View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

ARTICLE

Classification of Human Chromosome 21 Gene-Expression Variations in Down Syndrome: Impact on Disease Phenotypes

E. Aït Yahya-Graison, J. Aubert, L. Dauphinot, I. Rivals, M. Prieur, G. Golfier, J. Rossier, L. Personnaz, N. Créau, H. Bléhaut, S. Robin, J. M. Delabar, and M.-C. Potier

Down syndrome caused by chromosome 21 trisomy is the most common genetic cause of mental retardation in humans. Disruption of the phenotype is thought to be the result of gene-dosage imbalance. Variations in chromosome 21 gene expression in Down syndrome were analyzed in lymphoblastoid cells derived from patients and control individuals. Of the 359 genes and predictions displayed on a specifically designed high-content chromosome 21 microarray, one-third were expressed in lymphoblastoid cells. We performed a mixed-model analysis of variance to find genes that are differentially expressed in Down syndrome independent of sex and interindividual variations. In addition, we identified genes with variations between Down syndrome and control samples that were significantly different from the gene-dosage effect (1.5). Microarray data were validated by quantitative polymerase chain reaction. We found that 29% of the expressed chromosome 21 transcripts are overexpressed in Down syndrome and correspond to either genes or open reading frames. Among these, 22% are increased proportional to the gene-dosage effect, and 7% are amplified. The other 71% of expressed sequences are either compensated (56%, with a large proportion of predicted genes and antisense transcripts) or highly variable among individuals (15%). Thus, most of the chromosome 21 transcripts are compensated for the gene-dosage effect. Overexpressed genes are likely to be involved in the Down syndrome phenotype, in contrast to the compensated genes. Highly variable genes could account for phenotypic variations observed in patients. Finally, we show that alternative transcripts belonging to the same gene are similarly regulated in Down syndrome but sense and antisense transcripts are not.

Down syndrome (DS [MIM #190685]) results from the triplication of chromosome 21 and is the most common genetic cause of mental retardation in humans, occurring in ~1 in 800 newborns. The phenotype of DS is characterized by >80 clinical features, including cognitive impairments, muscle hypotonia, short stature, facial dysmorphisms, congenital heart disease, and several other anomalies.[1] These clinical features can vary considerably in number and in severity,[2] and certain abnormalities, such as acute megakaryoblastic leukemia and Hirschsprung disease, occur at higher frequencies in patients with DS than in the general population.

Trisomy 21 has been known to be the cause of DS since 1959, when Lejeune and colleagues demonstrated the presence in three copies of chromosome 21 in persons with DS.[3] The phenotype of DS is thus thought to be the result of gene-dosage imbalance. However, the molecular mechanisms by which such dosage imbalance causes abnormalities remain poorly understood. Two different hypotheses have been proposed to explain the phenotype of DS: "developmental instability" (loss of chromosomal balance) and "gene-dosage effect." According to the developmental instability hypothesis, the presence of a supernumerary chromosome globally disturbs the correct balance of gene expression in DS cells during development.[4,5] However, this hypothesis is weakened by the fact that other autosomal trisomy syndromes do not lead to the same clinical pattern.[6] Moreover, correlations between genotype and phenotype in patients with partial trisomies indicate that a restricted region in 21q22.2 is associated with the main features of DS, including hypotonia, short stature, facial dysmorphies, and mental retardation.[7–9] This DS chromosomal region (DCR) supports the alternative gene dosage–effect hypothesis, which postulates that the restricted number of genes from chromosome 21 that are overexpressed in patients with segmental trisomies contributes to the phenotypic abnormalities.

To determine which hypothesis applies to the etiology of DS, several gene-expression studies of human DS cells or tissues have been conducted.[10–17] Most of these studies have shown a global up-regulation of the three-copy genes mapping to the trisomic chromosome, but the limited number of studied DS cases restricted the statistical analysis and did not allow the identification of precise gene deregulation. Moreover, these studies were performed using a small number of three-copy genes. Several other ex-

From the Neurobiologie et Diversité Cellulaire, Unité Mixte de Recherche 7637 du Centre National de la Recherche Scientifique et de l'Ecole Supérieure de Physique et de Chimie Industrielles de la Ville de Paris (E.A.Y.-G.; L.D.; G.G.; J.R.; M.-C.P.), Université Paris Diderot-Paris (E.A.Y.-G.; N.C.; J.M.D.), Unité Mixte de Recherche de l'Ecole Nationale du Génie Rural, des Eaux et des Forêts, de l'Institut National Agronomique Paris-Grignon et de l'Institut National de la Recherche Agronomique (J.A.; S.R.), Equipe de Statistique Appliquée (I.R.; L.P.), Service de Cytogénétique, Hôpital Necker Enfants Malades (M.P.), and Institut Jérôme Lejeune (H.B.), Paris

Received March 13, 2007; accepted for publication May 21, 2007; electronically published July 19, 2007.

Address for correspondence and reprints: Dr. M.-C. Potier, Neurobiologie et Diversité Cellulaire, CNRS UMR 7637, ESPCI, 10 rue Vauquelin, 75005 Paris, France. E-mail: marie-claude.potier@espci.fr

Am. J. Hum. Genet. 2007;81:475–491. © 2007 by The American Society of Human Genetics. All rights reserved. 0002-9297/2007/8103-0006$15.00
DOI: 10.1086/520000

Table 1. Experimental Design

	Men with DS							Women with DS		
Controls	1	2	3	4	5	6	7	8	9	10
Men:										
11	−1	1		−1		1			−1	1
12				1	−1			1		−1
13		−1	1			−1		1	1	−1
14					1			−1	−1	1
Women:										
15		1						−1		
16					−1				1	
17					1		−1	1		−1
18			−1			1		−1	1	
19	1								−1	
20		−1							1	
21						−1	1		−1	1

NOTE.—Microarray experiments were performed using LCLs from individuals with DS and control individuals in accordance with a mixed model (see the "Material and Methods" section). Each "1" indicates one experiment. "+1" means that DS and control samples were labeled with Cy5 and Cy3, respectively. "−1" means that DS and control samples were labeled with Cy3 and Cy5, respectively.

periments have been done on animal models of DS with a greater number of chromosome 21 gene orthologs by use of microarray and quantitative PCR experiments.[18–21] In these studies, the three-copy genes were overexpressed, with a mean ratio of 1.5, which is proportional to the gene-dosage imbalance. However, some of these triplicated genes appeared to escape the "1.5-fold rule." Yet, these animal models are not trisomic for all chromosome 21 orthologs. Thus, a comprehensive classification of all human genes on chromosome 21, according to their level of expression in DS, does not yet exist.

The goal of the present study was to fill this knowledge gap and to find the genes that are likely to be involved in DS phenotypes through their transcriptional dysregulation.[22] For this purpose, we designed an oligonucleotide microarray containing all chromosome 21 genes, ORFs, antisense transcripts, and predicted genes listed in the most common databases (NCBI Gene Database, Eleanor Roosevelt Institute, and Max Planck Institute), except for the 53 genes of the keratin-associated protein cluster. Gene expression was measured on lymphoblastoid cell lines (LCLs) from 10 patients with DS and 11 control individuals. LCLs are easy to obtain and are widely used to study genotype-phenotype correlations.[23] To our knowledge, this is the most comprehensive study so far that has been done using triplicated genes in DS human cells. In addition, we analyzed data with a mixed-model analysis of variance, to find genes that are differentially expressed in DS independent of sex and interindividual variations. Our data show a global gene dosage–dependent expression of chromosome 21 genes in LCLs, with no effect of sex. In addition, by use of our data-analysis protocol, chromosome 21 genes can now be classified into four classes: class I genes are overexpressed with a mean ratio very close to 1.5, proportional to the gene-dosage effect of trisomy 21; class II genes are overexpressed with ratios significantly >1.5, reflecting an amplification mechanism; class III genes have ratios significantly <1.5, corresponding to compensated genes; and class IV genes have expression levels that are highly variable between individuals. This classification should have an impact on the search for genes that are involved in the DS phenotype.

Material and Methods

Cell Lines and Culture Conditions

LCLs were derived from the B lymphocytes of 10 patients with DS collected from the cytogenetic service of the hospital Necker Enfants Malades and the Institut Jérôme Lejeune. Parents of patients from the Institut Jérôme Lejeune gave their informed consent, and the French biomedical ethics committee gave its approval for this study (Comité de Protection des Personnes dans la Recherche Biomédicale number 03025). Written informed consent was obtained from the participants or from their families by the cytogenetic service of Hôpital Necker Enfants Malades. Cell lines from 11 control individuals were also obtained with their written informed consent, for comparison of chromosome 21 gene-expression profiles. Culture media consisted of Opti-MEM with GlutaMax (Invitrogen) supplemented with 5% fetal bovine serum from a unique batch and 1% penicillin and streptomycin mix (10,000 U/ml). Cell lines were grown at 37°C in humidified incubators, in an atmosphere of 5% CO_2. Each culture was grown to at least 60×10^6 cells. All cell lines were karyotyped, to confirm their trisomic or euploid status and also to verify that immortalization by the Epstein-Barr virus (EBV) did not produce any visible chromosomal rearrangement other than trisomy 21. Cells

