H, Cherry S, Klinedinst D, DeLeon D, Redig J, Reshey B, et al. Genetic modifiers predisposing to congenital heart disease in a sensitized population. *Circ Cardiovasc Genet.* 2012;5((3)):301-8. doi: 10.1161.
9. Locke AE, Dooley KJ, Tinker SW, Cheong SY, Feingold E, Allen EG, et al. Variation in folate pathway genes contributes to risk of congenital heart defects among individuals with Down syndrome. *Genet Epidemiol.* 2010;34(6):613-23. Epub 2010/08/19. doi: 10.1002/gepi.20518 [doi]. PubMed PMID: 20718043.
10. Sturgeon X, Gardiner KJ. Transcript catalogs of human chromosome 21 and orthologous chimpanzee and mouse regions. *Mamm Genome.* 2011;22(5-6):261-71. Epub 2011/03/15. doi: 10.1007/s00335-011-9321-y [doi]. PubMed PMID: 21400203.
11. Potier MC, Rivals I, Mercier G, Ettwiller L, Moldrich RX, Laffaire J, et al. Transcriptional disruptions in Down syndrome: a case study in the Ts1Cje mouse cerebellum during post-natal development. *J Neurochem.* 2006;97 Suppl 1:104-9. PubMed PMID: 16635258.
12. Gardiner K, Herault Y, Lott IT, Antonarakis SE, Reeves RH, Dierssen M. Down syndrome: from understanding the neurobiology to therapy. *J Neurosci.* 2010;30(45):14943-5. Epub 2010/11/12. doi: 10.1523/JNEUROSCI.3728-10.2010 [doi]. PubMed PMID: 21068296.

13. Salehi A, Faizi M, Belichenko PV, Mobley WC. Using mouse models to explore genotype-phenotype relationship in Down syndrome. *Ment Retard Dev Disabil Res Rev.* 2007;13(3):207-14. PubMed PMID: 17910089.
14. Herault Y, Delabar JM, Fisher EMC, Tybulewicz VLJ, Yu E, Brault V. Rodent models in Down syndrome research: impact and future opportunities. *Dis Model Mech.* 2017;10(10):1165-86. Epub 2017/10/11. doi: 10.1242/dmm.029728. PubMed PMID: 28993310.
15. Nordquist SK, Smith SR, Pierce JT. Systematic Functional Characterization of Human 21st Chromosome Orthologs in *Caenorhabditis elegans*. *G3 (Bethesda, Md).* 2018. Epub 2018/01/26. doi: 10.1534/g3.118.200019. PubMed PMID: 29367452.
16. Lopez-Rivera E, Liu YP, Verbitsky M, Anderson BR, Capone VP, Otto EA, et al. Genetic Drivers of Kidney Defects in the DiGeorge Syndrome. *N Engl J Med.* 2017;376(8):742-54. Epub 2017/01/26. doi: 10.1056/NEJMoa1609009. PubMed PMID: 28121514.
17. Carvalho CM, Vasanth S, Shinawi M, Russell C, Ramocki MB, Brown CW, et al. Dosage changes of a segment at 17p13.1 lead to intellectual disability and microcephaly as a result of complex genetic interaction of multiple genes. *Am J Hum Genet.* 2014;95(5):565-78. Epub 2014/12/03. doi: 10.1016/j.ajhg.2014.10.006. PubMed PMID: 25439725; PubMed Central PMCID: PMC4225592.
18. Golzio C, Katsanis N. Genetic architecture of reciprocal CNVs. *Curr Opin Genet Dev.* 2013;23(3):240-8. Epub 2013/06/12. doi: 10.1016/j.gde.2013.04.013. PubMed PMID: 23747035; PubMed Central PMCID: PMC3740179.
19. Dauber A, Golzio C, Guenot C, Jodelka FM, Kibaek M, Kjaergaard S, et al. SCRIB and PUF60 are primary drivers of the multisystemic phenotypes of the 8q24.3 copy-number variant. *Am J Hum Genet.* 2013;93(5):798-811. Epub 2013/10/22. doi: 10.1016/j.ajhg.2013.09.010. PubMed PMID: 24140112; PubMed Central PMCID: PMC3824129.
20. Golzio C, Willer J, Talkowski ME, Oh EC, Taniguchi Y, Jacquemont S, et al. KCTD13 is a major driver of mirrored neuroanatomical phenotypes of the 16p11.2 copy number variant. *Nature.* 2012;485(7398):363-7. Epub 2012/05/19. doi: 10.1038/nature11091. PubMed PMID: 22596160; PubMed Central PMCID: PMC3366115.
21. Gardiner K, Fortna A, Bechtel L, Davisson MT. Mouse models of Down syndrome: how useful can they be? Comparison of the gene content of human chromosome 21 with orthologous mouse genomic regions. *Gene.* 2003;318:137-47. PubMed PMID: 14585506.
22. Rupp RA, Snider L, Weintraub H. *Xenopus* embryos regulate the nuclear localization of XMyoD. *Genes Dev.* 1994;8(11):1311-23. Epub 1994/06/01. PubMed PMID: 7926732.
23. Motenko H, Neuhauser SB, O'Keefe M, Richardson JE. MouseMine: a new data warehouse for MGI. *Mamm Genome.* 2015;26(7-8):325-30. Epub 2015/06/21. doi: 10.1007/s00335-015-9573-z. PubMed PMID: 26092688; PubMed Central PMCID: PMC4534495.
24. Mi H, Dong Q, Muruganujan A, Gaudet P, Lewis S, Thomas PD. PANTHER version 7: improved phylogenetic trees, orthologs and collaboration with the Gene Ontology Consortium. *Nucleic Acids Res.* 2010;38(Database issue):D204-10. Epub 2009/12/18. doi: 10.1093/nar/gkp1019. PubMed PMID: 20015972; PubMed Central PMCID: PMC2808919.
25. Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, et al. The zebrafish reference genome sequence and its relationship to the human genome. *Nature.* 2013;496(7446):498-503. Epub 2013/04/19. doi: 10.1038/nature12111. PubMed PMID: 23594743; PubMed Central PMCID: PMC3703927.

