Meta-analysis of heterogeneous Down Syndrome data reveals consistent genome-wide dosage effects related to neurological processes

Mireia Vilardell1,2*, Axel Rasche1, Anja Thormann1, Elisabeth Maschke-Dutz1, Luis A Pérez-Jurado2,3, Hans Lehrach1 and Ralf Herwig1*

Abstract

Background: Down syndrome (DS; trisomy 21) is the most common genetic cause of mental retardation in the human population and key molecular networks dysregulated in DS are still unknown. Many different experimental techniques have been applied to analyse the effects of dosage imbalance at the molecular and phenotypical level, however, currently no integrative approach exists that attempts to extract the common information.

Results: We have performed a statistical meta-analysis from 45 heterogeneous publicly available DS data sets in order to identify consistent dosage effects from these studies. We identified 324 genes with significant genome-wide dosage effects, including well investigated genes like SOD1, APP, RUNX1 and DYRK1A as well as a large proportion of novel genes (N = 62). Furthermore, we characterized these genes using gene ontology, molecular interactions and promoter sequence analysis. In order to judge relevance of the 324 genes for more general cerebral pathologies we used independent publicly available microarray data from brain studies not related with DS and identified a subset of 79 genes with potential impact for neurocognitive processes. All results have been made available through a web server under http://ds-geneminer.molgen.mpg.de/.

Conclusions: Our study represents a comprehensive integrative analysis of heterogeneous data including genome-wide transcript levels in the domain of trisomy 21. The detected dosage effects build a resource for further studies of DS pathology and the development of new therapies.

Background

Down syndrome (DS) is the most frequent genomic aneuploidy with an incidence of approximately 1 in 700 live-newborn [1] resulting from the presence of an extra copy of human chromosome 21 (HSA21). DS is characterized by a complex phenotype with features that are not fully penetrant. The most frequent manifestations, which are virtually always present, include mental retardation, morphological abnormalities of the head and limbs, short stature, hypotonia and hyperlaxity of ligaments. Other features occur with less frequency such as organ malformations, particularly of the heart (50% of DS newborns), several types of gastrointestinal tract obstructions or dysfunctions (4-5% of DS newborns), increased risk of leukaemia (20 × higher compared to the normal population), and early occurrence of an Alzheimer-like neuropathology [2,3]. DS has been investigated with multiple functional genomics studies aiming to understand the molecular basis underlying the various aspects of the disease [4-7].

The most commonly accepted pathogenetic hypothesis is that the dosage imbalance of genes on HSA21 is responsible for the molecular dysfunctions in DS, meaning that genes on the triplicated chromosome are overexpressed due to an extra chromosome 21, as demonstrated for selected genes like SOD1 and DYRK1A [8]. Recent global transcriptome studies with microarrays, however, have generated a more complex picture in the sense that not all HSA21 genes have an elevated expression level as expected [9,10]. An alternative hypothesis is that the phenotype is due to an unstable environment resulting from the dosage imbalance of the hundreds of genes on HSA21.
which determines a non-specific disturbance of genomic regulation and expression. The significantly higher inter-individual variability in DS, as compared to euploid, individuals supports this hypothesis [11]. Moreover, the two hypotheses could be coexistent [3]. In both hypotheses it is understood that besides alterations of gene expression of HSA21 genes there are numerous genome-wide effects that lead to the dysregulation of many non-HSA21 genes through molecular pathways and interactions.

Many studies on the transcriptome and proteome levels have been conducted to understand the causal relationship between genes at dosage imbalance and DS phenotypes [12]. Gene expression profiles have been analysed from DS fetal [13] and adult human tissues [6]. Additionally, two classes of mouse models [14] have been developed for investigating the molecular genetics of DS, either mouse models with partial trisomies of the syntenic regions of HSA21 in mouse chromosomes 10, 16 or 17, such as Ts16 [15], Ts65Dn [16] and Ts1Cje mice [17], or transgenic mice for specific genes such as SOD1 [18]. Studies of gene expression profiles in human DS samples and mouse models have shown high genome-wide variability [11,19-22]. Furthermore, differences due to the applied experimental platforms, specific tissues, developmental stages or the triplicated segments under study introduce a high variation to the assessment of genome-wide effects of DS. Here, integrative and comparative studies are pivotal for the analysis of the complex nature of gene expression and regulation in DS at a more general level [2,23].

Meta-analysis was proven to be a valid strategy to extract consistent information from heterogeneous data, in particular with respect to complex phenotypes for example cancer [24], Alzheimer [25] and type-2 diabetes mellitus [26]. The purpose of meta-analysis is to compensate experiment-specific variations and to reveal consistent information across a wide range of experiments. To date, such a meta-analysis of DS data is missing.

In this paper we describe a comprehensive meta-analysis from 45 different DS studies on human and mouse on the transcriptome and proteome level including quantitative data such as Affymetrix microarrays, RT-PCR and MALDI studies as well as qualitative data such as SAGE and Western blot analyses. We applied an established computational framework [26] and identified 324 genes with consistent dosage effects in many of these studies. As expected, we observed a high fraction of HSA21 genes (N = 77) but also a large amount of non-HSA21 genes (N = 247). Besides well investigated genes in the context of DS we detected a significant proportion of novel ones (N = 62). The 324 genes were further investigated using functional information, molecular interactions and promoter analysis revealing over-represented motifs of four transcription factors: RUNXI, E2F1, STAF/PAX2 and STAT3. In order to test the relevance of the 324 genes for more general brain phenotypes we used independent publicly available data on cerebral pathologies not related to DS and identified a subset of 79 DS genes that were differentially expressed in these studies. The detected dosage effects can be used as a resource for further studies of DS pathology, functional experiments and the development of therapies. All data have been agglomerated and made available through a web server that tracks results of the meta-analysis http://ds-geneminer.molgen.mpg.de/ and that enables the community to validate any gene of interest in the light of the experimental data.

Results
Genome-Wide Dosage Effects
Genome-wide dosage effects were computed with the numerical scoring method described in Material and Methods. In total, 45 case-control experiments were interrogated (Additional file 1, Table S1), the alteration for each gene between the trisomic and normal states was scored in each experiment, gene scores were summarised across all experiments and the significance of the summarised scores was judged with a Bootstrap approach. This procedure resulted in a cut-off score value of 3.67 and identified 324 genes as being predominantly affected by DS. The thirty genes with the highest dosage effects, either on HSA21 or on other chromosomes, are listed in Table 1. The entire gene list is given in Additional file 1, Table S2. The meta-analysis identified genes that showed consistent changes in many of the different experiments rather than genes that were affected by a single (or few) experiment(s) (Figure 1A). This is an important fact since, for example, different mouse models have different coverage of triplicated HSA21 genes, and, thus, might introduce model-specific bias [14]. The consistency of the dosage effect was measured for each gene with an entropy criterion (see Materials and Methods) and Figure 1A reveals a strong preference for the selection of high-entropy genes. Highest scores were assigned to HSA21 genes (Figure 1B) what indicates that the meta-analysis scores reflect the effect of an extra chromosome 21 on gene expression (Table 1). While proportionally most dosage effects were identified for HSA21 genes (77 out of 324), the majority of genes (247 out of 324) was located on other chromosomes highlighting the genome-wide impact of DS (Figure 1C).

