Global Analysis of Gene Expression in the Developing Brain of Gtf2ird1 Knockout Mice

Jennifer O’Leary¹, Lucy R. Osborne¹,²*
¹ Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada, ²Department of Medicine, University of Toronto, Toronto, Ontario, Canada

Abstract

Background: Williams-Beuren Syndrome (WBS) is a neurodevelopmental disorder caused by a hemizygous deletion of a 1.5 Mb region on chromosome 7q11.23 encompassing 26 genes. One of these genes, GTF2IRD1, codes for a putative transcription factor that is expressed throughout the brain during development. Genotype-phenotype studies in patients with atypical deletions of 7q11.23 implicate this gene in the neurological features of WBS, and Gtf2ird1 knockout mice show reduced innate fear and increased sociability, consistent with features of WBS. Multiple studies have identified in vitro target genes of GTF2IRD1, but we sought to identify in vivo targets in the mouse brain.

Methodology/Principal Findings: We performed the first in vivo microarray screen for transcriptional targets of Gtf2ird1 in brain tissue from Gtf2ird1 knockout and wildtype mice at embryonic day 15.5 and at birth. Changes in gene expression in the mutant mice were moderate (0.5 to 2.5 fold) and of candidate genes with altered expression verified using real-time PCR, most were located on chromosome 5, within 10 Mb of Gtf2ird1. siRNA knock-down of Gtf2ird1 in two mouse neuronal cell lines failed to identify changes in expression of any of the genes identified from the microarray and subsequent analysis showed that differences in expression of genes on chromosome 5 were the result of retention of that chromosome region from the targeted embryonic stem cell line, and so were dependent upon strain rather than Gtf2ird1 genotype. In addition, specific analysis of genes previously identified as direct in vitro targets of GTF2IRD1 failed to show altered expression.

Conclusions/Significance: We have been unable to identify any in vivo neuronal targets of GTF2IRD1 through genome-wide expression analysis, despite widespread and robust expression of this protein in the developing rodent brain.

Introduction

Williams-Beuren Syndrome (WBS; OMIM 194050) is an autosomal dominant microdeletion disorder, which occurs at a frequency of about 1/7,500 live births, as a result of the deletion of a 1.55 Mb region from chromosome 7q11.23 [1,2]. The region is flanked by low copy repeats that, in rare instances, will undergo unequal recombination during meiosis, resulting in chromosomes with a deletion or duplication of the region [2]. The clinical symptoms of WBS are numerous, but commonly include supravalvular aortic stenosis (SVAS), dysmorphic facial features, hypersensitivity to sound, retarded growth, stellate patterning of the iris, renal and dental abnormalities and infantile hypercalcemia [2]. The complex phenotype also includes behavioural and cognitive components. WBS is characterized by mild to moderate intellectual disability with relative strengths in verbal short-term memory and in language and extreme weakness in visuospatial construction [3,4], as well as anxiety, attention deficit hyperactivity disorder (ADHD) and overfriendliness [5,6].

Of the more than 25 genes that are commonly deleted in WBS, only one has been definitively linked to a phenotype. Hemierysmyosity for, or point mutations in, the elastin gene (ELN) cause SVAS [7–9]. Studies of individuals with atypical deletions in the WBS region have implicated genes toward the telomeric end of the deletion [10,11], and in particular the two transcription factor genes, General Transcription factor 2 I (GTF2I) and GTF2I Repeat Domain containing protein 1 (GTF2IRD1) [12–18] with the majority of symptoms, including the cognitive and behavioural phenotypes. We have previously generated Gtf2ird1¹⁰⁻¹⁰ knockout mice (Gtf2ird1¹⁰⁻¹⁰) that show low innate anxiety, decreased aggression and increased social interaction [19], a phenotype that shares similarities to the high sociability and disinhibition seen in individuals with WBS [5]. Although no overt differences in brain morphology were noted, adult mice had selectively enhanced serotonin 1A receptor-mediated responses in layer V pyramidal neurons of the pre-frontal cortex, suggesting that brain function may be significantly impaired [20]. GTF2IRD1 is preferentially expressed in layer V of the adult cortex [20,21] but shows a more widespread distribution and higher expression during pre- and early post-natal brain development. GTF2IRD1 is expressed throughout pre- and post-implantation embryogenesis [22] [Allen Brain Atlas, http://www.brain-map.org/] and analysis of mice expressing a LacZ knock-in reporter showed expression throughout the midbrain and hindbrain during the second half of gestation [23].
Although GTF2IRD1 has been proposed to function as a transcription factor [24–26], few direct targets have been identified, and to date these have not included plausible candidates for the cognitive and behavioural phenotype seen in either individuals with WBS, or Gtf2ird1<sup>tm1Lro/tm1Lro</sup> mice. Using yeast-one hybrids and/or in vitro expression analysis, GTF2IRD1 has been implicated in the regulation of Tnni1, Hoxa8, Gsc, and Vegfa2 [24,25,27–29]. Genome-wide analysis in transformed mouse embryonic fibroblast (MEF) lines identified 4,678 genes whose expression was altered upon over-expression of GTF2IRD1 [30,31], although only a very small subset of these was validated by quantitative PCR [31]. In addition, GTF2IRD1 has been shown to bind to two distinct DNA consensus sequences, GTCGAGATTAGBGA and CWGCCAYA, both of which are found in the promoter regions of some of the proposed target genes [26,32,33].

Given the presumptive role of GTF2IRD1 in the cognitive and behavioural aspects of the WBS phenotype, the identification of neural targets of this transcription factor would shed light on the molecular pathways that are perturbed in WBS. Thus, we undertook an in vivo screen to identify targets of GTF2IRD1 by performing genome-wide microarray analyses of brain from Gtf2ird1<sup>+/−</sup> mice at two different developmental time points. We were not able to identify any genes whose expression was altered by lack of GTF2IRD1, nor were we able to validate altered in vitro alterations in expression of target genes previously identified in vitro.