26. Powell GT, Wright GJ. Jamb and janc are essential for vertebrate myocyte fusion. *PLoS Biol.* 2011;9(12):e1001216. Epub 2011/12/20. doi: 10.1371/journal.pbio.1001216
PBIOLGY-D-11-02322 [pii]. PubMed PMID: 22180726.
27. Li H, Edie S, Klinedinst D, Jeong JS, Blackshaw S, Maslen CL, et al. Penetrance of Congenital Heart Disease in a Mouse Model of Down Syndrome Depends on a Trisomic Potentiator of a Disomic Modifier. *Genetics.* 2016;203(2):763-70. Epub 2016/04/01. doi: 10.1534/genetics.116.188045. PubMed PMID: 27029737; PubMed Central PMCID: PMC4896192.
28. Takeda H, Matsuzaki T, Oki T, Miyagawa T, Amanuma H. A novel POU domain gene, zebrafish pou2: expression and roles of two alternatively spliced twin products in early development. *Genes Dev.* 1994;8(1):45-59. Epub 1994/01/01. PubMed PMID: 8288127.
29. Thisse B, Thisse C. Fast Release Clones: A High Throughput Expression Analysis. *ZFIN Direct Data Submission* (<http://zfin.org>). 2004.
30. Ohata S, Kinoshita S, Aoki R, Tanaka H, Wada H, Tsuruoka-Kinoshita S, et al. Neuroepithelial cells require fucosylated glycans to guide the migration of vagus motor neuron progenitors in the developing zebrafish hindbrain. *Development.* 2009;136(10):1653-63. Epub 2009/04/17. doi: dev.033290 [pii]
10.1242/dev.033290. PubMed PMID: 19369395.
31. Reymond A, Marigo V, Yaylaoglu MB, Leoni A, UCLA C, Scamuffa N, et al. Human chromosome 21 gene expression atlas in the mouse. *Nature.* 2002;420(6915):582-6.
32. Armit C, Venkataraman S, Richardson L, Stevenson P, Moss J, Graham L, et al. eMouseAtlas, EMAGE, and the spatial dimension of the transcriptome. *Mamm Genome.* 2012;23(9-10):514-24. Epub 2012/08/01. doi: 10.1007/s00335-012-9407-1. PubMed PMID: 22847374.
33. Ferencz C, Neill CA, Boughman JA, Rubin JD, Brenner JJ, Perry LW. Congenital cardiovascular malformations associated with chromosome abnormalities: an epidemiologic study. *The Journal of pediatrics.* 1989;114(1):79-86. PubMed PMID: 2521249.
34. Gitton Y, Dahmane N, Baik S, Ruiz i Altaba A, Neidhardt L, Scholze M, et al. A gene expression map of human chromosome 21 orthologues in the mouse. *Nature.* 2002;420(6915):586-90.
35. Brown SD, Moore MW. The International Mouse Phenotyping Consortium: past and future perspectives on mouse phenotyping. *Mamm Genome.* 2012;23(9-10):632-40. Epub 2012/09/04. doi: 10.1007/s00335-012-9427-x. PubMed PMID: 22940749; PubMed Central PMCID: PMC3774932.
36. Roper R, Reeves R. Understanding the basis for Down syndrome phenotypes. *PLoS Genetics.* 2006;2:231-6.
37. Das I, Reeves RH. The use of mouse models to understand and improve cognitive deficits in Down syndrome. *Dis Model Mech.* 2011;4(5):596-606. Epub 2011/08/06. doi: dmm.007716 [pii]
10.1242/dmm.007716 [doi]. PubMed PMID: 21816951.
38. Li H, Cherry S, Klinedinst D, DeLeon V, Redig J, Reshey B, et al. Genetic modifiers predisposing to congenital heart disease in the sensitized Down syndrome population. *Circ Cardiovasc Genet.* 2012;5(3):301-8. Epub 2012/04/24. doi: CIRCGENETICS.111.960872 [pii]
10.1161/CIRCGENETICS.111.960872 [doi]. PubMed PMID: 22523272; PubMed Central PMCID: PMC3386785.