Genome-wide dosage effects underlined the severe phenotypic consequences of DS caused by genes with a major role in human development (Additional file 2, Figure S1). Of the 247 non-HSA21 genes, 72 were associated with development, in particular with respect to organ development (62 genes, GO:0048513), tissue
Table 1: Top thirty DS dosage effects on A) HSA21 and B) other chromosomes

A) Direct effects

| Ensembl | HUGO | Score | Entropy | Chromo-some | Start position | End position | Band | CNV |
|---------|------|-------|---------|-------------|----------------|--------------|------|-----|
| ENSG00000154734 | ADAMTS1 | 18.487 | 4.083 | chr21 | 28208606 | 28217728 | q21.3 |
| ENSG00000159228 | CBR1 | 17.518 | 4.509 | chr21 | 3744239 | 37445464 | q22.12 |
| ENSG00000159140 | SON | 15.920 | 4.712 | chr21 | 34919424 | 34949812 | q22.11 |
| ENSG00000142168 | SOD1 | 15.817 | 4.372 | chr21 | 33031935 | 33041244 | q22.11 |
| ENSG00000182670 | TTC3L, TTC3 | 15.637 | 4.542 | chr21 | 38445526 | 38575413 | q22.13 |
| ENSG00000142192 | APP | 15.489 | 4.412 | chr21 | 27252861 | 27543446 | q21.3 |
| ENSG00000159128 | IFNGR2 | 15.006 | 4.640 | chr21 | 34757299 | 34851655 | q22.11 |
| ENSG00000182240 | BACE2 | 14.156 | 4.140 | chr21 | 42539728 | 42648524 | q22.2 |
| ENSG00000156256 | USP16 | 13.713 | 4.378 | chr21 | 30939050 | 30949812 | q22.11 |
| ENSG00000159131 | GART | 13.570 | 4.564 | chr21 | 34870940 | 34915797 | q22.11 |
| ENSG00000157540 | DYRK1A | 13.163 | 4.405 | chr21 | 38739236 | 38887680 | q22.13 |
| ENSG00000157557 | ETS2 | 13.088 | 4.440 | chr21 | 40177231 | 40196879 | q22.2 |
| ENSG00000159231 | CBR3 | 12.185 | 3.954 | chr21 | 37507210 | 37518864 | q22.12 | YES |
| ENSG00000159082 | SYNJ1 | 11.880 | 4.284 | chr21 | 33997269 | 34100359 | q22.11 |
| ENSG00000142188 | TMEM50B | 11.516 | 3.803 | chr21 | 3480792 | 34853499 | q22.11 | YES |
| ENSG00000159110 | IFNAR2 | 11.280 | 4.360 | chr21 | 34602206 | 3465082 | q22.11 |
| ENSG00000159119 | IFNAR2 | 11.280 | 4.360 | chr21 | 34602206 | 3465082 | q22.11 |
| ENSG00000159119 | IFNAR2 | 11.280 | 4.360 | chr21 | 34602206 | 3465082 | q22.11 |
| ENSG00000159119 | IFNAR2 | 11.280 | 4.360 | chr21 | 34602206 | 3465082 | q22.11 |

B) Indirect effects

| Ensembl | HUGO | Score | Entropy | Chromo-some | Start position | End position | Band | CNV |
|---------|------|-------|---------|-------------|----------------|--------------|------|-----|
| ENSG00000117289 | TXNIP | 8.790 | 3.281 | chr1 | 14543869 | 14543869 | q21.3 | YES |
| ENSG00000133110 | POSTN | 8.301 | 2.437 | chr13 | 38136722 | 38136722 | q21.3 | YES |
| ENSG00000118785 | SPP1 | 7.232 | 3.159 | chr1 | 88896802 | 88904563 | q22.1 |
| ENSG00000113140 | SPARC | 7.035 | 3.338 | chr5 | 15104057 | 15106672 | q3.11 |
| ENSG00000125968 | ID1 | 6.987 | 3.164 | chr20 | 30193086 | 30193086 | q11.21 |
| ENSG00000162407 | PPAP2B | 6.350 | 3.343 | chr1 | 5690419 | 5690419 | q22.11 |
| ENSG00000171951 | SCG2 | 6.747 | 2.950 | chr2 | 22446158 | 22446158 | q3.61 |
| ENSG00000135821 | GLUL | 6.604 | 3.702 | chr1 | 18235083 | 18235083 | q25.3 |
| ENSG00000123610 | TNFAIP6 | 6.575 | 2.377 | chr2 | 152214106 | 152214106 | q23.3 |
| ENSG00000118523 | CTGF | 6.567 | 2.996 | chr6 | 132269316 | 132269316 | q23.2 |
| ENSG00000168209 | DDIT4 | 6.477 | 3.318 | chr10 | 74033678 | 74033678 | q22.1 |
| ENSG00000126407 | PPAP2B | 6.350 | 3.343 | chr1 | 5690419 | 5690419 | q22.11 |
| ENSG00000138427 | VCAN | 6.240 | 2.958 | chr5 | 82767284 | 82767284 | q14.2 |
| ENSG00000151491 | EPS8 | 6.194 | 3.143 | chr12 | 15773076 | 15773076 | p1.23 |
| ENSG00000189067 | LITAF | 6.185 | 3.255 | chr16 | 11641582 | 11641582 | p1.33 |
| ENSG00000164692 | COL1A2 | 6.148 | 2.852 | chr7 | 94023873 | 94023873 | q21.3 |
development (34 genes GO:0009888) and cell development (30 genes, GO:0048468). Amongst these genes were known interactors of HSA21 genes, for example REST (RE1-silencing transcription factor). REST modulates expression of genes encoding fundamental neuronal functions including ion channels, synaptic proteins and neurotransmitter receptors and has been linked to an inherited form of mental retardation. Recently, Cзонетта et al. [5] demonstrated that the region capable of affecting REST levels, in both mouse and human cells, could be assigned to the DYRK1A locus on HSA21 which was found among the top-scoring HSA21 genes (Table 1).

**Table 1 Top thirty DS dosage effects on A) HSA21 and B) other chromosomes (Continued)**

| Ensembl ID | Symbol | Log2 Fold Change | Significance | Chromosome | Genomic Location |
|------------|--------|-----------------|--------------|------------|-----------------|
| ENSG00000204388 | HSPA1B | 6.109 | 2.550 | chr6 | 31795688-31798031 |
| ENSG00000162692 | VCAM1 | 6.012 | 2.533 | chr1 | 119288888-119293854 |
| ENSG00000135919 | SERPINE2 | 5.974 | 2.344 | chr11 | 119288888-119293854 |
| ENSG00000172201 | ID4 | 5.887 | 3.037 | chr6 | 19837617-19840915 |
| ENSG00000114315 | HES1 | 5.884 | 2.874 | chr3 | 193853934-193856521 |
| ENSG00000172893 | DHC7 | 5.883 | 3.441 | chr11 | 71145457-71159477 |
| ENSG00000204262 | COLSA2 | 5.857 | 3.174 | chr2 | 19837617-19840915 |
| ENSG00000149257 | SERPINH1 | 5.846 | 3.146 | chr11 | 75273170-75283844 |
| ENSG00000176697 | BDNF | 5.805 | 2.416 | chr11 | 27676440-27743605 |
| ENSG00000182551 | ADI1 | 5.782 | 3.081 | chr2 | 3501693-3523507 |
| ENSG000001772689 | CHST2 | 5.782 | 3.081 | chr2 | 3501693-3523507 |
| ENSG0000018821 | COL1A1 | 5.527 | 3.004 | chr17 | 48260650-48278993 |
| ENSG00000187498 | COL4A1 | 5.514 | 3.396 | chr13 | 110801318-110959496 |

Dosage Effects on HSA21

Proportionally HSA21 contributed mostly to the detected dosage effects (Figure 1C). On the other hand, it is remarkable that only a third of all HSA21 genes (77 out of 255 studied here using the Ensembl genome annotation [31]) showed consistent effects across the different experiments (see also Discussion). While 57 genes had a positive score below the significance threshold of 3.67 indicating relevance with respect to specific experiments only, 121 genes had a score near zero indicating that dosage effects were either compensated or not detected with the selected experimental data (Figure 1B).