Table 2. List of Oligonucleotide Primers Used in the QPCR Experiments

Primer	Sequence (5′→3′)
CHAF1B_UP	CCATCATATGGGATGTCAGCAA
CHAF1B_LOW	CTTCATGCTGTCGTCGTGAAAC
CSTB_UP	GCCACCGCCGAGACCCAGCA
CSTB_LOW	TGGCTTTGTTGGTCTGGTAG
DSCR1_UP	GCACAAGGACATTTGGGACT
DSCR1_LOW	TTGCTGCTGTTTTCACAACC
DYRK1A_UP	ATCCGACGCACCAGCATC
DYRK1A_LOW	AATTGTAGACCCTTGGCCTGGT
GART_UP	CTGGGATTGTTGGGAACCTGAG
GART_LOW	ACCAAAGCAGGGAAGTCTGCAC
H2BFS_UP	CAGAAGAAGGACGGCAGGAA
H2BFS_LOW	GAAGCCTCACCTGCGATGCG
MX1_UP	GCCAGTATGAGGAGAAGGTGCG
MX1_LOW	GTTTCAGCACCAGCGGGCATCT
SNF1LK_UP	GCCGCTTCCGCATCCCCTTCTT
SNF1LK_LOW	CTCATCGTAGTCGCCCAGGTTG
SOD1_long_UP	TCGCCCAATAAACATTCCCTTG
SOD1_long_LOW	AAGTCTGGCAAAATACAGGTCATTG
STCH_UP	GGACGTGGCCTTTCTGATAA
STCH_LOW	CTTGACGGATCCGAGGAATA
TMEM1_UP	CGTGCAGGAACTGAAGCTCTTA
TMEM1_LOW	TCTGAGCTGTGTTGGCTGTTTC
L13852_UP	CTCCAATCTCAGCCGTCAGT
L13852_LOW	AGCCACACCATCCACACGGG
AB000468_UP	CAAGAAAGCGTCGTGGTGGA
AB000468_LOW	ATCGTCACTGCTCACCACAC

Table 3. HSA21 Oligoarray Content

Putative Expressed Sequences	HSA21 Content[a]	HSA21 Oligoarray Content	
		No. of Sequences[b]	HSA21 Coverage[c] (%)
Genes	182	145	82.42
ORFs	93	58	62.37
Predictions	Not represented	118[d]	NA
Antisense transcripts	Only 1 represented	18	NA

[a] NCBI Gene Database build 36.2 was used to estimate HSA21 gene content. Only current sequences were considered, with the exception of pseudogenes and hypothetical proteins.

[b] The number of HSA21 sequences represented by at least one probe on the HSA21 oligoarray.

[c] The percentage of HSA21 sequences currently annotated in NCBI Gene Database that are represented on the microarray. NA = not available.

[d] Of the 118, 20 are represented with their reverse sequence.

were harvested by centrifugation, were washed in 5 ml PBS, followed by another centrifugation, and were stored at −80°C.

Human Chromosome 21 (HSA21) Oligoarray

A dedicated oligonucleotide microarray—named "HSA21 oligoarray," containing 664 50-mer amino-modified oligonucleotides representing 145 genes, 58 ORFs, 118 predictions (20 of them represented in both orientations), and 18 antisense transcripts assigned to chromosome 21—was used in the present study. Predictions represented on the array included cDNAs and exons from the *CBR-ERG* region on 21q deduced from cDNA isolation and exon-trapping experiments[8] and gene or exon predictions produced from *in silico* analysis of the complete sequence of human chromosome 21.[24] Nonredundant transcript sequences and antisense transcripts were also included in this oligoarray.[25–27] Thirty-nine genes assigned to chromosomes other than chromosome 21, represented by 58 oligonucleotides—showing a wide range of expression levels according to UniGene and no variations between DS and control samples as demonstrated by the first version of the HSA21 oligoarray (data not shown)—were added for data normalization. All probes present on the array were designed using the SOL software (G.G., S. Lemoine, A. Bendjoudi, J.R., S. Lecrom, and M.-C.P., unpublished data). Sequences were then synthesized by EuroGentec and were spotted onto CodeLink activated glass slides (Amersham Biosciences) by use of a MicroGrid II spotter (Biorobotics). Each array contained two matrices with eight blocks each, in which probes were present in duplicates so that each oligonucleotide was present in four replicates on each slide.

Table 4. Classification of HSA21 Genes by Statistical Analysis

DS/Control Ratio	Class by DS/Control Ratio	
	Not Significantly Different from 1	Significantly Different from 1
Significantly >1.5	…	II
Not significantly different from 1.5	IV	I
Significantly <1.5	III	…

Experimental Design

The experiment comprised 10 patients with DS (7 men and 3 women) and 11 controls (4 men and 7 women). Samples from the same individual were used in different hybridizations.

For each gene, we used the linear model

$$y_{ijklm} = \mu + D_i + S_j + A_l + F_m + DS_{ij} + DF_{im} + SF_{jm} + I(DS)_{ijk} + \varepsilon_{ijklm}\ , \qquad (1)$$

where y_{ijklm} is the normalized expression of the gene in $\log_2$ for factor i (DS or control sample), sex is j, patient number is k ($k = 1, \ldots, 21$), dye label is m (m is red, for Cy5, or green, for Cy3) on the HSA21 oligoarray l. The symbols D, S, A, and F represent the fixed effects due to the disease, sex, array, and fluorochrome, respectively. For example, D represents the modifications of the gene expression level due to the disease. A and F are both nuisance parameters that account for potential technological biases. DS, DF, and SF correspond to interacting effects: disease and sex, disease and fluorochrome, and sex and fluorochrome, respectively. The symbol $I(DS)$ refers to the patient (nested within disease and sex) random effect. This last effect accounted for the correlation between samples used in different hybridizations but collected from the same patient.

We assumed the independence between all $I(DS)_{ijk}$ and E_{ijklm}. We also assumed that $I(DS)_{ijk}$ was independent with a distribution $N(0,s^2)$ and that E_{ijklm} was independent with distribution $N(0,\sigma^2)$.

Model (1) can be rewritten under the matrix form

$$\mathbf{Y} = \mathbf{X}\theta + \mathbf{ZU} + \mathbf{E}\ , \qquad (2)$$

where θ is the vector of fixed effects (D, S, A, F, DS, DF, and SF), $\mathbf{U}$ is the vector of $I(DS)_{ijklm}$, and $\mathbf{E}$ is the vector of E_{ijklm}. $\mathbf{Y} \sim N(X\theta,\Sigma)$, where $\Sigma = 2\sigma_g^2 Id + s_g^2 ZZ^T$. $\mathbf{Y}$ has n rows (n is the total number of samples) and one column, $\mathbf{X}$ is the matrix describing the status of the patient (disease and sex) from which the sample was collected. $\mathbf{Z}$ has n rows and I columns (I is the total number of patients) and describes the correspondence between samples and patients.

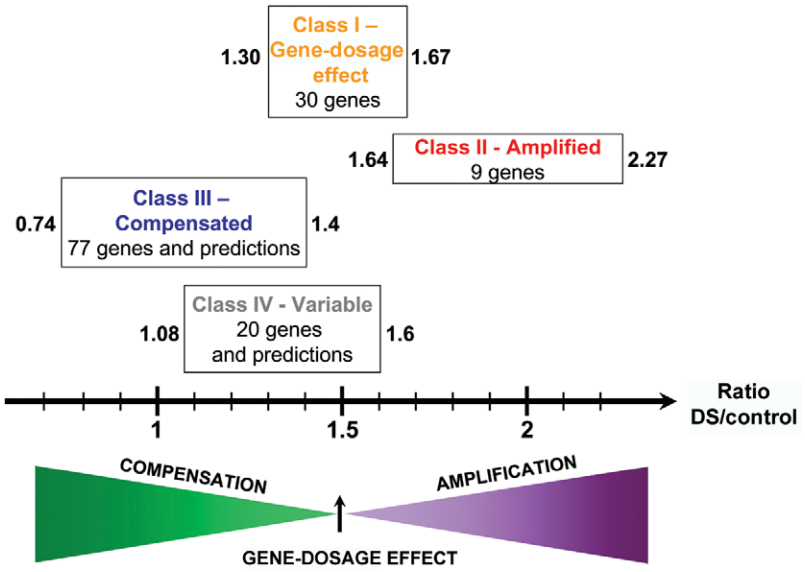

Figure 1. Classification of HSA21 genes according to the expression ratio between DS and control LCLs. The sum of classified genes is 136 genes minus 2 (*C21ORF108* and *PRMT2*) that appear twice, depending on the oligonucleotide probe considered (see the "Results" section for details).

On each array *l*, we actually observed the differential expression (red signal minus green signal)

$$\begin{aligned} y_{ii'jj'kk'lmm'} = y_{ijklm} - y_{i'j'k'lm'} &= (D_i - D_{i'}) + (S_j - S_{j'}) \\ &+ (DS_{ij} - DS_{i'j'}) + (F_m - F_{m'}) \\ &+ (DF_{im} - DF_{i'm'}) + (SF_{im} - SF_{i'm'}) \\ &+ [I(DS)_{ijk} - I(DS)_{i'j'k'}] + (\varepsilon_{ijklm} - \varepsilon_{i'j'k'lm'}) \ . \end{aligned} \quad (3)$$

This model can be rewritten under another matrix, Δ, describing the comparisons performed on each array. This matrix had L rows (L is the total number of arrays) and n columns. The lth row of Δ was zero except for the value 1 in the column corresponding to the sample labeled in red and -1 in the column corresponding to the sample labeled in green. The model for the differential expression was obtained by multiplying all terms of equation (2) by Δ:

$$\Delta\mathbf{Y} = \Delta\mathbf{X}\theta + \Delta\mathbf{Z}\mathbf{U} + \Delta\mathbf{E} \ .$$

The vector of differential expression $\Delta\mathbf{Y}$ has distribution $N(\Delta\mathbf{X}\theta, \Delta\Sigma\Delta^{\mathrm{T}})$.