39. Garrett AM, Tadenev AL, Burgess RW. DSCAMs: restoring balance to developmental forces. *Front Mol Neurosci*. 2012;5:86. Epub 2012/08/23. doi: 10.3389/fnmol.2012.00086. PubMed PMID: 22912601.
40. Grossman TR, Gamliel A, Wessells RJ, Taghli-Lamalle O, Jepsen K, Ocorr K, et al. Over-expression of DSCAM and COL6A2 cooperatively generates congenital heart defects. *PLoS Genet*. 2011;7(11):e1002344. Epub 2011/11/11. doi: 10.1371/journal.pgen.1002344 [doi] PGENETICS-D-11-01353 [pii]. PubMed PMID: 22072978; PubMed Central PMCID: PMC3207880.
41. Zhu K, Xu Y, Liu J, Xu Q, Ye H. Down syndrome cell adhesion molecule and its functions in neural development. *Neurosci Bull*. 2011;27(1):45-52. Epub 2011/01/29. doi: 10.1007/s12264-011-1045-1. PubMed PMID: 21270903.
42. de la Torre R, de Sola S, Hernandez G, Farre M, Pujol J, Rodriguez J, et al. Safety and efficacy of cognitive training plus epigallocatechin-3-gallate in young adults with Down's syndrome (TESDAD): a double-blind, randomised, placebo-controlled, phase 2 trial. *The Lancet Neurology*. 2016;15(8):801-10. Epub 2016/06/16. doi: 10.1016/s1474-4422(16)30034-5. PubMed PMID: 27302362.
43. De la Torre R, De Sola S, Pons M, Duchon A, de Lagran MM, Farre M, et al. Epigallocatechin-3-gallate, a DYRK1A inhibitor, rescues cognitive deficits in Down syndrome mouse models and in humans. *Mol Nutr Food Res*. 2014;58(2):278-88. Epub 2013/09/17. doi: 10.1002/mnfr.201300325 [doi]. PubMed PMID: 24039182.
44. Currier DG, Polk RC, Reeves RH. A Sonic hedgehog (Shh) response deficit in trisomic cells may be a common denominator for multiple features of Down syndrome. *Progress in brain research*. 2012;197:223-36. Epub 2012/05/01. doi: 10.1016/b978-0-444-54299-1.00011-x. PubMed PMID: 22541295; PubMed Central PMCID: PMC4405118.
45. Das I, Park JM, Shin JH, Jeon SK, Lorenzi H, Linden DJ, et al. Hedgehog agonist therapy corrects structural and cognitive deficits in a Down syndrome mouse model. *Science translational medicine*. 2013;5(201):201ra120. Epub 2013/09/06. doi: 10.1126/scitranslmed.3005983. PubMed PMID: 24005160; PubMed Central PMCID: PMC4006719.
46. Roper RJ, VanHorn JF, Cain CC, Reeves RH. A neural crest deficit in Down syndrome mice is associated with deficient mitotic response to Sonic hedgehog. *Mech Dev*. 2009;126(3-4):212-9. PubMed PMID: 19056491.
47. Katzen F. Gateway((R)) recombinational cloning: a biological operating system. *Expert Opin Drug Discov*. 2007;2(4):571-89. Epub 2007/04/01. doi: 10.1517/17460441.2.4.571. PubMed PMID: 23484762.
48. Westerfield M. *The Zebrafish Book: a guide for the laboratory use of zebrafish (Danio rerio)*. 5 ed. Eugene, OR: University of Oregon Press; 2000 2007.
49. Cunningham SA, Arrate MP, Rodriguez JM, Bjerkke RJ, Vanderslice P, Morris AP, et al. A novel protein with homology to the junctional adhesion molecule. Characterization of leukocyte interactions. *J Biol Chem*. 2000;275(44):34750-6. Epub 2000/08/18. doi: 10.1074/jbc.M002718200. PubMed PMID: 10945976.