HSA21 dosage effects included, for example APP (beta-amyloid precursor protein) involved in senile plaque formation in DS and Alzheimer’s disease [3], SOD1 (superoxide dismutase 1), a key enzyme in the metabolism of oxygen-derived free radicals [3], DYRK1A (dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1A) involved in neuroblast proliferation, crucial for brain function, learning and memory [32], RUNX1 (runt-related transcription factor 1) which plays a critical role in normal hematopoiesis [33], or GABPA (GA binding protein transcription factor, alpha subunit 60 kDa) encoding a DNA binding domain with a huge variety of targets including genes from different cell/tissue specificities and functions [34]. HSA21 genes were mostly up-regulated in gene expression studies (69 out of 77) with the exception of eight genes that were either variable or down-regulated (SLCSA3, MRPS6, B3GALT6, CBS, KCNJ6, KCNJ15, CLDN14, COL1A1). Possible explanations for this observation might be tissue-specificity of gene expression as in the case of MRPS6 which was mostly up-regulated in brain samples and down-regulated in other tissues like heart or kidney, or differences in human and mouse gene expression as in the case of CBS which was up-regulated in human but down-regulated in mouse experiments what might be caused by differential tissue specificity of the orthologous mouse gene [35].

Three genomic regions on HSA21 were enriched with the significant genes using the MSigDB c1 positional database: chr21q22, chr21q21 and chr21q11, located on the q-terminal arm (Figure 1D). This contradicts the
Figure 1 Characterization of dosage effects. A) Entropy (Y-axis) vs. score of dosage effect (X-axis) for all genes, B) Histogram of scores for all 255 HSA21 genes accessible with the experiments under study, C) Distribution of genomic locations of the 324 candidate genes, D) Cytogenetic location of 77 HSA21 genes that show significant dosage effects for all experiments (blue line). Additionally, the same meta-analysis approach has been conducted with human (green line) and mouse (red line) data separately. The yellow line plots the relative number of HSA21 genes within each band (gene density). Y-axis shows percentage of significant genes with respect to all genes annotated for the chromosomal band.
hypothesis that a single region on HSA21 could be responsible for the molecular and phenotypic consequences of DS with only a few responsive genes [36,37]. Rather our findings support studies that identified more than one HSA21 region causative for DS phenotypes so that the dosage effects were not uniformly distributed along the chromosome but rather enriched in certain regions on HSA21 similar to the results in [38,39].

Functional Annotation Using Gene Enrichment Analysis

Functional annotation of biological pathways was retrieved from the ConsensusPathDB [40], a meta-database that summarizes the content of 22 human interaction databases. A total of 1,695 pre-defined pathways were screened with the 324 genes using gene set enrichment analysis [41]. A total of 277 pathways were found significantly enriched (family-wise error rate (FWER) <0.01) of which several pathways were associated with neurological and neuropathological processes (Table 2). These pathways referred mainly to (i) neurodegeneration (e.g. Huntington's disease, Alzheimer's disease or Parkinson's disease) and (ii) defects in synapsis (e.g. Axon guidance, NGF signalling). Furthermore, the results emphasized the role of tyrosine-kinase receptors in DS pathology (for example P75(NTR)- mediated signalling or NGF signalling via TRKA) which interact directly with BDNF (brain-derived neurotrophic factor). Moreover, our results showed gene dosage effects caused either directly by genes located on HSA21 (e.g. SOD1, APP, DONSON, TIAM1, COL6A2, ITSN1 and BACE2) or indirectly by HSA21 interactors, highlighting the intrinsic complexity of the DS pathology. For example, PIK3R1 de-regulation impacts on many of these pathways and is a direct interactor of IFNAR1, a significant DS gene. A similar effect can be observed for TPJ1A that has interactions with HSA21 genes JAM2 and CDLN8 both showing consistent dosage effects (cf. Figure 2A).

| PATHWAY (Source Database)         | Pathway size | P-value | FWER P-value | Genes on HSA21 | HSA21 Interactors | Others                  |
|-----------------------------------|--------------|---------|--------------|---------------|-------------------|------------------------|
| HUNTINGTONS DISEASE (KEGG)       | 159          | 0       | 0            | SOD1; DONSON  | REST              | BDNF; SOD2             |
| ALZHEIMERS DISEASE (KEGG)        | 147          | 0       | 0            | APP; BACE2; DONSON | PPP3CA; GSK3B | CAPN2                  |
| SIGNALLING BY NGF (REACTOME)     | 209          | 0       | 0            | ITSN1; TIAM1 | PIK3R1; GSK3B | RPS6KA2; RAP1A; KRAS  |
| AXON GUIDANCE (REACTOME)         | 256          | 0       | 0            | COL6A2       | GSK3B; COL1A1; COL1A2; COL4A1; COL4A2 | COL5A2; DPYSL3; RPS6KA2; LAMB1; COL3A1; COL5A1; ALCAM; KRAS |
| PARKINSONS DISEASE (KEGG)        | 105          | 0       | 0            | DONSON       |                  |                        |
| P75(NTR)-MEDIATED SIGNALING (PID) | 68           | 0       | 0            | APP          | PIK3R1           | BDNF                   |
| NOTCH (NETPATH)                  | 61           | 0       | 0            | APP          | PIK3R1; GSK3B   |                        |
| NEUROTROPHIN SIGNALING PATHWAY (KEGG) | 121       | 0       | 0            | PIK3R1; GSK3B |                  |                        |
| NGF SIGNALLING VIA TRKA FROM THE PLASMA MEMBRANE (REACTOME) | 127 | 0 | 0 | PIK3R1; GSK3B | RPS6KA2; RAP1A; KRAS |
| MEMBRANE TRAFFICKING (REACTOME)  | 87           | 0       | 0            | TJP1         |                  | GJA1; COPG             |
| NEUROTROPHIC FACTOR-MEDIATED TRK RECEPTOR SIGNALING (PID) | 60 | 0 | 0 | TIAM1 | PIK3R1 | BDNF; RAP1A; KRAS |
| EPO SIGNALLING (INOH)            | 180          | 0       | 0            | PIK3R1; GSK3B |                  |                        |
| CDC42 SIGNALING EVENTS (PID)     | 68           | 0       | 0            | TIAM1        | PIK3R1; GSK3B   | EP58; YES1             |
| LICAM INTERACTIONS (REACTOME)    | 93           | 0       | 0            |              |                  |                        |

Dosage Effects on Transcriptional Regulation

Dysregulation of transcriptional regulation is widely reported in DS [34]. Among the 324 significant genes were 13 transcription factors (TFs) (PSIP1, RBP1, TCF4, HES1, ETS2, BACH1, RUNX1, GABPA, SNAI2, REST, LITAF, EGR1, FOS), 6 TFs (PSIP1, HOXC8, DLX5, HIVEP3, ZNF187, ATF6) had significant enrichment of their targets as retrieved by the TRANSFAC [42] database. Additionally, 57 TFs had significant enrichment of their interacting proteins when judged with physical interactions retrieved from the ConsensusPathDB [40]. In total, 70 different TFs were identified as being (directly or indirectly) affected by dosage imbalances. The list of TFs and their associated functional categories is given in Additional file 1, Table S3. GO categories indicate a broad impact of transcriptional regulation for neurological development, the central nervous system development (RUNX1 and TP53), nervous system development (DLX5, FOS, HES1, STAT3 and EP300), axonogenesis (DLX5, NOTCH1 and CREB1), neuron differentiation (HOXC8, NOTCH1 and...
RUNXI), negative regulation of neuron differentiation (HES1, NOTCH1 and REST) and regulation of long-term neuronal synaptic plasticity and learning or memory (EGR1 and JUN). Other prominent categories refer to organ development (RBPJ, ETS2, GABPA and SNAI2) and stress response (ATF6, FOS and RELA).