Results

Global expression analysis of Gtf2ird1<sup>+/−</sup> mouse brains

Microarray analysis of whole mouse brain was performed at two different time points, embryonic day (E) 15.5 and post-natal day (P) 0, to identify direct and indirect targets of GTF2IRD1 in an unbiased manner. For P0 mice, cRNA was prepared from three pools each of total RNA (nine mice per pool) extracted from Gtf2ird1<sup>+/−</sup> and WT mice, and hybridized to the GeneChip Mouse Genome 430 2.0 Array (Affymetrix, Santa Clara, CA). Signals obtained were normalized using Robust Multiarray Analysis (RMA) [34], and differences in gene expression were detected using a second software program, Significance Analysis of Microarrays (SAM) [35]. The data was submitted to the ArrayExpress database under accession number E-MTAB-622.

Relatively few transcripts showed altered expression, and the magnitudes of these changes were generally small (Table S1). Using a false discovery rate (FDR) cut-off of 10%, eight genes were identified as having changes in expression greater than 2-fold in the Gtf2ird1<sup>+/−</sup> mice compared to WT littermates. An additional 79 genes had changes in expression between 1.2 and 2-fold.

For E15.5 mice, cRNA was prepared from total RNA extracted from the heads of five Gtf2ird1<sup>+/−</sup> embryos and five WT littermates, and hybridized to the Mouse WG-6 v2.0 Expression BeadChip (Illumina, San Diego, CA). Differentially expressed genes were detected using LIMMA [36], following normalization of the log2 scale transformed data using the quantile normalization method [37]. The data was submitted to the ArrayExpress database under accession number E-MTAB-626.

Eighteen transcripts were found to have altered expression in Gtf2ird1<sup>+/−</sup> mice, with an adjusted p-value of <0.1 (Table S2). Similar to the results from array analysis of P0 mouse brain, the changes in expression detected in E15.5 embryos were generally small with approximately half of the identified genes showing a change of less than 2-fold.

\[\text{M}ospd3\] and \[\text{Auts2}\] were the only genes identified in both microarray experiments, however, in P0 mouse brains Auts2 was found to be increased by 1.3 fold, while in E15.5 embryos it was found to be decreased by 1.5 fold.

Validation of microarray results

Quantitative reverse transcription PCR (qRT-PCR) was used to validate alterations in gene expression detected in Gtf2ird1<sup>+/−</sup> mouse brain using genome-wide microarray analysis. Where possible, qRT-PCR primer pairs in which one of the primers overlapped with the microarray probe sequence were used. The majority of expression changes in protein coding genes could not be validated. Where gene expression was found to be significantly different between genotypes, the changes in expression were generally small, in concordance with the microarray results (Figure 1A). Genes that showed the largest changes in expression were Picl, Sst3 and Mphp16 with alterations of 3.57, 2.95, and 2.65 fold respectively, and AI506816 which was not expressed at all in Gtf2ird1<sup>+/−</sup> mice.

qRT-PCR also failed to confirm altered expression of the majority of genes found to be differentially expressed in the E15.5 embryos (Figure 1B). Significant differences in expression between genotypes were only detected for seven genes, with the largest changes in expression seen in \[\text{Actb6b}, \text{Taf6}\] and \[\text{Zfp68}\], with alterations of 1.97, 1.63 and 1.39 respectively.

Altered expression detected in Gtf2ird1<sup>+/−</sup> mice was due to a large genetic background effect

Many of the genes identified by microarray as being altered in the P0 or E15.5 mice were located on distal chromosome 5, most within 10 Mb of the Gtf2ird1 locus, and the genes with the largest changes in expression in E15.5 embryos, \[\text{Actb6b}, \text{Taf6}\], and \[\text{Zfp68}\], were all within 5 Mb of the Gtf2ird1 locus. The initial targeting of Gtf2ird1 was performed in R1 embryonic stem (ES) cells (which are derived from a 129X1/SvJ and 129S1 cross) and the mice were backcrossed onto CD1 genetic background.

Analysis of strain-specific SNPs within the 3’ UTR of \[\text{Zfp68}\] revealed that Gtf2ird1<sup>+/−</sup> mice were homozygous for 12 of 13 129S1/SvImJ SNPs, while WT mice were only similar to the 129S1/SvImJ mice at 2 of 13 SNPs (Table 1). Although the 129S1/SvImJ mice are genetically very similar to the R1 line, they are not identical, which is likely why one of the 129S1/SvImJ SNP alleles was not present in the Gtf2ird1<sup>+/−</sup> mice. These results indicate that a segment of chromosome 5 from the ES cell line strain spanning \[\text{Zfp68}\] and Gtf2ird1, and a minimum of 4.5 Mb, had been retained in the Gtf2ird1<sup>+/−</sup> mice used for the E15.5 expression array.

qRT-PCR was then performed on whole brains from P0 mice to determine if the expression differences detected between Gtf2ird1<sup>+/−</sup> and WT mice were actually the result of strain specific differences in expression between the CD1 and 129 genetic backgrounds. Expression of genes on chromosome 5 was found to be the same in Gtf2ird1<sup>+/−</sup> mice as in WT 129S1/SvImJ mice, and significantly different from that seen in CD1 WT mice (Figure 2). Sequencing of the PCR amplicons was performed to rule out differences in PCR efficiencies caused by strain specific SNPs that might have existed within the primer binding sequences. These results demonstrated that differences in expression detected in Gtf2ird1<sup>+/−</sup> mouse brain, relative to WT littermates, were not dependent upon Gtf2ird1<sup>+/−</sup> genotype and were not, therefore, the result of a lack of GTF2IRD1.

Expression of AI506816 and \[\text{Pesl}\] had previously been validated as significantly different between P0 Gtf2ird1<sup>+/−</sup> and WT CD1 mice, but upon qRT-PCR expression analysis to compare background differences, no change in expression between
Gtf2ird1<sup>−/−</sup> and WT mice were identified. However, the altered expression values in Gtf2ird1<sup>−/−</sup> mice from the original validation, mirrored those seen in WT 129S1/SvImJ mice, suggesting that although not located close to Gtf2ird1, the portion of chromosome 5 containing these adjacent genes was derived from the R1 ES cell line in the mice used for the microarray experiment.