The experimental design was defined by the three matrices $\mathbf{X}$, $\mathbf{Z}$, and Δ. $\mathbf{X}$ and $\mathbf{Z}$ basically depend on the number of samples for each patient. Because the microarrays used were two-color assays, the total number of samples was twice the number of slides.

The important remaining choice was the comparison to be made—that is, the choice of Δ. Consideration of the differential expression between two patients analyzed on the same array eliminated the array effect and the corresponding constants. We were not interested in other technical effects, such as fluorochrome or interactions of fluorochrome with other effects. To eliminate DF_{im} and SF_{jm} effects, we proposed a balanced design for the fluorochrome effect. Finally, data were normalized across genes, to remove the dye effect F_m.

The last criterion was Δ—the precision of the estimated effects (gathered into the vector $\hat{\theta}$). According to the mixed linear model theory, this precision is given by its variance matrix,

$$V(\hat{\theta}_A) = [\Delta^T\mathbf{X}^T(\Delta\Sigma\Delta^T)^{-1}\Delta\mathbf{X}]^{-1} \ ,$$

which depends on Δ and the ratio σ^2/s^2. The diagonal contained all the information about the quality of the estimates. It gave the variance of the estimates of all effects of interest to D_i, S_j, and DS_{ij}. This matrix was the ultimate tool for comparing the designs.

We calculated the variance of the disease effect (which is of primary interest) and the determinant of the variance covariance matrix (which gave a global measure of the precision of the estimates) for a certain number of designs Δ. To do that, we used a value given a priori for the ratio σ^2/s^2 equal to 2.

We finally chose a design involving 40 arrays. Each patient appeared in two to eight different experiments, and samples from the same patient were marked the same number of times with each fluorescent dye (Cy3 or Cy5). On each array, a DS LCL and a control LCL were compared, to increase the precision of the disease effect. Ten arrays compared (i) a man with DS and an unaffected man, (ii) a female with DS and an unaffected female, (iii) a man with DS and an unaffected female, or (iv) a female with DS and an unaffected male. The design is described in table 1.

mRNA Extraction, HSA21 Oligoarray Hybridization, Data Filtering, and Normalization

mRNA was extracted from frozen individual cell samples by use of Fast Track 2.0 mRNA Isolation kit (Invitrogen) in accordance with the manufacturer's instructions. To eliminate DNA contamination, the appended DNase protocol of RNeasy mini kit (Qia-

gen) was used in accordance with the manufacturer's protocol. Samples were further tested for purity and quantity with RNA 6000 NanoChips by use of the Agilent 2100 Bioanalyzer (Agilent Technologies). By use of the Reverse-iT RTase Blend kit (ABGene), 2 μg of mRNA was converted into Cy3- or Cy5-labeled cDNA by incorporation of fluorescent dUTP (Amersham). Labeled targets were then purified on Qiaquick columns in accordance with the manufacturer's protocol (Qiagen). Hybridization of sample pairs on HSA21 oligoarrays (one DS sample and one control sample), according to the experimental design, was performed using hybridization buffer (50% formamide, 4× saline sodium citrate [SSC], 0.1% SDS, and 5× Denhart) at 42°C overnight. Slides were washed in 2× SSC and 0.1% SDS three times for 5 min, in 0.2× SSC for 1 min, and in 0.1× SSC for 2 min. Data were acquired with GenePix 4000B scanner and by use of the GenePix Pro 6.0 software (Axon). For each array, the raw data comprised the median feature pixel intensity at wavelengths 635 nm and 532 nm for Cy5 and Cy3 labeling, respectively. After subtraction of the background signal, LOWESS normalization[28] of the *M* values corresponding to Cy5/Cy3 signal ratios in $\log_2$ was applied to all oligonucleotides representing non-HSA21 genes and was used to calculate a correction factor applied to *M* values for HSA21 probes under The R Project for Statistical Computing. Normalized data from each slide was then filtered using two criteria: (i) for each oligonucleotide, at least two values among the four replicates had to be available, and (ii) SD of values corresponding to the geometric mean in $\log_2$ of Cy3 and Cy5 signal intensities (the *A* value) had to be <1. Arithmetic means of normalized and filtered *M* and *A* values were calculated for each oligonucleotide and were submitted to the statistical analysis. All microarray data used in this study were deposited in the Gene Expression Omnibus (GEO) database (accession number GSE6408).

Expression Data Analysis: Statistical Testing

Mixed model.—To determine differentially expressed genes for DS, sex, and DS × sex effects, we performed a mixed-model analysis of variance according to the experimental design. This method was chosen to distinguish between interindividual variability and experimental variability.[29] We used the mixed procedure of the SAS software with the restricted maximum likelihood (REML) method of estimation.[30] After the filtering and normalization steps, the number of observations per spot varied between 8 and 40, which was enough to calculate the variance for each gene.

We first tested the effects of the complete model (3). Since the sex and DS × sex effects were not significant for any gene, these two effects were dropped from the model. We finally analyzed the simplified model

$$y_{ii'jj'kk'l} = y_{ijkl} - y_{i'j'k'l} = (D_i - D_{i'}) + [I(DS)_{ijk} - I(DS)_{i'j'k'}] + (\varepsilon_{ijklm} - \varepsilon_{i'j'k'l'm'}) \ .$$

We deduced raw *P* values from comparison with 1 under the null hypothesis and adjusted them by the Benjamini-Hochberg procedure, which controls the false-discovery rate (FDR).[31] We then analyzed the simplified model by comparison with 1.5 under the null hypothesis and adjusted the raw *P* values by use of the method described by Storey et al.[32]

Principal-components analysis (PCA).—Results from microarray experiments were obtained as the differential expression between DS and control samples. *M* values corresponded to $\log_2(\text{DS}) - \log_2(\text{control})$, and *A* values to $[\log_2(\text{DS}) + \log_2(\text{control})]/2$. For each probe *p*, the mean value of its expression (in $\log_2$) in DS cell lines and in controls could thus be expressed as

$$\left\{ \begin{array}{ll} E_p^i = \frac{1}{N_i}\sum_{k\in S_i}\left(A_p^k + \frac{M_p^k}{2}\right) & \text{for } i = 1\text{–}10 \text{ (DS)} \\ E_p^i = \frac{1}{N_i}\sum_{k\in S_i}\left(A_p^k - \frac{M_p^k}{2}\right) & \text{for } i = 11\text{–}21 \text{ (controls)} \end{array} \right\} \ ,$$

where *i* denoted the index of the individual, S_i the set of slides on which all samples from the individual *i* have been hybridized, and N_i the size of this set. Since the overall expression level of the genes of one individual varied from one slide to another and to avoid normalization across slides, we made the simplification $A_p^k = 0$ and reconstructed relative mean values of expression *E* for each individual as

$$\left\{ \begin{array}{ll} E_p^i = \frac{1}{N_i}\sum_{k\in S_i}\left(\frac{M_p^k}{2}\right) & \text{for } i = 1\text{–}10 \text{ (DS)} \\ E_p^i = \frac{1}{N_i}\sum_{k\in S_i}\left(-\frac{M_p^k}{2}\right) & \text{for } i = 11\text{–}21 \text{ (controls)} \end{array} \right\} \ .$$

PCA of chromosome 21 genes and genes mapping to other chromosomes was performed separately using *E* values deduced from all expressed chromosome 21 probes and all expressed non–chromosome 21 probes, respectively.

PCR Experiments

To validate expression ratios between DS and control samples obtained from HSA21 oligoarray data, 100 ng of mRNA was reverse transcribed into cDNA by use of Reverse-iT RTase Blend kit (ABgene). Quantitative real-time PCR (QPCR) on diluted cDNA was conducted in the presence of 0.6 mM of each specific primer (designed by Primer3 software) and 1× Quantitect SYBR Green PCR master mix (Qiagen) containing 2.5 mM $MgCl_2$, Hotstart *Taq* polymerase, dNTP mix, and the fluorescent dye SYBR Green I. QPCR experiments were performed in a Lightcycler system (Roche Molecular Biochemicals) on 11 HSA21 genes: *CHAF1B, CSTB, DSCR1, DYRK1A, GART, H2BFS, MX1, SNF1LK, SOD1, STCH,* and *TMEM1*. The ubiquitin-activating enzyme E1 (NCBI Entrez accession number L13852) mRNA mapping to HSA3 and the zinc-finger protein (NCBI Entrez accession number AB000468) mRNA mapping to HSA4 were used as endogenous control genes, as described by Janel et al.[33] For each sample, the mean cycle threshold value, C_t, was corrected by subtracting the mean of the C_t obtained with the two reference genes. PCR primers are listed in table 2.