We further analyzed the promoter sequences of the 324 genes for enrichment of transcription factor binding sites using the AMADEUS software [43]. Significant enrichment was computed for 4 TF motifs, E2F1, RUNX1, STAF/PAX2 and STAT3 (Table 3). Enrichment was evident for RUNX1, which is among the most
studied genes implicated in DS. The implication of E2F1 in DS was also previously reported [34] and could be responsible for impaired cell proliferation documented for hippocampus, cerebellum and astrocytes of DS mouse models.

**Dosage Effects and Molecular Interactions**

Molecular interactions among the 324 significant genes on HSA21 and on other chromosomes exhibited a complex network supporting the important role of physical interactions as transmitter of dosage effects (Figure 2A). The consequences of HSA21 triplication on the interacting genes was fairly stable as Figure 2B demonstrates. For example, DNAJ1 (DnaJ (Hsp40) homolog, subfamily B, member 1) and PPP3CA (protein phosphatase 3, catalytic subunit, alpha isozyme, data not shown), both interacting with SOD1, were consistently and significantly down-regulated in the human microarray experiments as the fold-changes and P-values indicate. Opposite trends were observed for TJP1 and RHOQ.

**Assessing General Relevance of DS Dosage Effects for Neurological Processes**

We were further interested in identifying, among the 324 genes, those which were relevant for other brain disorders. To achieve this, we interrogated 19 independent genes, those which were relevant for other brain disorders including Alzheimer’s disease [34-36]. We were further interested in identifying, among the 324 genes, those which were relevant for other brain disorders (ii) genes that interacted with HSA21 genes, as well as (iii) genes that were associated with DS in the literature (Table 4). Remaining candidates (N = 62) comprised BDNF-related genes (SST), MAPK-pathway genes (KRAS, IGFR1, GNG11 and RAP1A), genes related with leukemia (SFRP1) and Rho-Proteins (DHCR7 and RAB21). SST was found as co-expressed in previous studies with TAC1 [44] which is also significant in our meta-analysis and both showed a strong correlation across DS studies (Figure 4A).

Novel candidates are associated with neurodegenerative disorders including Alzheimer’s disease (VSNL1), prion disease (SCRG1, HSPH1, HSPA5 (Figure 4B) and CTR9) and age-related degeneration (GAS6 and GNG11). Moreover, candidates could explain evident DS features (Additional file 1, Table S5): (i) genes related to neurogenesis and neurite outgrowth (LPARI [45], LIN7C, JARID2, GREM1, SERPINE2, IGFR1 and SPOCK1) that could be related with mental retardation or cognitive impairment, (ii) genes involved in synapsis (AGT, KRAS, ATP1A2, GNAI2, SST and LIN7C) (iii) cytoskeletal related proteins (KANK1 [46]; Figure 4C), CKAP2, CKAP4, HAT1, NEK7 and VAMP3, (iv) macular degeneration genes [47] or genes (HTRA1 and EFEMP1) associated with age-related visual problems [48], (v) genes (AGT, CNN3, FBN1, RBPF, PON2, POSTN, RAP1A, WNK1 and STK39) that were related with cardiac impairments and could be candidates to explain this DS characteristic [49], and (vi) genes related with cancer (BTBD3 [50], DNAJB4 [51], FIBP [52] and GSTZ1 [53]) [54].

**Table 3 Enriched TFBSs**

| TF          | Description                                      | Cromo-some | P-Value      | Binding motif                                      | Strand |
|-------------|--------------------------------------------------|------------|--------------|----------------------------------------------------|--------|
| E2F1        | E2F transcription factor 1                        | chr20      | 9.3*10^-18   | [C/A][C/G][C/A][G/C][G/C][G/C][C/T][G/C]          | -      |
| RUNX1       | runt-related transcription factor 1              | chr21      | 4.0*10^-18   | [C/A][T/A][G/C][G/C][A/T][A/G][C/A][A/G][C/A]    | +      |
| STAF/PAX2   | paired box 2                                     | chr10      | 8.4*10^-17   | [A/G][A/G][C/T][T/C][G/C][A/G][A/G][C/T][G/C]    | +      |
| STAT3       | signal transducer and activator of transcription 3 (acute-phase response factor) | chr17      | 8.4*10^-17   | GAA[A/T][C/T][G/T][C/T][G/A][A/T][C/T][G/A]      | +      |

Binding motifs have been represented using the IUPAC nomenclature and incorporating lower case for low frequency bases.
These examples show that the meta-analysis approach identified multiple additional genes that might be involved in DS pathology. In order to enable the community to check any particular gene of interest for DS relevance in the studies under analysis, we have agglomerated all information of the meta-analysis into a WEB-interface http://ds-geneminer.molgen.mpg.de/. Examples of possible views and information are shown in Figure 4.

**Discussion**

The statistical meta-analysis approach was described previously by Rasche et al. [26]. The score computed with meta-analysis correlates with entropy (Figure 1A) indicating the ability to identify general dosage effects across many experiments that might be of more phenotypic relevance than very specific ones. Additional file 4, Figures S3A and B provide an overview of the different
| Ensembl | HUGO | Score | Entropy | Chromosome | Start position | End position | Band | CNV |
|---------|------|-------|---------|------------|---------------|-------------|------|-----|
| ENSG00000133110 | POSTN | 8.301 | 2.437 | chr13 | 38136722 | 38172981 | q13.3 | YES |
| ENSG00000135919 | SERPIN2 | 5.904 | 3.048 | chr2 | 224839765 | 22490436 | q36.1 | YES |
| ENSG00000172893 | DCR7 | 5.883 | 3.441 | chr11 | 71145457 | 71159477 | q13.4 | YES |
| ENSG00000135744 | AGT | 5.467 | 2.799 | chr1 | 230838269 | 230850043 | q42.2 | YES |
| ENSG00000159176 | C5R1 | 5.424 | 3.136 | chr13 | 201452658 | 201478584 | q32.1 | YES |
| ENSG00000178695 | KCTD12 | 5.344 | 2.373 | chr13 | 77453412 | 77460540 | q22.3 | YES |
| ENSG00000183087 | GAS6 | 5.129 | 2.904 | chr13 | 114523524 | 114567046 | q34 | YES |
| ENSG00000164106 | SCR1 | 5.127 | 2.728 | chr4 | 174309300 | 174320617 | q34.1 | YES |
| ENSG00000166923 | GREM1 | 5.073 | 1.486 | chr13 | 77454312 | 77460540 | q22.3 | YES |
| ENSG00000183087 | GAS6 | 5.129 | 2.904 | chr13 | 114523524 | 114567046 | q34 | YES |
| ENSG00000159176 | CSRP1 | 5.424 | 3.136 | chr13 | 201452658 | 201478584 | q32.1 | YES |
| ENSG00000166923 | GREM1 | 5.073 | 1.486 | chr13 | 77454312 | 77460540 | q22.3 | YES |