Acknowledgments

We thank Valerie DeLeon, Deborah Andrew, and Steven Leach for their advice and support. This work is submitted in partial fulfillment of the requirements for a Ph.D. (SE).

Author Contributions

SE, RR, NK and AM conceived and designed the experiments. SE, NAZ, CCL, DK, JT and JL performed the experiments. SE, NAZ, RR and AM analyzed the data. SE and RR wrote the manuscript, which was reviewed and revised by all authors.

Figure legends:

Figure 1: Flowchart showing steps of making the Hsa21 Gene Expression Clone-set and the screen of the set in zebrafish

Figure 2: Examples of phenotypes observed in screen. Control embryos are on the left panel and injected embryos are on the right panel. Somites: *RWDD2B* 100pg injected embryos at 24 hpf with dashed lines to highlight somitic boundaries. Cyclopia: *C21ORF84* 100pg injected embryos at 5 dpf. Pigment cell migration: *CCT8* 100pg injected embryos at 4 dpf, arrows indicating melanocytes. Heart: *JAM2* 100pg injected embryos at 48 hpf.

Figure 3: Candidate genes (*SOD1*, *RWDD2B*, *CCT8*, and *JAM2*) coinjected with translational blocking morpholino. Hs-*RWDD2B*, Hs-*CCT8* and Hs-*SOD1* MOs were targeted against the human mRNA, while *jam2* MOs were targeted against the zebrafish ortholog, DR-jam2. 100pg RNA was injected alone, 2 ng MO alone, or both were coinjected. * p<0.05, ** p<0.01. *JAM2* data adapted from [27].

Figure 4: Pairwise combinatorial injections of Shh candidate genes. *C21ORF84* was coinjected with 6 other genes to look for synthetic effects. *C21ORF84* was injected individually at 100pg RNA, the other gene was injected individually at 100pg RNA and then the two were injected together, 100pg each, for a total of 200pg RNA.

Supplemental Figure 1: Human mRNAs did not show a dose response at the concentrations used. Phenotype penetrance is shown for the ten candidate genes at concentrations from 10-200 pg/embryo.

Supplemental Figure 2: SOD1 and RRP1 were examined for penetrance of the phenotype at very low RNA doses but did not show significant penetrance. Uninjected embryos for controls. * $p < 0.01$

Supplemental Figure 3: Pairwise injections of Shh related genes using *SOD1* as the reference gene. *SOD1* was coinjected with 6 other genes to look for synthetic effects. *SOD1* was injected individually at 50pg RNA, the other gene was injected individually at 100pg RNA and then the two were injected together, 50pg each, for a total of 100pg RNA.

1 Table 1: Twenty-four candidates from first pass of screen (see Supplemental Table 3)

Phenotype category	Description	# Candidates
U-shaped Somites	Somites with characteristic U shape	8
Cyclopia	Single large eye	6
Craniofacial abnormalities	Small/missing mandible; skull abnormalities	4
Pigment abnormalities	Floating melanocytes; reduced pigment in eye	3
Heart	Pericardial edema	1
Other	Tail/fin abnormalities, embryo lethality	2

2
3

Table 2: Final Candidate list of genes that produced a phenotype consistently (see also Supplemental Table 2).

Gene Symbol	Phenotype	Penetrance	Expression in mouse [31]
<i>SOD1</i>	U-somites	10-35%	Expressed ubiquitously at E10.5, strongly expressed in muscles at E14.5
<i>RWDD2B</i>	U-somites	15-31%	Expressed ubiquitously at E10.5
<i>RRP1</i>	U-somites	15-40%	Weakly expressed in somites at E10.5 [32]
<i>PCBP3</i>	U-somites	8-13%	Expressed in brain and spinal cord at E14.5
<i>YBEY</i>	U-somites	12-19%	Expressed ubiquitously at E10.5
<i>C21orf84</i>	U-somites/ Cyclopia	9-23% U-somites 0-7% cyclopia	human specific [21]
<i>POFUT2</i>	Cyclopia	0-4%	Strong in face and pharyngeal arches at E9.5 [32]
<i>CBR3</i>	Craniofacial	7-10%	Expressed ubiquitously at E10.5, strong in cartilage at E14.5
<i>CCT8</i>	Pigment	15-40%	Expressed ubiquitously at E10.5
<i>JAM2</i>	Pericardial edema	20-60%	High expression in human heart [49]
<i>ERG</i>	Embryonic lethal	64-93%	Expressed in the pharyngeal arches and limb buds at E10.5 [17]