Results

Design of a Comprehensive HSA21 Oligoarray

The HSA21 oligoarray was designed for the exhaustive study of human chromosome 21 gene expression in DS. This microarray contained 664 sequences representing 145 genes, 58 ORFs, 118 predictions (plus the reverse sequences for 20 of them), and 18 antisense transcripts, allowing expression analysis of all putative genes mapping to chromosome 21 and related to DS. To increase the

Table 5. List of Genes Classified According to Their Expression in DS LCLs

Gene Symbol	GenBank Accession Number	Ratio	*A*	*M*	Var(*M*)	Class[a]
as-TTC3	BF979681.2	1.56	9.07	.64	.64	I
C21orf108 (exon 39)	AF231919.1	1.30	7.84	.38	.18	I
C21orf33	BI824121.1	1.52	10.54	.60	.20	I
C21orf59	AF282851.1	1.35	9.61	.44	.21	I
C21orf66	AY033903.1	1.51	9.50	.59	.12	I
C21orf7	AY171599.2	1.43	8.00	.52	.49	I
C21orf91	BC015468.2	1.59	9.50	.67	.26	I
CBS	AF042836.1	1.61	8.13	.69	.48	I
CCT8	BC095470.1	1.65	12.64	.72	.32	I
CRYZL1	BC033023.2	1.52	9.45	.60	.12	I
DONSON	AF232673.1	1.42	9.73	.50	.11	I
DYRK1A	D86550.1	1.41	10.20	.49	.17	I
HMGN1	M21339.1	1.38	11.82	.46	.10	I
IFNAR1	AY654286.1	1.47	8.10	.56	.18	I
IFNAR2	BC013156.1	1.67	8.38	.74	.14	I
IFNGR2	AY644470.1	1.45	10.35	.54	.24	I
IL10RB	BT009777.1	1.66	10.18	.73	.21	I
ITGB2	BC005861.2	1.61	11.34	.68	.40	I
MCM3AP	AY590469.1	1.45	11.69	.54	.19	I
MRPL39	AF109357.1	1.47	9.82	.55	.15	I
PFKL	X15573.1	1.50	12.44	.58	.23	I
PIGP	AF216305.1	1.60	8.14	.68	.25	I
PTTG1IP	NM_004339.2	1.52	9.89	.60	.20	I
SFRS15	AF023142.1	1.39	10.42	.47	.14	I
SLC5A3	L38500.2	1.57	8.92	.66	.18	I
SON	AY026895.1	1.51	12.64	.60	.14	I
SUMO3	BC008420.1	1.41	10.82	.50	.10	I
USP16	AY333928.1	1.46	9.94	.54	.10	I
USP25	AF170562.1	1.58	9.93	.66	.10	I
ZNF294	NM_015565.1	1.51	8.77	.60	.11	I
BTG3	D64110.1	1.82	10.63	.86	.21	II
C21orf57	AY040875.1	1.74	9.49	.80	.28	II
MRPS6	AB049942.1	1.64	10.26	.72	.13	II
PDXK	AY303972.1	1.71	10.06	.77	.23	II
SAMSN1	AF222927.1	2.27	10.47	1.18	.72	II
SLC37A1	AF311320.1	1.72	9.33	.78	.32	II
SNF1LK	AB047786.1	2.14	9.47	1.10	1.15	II
STCH	U04735.1	1.97	9.86	.98	.42	II
TTC3	D84296.1	1.79	10.59	.84	.20	II
aa071193	AA071193.1	.97	7.34	−.05	.48	III
AIRE	AB006682.1	.82	7.28	−.28	.29	III
AL041783	AL041783.1	1.06	7.10	.09	.36	III
as-C21orf56	BC084577.1	1.07	11.75	.10	.08	III
as-KIAA0179	AA425659.1	1.16	8.08	.22	.67	III
ATP5J	BC001178.1	1.35	7.99	.44	.11	III
B184	AL109967.2	1.06	9.10	.08	.62	III
B27 inverse	AP000034.1	.83	8.40	−.27	.31	III
C21orf108 (exon 26)	AF231919.1	1.09	8.86	.13	.09	III
C21orf12	AP001705.1	.97	7.49	−.05	.72	III
C21orf2	NM_004928.1	1.30	8.80	.38	.14	III
C21orf21	AA969880	1.16	7.33	.22	.28	III
C21orf25	AB047784.1	1.19	8.61	.25	.31	III
C21orf29	AJ487962.1	.98	7.27	−.03	.22	III
C21orf34	AF486622.1	.93	8.17	−.11	.28	III
C21orf42	AY035383.1	1.14	10.05	.19	.22	III
C21orf45	*AF387845.1*	1.23	9.39	.30	.12	III
C21orf49	BC117399.1	1.14	7.80	.19	.32	III
C21orf51	AY081144.1	1.26	9.47	.33	.09	III
C21orf54	AA934973.1	1.01	7.26	.01	.38	III
C21orf58	BC028934.1	1.11	7.73	.15	.21	III
C21orf6	BC017912.1	1.20	9.63	.26	.22	III
CHAF1B	U20980.1	1.29	8.35	.37	.13	III
CLIC6	AF448438.1	.74	9.26	−.43	.98	III

(*continued*)

Table 5. (continued)

Gene Symbol	GenBank Accession Number	Ratio	*A*	*M*	Var(*M*)	Class[a]
CXADR	AF200465.1	.90	8.16	−.15	.32	III
D21S2056E	U79775.1	1.33	10.55	.41	.18	III
DCR1-17.0	AJ001875.1	1.13	7.16	.17	.71	III
DCR1-19.0	AJ001906.1	.94	7.13	−.09	.19	III
DCR1-20.0-reverse	AJ001893.1	.95	7.14	−.07	.36	III
DCR1-25.0-reverse	AJ001905.1	.96	7.34	−.06	.31	III
DCR1-7.0	AJ001861.1	1.09	7.48	.12	.27	III
DCR1-7.0-reverse	AJ001861.1	1.21	7.13	.28	.28	III
DCR1-8.0	AJ001862.1	.98	7.58	−.03	.42	III
DCR1-8.0-reverse	AJ001862.1	1.00	8.61	.00	.20	III
DSCAM_Intronic_Model	BG221591.1	1.21	8.39	.28	1.39	III
DSCR1	AY325903.1	.93	7.13	−.10	.74	III
DSCR10	AB066291.1	.95	8.60	−.07	.25	III
DSCR2	AY463963.1	1.25	10.98	.32	.18	III
DSCR3	D87343.1	1.40	10.09	.48	.12	III
DSCR6	AB037158.1	1.03	7.37	.04	.41	III
DSCR9	BC066653.1	1.05	7.69	.07	.28	III
ETS2	J04102.1	1.40	7.18	.49	.43	III
GABPA	BC035031.2	1.40	8.88	.49	.08	III
GART	X54199.1	1.17	9.55	.23	.15	III
H2BFS	AB041017.1	1.13	13.19	.17	.57	III
HLCS	AB063285.1	1.32	8.48	.40	.18	III
ICOSLG	AF289028.1	1.23	9.99	.30	.28	III
JAM2	AY016009.1	.98	8.07	−.03	.82	III
KCNE1	BC046224.1	1.12	7.18	.16	.40	III
KIAA0179	D80001.1	1.21	10.15	.28	.24	III
MORC3	BC094779.1	1.34	8.51	.42	.15	III
n74695	N74695	.93	9.83	−.10	.18	III
PKNOX1	AY196965.1	1.04	7.81	.05	.12	III
POFUT2	NM_015227.3	1.35	7.46	.43	.49	III
PRED21	AP001693.1	.95	7.75	−.07	.66	III
PRED24	AP001695.1	1.00	7.41	.00	.65	III
PRED41	AP001726.1	.93	7.15	−.11	.64	III
PRED59	AL163301.2	.94	7.94	−.09	.17	III
PRED63	AP001759.1	.99	7.56	−.01	.18	III
PRED65	AL163202.2	1.06	7.10	.08	.25	III
PRMT2 (exon 5/6)	U80213.1	1.34	8.67	.42	.14	III
PWP2H	U56085.1	1.34	9.43	.42	.11	III
RUNX1	D43968.1	.85	7.91	−.23	.76	III
SETD4	AF391112.1	1.09	9.24	.13	.21	III
SH3BGR	X93498.1	1.07	7.19	.10	.83	III
SLC19A1	AF004354.1	1.20	7.80	.26	.14	III
SOD1[b]	AY835629.1	1.15	9.55	.21	.21	III
SYNJ1	AF009039.1	1.29	8.07	.37	.13	III
TFF3	BC017859.1	.97	8.00	−.04	.46	III
TMEM1	BC101728.1	1.27	10.52	.34	.14	III
TMEM50B	AF045606.2	1.38	10.07	.46	.20	III
U2AF1	M96982.1	1.27	12.15	.35	.06	III
UBASH3A	AJ277750.1	1.13	7.32	.17	.21	III
UBE2G2	AF032456.1	1.17	12.08	.22	.18	III
W90635	W90635.1	1.09	7.39	.12	.45	III
WRB	BC012415.1	1.21	8.21	.27	.29	III
ZNF295	BC063290.1	1.19	8.78	.25	.27	III
ABCG1	AY048757.1	1.25	8.11	.32	.79	IV
ADARB1	AY135659.1	1.26	7.28	.34	3.06	IV
as-MCM3AP-C21orf85	AW163084.1	1.41	7.71	.50	1.02	IV
BRWD1	AB080586.1	1.44	9.14	.53	.38	IV
C21orf22	AY040089.1	1.27	7.56	.34	.88	IV
C21orf8	AA843704.1	1.09	7.55	.12	1.32	IV
CBR1	AB124848.1	1.47	8.20	.55	.76	IV
COL6A1	NM_001848.2	1.60	7.34	.68	.62	IV
CSTB	AF208234.1	1.35	12.79	.43	.38	IV

(*continued*)

Table 5. (continued)

Gene Symbol	GenBank Accession Number	Ratio	*A*	*M*	Var(*M*)	Class[a]
DCR1-12.0	AJ001868.1	1.28	7.37	.36	.76	IV
DCR1-12.0-reverse	AJ001868.1	1.44	8.01	.53	1.06	IV
DCR1-13.0	AJ001869.1	1.25	9.20	.32	1.26	IV
DCR1-13.0-reverse	AJ001869.1	1.14	8.98	.19	.91	IV
DCR1-15.0	AJ001872.1	1.28	7.08	.36	1.00	IV
DSCR4	DQ179113.1	1.32	7.41	.40	.53	IV
MX1	AF135187.1	1.49	13.61	.58	.67	IV
MX2	M30818.1	1.33	11.94	.41	.48	IV
PRDM15	AF426259.1	1.38	8.88	.47	.51	IV
PRMT2 (exon 8/9)	U80213.1	1.46	8.73	.55	.58	IV
TRPM2	AY603182.1	1.08	7.50	.11	1.82	IV

NOTE.—The value *A* corresponds to $[\log_2(\text{DS}) + \log_2(\text{control})]/2$ for the corresponding gene across the 40 hybridizations, *M* corresponds to the mean of $\log_2(\text{DS}) - \log_2(\text{control})$ for the corresponding gene across the 40 hybridizations, Var (*M*) is the variance of *M,* and the DS/control ratio is equal to 2^M.