Differences in genetic background could not explain the expression differences detected in Mrpl16 and Stx3 as they showed significantly decreased expression in Gtf2ird1<sup>−/−</sup> mice relative to both WT CD1 mice and WT 129S1/SvImJ mice (Figure 2).

siRNA knockdown of Gtf2ird1 in neuronal cell lines does not affect expression of genes identified from microarray analysis.

To confirm whether expression differences detected for genes on chromosome 5 were due to the presence of flanking DNA from

---

**Table 1.** A comparison of SNPs in the 3′ UTR Zfp68 in Gtf2ird1<sup>−/−</sup> mice and WT mice relative to 129S1/Sv1mJ mice.

| SNP ID (dbSNP build 128) | Map Position (NCBI Build 37) | 129S1/Sv1mJ | Allele Summary (all strains) | Gtf2ird1<sup>−/−</sup> | WT |
|--------------------------|-----------------------------|-------------|-----------------------------|-------------------------|-----|
| rs38291282               | Chr5: 139045176             | T           | A/T                         | T                       | A   |
| rs37309562               | Chr5: 139045486             | T           | C/T                         | T                       | C   |
| rs36257731               | Chr5: 139045921             | G           | G/T                         | G                       | T   |
| rs36653430               | Chr5: 139045957             | C           | C/T                         | C                       | T   |
| rs34458231               | Chr5: 139045989             | G           | G/T                         | G                       | G   |
| rs39288035               | Chr5: 139046020             | T           | G/T                         | T                       | G   |
| rs39363663               | Chr5: 139046416             | C           | C/G                         | C                       | G   |
| rs36967976               | Chr5: 139046738             | T           | C/T                         | T                       | T   |
| rs36713208               | Chr5: 139046762             | G           | C/G                         | C                       | C   |
| rs33378677               | Chr5: 139047295             | T           | C/T                         | T                       | C   |
| rs33527435               | Chr5: 139047340             | A           | A/G                         | A                       | G   |
| rs33143922               | Chr5: 139047382             | T           | A/T                         | T                       | T   |
| rs29551605               | Chr5: 139047411             | A           | A/G                         | A                       | G   |

---

Figure 1. qRT-PCR validation of expression changes identified by microarray. Expression values were normalized to the housekeeping gene Sdha, and are depicted as a ratio of expression in Gtf2ird1<sup>−/−</sup> mice relative to WT littermates. * p < 0.05, ** p < 0.005 using Student’s t-test.

A. Expression in whole brain from P0 mice. RNA from 9 mice of the same genotype was pooled together to make cDNA (n = 3 pools/genotype).

B. Expression in heads of E15.5 mouse embryos (n = 5 genotype).

doi:10.1371/journal.pone.0023868.g001

**Figure 1.** qRT-PCR validation of expression changes identified by microarray. Expression values were normalized to the housekeeping gene Sdha, and are depicted as a ratio of expression in Gtf2ird1<sup>−/−</sup> mice relative to WT littermates. * p < 0.05, ** p < 0.005 using Student’s t-test.

A. Expression in whole brain from P0 mice. RNA from 9 mice of the same genotype was pooled together to make cDNA (n = 3 pools/genotype).

B. Expression in heads of E15.5 mouse embryos (n = 5 genotype).

doi:10.1371/journal.pone.0023868.g001
the R1 ES cell line, and not specifically related to the loss of GTF2IRD1, siRNA knockdown of Gtf2ird1 was performed in two neuroblastoma derived cell lines: Neuro2A and N1E-115. Eight Gtf2ird1 siRNAs (Dharmacon, Lafayette, CO) were tested, and those that resulted in the highest levels of knockdown were combined into three different pools (A, B & C) to maximize the level of knockdown. Each pool was then tested on both N2A and N1E-115 cell lines in duplicate. Each of the three Gtf2ird1 siRNA pools specifically knocked down Gtf2ird1 expression by ~60% in Neuro2A cells and by ~80% in N1E-115 cells (Figure 3A). Treatment with a non-targeting siRNA, or an siRNA targeting Gapdh had no effect on the expression of Gtf2ird1. Gapdh siRNA treated cells showed a specific 90% reduction in Gapdh expression (data not shown).

Expression levels of 4 candidate genes identified in the microarrays and verified using qRT-PCR were analyzed in Gtf2ird1 siRNA treated cells. As there were no differences in the effects between Gtf2ird1 pools A, B, and C, gene expression was measured in cells treated with each pool (in duplicate) and the expression values averaged together (n = 6). No significant differences in the expression of Actb6, Taf6, Kin or Mrpl16 could be detected in Gtf2ird1 siRNA treated cells when compared with either non-targeting siRNA treated cells or untreated cells (Figure 3B). These results indicate that GTF2IRD1 does not play a role in the transcriptional regulation of these genes in the cell types examined, and that the expression differences detected by microarray analysis were not the result of the disruption of Gtf2ird1. Stx3 and Mrpl16 were not expressed in these cell lines and the effect of Gtf2ird1 knockdown on their expression could not be measured.

Stx3 and Mrpl16 show highly variable expression

When Stx3 and Mrpl16 expression was analyzed in P0 mice, their expression in Gtf2ird1−/− and WT mice was highly variable. These genes are located in a tail-to-tail orientation on mouse chromosome 19, with only ~150bp separating their 3’UTRs, an orientation that is conserved in humans. To determine whether Stx3 and/or Mrpl16 expression was dependent on Gtf2ird1 genotype or whether it was different in the initial Gtf2ird1−/− experimental group by chance, we examined expression in a larger cohort of mice. Expression of both genes was measured in the whole heads of five Gtf2ird1−/− and five WT E15.5 embryos. Expression of each gene was measured using two different primer sets: one pair that had been used previously and that amplified a 400bp sequence from the 3’UTR, and one pair located further upstream within an exon.