Table 4 Novel DS dosage effects

Vilardell et al. BMC Genomics 2011, 12:229
http://www.biomedcentral.com/1471-2164/12/229
Page 10 of 16
Table 4 Novel DS dosage effects (Continued)

| Ensembl | Gene Symbol | Fold Change | Gene Symbol | Fold Change | Gene Symbol | Fold Change | Gene Symbol | Fold Change | Gene Symbol | Fold Change | Gene Symbol | Fold Change |
|---------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| ENSG00000162734 | PEA15 | 3.747 | 3.418 | chr1 | 160175127 | 160185166 | q23.2 |
| ENSG00000103187 | COTL1 | 3.731 | 3.304 | chr16 | 84599200 | 84651683 | q24.1 | YES |
| ENSG00000198648 | STK39 | 3.722 | 3.439 | chr2 | 168810530 | 169104651 | q24.3 |
| ENSG00000100577 | GSTZ1 | 3.713 | 2.759 | chr14 | 77787230 | 77797939 | q24.3 |
| ENSG00000080371 | RAB21 | 3.707 | 3.312 | chr12 | 72146858 | 72181150 | q21.1 | YES |
| ENSG00000136108 | CATION2 | 3.688 | 2.960 | chr13 | 53029495 | 53050485 | q14.3 |
| ENSG00000065683 | ISOC1 | 3.686 | 2.655 | chr5 | 128430442 | 128449721 | q24.3 |
| ENSG00000143420 | ENSA | 3.681 | 3.276 | chr1 | 150573327 | 150602098 | q21.3 |
| ENSG00000114363 | GNAI2 | 3.680 | 3.138 | chr3 | 50263724 | 50296787 | p21.3 | YES |
| ENSG00000140105 | WARS | 3.671 | 2.994 | chr14 | 10800125 | 10842680 | q32.2 |
| ENSG00000018625 | ATP1A2 | 3.670 | 2.733 | chr1 | 160085549 | 160113838 | q23.2 |

sources of data, including two organisms (human and mouse), different tissues (brain, heart and others), different stages of development (adult, postnatal, embryonic) and different mouse models (Ts65DN, Ts1Cje, Tc1). It is per se interesting that, in spite of such heterogeneity, common dosage effects could be identified at all and it should be highlighted that whole-genome data was fairly robust across experiments. Additional file 4, Figure S3D shows the overall correlation of the quantitative values of PCR and microarray values averaged from all experiments with only few genes in the non-concordant sectors of the graph (red points).

The score used in this analysis allows detecting genes that could be either up- or down-regulated in different studies. An overview of the fold-changes for the genes across the different experiments is given in Additional file 1, Table S6. Because genes might change their expression level depending on the developmental state, tissue or because of other variables, we expected that this flexibility allows checking the hypothesis of random disturbances as well as the hypothesis of increased expression of HSA21 genes. We detected a clear enrichment of up-regulated genes on the q-terminal part of HSA21 (Figure 1D and Additional file 3, Figure S2). However, not a single region was identified but rather several smaller regions on HSA21 that agglomerate a large amount of significant dosage effects. This finding was also elaborated before (Korbel et al. [38] and Lyle R et al. [39]) using two independent data sets to characterize the molecular HSA21 regions in a set of DS-patients with partial duplications.

We studied 255 HSA21 genes matched with the probe sets from the microarrays. Of these only 77 showed consistent dosage effects (Figure 1). While 165 HSA21 genes had score values different from zero indicating response in some of the microarray studies, 90 HSA21 genes were not responsive at all and provide evidence for a strong mechanism of dosage compensation. On the other hand, these figures could also reflect the limitation of detecting reliable fold-changes of low magnitude with microarray technology. Furthermore, experiments covered only a limited amount of tissues so that it is likely that some genes were missed simply because they were not responsive in the tissues under analysis. However, having brain as the dominant whole-genome sample source this should ensure expression of most of the genes. Microarray data was focused on the Affymetrix platform in order to reduce variance arising from platform inconsistencies. We have also compared our results with additional studies including own previous research [9] and others [55] and found relevance of selected dosage effects with respect to other tissues as well (data not shown). Additional cross-validation was performed with an independent microarray data set [10]. These authors compared human lymphoblastoid cell lines derived from DS patients and normal controls with a custom-made HSA21 array. Yahya-Graison et al. [10] divided the expression ratios in four classes: class I and class II genes were significantly up-regulated, while class III and IV genes were either compensated or showed variable response. Our meta-analysis revealed a high-degree of concordance taking into account that the cell model, platform and the methodology used were completely different. The meta-analysis scores were significantly higher for class I and II genes than for class III and IV genes (P-value <0.01, Additional file 5, Figure S4). 25 out of 39 class I-II genes revealed a significant score in our meta-analysis (75%).

In this study we monitored molecular interactions of HSA21 genes that might function as drivers of dosage effects (Figure 2A). For example, (i) TIP1 (Tight junction protein ZO-1) interacts with two HSA21 genes, JAM2 and CLDN8, (ii) FO5 (FB) murine osteosarcoma viral oncogene homolog) interacts with HSA21 genes ETS2, SUMO3, RUNXI and indirectly with ERG, (iii) RHQ (ras homolog gene family, member Q) interacts directly with ITSN1 and TIA1 and indirectly with SYN1, and (iv) P1K3R1 interacts directly with IFNAR1 and indirectly with IFNAR2. It should be emphasized that current information on molecular interactions is far from
complete, thus we either might miss important interactions not yet detected and/or we might count false positive interactions due to the high error rates of current annotations of interactions.

Several of the DS genes (N = 79) extrapolated to more general neurological phenotypes (Figure 3A). The dendrogram (Figure 3D) shows further interesting profiles of these genes in the DS samples under analysis: (i) differential gene expression in the cerebellum region versus whole “brain” or cerebrum areas which has been reported in other studies (e.g. Moldrich et al. [56]), (ii) different patterns of gene expression associated to particular developmental stages (P0, P15 and P30); these changes were reported before by Dauphinot et al. [57], and (iii) differences in ES studies.

We further analyzed human and mouse studies separately and found 182 significant dosage effects using only human and 107 dosage effects using only mouse data. The Venn diagram in Additional file 4, Figure S3C clearly shows the benefit in detecting additional dosage effects when mixing the two species. Overlapping dosage effects were detected for 29 genes with both analyses (Additional file 1, Table S7). Results for the human and mouse specific analyses can be found in Additional file 1, Tables S8 and S9. It should be noted here that comparisons between human and mouse using microarrays are inherently difficult and have limitations since the probes for the orthologous mouse and human genes do not correspond well. Furthermore, gene expression variation is generally higher in human individuals compared to mouse inbred strains. Nonetheless, the 107 genes found in the analysis of mouse data (derived from the different mouse models for trisomy 21) represent a core set of genes responsive across different DS mouse models and, thus, could be highly relevant for DS pathogenesis.