[a] Class I corresponds to genes expressed proportionally to the gene-dosage effect in DS cell lines, class II contains genes that are amplified, class III contains genes that are compensated, and class IV contains genes that are highly variable between individuals.

[b] Oligonucleotide probes mapped to the long isoform of *SOD1*. See details in the "Discussion" section.

strength of the results, where possible, at least two probes per gene were designed (~80% of the HSA21 oligoarray content). The description of the HSA21 oligoarray content according to BLAST results performed on the latest version of the human genome sequence (NCBI Gene Database build 36.2) is summarized in table 3. Oligonucleotide sequences spotted on the array have been designed on the basis of four main criteria: specificity for the represented sequence, GC content equal to 50%, melting temperature allowing an optimal match between probe and target at the hybridization temperature, and no stable predicted secondary structure.

By use of this new specific high-content HSA21 oligoarray, 40 differential hybridizations comparing DS and control LCLs were performed. The mean signal intensities (represented in $\log_2$ by the *A* value) of each array spot indicated the expression levels of chromosome 21 genes. A total of 134 genes gave signal intensities above the background cutoff (mean $A > 7$).

Biological Material from Patients with DS and Control Individuals

LCLs were obtained after immortalization by EBV of B lymphocytes collected from blood samples of individuals with DS and control individuals. To make sure that EBV transformation did not induce any chromosomal rearrangement, all cell lines were karyotyped after immortalization. Cell lines were always maintained in exponential growth phase. No significant difference in cell morphology or cell proliferation was observed between DS and control LCLs.

For three individuals with DS, transcriptome comparisons between fresh blood samples and LCLs obtained from the same individuals were conducted on pangenomic microarrays.[34] From these experiments, no major alteration of the transcriptome by the EBV transformation could be detected; only 0.5% of the genes exhibited significant differential expression ($P = .01$) (L.D., E.A.Y.-G., and M.-C.P., unpublished data).

Experimental Design and Statistical Analysis

The main objective was to detect differentially expressed genes between DS and control samples, taking into account the sex of and variability between individuals. The aim of the experimental design was to adapt to the experimental constraints of the study (see table 1 and the "Material and Methods" section). Forty experiments were thus programmed.

First, a mixed model was constructed to highlight the effects of DS, sex, and DS × sex and to take into account the gene-expression variability between individuals. We used the Benjamini-Hochberg procedure to adjust the *P* values obtained and to limit false-positive results due to multiple testing.[31] The FDR was set at 0.05. The list of significant genes was thus expected to contain 5% false-positive results.

Chromosome 21 genes did not have significant *P* values when sex or DS and sex combined (DS × sex) were tested. In other words, chromosome 21 gene expression was not significantly different between men and women. In addition, DS effects on gene expression were not dependent on sex. The effects of sex and DS × sex were thus dropped from the model, and a simplified mixed model testing the effects of DS on HSA21 gene expression was ultimately used. We first selected genes that had DS/control ratios significantly different from 1 (FDR 0.05), using the Benjamini-Hochberg procedure.[31] Among the 136 expressed transcripts (134 genes), about half (58) had DS/control

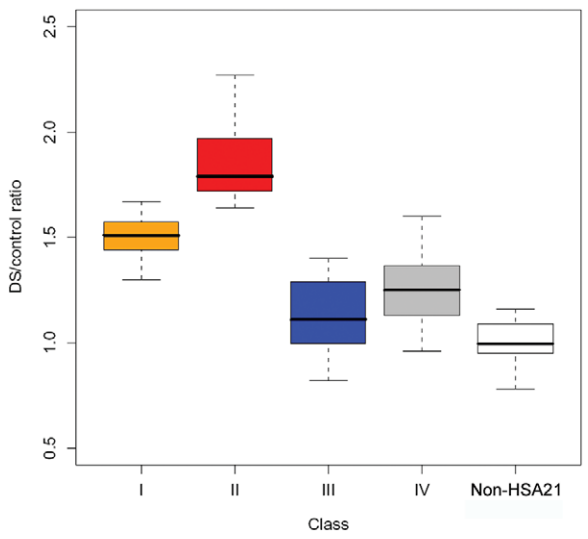

Figure 2. Distribution of DS/control ratios for class I, II, III, and IV genes and non-HSA21 reference genes. The plot represents the minimum and maximum values (*whiskers*), the first and third quartiles (*box*), and the median value (*midline*) of DS/control ratios for each class of genes.

ratios different from 1 and always >1, indicating that these genes were significantly overexpressed in DS LCLs. Expression ratios of these 58 genes ranged from 1.25 to 2.27, with a mean of 1.5, corresponding to the gene-dosage effect in DS. In parallel, among the 134 expressed genes, we selected those that deviated from this gene-dosage effect with a ratio significantly different from 1.5. Surprisingly, the majority of expressed transcripts (86) had DS/control ratios significantly different from 1.5 (FDR 0.05). Because of this high number, the method described by Storey et al.[32] for adjustment of the FDR had to be applied. Results showed that 86 genes had DS/control ratios significantly different from 1.5. Of these 86 genes, 9 had DS/control ratios >1.5, in the range 1.64–2.27, and 77 had DS/control ratios <1.5, in the range 0.74–1.4.

On the basis of this statistical analysis, we classified genes into four categories according to their variation of expression between DS and control LCLs, as described in table 4 and represented in figure 1. Class I contained 30 genes with DS/control ratios significantly different from 1 but not significantly different from 1.5, in the range 1.3–1.67. Class II contained nine genes that were significant in both statistical tests, with DS/control ratios significantly different from 1, significantly different from 1.5, and >1.5 (range 1.64–2.27). Class III comprised 77 genes that had DS/control ratios significantly different from 1.5 and <1.5 (range 0.74–1.4). The majority of gene predictions and antisense transcripts (77%) belonged to this class. In addition, gene expression levels of the transcripts belonging to class III were significantly lower than those belonging to class I ($P = 1.32 \times 10^{-5}$) and class II ($P = 5.7 \times 10^{-7}$). Class IV included the remaining 20 genes with DS/control ratios not significantly different from 1 or from 1.5, in the range 1.08–1.6. Table 5 gives the complete list of genes. Distributions (box plots) of DS/control ratios for each class and for non–chromosome 21 reference genes are shown in figure 2.

The goal of this study was also to demonstrate whether chromosome 21 gene-expression profiles could differentiate DS from control samples. We therefore performed two distinct PCAs on the 134 chromosome 21–expressed genes and the 39 non–chromosome 21 genes used as references (see the "Material and Methods" section). PCA could clearly distinguish individuals with DS from control individuals, suggesting that the effects of DS on chromosome 21 gene expression prevails over any other effect, including biological variability (fig. 3*A*). In addition, no distinction could be obtained between individuals with DS and control individuals when PCA was conducted on non–chromosome 21 genes (fig. 3*B*). Non–chromosome 21 genes had a mean DS/control expression ratio of 1 (fig. 2).

Most of the genes present on the HSA21 oligoarray were represented by two probes. When the two probes were found to be expressed, they belonged to the same class, except for *C21ORF108* and *PRMT2*. Concerning *C21ORF108*, one probe (*B1+KIAA0539.eri10102_a*) mapping to exon 39 of *C21ORF108* belonged to class I (DS/control ratio 1.3). The other probe (*KIAA0539.gff6561_b*), mapping to exon 26 of *C21ORF108*, was in class III (DS/control ratio 1.09). This difference could result from the existence of two alternative transcripts containing either exon 26 or exon 39. Similarly, the two probes representing *PRMT2* (*HRMT1L1.gff2216_a* and *HRMT1L1.gff2216_b*) belonged to classes IV and III, respectively. This difference could also be explained by the existence of two alternative transcripts containing either exons 5/6 or exons 8/9 described in the ENSEMBL database.

QPCR Validation Experiments

To confirm variations in gene expression and to validate the classification of chromosome 21 genes, we performed QPCR on 11 genes belonging to class I (1 gene), class II (2 genes), class III (6 genes), and class IV (2 genes). QPCR was conducted on all LCLs from individuals with DS and control individuals. Ratios obtained by QPCR confirmed the classification of chromosome 21 genes deduced from HSA21 oligoarrays, except for *SOD1*. *SOD1* belonged to class III and had a ratio of 1.57 by QPCR. However, this 1.57-fold difference between LCLs from individuals with DS and control individuals was not significant ($P = .15$). DS/control ratios from QPCR were in agreement with ratios obtained from HSA21 oligoarrays (table 6), with a correlation coefficient of 0.82. Our HSA21 oligoarray was thus a comprehensive, reproducible, and sensitive tool for studying gene expression in DS.