Expression of Stx3 and Mrpl16 exon sequences did not vary greatly between mice regardless of genotype, however expression of 3’UTR sequences were variable and did not correlate with genotype. Within each genotype, mice were detected that displayed either high or low expression of both Mrpl16 and Stx3 (Figure 4A). Mice with low expression in the Mrpl16 3’UTR also had low expression in the Stx3 3’UTR and vice versa, likely due to their overlapping tail-tail orientation.
Expression of Stx3 was then examined in a second cohort to determine if expression level was associated with Gtf2ird1 genotype. Expression was measured in whole brains from P0 (WT, n = 15; Gtf2ird1<sup>+/−</sup>, n = 13) and adult (WT, n = 12; Gtf2ird1<sup>+/−</sup>, n = 12) mice. While the differences in expression between WT and Gtf2ird1<sup>+/−</sup> P0 mice were marginally significantly different (p = 0.047), expression of Stx3 was still highly variable in both genotypes (Figure 4A). Variability was also detected in Stx3 expression in adult mice (Figure 4B), and the differences in expression level between these genotypes were not significant (p = 0.899).

It is likely that the differences in both Stx3 and Mrp16 expression that were detected represent natural expression variation of these genes. As the expression differences were only detected in the 3′ UTR and not in upstream coding sequences they may be the result of alternative polyadenylation site selection and/or splicing in the 3′ UTR.

In vivo expression of GTF2IRD1 in vitro target genes was not altered

Previous studies have used both a candidate gene approach and unbiased global analyses to identify target genes of GTF2IRD1, in vitro. We examined expression of some of these targets in vivo, in tissues from Gtf2ird1<sup>−/−</sup> mice using qRT-PCR. Expression of bone morphogenetic protein receptor, type 1b (Bmpr1b) and fibroblast growth factor 15 (Fgf15) were found to be highly upregulated following Gtf2ird1 knockdown in C2C12 cells [33]. To determine if GTR2IRD1 was involved in the regulation of Fgf15 and Bmpr1b in vivo, we measured expression of these genes in the brains of Gtf2ird1<sup>−/−</sup> and WT mice, at P0, a developmental timepoint when both genes were expressed. No genotype-specific differences in gene expression were identified (Figure 5A), suggesting that GTF2IRD1 does not play a role in the regulation of Bmpr1b and Fgf15 in the newborn brain.

qRT-PCR was also used to look at the in vivo expression of GTF2IRD1 candidate target genes identified in transformed MEFs [30,31]. Primary MEFs were cultured from Gtf2ird1<sup>−/−</sup> (n = 1), Gtf2ird1<sup>+/−</sup> (n = 2) and WT (n = 3) E15.5 littermate embryos, and the expression levels of seven GTF2IRD1 in vitro target genes were quantified (cyclin D3 (Cnd3); cofilin 1 (Cfl1); enhancer of zeste homolog 2 (Ezh2); fibroblast growth factor 11 (Fgf11); enhancer of polycomb homolog 1 (Epc1); secreted phosphoprotein 1 (Spp1) or Omp; transforming growth factor, beta 2 (Tgfβ2)). There were no significant differences in expression between WT and heterozygous MEFs for any of the genes studied (Figure 5B). MEFs were only cultured from one Gtf2ird1<sup>−/−</sup> embryo and so statistical analyses could not be completed, however with the possible exception of Tgfβ2, expression levels in this embryo did not differ from WT or heterozygous littersmates. Based on these results, GTF2IRD1 is unlikely to regulate expression of these genes in MEFs under the culture conditions used. Expression of five target genes (LIM homeobox protein 1 (Lhx1); Gtf1; Fgf11; Omp; Tgfβ2) was also determined in whole brain from E18.5 and adult mice but no significant differences were detected between genotypes for any of the genes (Figure 5C).

Comparison with other analyses of global gene expression

The number of genes identified in each of our experiments, and the magnitudes of the changes in expression were both smaller than expected, given that more than 4600 genes with altered expression were identified in MEFs that over-expressed GTF2IRD1 [30,31]. This microarray experiment was performed on transformed MEF cell lines which expressed Gtf2ird1 mRNA at >7-fold above endogenous levels, and appeared to show even higher levels of protein according to Western blot analysis [31]. Transformed MEFs likely have a very different expression profile than the developing brain, since they are composed of a single cell type and are undergoing rapid cell division. The different endogenous global expression profiles, coupled with contrasting levels of Gtf2ird1 expression (considerable over-expression in MEFs vs. no expression in mice) may help explain why the results of our global analysis of gene expression varied so greatly from that in MEFs. In addition, only two biological replicates were used in the analysis of MEFs, and >10-fold differences in expression were reported between these replicates for some genes [31]. We used three biological replicates of pooled samples in our first array.
experiment and five individual biological replicates in our second array experiment, to account for inter-animal differences.

A second global analysis of gene expression in vivo has been reported recently [38]. This analysis was performed on a different Gtf2ird1 knockout mouse model, Gtf2ird1(G2X/E465)Brg(G2X/E465)Brg (Gtf2ird1(G2X/E465)), generated from a gene trap ES cell line. This study found 536 genes with altered expression in E9.5 Gtf2ird1Gt/Gt embryos, however there are several caveats to the interpretation of these data. Firstly, Gtf2ird1Gt/Gt embryos die between E8.5 and E12.5, with most showing signs of being actively resorbed by E9.5 [38]. Thus, it is likely that much of the altered gene expression may have been due to processes involved in embryonic death and resorption. Secondly, the Gtf2ird1Gt/Gt mouse has a far more severe phenotype than the other four published Gtf2ird1 mouse models [19,23,39,40]. In all these other models, homozygous mice are healthy and fertile, with milder phenotypes such as behavioural and cognitive deficits or craniofacial abnormalities. As discussed previously [40,41], the embryonic lethality observed in the Gtf2ird1Gt/Gt mice likely stems from the use of the gene trap ES cell line with an insertion into intron 22 of Gtf2ird1. The resulting transcript would lead to translation of a fusion protein encompassing most of GTF2IRD1, which may still interact with its usual protein partners but be incapable of carrying out its normal function. If indeed the case, the downstream effects on global gene expression would be likely to include effects on genes that are not normally either direct or indirect GTF2IRD1 targets.