In addition to genes commonly related to DS, we have identified novel genes that can be associated with DS phenotypes, in particular with neural development and neurodegeneration. To our best knowledge, this study is the first meta-analysis of genome-wide transcript levels along with other data domains in DS research. The agglomerated data can be accessed through the WEB server at http://ds-geneminer.molgen.mpg.de and the identified dosage effects are a resource for further functional testing and therapeutic development.
Conclusions
We have identified a set of 324 genes with consistent dosage effects from 45 different experiments related to DS. Since the meta-analysis was enriched with brain experiments, we were able to detect a high fraction of genes related to neuro-development, synopsis and neuro-degeneration. Moreover, our results give more information about known and new pathways related to DS and also about 62 novel candidates. The results of the meta-analysis as well as the source data have been made accessible for the community through a WEB interface.

Material and methods
Selection and integration of DS resources
Data sets were selected from heterogeneous technical platforms, different model systems (human cell lines, human tissues, mouse models) and different developmental stages (Additional file 1, Table S1). For each gene and for each source we computed a numerical value that measures its dosage effect. Data categories were either qualitative or quantitative. Qualitative data incorporated a total of 30 published manuscripts including reviews and semi-quantitative studies as well as two SAGE studies [21,58] and were summarised to one score point in order to avoid over-scoring. Here, a “1” referred to the case that the gene was found as DS relevant in one (or more) studies. Quantitative data from differential gene expression studies such as Affymetrix microarrays, RT-PCR, MALDI and other quantitative techniques were evaluated in order to extract comparable information across the different studies. We considered Affymetrix studies that provided the raw data (CEL file level). Raw data were extracted from Gene Omnibus Expression (GEO, [59]), Array Express [60] or were retrieved from the author’s web pages (in total 16 data sets including human tissues and four different mouse models (Ts65Dn, Ts1Cje, Tc1 and Ts + HSA21). Furthermore, we incorporated 18 RT-PCR and MALDI data sets for which the authors displayed the information for all genes under study (either significant or not).

Mapping of gene IDs
A central pre-requisiste of any meta-analysis approach is the consolidation of the different ID types, for example coming from different organisms and from different versions of chips. We used the Ensembl database (version 56) as the backbone annotation for all studies. IDs were mapped to human Ensembl gene IDs. Mapping and merging of the information was done within R and the BioConductor package. In total, information on 19,388 ENSEMBL genes was mapped.

Mapping SAGE IDs
Differential expressed tags were extracted from additional files of the studies. Identifiers (based on sequences) were cross-tagged with the information displayed in the updating tables (SAGEmap_Hs and SAGEmap_Mn) from the SAGE site ftp://ftp.ncbi.nlm.nih.gov/pub/sage/mappings.

Transcriptome data pre-processing and normalization
We incorporated only case-control studies in the meta-analysis in order to derive expression fold-changes. Affymetrix gene chip annotations were adapted from the latest genome annotation (version 12). Affymetrix data were normalized with GC RMA. For transcriptome case-control studies three pieces of information were stored for each gene; (i) the fold-change (DS vs. controls), (ii) the standard error of the fold-change from the replicated experiments in that study and (iii) the expression p-value (presence-call) that indicates whether or not the gene is expressed in the target samples under study. For RT-PCR and MALDI experiments we computed the fold-change of the mean expression (DS vs. controls) as well as the reported standard error of the ratio. When mean and standard variation for each group (DS and controls) was provided we calculated the ratios as well as their associated standard errors.

Scoring DS dosage effects across studies
In order to score the different categories of information such as binary counts and quantitative gene expression values, we summarized the scores of the individual experiments for each category. For microarray studies, the score of the i-th gene in the j-th study, sij, was computed as described in Rasche et al. [26]:

\[
    s_{ij} = \begin{cases} 
    \log_2 \left( \frac{r_{ij}}{p_{ij}} \right) \cdot (1 - p_{ij}) \cdot (1 - e_{ij}/r_{ij}) & \text{if } p_{ij} \leq 0.1, e_{ij}/r_{ij} \leq 1 \\
    0 & \text{else}
    \end{cases}
\]

Here \( r_{ij} \) is the fold-change, \( p_{ij} \) is the average detection p-value and \( e_{ij} \) is the standard error of the ratio derived from the experimental replicates of the study. Thus, the fold-change is weighted with its reproducibility across the experimental replicates and with the likelihood of the gene being expressed in the study’s case or control samples.

For RT-PCR and MALDI studies we applied the following equation:

\[
    s_{ij} = \begin{cases} 
    \log_2 \left( \frac{r_{ij}}{e_{ij}} \right) \cdot (1 - e_{ij}/r_{ij}) & \text{if } e_{ij}/r_{ij} \leq 1 \\
    0 & \text{else}
    \end{cases}
\]

Here \( r_{ij} \) is the fold-change and \( e_{ij} \) is the standard error of the ratio.
The total score of the gene was computed as the sum across all individual study scores.

### Sampling for significance

In order to assess the significance of the overall gene scores we generated random scores by re-sampling the scores 50,000 times with replacement within the same study. Using the random distribution as background we assigned as significant those genes that were above the 99.9 percentile of the background distribution.

### Judging consistency of dosage effects

For each gene, entropy of the score distribution was computed in order to quantify the relevance of the gene across many experiments. Let \( s_{ij} \) be the score of the \( i \)th gene in the \( j \)th study, then \( E_i \) is a measure for the uniformity of the score distribution over the individual experiments:

\[
E_i = - \sum_j f_{ij}
\]

\[
f_{ij} = \begin{cases} s_{ij} \log_2 \left( \frac{s_{ij}}{\sum_k s_{ik}} \right), & s_{ij} > 0 \\ 0, & s_{ij} = 0 \end{cases}
\]

High entropy is assigned to a gene if many experiments contribute to the overall score whereas low entropy is assigned if only a few experiments contribute to the overall score.

### Enrichment analysis

Gene Set Enrichment Analysis (GSEA, [41]) of the 324 genes was performed with respect to pre-defined human pathways agglomerated from 22 pathway resources from the ConsensusPathDB ([40], http://cpdb.molgen.mpg.de). Over-representation analysis of TF target sets was performed with Fisher’s test based on annotation from TRANSFAC [42]. Motif enrichment analyses were performed using AMADEUS [43] with significant genes as background set.

### Selection of independent brain experiments

In order to prove general brain relevance of the 324 genes, we collected DS-independent gene expression studies to decipher brain features, performed with Affymetrix technology and, with experiments deposited in GEO or ArrayExpress (Additional file 1, Table S4). Mostly, these experiments were performed in mouse tissues. For each study we collected one or more resulting gene lists that were evaluated using Gene Set Enrichment Analysis (GSEA, [41]) against the complete list of 19,388 genes ranked by score.

### Additional material

#### Additional file 1: Supplementary tables

- **Table S1.** Data sources used for meta-analysis. Table S2. The 324 candidate genes detected in the meta-analysis study. Table S3. Transcription factors and associated GO terms. Table S4. Cross-validation studies. Table S5. Functional annotation of novel candidates. Table S6. Fold-changes and qualitative data. Table S7. Human and mouse data overlap. Table S8. DS genes derived from meta-analysis of human data. Table S9. DS genes derived from meta-analysis of mouse data.

#### Additional file 2: Figure S1

Enrichment of GO categories for organ, tissue and cell development with respect to the significant HSA21 genes (red bars), the significant non-HSA21 genes (green bars) and the non-significant genes (blue bars).

#### Additional file 3: Figure S2

Genomic location of DS dosage effects in human B) mouse. Significant genes are marked in red, non-significant genes in white.