Discussion

The aim of the study was to analyze chromosome 21 gene expression in LCLs from individuals with DS and control individuals. Forty differential hybridizations comparing DS LCLs with control LCLs were performed on a dedicated HSA21 oligoarray designed from the complete human chromosome 21 gene catalogue (359 genes). Approximately one-third (134) of all chromosome 21 genes, ORFs, and predictions were expressed in LCLs.

On the basis of the expression levels of chromosome 21 genes, DS samples were clearly distinct from control samples, thus reflecting the prevalent effect of DS on chromosome 21 gene expression. On the contrary, reference genes mapping to chromosomes other than 21 could not distinguish DS LCLs from control LCLs. Using the mixed-model analysis, we were able to detect genes that are significantly overexpressed in DS cell lines (58) and also genes that deviate from the gene-dosage effect, with DS/control expression ratios significantly different from 1.5.

Classification of HSA21 Genes

By use of this new data analysis protocol, human chromosome 21 genes can now be ranked into four classes by their expression levels in DS cell lines relative to controls. This protocol could be applied to expression data obtained from other human tissues, to validate the classification.

Class I contains 30 genes with expression ratio of DS/control close to 1.5 (range 1.3–1.67), correlated to the presence of three genomic copies (table 4 and fig. 1). These class I genes could be responsible for the phenotype observed in DS, either directly or indirectly through a secondary effect of *cis*- or *trans*-acting genes.

Class II contains nine genes with expression ratio of DS/control >1.64, corresponding to an amplification of the initial gene dosage (table 4 and fig. 1). Among these genes, *SAMSN1, SNF1LK, STCH,* and *BTG3* show the highest expression ratio, in the range 1.67–2.27.

Gene-dosage amplification could result from a cascading effect through regulation networks involving *trans*- or *cis*-acting genes.[35] Pellegrini et al.[36] identified *in silico* a putative mitogen-activated kinase cascade with chromosome 21 kinases involved in various signaling pathways: *DYRK1A, SNF1LK, RIPK4,* and *DSCR3*. In our study, *DYRK1A, SNF1LK,* and *DSCR3* were expressed in LCLs, whereas *RIPK4* was not. Thus, four replicates were chosen for each patient.

DYRK1A is under the gene-dosage effect and *DSCR3* is compensated, whereas *SNF1LK* is amplified from the initial gene dosage. On the basis of this putative mitogen-

Table 6. Comparison between QPCR Results and Microarray Data

HSA21 Gene	GenBank Accession Number	Data from HSA21 Oligoarray		DS/Control Ratio by QPCR[a]
		DS/Control Ratio[b]	Class	
SNF1LK	AB047786.1	2.14	II	3.36
STCH	U04735.1	1.97	II	2.06
MX1	AF135187.1	1.49	IV	1.78
DYRK1A	D86550.1	1.41	I	1.77
CSTB	AF208234.1	1.35	IV	1.49
CHAF1B	U20980.1	1.29	III	1.38
TMEM1	BC101728.1	1.27	III	1.27
GART	X54199.1	1.17	III	1.38
SOD1	AY835629.1	1.15	III	1.57
H2BFS	AB041017.1	1.13	III	1.05
DSCR1	AY325903.1	.93	III	1.11

[a] The mean expression ratio for the corresponding gene between DS and control cell lines.

[b] DS/control ratio by QPCR was calculated from normalized C_t obtained for DS cell lines relative to control cell lines.

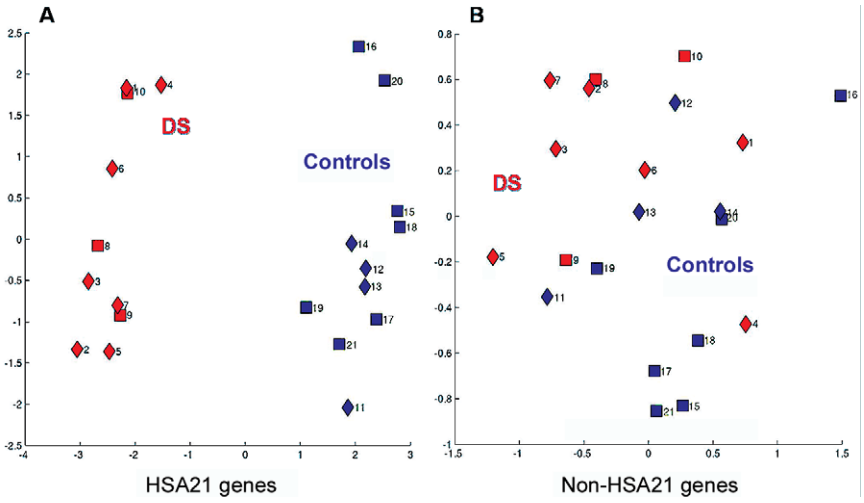

Figure 3. PCA of HSA21 genes (*A*) and non-HSA21 genes (*B*). Red and blue symbols represent DS and control samples, respectively. Squares represent samples extracted from females, and diamonds represent samples extracted from males.

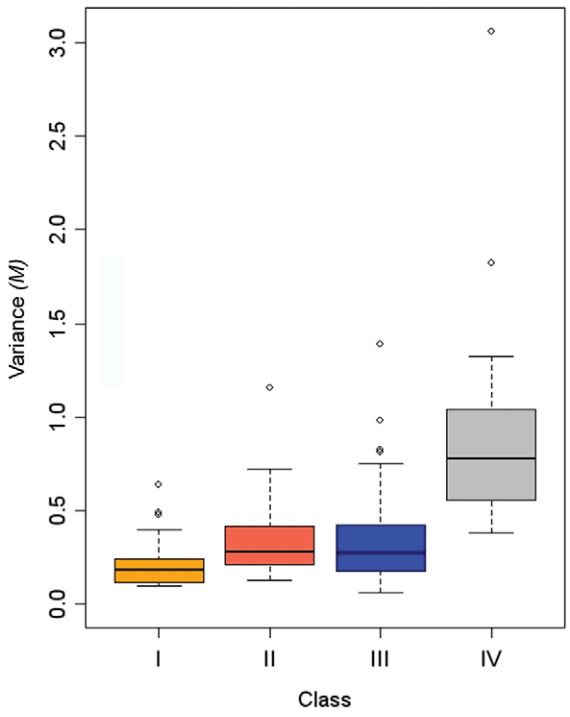

Figure 4. Distribution of the variance of *M* for class I, II, III, and IV genes. The plot represents the minimum and maximum values (*whiskers*), the first and third quartiles (*box*), and the median value (*midline*) of the variances of *M* for each class of genes, where *M* is the mean of $\log_2(\text{DS}) - \log_2(\text{control})$.

activated kinase cascade, amplification of *SNF1LK* gene expression could thus result from the overexpression of DYRK1A acting as a regulatory factor on *SNF1LK* in the cascade, and DSCR3 could act as a scaffolding protein.

Class III is the most abundant and contains 77 genes, with a large proportion of gene predictions and antisense transcripts with DS/control expression ratio <1.4 (table 4 and fig. 1). These class III genes are likely to be compensated in DS. Compensation mechanisms in trisomic conditions have been described in maize and *Drosophila*[37–39] and have been suggested in previous transcriptome studies, both in patients with DS[10,12,16] and in mouse models.[4,19–21] For example, Lyle et al.[20] found that 45% of the triplicated genes analyzed in their study were compensated. Compensation is most likely due to negative feedbacks that would modulate transcriptional activity or mRNA stability of class III genes. Thus, expression of compensated genes could be regulated by mechanisms that are not impaired in DS. For example, *trans*-inhibitors could act directly on the level of expression of these genes. Alternatively, *trans*-activators would activate inhibitors present in three copies on chromosome 21 and would reduce the expression level of target genes that could be also be present on chromosome 21.[40] However, the existence of polymorph alleles correlated to different levels of expression should not account for either gene compensation or amplification in a representative population of patients with DS. A recent study has demonstrated that two CpG islands from human chromosome 21 can be methylated monoallelically.[41] One of those maps to *DSCR3,* the other to *C21orf29.* Both are class III genes in LCLs.

Six class III genes were tested by QPCR, and all were validated, except *SOD1. SOD1* is a well-characterized gene that has been shown elsewhere to be overexpressed in DS tissues and cells at the RNA and protein levels.[13,16,42] In LCLs, the *SOD1* gene is transcribed into two variants, a long and a short isoform.[43] Since *SOD1* probes from the HSA21 oligoarray mapped to the long isoform only, the classification (class III compensated) (table 5) corresponded to this long isoform. The ratio deduced from the HSA21 oligoarray (1.15) was found to be slightly lower than the one obtained by QPCR for the long isoform (1.57). However, by use of QPCR primers amplifying both isoforms of *SOD1,* with the short isoform the most abundant in LCLs, we found that the ratio between DS and control LCLs was 1.96 (data not shown). These results suggest that, in DS LCLs, *SOD1* is overexpressed.

Class IV contains 15 genes and 5 gene predictions that have DS/control expression ratio not different from either 1 or 1.5. These class IV genes are thus highly variable between individuals with DS and control individuals. Indeed, figure 4 shows that the variance distribution of expression ratios is the highest for class IV genes. Three class IV genes (*CBR1, PRDM15,* and *ADARB1*) were shown elsewhere to be highly variable among unaffected individuals.[44]

Using the mixed-model analysis, we have been able to distinguish between gene expression differences resulting from DS and those from interindividual variations. Interindividual variations have been assessed in normal LCLs.[45–47] In the present study, we used lymphoblastoid cells established from individuals with DS and control individuals all belonging to Indo-European populations, thus limiting the variations due to ethnic groups.