Strain-dependent differences in gene expression predominate in array analysis of Gtf2ird1−/− mice

A number of the genes we identified as having decreased expression in brains using microarray analysis were good candidate genes for the behavioural phenotype seen in Gtf2ird1−/− mice. ACTL6B is a member of a post-mitotic neuron-specific chromatin remodelling complex, and is known to be involved in dendritic growth and development [42]. STX3 is known to be involved in neuronal growth [43] and synapse function [44]. However, while expression of these genes was verified by qRT-PCR, the expression differences are unlikely to be linked to the disruption of GTF2IRD1.

In the case of Actl6b, we believe that its altered expression is related to strain background, rather than to Gtf2ird1 genotype. All genes identified through microarray analysis of E15.5 embryos, and many identified through microarray analysis of P0 mice, were located on the distal portion of chromosome 5, most within 10 Mb of the Gtf2ird1 locus at 135 Mb. This clustering of candidate target genes around Gtf2ird1 could have resulted from a disruption of transcriptional regulation due to the targeting itself, perhaps due to the presence of the neomycin cassette used for ES cell selection, however genes immediately adjacent to Gtf2ird1 showed unchanged expression [19], suggesting that the targeting itself had minimal effect. Our mice were generated from the R1 ES cell line [129X1/SvJ-129S1 hybrid] and backcrossed onto CD1. By selecting for the targeted Gtf2ird1 allele, we were also selecting for mice where the surrounding region of chromosome 5 originated in the ES cell line.

The phenomenon of retained ES cell-derived DNA flanking a targeted locus confounding the results of microarray experiment is a recognized problem [45–47]. Many polymorphisms exist which result in altered levels of gene expression between different mouse strains [48], thus, a mouse with a targeted allele may express the genes surrounding the targeted locus at different levels than the WT mouse to which it is being compared, even if littermates are used. This flanking gene effect has been shown to persist after 11 generations of back crossing and extend up to 40 MB from the targeted locus [47].

To further confirm that in vivo differences in gene expression were unrelated to Gtf2ird1 genotype, we used siRNA to knockdown Gtf2ird1 expression in vitro in two neuronal cell lines. This allowed the expression of candidate genes to be studied in the absence of physical disruption of chromosome 5. Even with Gtf2ird1 knockdown at 80%, we were unable to validate any of the genes identified in the microarray experiments. Since we were able to detect differences in candidate gene expression even in Gtf2ird1−/− mice (data not shown), we would have expected to see a significant change in expression in the siRNA treated cells, were the candidate genes being either directly or indirectly regulated by GTF2IRD1.

Why in vivo differences in global gene expression were not found in Gtf2ird1−/− mice

Given the large number of potential GTF2IRD1 targets previously identified in vitro, and the widespread expression of
Gtf2ired1 in the developing mouse brain, it is surprising that we did not identify any *bona fide* changes in gene expression using microarray analysis of both newborn and mid-gestation knockout mice.

There are a number of possible explanations as to why no *in vivo* targets of GTF2IRD1 were identified. It may be that GTF2IRD1 does not regulate gene expression at the time points examined, although this is unlikely given its expression pattern. Regulation of target genes may only occur in very specific cell populations in the brain, and by examining whole brain, any effect was diluted out, but again, the widespread distribution of GTF2IRD1 in brain suggests otherwise.

Another possibility is that the absence of GTF2IRD1 does affect gene expression in *Gtf2ired1*/*--* mice, but that the changes in expression are small, and so fall below the threshold of detection. This was the case when gene expression was analyzed by microarray in mice deficient for the general transcriptional repressor, methyl CpG binding protein 2 (*Mecp2*) [49]. Microarray analysis performed on brain from *Mecp2*/*--* mice at multiple time points did not detect any significant changes in gene expression. However, mutant mice could be differentiated from WT littermates by looking at very subtle changes in gene expression that occurred in a number of genes in parallel. It is possible that deficiency of GTF2IRD1 also only causes subtle changes in gene expression, which are sufficient to cause the behavioural phenotype seen in *Gtf2ired1*/*--* mice but hard to detect using microarray.

An alternate theory to explain the lack of *in vivo* transcriptional targets for GTF2IRD1 is that the main function of this protein is not transcriptional activation or repression. GTF2IRD1 may not regulate gene expression *in vivo*, but instead may be involved in protein-protein interactions in the absence of biologically relevant DNA binding. A cytoplasmic role for GTF2 has been demonstrated. GTF2I belongs to the same protein family as GTF2IRD1 and can negatively regulate agonist-induced calcium entry into cells by indirectly inhibiting the localization of subunits of the transient receptor potential cation channel, TRPC3, to the plasma membrane [50]. Proper regulation of intercellular calcium levels via TRPC channels is essential for many neuronal functions including axon guidance [51] and membrane depolarization [52], and disruption of another member, TRPC3, reduces innate fear [53]. It is possible that GTF2IRD1 may also play a cytoplasmic role in developing brain, or other tissues. Reports of the localization of GTF2IRD1 within the cell are conflicting, in part due to the lack of a specific antibody, necessitating the over-expression of tagged protein [29,54,55], so it may yet be shown to have a role outside the nucleus.