#### Additional file 4: Figure S3

A) Categorization of the 35 qualitative studies, B) Categorization of the 34 quantitative studies. C) Venn diagram of dosage effects detected with mouse and human data alone and with the combination of all data, D) correlation between average PCR and microarray values for the detected 324 dosage effects.

#### Additional file 5: Figure S4

Cross-validation with DS dosage effects detected with an HSA21 microarray [54]. Box-plots of meta-analysis scores (Y-axis) for class I and II (dosage effects) and class III and IV (compensation and variable expression) genes as judged by the authors.

### Abbreviations

- DS: Down Syndrome; HSA21: human chromosome 21; TF: Transcription Factor; PCR: Polymerase Chain Reaction; RT-PCR: real-time Polymerase Chain Reaction; MALDI: Matrix-Assisted Laser Desorption/Ionization; SAGE: Serial Analysis of Gene Expression; GO: Gene Ontology; ES: Embryonic Stem Cells; IDentifier: GEO: Gene Omnibus Expression; GSEA: Gene Set Enrichment Analysis; CNV: Copy Number Variation; TFBS: Transcriptor Factor Binding Site

### Acknowledgements

We want to express our gratitude to all researchers that made DS data available for the community. The free access to high quality experimental data is the necessary pre-requisite for all integrative studies. Furthermore, we apologize for all data sets that could not be integrated into the analysis because of specific constraints such as chip platforms, access to raw data etc. We thank Bernhard Herrmann for giving access to the in situ mouse brain images shown in Figure 3. We thank Marie-Laure Yaspo for discussions, James Adjaye for proof-reading of the manuscript and Reha Yildirimman and Atanas Kamburov for computational support. This work was funded by the European Commission within its 6th Framework Programme with the grant AnEUploidy (LSHG-CT-2006-037627), by the Max Planck Society and the Beatniu de Pinos postdoctoral fellowship (2008 BP-A 00184).

### Author details

1. Department of Vertebrate Genomics, Max-Planck-Institute for Molecular Genetics, Ihnesr. 63-73, D-14195 Berlin, Germany. 2. Unitat de Genètica, Universitat Pompeu Fabra, y CIBER de Enfermedades Raras (CIBERER), Parc de Recerca Biomèdica de Barcelona, C/Di Aiguader, 88, 08003, Barcelona, Spain. 3. Programa de Medicina Molecular I Genética, Hospital Vall d’Hebron, 08035 Barcelona, Spain.

### Authors’ contributions

MV carried out the systematic revisions, collected the data for the meta-analysis and for the related studies. AR wrote the general code for the meta-analysis. AR and MV adjusted the code for DS study. AT created the browser which allows the results visualization, EMD carried out the transcription factor analysis. MV performed the promoter sequences’ analysis and the
References

1. Patterson D: Genetic mechanisms involved in the phenotype of Down syndrome. Ment Retard Dev Disabil Res Rev 2007, 13(3):199-206.

2. Antonicakis SE, Epstein CJ: The challenge of Down syndrome. Trends Mol Med 2006, 12(10):475-479.

3. Rachidi M, Lopes C: Mental retardation in Down syndrome: from gene dosage imbalance to molecular and cellular mechanisms. Neurosci Res 2007, 59(4):349-369.

4. Gitton Y, Dahmane S, Rof S, O'Doherty A, Lyle R, Borel C, Lin-Marc N, Delom F, Groet J, Scholze M, Martin JE, Shipley J, Denyer G, Epplen CJ, Filliat C, Estivill X, Tybulewicz VL, Fisher EM, Antonicakis SE, Naert D: DYSRK1a-dosage imbalance perturbs NFkappaB: REST levels, deregulating pluripotency and embryonic stem cell fate in Down syndrome. Am J Hum Genet 2008, 83(3):388-400.

5. Lockstone HE, Harris LW, Swatton JE, Wayland AJ, Holland AJ, Bahn S: Gene expression profiling in the adult Down syndrome brain. Genomics 2007, 90(6):647-660.

6. Antonicakis SE, Lyle R, Dermitzakis ET, Reynoard A, Deutsch S: Chromosome 21 and Down syndrome: from genetics to pathophysiology. Nat Rev Genet 2004, 5(10):725-38.

7. Wiseman FK, Allford KA, Tybulewicz VL, Fisher EM: Down syndrome--recent progress and future prospects. Hum Mol Genet 2009, 18(11): R75-83.

8. Kahlper M, Sultan M, Hervig R, Steinfath M, Balzereit D, Eppens B, Saran NG, Pletcher MT, South ST, Stetten G, Lehrach H, Reeves RH, Yaspo ML: Transcript level alterations reflect gene dosage effects across multiple tissues in a mouse model of down syndrome. Genome Res 2004, 14(7):1258-1267.

9. Ait Yahya-Graison E, Aubert J, Dauphinot L, Rivals I, Prieur M, Golfier G, Martin JE, Shipley J, Denyer G, Epplen CJ, Filliat C, Estivill X, Tybulewicz VL, Fisher EM, Antonicakis SE, Naert D: DYSRK1A-dosage imbalance perturbs NFkappaB:REST levels, deregulating pluripotency and embryonic stem cell fate in Down syndrome. Am J Hum Genet 2008, 83(3):388-400.

10. Chou CY, Liu LY, Chen CY, Tsai CH, Hwa HL, Chang LY, Lin YS, Hsieh FJ: Gene expression variation increase in trisomy 21 tissues. Mamm Genome 2008, 19(6):398-405.

11. Gardiner K, Davison MT, Covic LS: Building protein interaction maps for Down's syndrome. Brief Funct Genomic Proteomic 2004, 3(2):142-156.

12. Shin JH, Guleserian T, Wetzdoerfer R, Fountoulakis M, Lubec G: Derangement of hypothetical proteins in fetal Down's syndrome brain. Neurochem Res 2004, 29(6):1307-1316.

13. Seregaa Z, Robbertoux FL, Jaramon M, Sournire-Mourat B: Mouse models of trisomy disorders in trisomy 21: a review. Behav Genet 2006, 36(3):387-404.

14. Gropp A, Kolbus U, Giers D: Systematic approach to the study of trisomy in the mouse. II. Cytogenet Cell Genet 1975, 14(1):42-62.

15. Reeves RH, Irving NG, Moran TH, Wohr A, Kit J, Sisodia SS, Schmidt C, Bronson RT, Davison MT: A mouse model for Down syndrome exhibits learning and behaviour deficits. Nat Genet 1995, 11(2):177-184.

16. Saran NG, Pletcher MT, Natale JE, Cheng Y, Reeves RH: Global disruption of the cerebellar transcriptome in a Down syndrome mouse model. Hum Mol Genet 2003, 12(16):2013-2019.

17. Amano K, Sato H, Uchikawa C, Suzuki T, Kotirowa SE, Nukina N, Epstein CJ, Yamakawa K: Dosage-dependent over-expression of genes in the trisomic region of Ts1Cje mouse model for Down syndrome. Hum Mol Genet 2004, 13(13):1333-1340.

18. Rhodes DR, Chintiramay AM: Integrative analysis of the cancer transcriptome. Nat Genet 2005, 37 Suppl: S31-37.

19. Bertram L, McQueen MB, Mullin K, Blacker D, Tanz RE: Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. Nat Genet 2007, 39(1):17-23.

20. Rasche A, Al-Husaini H, Hervig R: Meta-analysis approach identifies candidate genes and associated molecular networks for type-2 diabetes mellitus. BMC Genomics 2008, 9:110.