Copy-number variations have also been described in LCLs.[48] They should not have an impact on the results unless their frequencies are different in individuals with DS and control individuals, which is unlikely. Moreover, we could not find any correlation between the type of copy-number variation (gain or loss) described for particular genes and their gene class. For example, two genes mapping to the same copy-number variant (variation 5162[48]) belonged to class II (*PDXK*) and class IV (*CSTB*).

Comparison with Expression Data Obtained from DS Tissues

Mao et al.[16] studied transcriptome modifications in DS fetal heart, cerebellum, and astrocyte cells, using a pangenomic Affymetrix U133A chip. Of the 200 genes assigned to HSA21, 23 were significantly changed in DS tissues and 17 were in common with the 58 HSA21 genes that were significantly changed in our study. Our results

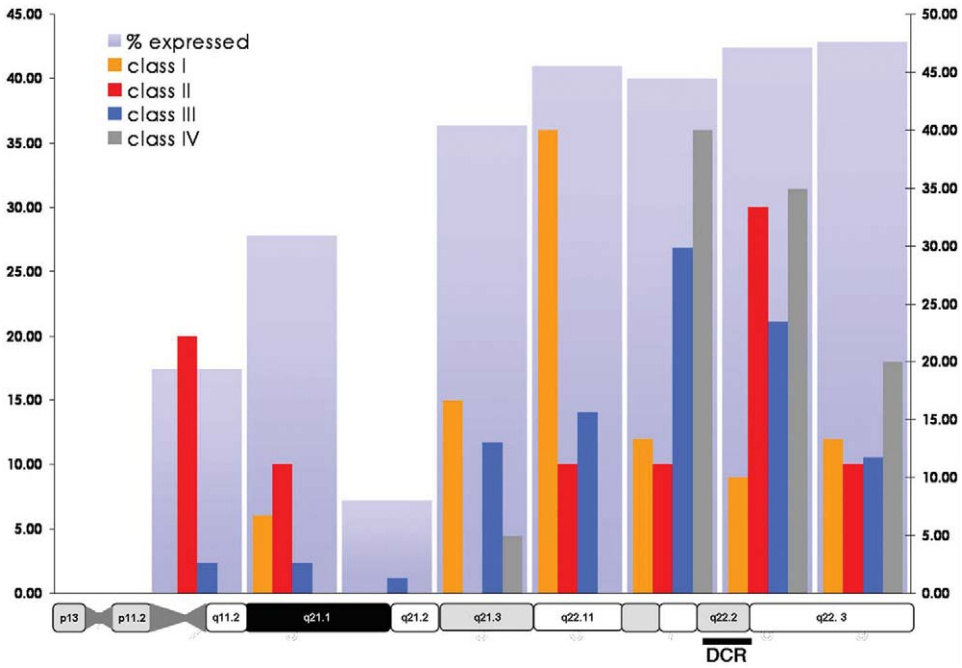

Figure 5. Distribution of expressed, class I, II, III, and IV genes along HSA21. The left *Y*-axis indicates the proportion of expressed genes in each 5-Mb interval, and the right *Y*-axis indicates the proportion of class I, II, III, and IV genes in each 5-Mb interval.

are also in agreement with another gene-expression study performed on DS fetal heart cells[17] that showed that 16 HSA21 genes are significantly overexpressed in fetal hearts. Among these 16 genes, 8 were significantly changed in our study. Class I, II, and III genes were present in all tissues, suggesting that gene-dosage effect, amplification, and compensation are general phenomena and that LCLs are a good model for studying gene-dosage effects. The differences observed between our study and the others suggest tissue-specific regulations that have been described elsewhere for the control of *GABPα* expression.[49]

Distribution of Gene-Expression Modifications along HSA21

We have analyzed the distribution of expressed genes, as well as individual gene classes along HSA21. Figure 5 shows that expressed genes map preferentially to the distal part of HSA21, reflecting the nonuniform gene density along HSA21q.[27] The most telomeric region of HSA21 has a high proportion of class III genes, perhaps because of the presence of a higher proportion of gene predictions that are localized inside gene introns.

DS Effects on Alternative Transcripts

To search for differential effects of DS on alternatively spliced transcripts, we analyzed genes for which oligonucleotide probes present on the HSA21 oligoarray could differentiate between alternative transcripts. Seventeen genes had probes specific to alternative transcripts (table 7). For seven of those genes, all the probes gave a very low signal, indicating that these transcripts are not expressed in LCLs. Three genes (*C21orf33, C21orf34,* and *MRPL39*) were expressed in LCLs as a unique transcript and belonged to class I genes, which are overexpressed with a ratio close to 1.5. For the last seven genes (*ADARB1, C21orf66, DYRK1A, GART, PKNOX1, RUNX1,* and *TMEM1*), oligonucleotide probes could distinguish between splicing variants that had very similar DS/control ratios. Only two of these genes (*C21orf66* and *DYRK1A*) were significantly overexpressed in DS LCLs (i.e., were class I genes), whereas the others were compensated. These results suggest that most of the transcripts belonging to the same gene and expressed in LCLs are similarly regulated in DS.

DS Effects on Antisense Transcripts

The HSA21 oligoarray was also designed to analyze the effects of DS on the expression of antisense transcripts. Fourteen antisense transcripts are present with their nesting genes on the HSA21 oligoarray (table 8). Among them, 10 have been extracted from the HSA21 database established by Kathleen Gardiner at the Eleanor Roosevelt Institute.[50] The four remaining antisense transcripts corresponded to transcribed sequences in the DCR that have been generated from various cDNA mapping and exon-trapping experiments.[8,51,52] Seven genes (*C21orf25, CHAF1B, DYRK1A, HLCS, KIAA0179, MCM3AP,* and *TTC3*) are expressed in LCLs. Four of the corresponding antisense transcripts (*as-DYRK1A, as-HLCS, as-KIAA0179,* and *as-TTC3*) were also found to be expressed in LCLs, thus confirming

Table 7. Alternative Transcripts of HSA21 Genes

Gene Symbol and Probe	GenBank Accession Number	*A*	DS/Control Ratio	*M*	Var(*M*)	Recognized Variants[a]	Class
ADARB1:							
ADARB1.alt23565_a	AY135659.1	5.23	NE	NE	NE	1, 2, 3, 4	NE
ADARB1.alt23565_b	AY135659.1	4.89	NE	NE	NE	1, 2, 3, 4	NE
ADARB1.alt3788_a	AY135659.1	7.36	1.22	.29	2.69	1, 2, 4	IV
ADARB1.alt3788_b	AY135659.1	7.21	1.3	.38	3.46	1, 2, 4	IV
C21orf33:							
HES1.gff1583_b	BC003587.1	5.46	NE	NE	NE	1	NE
bi824121	BI824121.1	10.2	1.5	.58	.16	1, 2	I
HES1.gff1583_a	BC003587.1	10.86	1.53	.62	.25	1, 2	I
C21orf34:							
orf34+35.eri594_a	AF486622.1	5.5	NE	NE	NE	1	NE
C21orf34.gff397_a	AF486622.1	5.87	NE	NE	NE	1, 2	NE
C21orf34.gff397_b	AF486622.1	5.83	NE	NE	NE	1, 2	NE
orf34+35.alt629_b	AF486622.1	6.15	NE	NE	NE	1, 2	NE
orf34+35.eri594_b	AF486622.1	5.69	NE	NE	NE	1, 2	NE
C21orf35.gff251_a	AF486622.1	5.95	NE	NE	NE	1, 2, 3	NE
orf34+35.alt2559_a	AF486622.1	8.15	.95	−.07	.30	1, 2, 3	III
orf34+35.alt629_a	AF486622.1	8.2	.9	−.15	.27	1, 2, 3	III
C21orf66:							
B3+GCFC.eri2361_a	AY033903.1	5	NE	NE	NE	1, 2, 3	NE
B3+GCFC.eri2361_b	AY033903.1	5.48	NE	NE	NE	1, 2, 3, 4	NE
GCFC.eri2361_a	AY033903.1	10.6	1.56	.64	.09	1, 2, 3	I
GCFC.eri2361_b	AY033903.1	8.93	1.51	.6	.04	1, 2, 3, 4	I
GCFC.gff1083_a	AY033903.1	10.19	1.48	.57	.22	1, 2, 3	I
GCFC.gff1083_b	AY033903.1	8.29	1.47	.56	.12	1, 2, 3	I
DYRK1A:							
DYRK1.alt2571_a	D86550.1	10.33	1.4	.48	.13	1, 2, 3, 4, 5	I
DYRK1.alt2571_b	D86550.1	10.71	1.4	.49	.22	1, 2, 3, 4, 5	I
DYRK1.gff5318_a	D86550.1	9.47	1.41	.5	.19	1, 2, 3, 4, 5	I
DYRK1.gff5318_b	D86550.1	11.02	1.41	.5	.22	1, 2, 3, 4, 5	I
GART:							
GART.gff3271_a	X54199.1	8.79	1.18	.24	.10	1	III
GART.gff3271_b	X54199.1	10.35	1.17	.22	.20	1, 2	III
MRPL39:							
PRED22.eri707_a	AF109357.1	9.46	1.41	.5	.14	1, 2	I
PRED22.eri707_b	AF109357.1	10.56	1.47	.55	.17	1, 2	I
PRED22.gff1072_a	AF109357.1	10.33	1.47	.56	.13	1, 2	I
PRED22.gff1072_b	AF109357.1	10.33	1.46	.55	.18	1, 2	I
PRED66.eri187_a	AF109357.1	10.28	1.51	.6	.12	1, 2	I
PRED66.eri187_b	AF109357.1	9.36	1.51	.59	.16	1, 2	I
PRED66.gff641_b	AF109357.1	8.31	1.42	.5	.17	1, 2	I
PRED66.gff641_a	AF270511.1	6.58	NE	NE	NE	2	NE
PKNOX1:							
PKNOX1.gff3279_a	AY196965.1	7.64	1.05	.07	.11	1	III
PKNOX1.gff3279_b	AY196965.1	7.98	1.03	.04	.14	1, 2	III
RUNX1:							
RUNX1.alt25714_a	D43968.1	7.23	.86	−.21	.84	1, 2	III
RUNX1.alt7267_a	D43968.1	7.37	.8	−.32	.74	1, 2	III
RUNX1.alt7267_b	D43968.1	5.56	NE	NE	NE	2	NE
RUNX1.gff2722_a	D43968.1	7.92	.85	−.23	.74	2	III
RUNX1.gff2722_b	D43968.1	9.03	.89	−.17	.76	2	III
TMEM1:							
TMEM1.gff5126_a	BC101728.1	10.53	1.25	.32	.14	1	III
TMEM1.gff5126_b	BC101728.1	10.5	1.28	.36	.14	1, 2	III