**Why GTF2IRD1 gene targets identified *in vitro* were not validated *in vivo***

Having failed to identify any genotype-dependent changes in gene expression using an unbiased genome-wide approach, we investigated direct targets of GTF2IRD1 previously identified *in vitro*, in our mouse model. Of the ten genes that we analyzed, none showed altered expression in either developing or adult brain, or in MEFs. Expression of *Bone1b* and *Fgf13* had been shown to increase by 600- and 6900-fold, respectively, in a *Gtf2ired1*-silRNA treated C2C12 myoblast cell line, and GTF2IRD1 was shown to bind to the *Fgf13* promoter using chromatin immunoprecipitation (ChIP) [33]. The absence of any detectable effect on the expression of either of these genes in mice that lack GTF2IRD1 suggests that if GTF2IRD1 does play a role in the regulation of these genes, it may be dependent upon co-factors that are expressed in C2C12 cells but not in brain.

We also investigated *in vivo* expression of eight other target genes previously identified *in vitro*, but did not find any altered expression in either primary MEFs or brain tissue from *Gtf2ired1*/*--* or *Gtf2ired1*/*--* mice. These genes had been identified through GTF2IRD1 over-expression studies in MEFs [31], GTF2IRD1 had been shown to bind to the promoter regions of six of them using ChIP, and each of them demonstrated at least a two-fold decrease in expression with double knockdown of *Gtf2ired1* and *Gtf2* using siRNA (with the exception of Fgf11 which showed 1.3-fold up-regulation) [30].

It is possible that binding of GTF2IRD1 to the promoters of these genes, and subsequent regulation of gene expression is dependent upon co-factors which are not found in the brain, or that GTF2IRD1 regulates their transcription in a specific type or subpopulation of cells. However, recently, it has been shown that GTF2IRD1 may need at least two, closely adjoining, consensus binding sequences to be present within a gene promoter in order to facilitate robust protein binding in vivo [40]. This finding contradicts the *in vitro* evidence for GTF2IRD1 binding to the promoter of genes such as Tmx1 [56,57], Hoxc8 [27] and Gsc [29], as well as the genes tested in this study, since they all contain a single GTF2IRD1 binding consensus sequence.

**Summary**

In conclusion, this study failed to identify new *in vivo* targets or to validate previously identified *in vitro* targets of GTF2IRD1 in mice with homozygous disruption of this gene. The biological mechanism by which loss or reduction of GTF2IRD1 leads to the behavioural phenotype in mice or in people with WBS, therefore remains unknown. The robust phenotype seen even in mice missing a single copy of *Gtf2ired1* [19] suggests that this protein does play an important role in proper brain development and/or function but additional experiments will be needed to shed light on the precise role of GTF2IRD1. These could include the identification of interacting proteins in the tissues and time-points of particular relevance as well as the isolation of specific cell populations that express GTF2IRD1.

**Materials and Methods**

**Isolation of tissues and RNA**

All procedures were approved by the University of Toronto Animal Care Committee and carried out in compliance with the Canadian Council on Animal Care guidelines (U of T Protocol # 20008762). Animals were housed with access to food and water *ad libitum* and were on a 12 hour light/dark cycle throughout the experiments. *Gtf2ired1*/*+*/*+* mice (*Gtf2ired1*/*--*) were generated as described previously [19]. Mice were maintained on a CD1 background and had reached the 4th generation of backcrossing at the time of these experiments. For all experiments using embryos, *Gtf2ired1*/*++*/*+* mice and WT littermates were generated through the intercrossing of *Gtf2ired1*/*++*/*--* mice. Male and female mice were housed together overnight, and the female was checked for a vaginal plug in the morning. 12:00 pm on the day the plug was found was considered to be E0.5. Data were pooled across all male and female pups within genotype groups.

P0 mice were sacrificed by decapitation, and adult mice by cervical dislocation. Whole brain was removed and immediately flash frozen in liquid nitrogen and tail tissue collected for DNA extraction and genotyping. For embryonic dissections, the mother was sacrificed using cervical dislocation, the uterus removed, embryos immediately dissected from the yolk-sacs, the heads removed and flash frozen in liquid nitrogen. Tissues were homogenized in TRIReagent (Sigma-Aldrich Canada, Oakville,
Microarray analysis using the Affymetrix mouse 430 2.0 gene chip

RNA from individual P0 mouse brains were pooled together in equal concentrations. Three pools containing RNA from nine WT mice were created, along with three pools containing RNA from nine Gtf2ird1+/− mice. Microarray analysis was performed by The Centre for Applied Genomics (TCAG) at the Hospital for Sick Children (Toronto, ON) using the GeneChip Mouse Genome 430 2.0 Array (Affymetrix) (which contains probes for over 39,000 transcripts) following the manufacturer’s protocol. The signals from the gene chips were normalized using Robust Multiarray Analysis (RMA) [34]. Differences in gene expression were detected using a second software program, Significance Analysis of Microarrays (SAM) [35], which uses q values as a measure of the false discovery rate.

Microarray analysis using the Illumina mouseWG-6 v2.0 BeadChip

RNA from five Gtf2ird1+/− mice and five WT littermates were used for microarray analysis. Microarray analysis was performed by TCAG at the Hospital for Sick Children (Toronto, ON) using the Illumina Mouse WG-6 v2.0 Expression BeadChip (which contains probes for over 45,200 transcripts) following the manufacturer’s protocol. Analysis of microarray data was performed by the Statistical Analysis Core Facility at TCAG. The data pre-processing included three steps: background correction was performed in the Beadstudio program (Illumina), the data was then transferred to log2 scale and quantile normalization was performed [37]. Differentially expressed genes were identified using LIMMA (linear models for microarray data) [36]. All data is MIAME compliant and that the raw data has been deposited in the ArrayExpress database under the accession numbers E-MTAB-622 and E-MTAB-626.