21. Sbai O, Devi TS, Melone MA, Feron F, Khechtatschi M, Singh LP, Perrone L: RAGE-TNXP1 axis is required for S100B-promoted Schwann cell migration, fibronectin expression and cytokine secretion. J Cell Sci 2010, 123(Pt24):4332-9.

22. Panikh H, Carlson E, Chukwaw LA, Johansson LE, Storgaard H, Poulsen P, Saxena R, Ladd C, Schulte PC, Mazzini MJ, Jensen CB, Krook A, Bjarnholdt M, Tormoyd H, Zierath JR, Riddlestråle D, Alshulker D, Lee RT, Vaag A, Groop LC, Mootha VK: TNXP1 regulates peripheral glucose metabolism in humans. PLoS Med 2007, 4(5):e158.

23. Yamawaki H, Haendeler J, Berk BC: Thioredoxin: a key regulator of cardiovascular homeostasis. Circ Res 2003, 93(11):1029-1033.

24. Austin C: Does oxidative damage contribute to the generation of leukemia? Leuk Res 2009, 33(10):1297.

25. Flicker P, Aiken BL, Ballester B, Beal K, Bragin E, Brent S, Chen Y, Clapham P, Coates G, Fairley S, Fitzgerald S, Fernandez-Banet J, Gordon L, Gral F, Haider S, Hammond M, Howe K, Jenkinson A, Johnson N, Kahán A, Khee D, Keenan S, Kinsella R, Kokocinski F, Koscielski G, Kulesza E, Lawson D, Longdon I, Massingham T, McLaren W, et al: Ensemble's 10th year. Nucleic Acids Res 2011, 38 Database: D557-562.

26. Dierssen M, De Lanagra MW: DYSRK1A (dual-specificity tyrosine-phosphorylated and -regulated kinase 1A): a gene with dosage effect during development and neurogenesis. ScientificWorldJournal 2006, 6:1911-1922.

27. Edwards H, Xie C, LaFiura KM, Dombkowski AA, Buck SA, Boerner JL, Taub JW, Matherly LH, Ge Y: RUNX1 regulates phosphoinositide 3-kinase/AKT pathway: role in chemotherapy sensitivity in acute megakaryocytic leukemia. Blood 2009, 114(13):2744-2752.

28. Gardiner K: Transcriptional dysregulation in Down syndrome: predictions for altered protein complex stoichiometries and post-translational modifications, and consequences for learning/behavior genes ELK, CREB, and the estrogen and glucocorticoid receptors. Behav Genet 2006, 36(3):439-453.

29. Butler C, Knox AJ, Bowersox J, Forbes S, Patterson D: The production of transgenic mice expressing human cystathionine beta-synthase to study Down syndrome. Behav Genet 2006, 36(3):429-38.

30. Belichenko PV, Krilchenkov AM, Salek A, Epstein CJ, Mobley WC: Synaptic and cognitive abnormalities in mouse models of Down syndrome: exploring genotype-phenotype relationships. J Comp Neurol 2007, 504(4):329-345.

31. Ronan A, Fagan K, Christie L, Conroy J, Nowak NJ, Turner G: Familial 4.3 Mb duplication of 21q22 sheds new light on the Down syndrome critical region. J Med Genet 2007, 44(7):448-457.

32. Korbel JO, Tirosh-Wagner T, Urban AE, Chen XN, Kasovics M, Dai L, Grubert F, Erdman C, Gao MC, Lange K, Sobel EM, Barlow GM, Aylsworth AS, Carpenter NI, Clark RD, Cohen MY, Doran E, Falk-Zaccari T, Lewin SD, Lott IT, McGillivray BC, Moechser JB, Pettenati MJ, Pueschel SM, Rao KW, Shaffer LG, Shohat M, Van Riper AJ, Warburton D, Weissman S, et al: The genetic architecture of Down syndrome phenotypes revealed.
by high-resolution analysis of human segmental trisomies. Proc Natl Acad Sci USA 2009, 106(29):12031-12036.

39. Lyle R, Béna F, Gago S, Gehing C, Lopez G, Schinzel A, Lespinasse J, Bottani A, Dahoun S, Tarie L, Diaco-Fenyi M, Cornillet-Lefèvre P, Peket A, Lyonnard S, Toutain A, Colleaux L, Horst J, Kenneke Mecht I, Wakamatsu M, Descartes M, Franklin JC, Florentin-Arar L, Kitisou S, Ait Yaha-Gasion E, Costantini M, Snet PM, Delabar JM, Antonarakis SE: Genotype-phenotype correlations in Down syndrome identified by array CGH in 30 cases of partial trisomy and partial monosomy chromosome 21. Eur J Hum Genet 2009, 17(4):454-66.

40. Kamburov A, Weirting C, Lehrach H, Henrig V: ConsensusPathDB—a database for integrating human functional interaction networks. Nucleic Acids Res 2009, 37 Database: D623-628.

41. Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP: GSEA-P: a desktop application for Gene Set Enrichment Analysis. Bioinformatics 2007, 23(23):3251-3253.

42. Wingender E, Kel AE, Kel OV, Karas H, Heinemeyer T, Dietze P, Knuppel R: transcription factors YY1 and activator protein 1. Synergistic activation of the tumor suppressor, HLJ1, by the prefrontal cortex: combined analysis of gene expression microarrays. Nucleic Acids Res 1997, 25(1):265-268.

43. Linhart C, Halperin Y, Shamir R: Transcription factor and microRNA motif discovery: the Amadeus platform and a compendium of metazoan target sets. Genome Res 2008, 18(7):1180-1189.

44. Choi KH, Elashoff M, Higgs BW, Song J, Kim S, Sabunciyan S, Diglisic S, Park KE, Mosley AN, Chun J: Expression of fibroblast growth factors YY1 and activator protein 1. Cancer Res 2007, 67(10):4816-4826.

45. Choi JW, Herr DR, Noguchi K, Yung YC, Lee CW, Mutoh T, Lin ME, Teo ST, Vilardell et al: Glutathione transferases. Nat Genet 2005, 37 Database: D868-872.

46. Kakinuma N, Zhu Y, Wang Y, Roy BC, Kiyama R: Regulation of the CD56 promoter and its association with proliferation, apoptosis and clinical features. Hum Mol Genet 2010, 19(17):3498-3513.

47. Allikmets R, Dean M: Bringing age-related macular degeneration into focus. Nat Genet 2008, 40(7):820-821.

48. Kamburov A, Weirting C, Lehrach H, Henrig V: ConsensusPathDB—a database for integrating human functional interaction networks. Nucleic Acids Res 2009, 37 Database: D623-628.

49. Damgaard T, Knudsen LM, Dahl IM, Gimsing P, Lodahl M, Rasmussen T: Regulation of the CD56 promoter and its association with proliferation, apoptosis and clinical features. Hum Mol Genet 2010, 19(17):3498-3513.

50. Damgaard T, Knudsen LM, Dahl IM, Gimsing P, Lodahl M, Rasmussen T: Regulation of the CD56 promoter and its association with proliferation, apoptosis and clinical features. Hum Mol Genet 2010, 19(17):3498-3513.

51. Chen T, Chilenski M, Chen S, Wang R, Ilagan M: The cerebellar transcriptome during postnatal development. J Neurosci 2002, 22(9):3698-3710.

52. Chen R, Chen L, Li W, Wang C, Juhn SK, Ondrey FG, Lin J: Meta-analysis of heterogeneous Down Syndrome data reveals consistent genome-wide dosage effects related to neurological processes. BMC Genomics 2011 12:229.