Expression analysis using quantitative real-time PCR

Following extraction, total RNA samples were treated with DNase (Turbo DNA free, Applied Biosystems/Ambion, Austin, TX) and 5 ug of RNA converted to cDNA using the SuperScript II First-Strand Synthesis System (Invitrogen Canada, Burlington, ON) and random hexamer primers. cDNA samples were diluted 1/100 with sterile water and subjected to real-time PCR analysis using the Power SYBR Green PCR Master mix (Applied Biosystems, Foster City, CA) and the ABI Prism 7900HT sequence detection system (Applied Biosystems) as previously described [19]. Primers used for expression analysis are listed in Table S3. Samples were run in triplicate, and each experiment was repeated at least twice with consistent results. Absolute quantification analysis was used; each plate included a no template control, serially diluted concentrations of control genomic DNA (range 0.63–10 ng/well) to generate a standard curve for transcript quantification and a no reverse transcriptase control to ensure the cDNA was free of genomic contamination. All test genes were normalized to the housekeeping gene succinate dehydrogenase (SdhA). Normalized values for each gene were then averaged for each genotype group. Comparative expression ratios (%) were calculated by dividing the averaged normalized values for each of the test genes in the Gtf2ird1−/− group by the normalized test gene values for the WT group.

siRNA knockdown of Gtf2ird1 in neuronal cell lines

siRNA knockdown of Gtf2ird1 was performed in two different neuroblastoma derived cell lines: Neuro2A (N2A; ATCC #CCL-131) and N1E-115 (ATCC # CRL-2263). Cell lines were maintained in D-MEM (Sigma-Aldrich) with 10% fetal bovine serum (FBS) (Invitrogen) and 1 × penicillin-streptomycin (Sigma-Aldrich). For siRNA transfection, cells were cultured in D-MEM +10% FBS without antibiotics. Cells were maintained at 37°C with 5% CO2.

siRNAs targeting Gtf2ird1 (Table S4), Gapdh (ON-TARGETplus GAPD Control Pool (Mouse)) and a non-targeting control (ON-TARGET plus Non-targeting siRNA #1) were obtained from Dharmacon. Transfections of siRNA into N2A and N1E-115 cells were conducted using Lipofectamine 2000 (Invitrogen) following the manufacturer’s protocol. Briefly, cells were transfected in 6-well plates at 50–60% confluency. Lipofectamine 2000 was diluted 1/50 in Opti-MEM Reduced Serum Medium (Invitrogen), and siRNAs were diluted similarly to a final concentration of 100 nM. Cells were harvested at 24 hrs following transfection and total RNA was extracted.

Preparation and culture of mouse embryonic fibroblasts (MEFs)

Embryos were harvested from the pregnant mother at E15.5 and yolk-sacs collected for genotyping. The embryos were dissected from the uterus into sterile PBS, the head, limbs and internal organs were removed from the embryos, and the carcasses washed three times with sterile Dulbecco’s Modified Eagles Medium (D-MEM) (Sigma-Aldrich). Embryos were then minced into small pieces, trypsinized, washed and plated in D-MEM +10% FBS and 1 × penicillin-streptomycin (Sigma-Aldrich). Cells were cultured at 37°C with 5% CO2 and passaged at least twice before use to ensure a homogenous population.

Supporting Information

Table S1: Genes found to have altered expression in the brains of P0 Gtf2ird1+/− mice according to microarray analysis. (DOCX)

Table S2: Genes found to have altered expression in the brains of E15.5 Gtf2ird1−/− mice using microarray analysis. (DOCX)

Table S3: Sequences of primers used in qRT-PCR. (DOCX)

Table S4: Sequences of siRNAs used to knockdown Gtf2ird1 expression. (DOCX)

Acknowledgments

We thank Elaine Tam for her assistance in maintaining the mouse colony.

Author Contributions

Conceived and designed the experiments: JO LRO. Performed the experiments: JO. Analyzed the data: JO LRO. Contributed reagents/materials/analysis tools: JO. Wrote the paper: JO LRO.
References

1. Stromme P, Bjornstad PG, Ramstad K (2002) Prevalence estimation of Williams syndrome. J Child Neurol 17: 269–271.
2. Pober BR (2010) Williams-Beuren syndrome. N Engl J Med 362: 239–252.
3. Bellugi U, Lichtenberger L, Jones W, Lai Z, St. George M (2000) I. The neurocognitive profile of Williams Syndrome: a complex pattern of strengths and weaknesses. J Cogn Neurosci 12 Suppl 1: 7–29.
4. Mervis CB, Robinson BF, Bertrand J, Morris CA, Klein-Tasman B, et al. (2006) Identification of mental retardation locus in Williams-Coggs region 7q11.23. Proc Natl Acad Sci U S A 103: 5469–5474.
5. Klein-Tasman B, Mervis CB (2003) Distinctive personality characteristics of 8–9, and 10–year-olds with Williams syndrome. Dev Neuropsychol 30: 369–290.
6. Lefter OT, Woodruff-Borden J, Klein-Tasman BP, Frick JS, Mervis CB (2006) Prevalence of attentional disorders in 4 to 6-year-olds with Williams syndrome. Am J Med Genet B Neuropsychiatr Genet 141B: 615–622.
7. Curran ME, Atkinson DL, Ewart AK, Morris CA, Leppert MF, et al. (1993) The elastin gene is transcribed by a translocation associated with supravalvular aortic stenosis. Cell 73: 159–161.
8. Tassabehji M, Metcalf K, Domaus D, Hurst J, Reardon W, et al. (1997) Elastin: genomic structure and point mutations in patients with supravalvular aortic stenosis. Hum Mol Genet 6: 1029–1036.
9. Li DV, Tisdal AE, Boak BB, Atkinson DL, Ersing GJ, et al. (1997) Elastin point mutations cause an obstructive vascular disease, supravalvular aortic stenosis. Hum Mol Genet 6: 1021–1028.
10. Botta A, Novelli G, Mari A, Novelli A, Sahani M, et al. (1999) Detection of an atypical 7q11.23 deletion in Williams syndrome patients which does not include the SINE and FZD3 genes. J Med Genet 36: 478–480.
11. Heller R, Rauch A, Lutten S, Schroder B, Winterpacht A (2003) Partial deletion of the critical 1.5 Mb interval in Williams-Beuren syndrome. J Med Genet 40: e99.
12. Hirota H, Matsuo R, Chen XN, Salandanan LS, Lincoln A, et al. (2003) Williams syndrome deficiencies in visual spatial processing linked to GTF2IRD1 and GTF2I on chromosome 7q11.23. Genet Med 5: 311–321.
13. Morrison CA, Mervis CB, Hobart HH, Gregg RG, Bertrand J, et al. (2003) GTF2I homoeobox implicated in mental retardation in Williams syndrome: genotype-phenotype analysis of five families with deletions in the Williams syndrome region. Am J Hum Genet 72A: 45–59.
14. Gaglardi C, Bonaglia MC, Selicorni A, Borgatti R, Giordano R (2003) Unusual cognitive and behavioural profile in a Williams syndrome patient with atypical 7q11.23 deletion. J Med Genet 40: 526–530.
15. Howald C, Merla G, Biglione MC, Amenta S, Lyle R, et al. (2006) Two high throughput technologies to detect segmental aneuploidies identify new Williams-Beuren syndrome patients with atypical deletions. J Med Genet 43: 266–273.
16. Dai L, Bellugi U, Chen XN, Paht-Korenberg AM, Jarvisen-Padely A, et al. (2006) Is it Williams syndrome? GTF2IRD1 implicated in visual-spatial construction and GTF2I in sociability revealed by high resolution arrays. Am J Hum Genet 18A: 302–314.
17. Ferrero GB, Holawka C, Micale L, Biamino E, Angelillo B, et al. (2010) An atypical 7q11.23 deletion in a normal IQ Williams-Beuren syndrome patient. Eur J Hum Genet 18: 33–44.
18. Antonell A, Del Campo M, Mageno LF, Kaufmann L, de la Iglesia JM, et al. (2007) Expression of BEN, a member of the TFII-I family of transcription factor, during mouse pre- and postimplantation development. Gene Expr 15: 101–114.
19. Young EJ, Lipina T, Tam E, Mandel A, Clapcote SJ, et al. (2008) Reduced fear and aggression and altered serotonin 5-HT(1A) currents in a mouse model of Williams-Beuren syndrome. J Neurosci 28: 1547–1557.
20. Antonell A, Del Campo M, Magano LF, Kaufmann L, de la Iglesia JM, et al. (2007) Expression of BEN, a member of TFII-I family of transcription factor. Nature 434: 894–898.
21. Ferrero GB, Holawka C, Micale L, Biamino E, Angelillo B, et al. (2010) An atypical 7q11.23 deletion in a normal IQ Williams-Beuren syndrome patient. Eur J Hum Genet 18: 33–44.
22. Tassabehji M, Hammond P, Karmiloff-Smith A, Thompson P, Thorpeiresson SS, et al. (2005) GTF2IRD1 in craniofacial development of humans and mice. Science 310: 1184–1187.
23. Vassena R, Guarneri M, Gatti A, Basile G, Battistuzzi A, et al. (2007) Expression profiling of BEN, regulating a set of general muscle genes in mouse embryonic fibroblasts. J Exp Zool B Mol Dev Evol 308: 209–224.
24. Vullhorst D, Buonanno A (2005) Multiple GTF2I-like repeats of general muscle genes responsible for the Williams-Beuren syndrome neocognitive profile. Faseb J 19: 1078–1086.
25. Ferrero GB, Holawka C, Micale L, Biamino E, Angelillo B, et al. (2010) An atypical 7q11.23 deletion in a normal IQ Williams-Beuren syndrome patient. Eur J Hum Genet 18: 33–44.
26. Thompson PD, Webb M, Beckett W, Hinesley T, Jowitt T, et al. (2007) GTF2IRD1 regulates transcription by binding an evolutionarily conserved DNA element in mouse beta-2-adrenergic receptor gene. Proc Natl Acad Sci U S A 104: 10516–10521.
27. Bayarsaihan D, Ruddle FH (2000) Isolation and characterization of BEN, a member of the TFII-I family of DNA-binding proteins containing distinct helix-loop-helix domains. Proc Natl Acad Sci U S A 97: 7342–7347.
28. Ring O, Ogata S, Meek L, Song J, Olive T, et al. (2002) The role of a Williams-Beuren syndrome-associated helix-loop-helix domain-containing transcription factor in actin/myosin signaling. Genes Dev 16: 2020–2023.
29. Ku M, Sokol SY, Wu J, Tussie L, Miller LJ, et al. (2003) Positive and negative regulation of the transforming growth factor beta/activin target gene gosoccid by the TFII-I family of transcription factors. Mol Cell Biol 23: 7157–7167.
30. Chiappe NO, Makeyev AV, Ruddle FH, Bayarsaihan D (2008) Identification of the TFII-I family target genes in the vertebrate genome. Proc Natl Acad Sci U S A 105: 9006–9010.
31. Chiappe NO, Munosugoh O, Ruddle F, Bayarsaihan D (2007) Expression profiling of BEN, regulating a set of general muscle genes in mouse embryonic fibroblasts. J Exp Zool B Mol Dev Evol 308: 209–224.
32. Vullhorst D, Buonanno A (2005) Multiple GTF2I-like repeats of general muscle genes responsible for the Williams-Beuren syndrome neocognitive profile. Faseb J 19: 1078–1086.
55. Tussie-Luna MI, Bayarsaikh G, Ruddle FH, Roy AL. (2001) Repression of
TFII-I-dependent transcription by nuclear exclusion. Proc Natl Acad Sci U S A
98: 7789–7794.
56. O’Mahoney JV, Guven KL, Lin J, Joya JE, Robinson CS, et al. (1998)
Identification of a novel slow-muscle-fiber enhancer binding protein,
MusTRD1. Mol Cell Biol 18: 6641–6652.
57. Calvo S, Vullhorst D, Venepally P, Cheng J, Karavanova I, et al. (2001)
Molecular dissection of DNA sequences and factors involved in slow muscle-
specific transcription. Mol Cell Biol 21: 8490–8503.