NOTE.—The value *A* corresponds to $[\log_2(\text{DS}) + \log_2(\text{control})]/2$ for the corresponding gene across the 40 hybridizations, *M* corresponds to the mean of $\log_2(\text{DS}) - \log_2(\text{control})$ for the corresponding gene across the 40 hybridizations, Var(*M*) is the variance of *M,* and the DS/control ratio is equal to 2^M. NE = not expressed.

[a] The number of transcript variants hybridizing to the oligonucleotide probe.

Table 8. Sense and Antisense Transcripts on Chromosome 21

Gene Symbol and Probe	GenBank Accession Number	Intragenic Location[a]	*A*	*M*	Var(*M*)	DS/Control Ratio	Class
C21orf56:							
C21orf56.gff691_a	BC084577.1	Exon 2	5.51	NE	NE	NE	NE
as-C21orf56:							
C21orf56.gff284_a	BC084577.1	Exon 4	10.68	.12	.08	1.08	III
C21orf56.gff284_b	BC084577.1	Exon 4	12.79	.08	.08	1.05	III
CHAF1B:							
CHAF1B.gff2194_a	U20980.1	Exon 14	8.66	.37	.12	1.29	III
CHAF1B.gff2194_b	U20980.1	Exon 14	8.03	.38	.14	1.30	III
as-CHAF1B:							
BF740066	BF740066.1	3′	5.08	NE	NE	NE	NE
HLCS:							
HLCS.gff6722_a	AB063285.1	Exon 12	6.37	NE	NE	NE	NE
HLCS.gff6722_b	AB063285.1	Exon 11	8.48	.40	.18	1.32	III
as-HLCS:							
DCR1-8.0_a	AJ001862.1	Intron 7	7.58	−.03	.42	.98	III
DCR1-8.0_b	AJ001862.1	Intron 7	5.83	NE	NE	NE	NE
TTC3:							
TTC3.gff9074_a	D84296.1	Exon 47	10.02	.84	.23	1.79	II
TTC3.gff9074_b	D84296.1	Exon 34	11.17	.83	.17	1.78	II
as-TTC3:							
bf979681	BF979681.2	Exon 41	9.07	.64	.64	1.56	I
DYRK1A:							
DYRK1.alt2571_a	D86550.1	Exon 13	10.33	.48	.13	1.40	I
DYRK1.alt2571_b	D86550.1	Exons 7/8	10.71	.49	.22	1.40	I
DYRK1.gff5318_a	D86550.1	Exon 13	9.47	.50	.11	1.41	I
DYRK1.gff5318_b	D86550.1	Exon 11	11.02	.50	.22	1.41	I
as-DYRK1A:							
DCR1-12.0_a	AJ001868.1	Intron 1	7.37	.36	.76	1.29	IV
DCR1-13.0-RC_a	AJ001869.1	Intron 1	8.33	.15	.95	1.11	IV
DCR1-13.0-RC_b	AJ001869.1	Intron 1	9.55	.22	.91	1.17	IV
KCNJ6:							
GIRK2(U52153)_a	U52153.1	Exon 3	5.60	NE	NE	NE	NE
GIRK2(U52153)_b	U52153.1	Exon 1	5.28	NE	NE	NE	NE
as-KCNJ6:							
DCR1-17DCR1-17_a	AJ001875.1	Intron 3	7.16	.17	.70	1.12	IV
ADAMTS5:							
ADAMTS5.gff5523_a	AF142099.1	Exon 8	5.39	NE	NE	NE	NE
ADAMTS5.gff5523_b	AF142099.1	Exon 8	5.39	NE	NE	NE	NE
as-ADAMTS5:							
r18879	R18879.1	Intron 3	5.69	NE	NE	NE	NE
as-C21orf25:							
aa575913	AA575913.1	3′	5.85	NE	NE	NE	NE
C21orf25:							
C21orf25.gff6217_a	AB047784.1	Exon 14	8.96	.17	.45	1.13	III
C21orf25.gff6217_b	AB047784.1	Exon 14	8.24	.32	.17	1.25	III
as-CBR3:							
bi836686	BI836686.1	Exon 3	5.90	NE	NE	NE	NE
CBR3:							
CBR3.gff878_a	AB124847.1	Exon 3	6.18	NE	NE	NE	NE
CBR3.gff878_b	AB124847.1	Exons 1/2	5.31	NE	NE	NE	NE
CLDN14:							
CLDN14.gff1693_a	AF314090.1	Exon 3	6.78	NE	NE	NE	NE
CLDN14.gff1693_b	AP001726.1	3′	6.05	NE	NE	NE	NE
as-CLDN14:							
w90592	W90592.1	3′	5.71	NE	NE	NE	NE
KIAA0179:							
KIAA0179.gff4984_a	D80001.1	Exon 16	5.26	NE	NE	NE	NE
KIAA0179.gff4984_b	D80001.1	Exon 16	10.15	.28	.24	1.22	III
as-KIAA0179:							
aa425659	AA425659.1	3′	8.08	.22	.67	1.17	III
MCM3AP:							
MCM3.gff6113_a	AY590469.1	Exon 27	11.69	.54	.19	1.45	I
MCM3APAS:							
af426262	AF426262.1	Introns 25–26	5.54	NE	NE	NE	NE
af426263	AF426263.1	Introns 20–21	5.73	NE	NE	NE	NE

NOTE—The value *A* corresponds to $[\log_2(\text{DS}) + \log_2(\text{control})]/2$ for the corresponding gene across the 40 hybridizations, *M* corresponds to the mean of $\log_2(\text{DS}) - \log_2(\text{control})$ for the corresponding gene across the 40 hybridizations, Var(*M*) is the variance of *M*, and the DS/control ratio is equal to 2^M. NE = not expressed.

[a] The exon or intron to which the oligonucleotide probe maps.

their existence. *TTC3* (class II) and its antisense transcript (class I) were overexpressed, whereas the other antisense sequences did not belong to the same class as their corresponding genes. In addition, two antisense sequences (*as-C21orf56* and *as-KCNJ6*) were expressed in LCLs, whereas their corresponding genes were not. The probe referred to as antisense transcript *as-KCNJ6* mapping in intron 3 of *KCNJ6,* on the opposite strand, corresponds to one of the transcribed sequences isolated in the DCR by exon-trapping experiments.[8] Since there is no evidence that this sequence is an antisense transcript of *KCNJ6,* it could thus belong to a gene locus that has not yet been identified and might map to the opposite orientation of *KCNJ6*.

According to the NCBI Gene Database, *C21orf56* (accession number 84221) currently maps on the negative strand of HSA21 but was previously annotated on the positive strand when the HSA21 oligoarray was designed. Thus, probe sequence representing the antisense transcript *as-C21orf56* could correspond to the actual sense transcript of *C21orf56*. Therefore, the expression of antisense transcripts is confirmed by our HSA21 oligoarray experiments. Sense and antisense transcripts are not always similarly changed in DS.

In conclusion, using our new high-content HSA21 oligoarrays combined with a new powerful statistical analysis protocol, we were able to classify HSA21 genes according to their level of expression in DS LCLs. We show that, among the expressed transcripts, 29% are sensitive to the gene-dosage effect or are amplified, 56% are compensated, and 15% are highly variable among individuals. Thus, most of the chromosome 21 genes are compensated for the gene-dosage effect. Gene-expression variations in DS are controlled by mechanisms involving *trans* and *cis* regulators acting either directly or through gene-regulation networks. Overexpressed genes are likely to be involved in the DS phenotype, in contrast to the compensated genes. Highly variable genes could account for phenotypic variations observed in patients. Finally, we show that alternative transcripts belonging to the same gene are similarly regulated in DS, whereas sense and antisense transcripts are not always similarly regulated. Studies of human tissues by use of the same analysis protocol will validate genes that are involved in the DS phenotype.

Acknowledgments

We thank the Banque de Cellules from the Cochin Hospital; Drs. N. Faucon and R. Meloni (CNRS UMR 9923, Hôpital Pitié Salpetrière, Paris), for advice; D. Leclerre, C. Mikonio, and F. Richard, for their help with spotting the HSA21 oligoarrays; I. Haddad, for the design of the first version of the HSA21 oligoarray; C. J. Epstein (University of California, San Francisco), for very important comments on the manuscript; and R. Veitia (Hôpital Cochin, Paris), J.-J. Daudin (Institut National d'Agronomie Paris-Grignon, Paris), Dr. Y. Hérault (CNRS, Orléans), and Dr. P. M. Sinet (Institut National de la Santé et de la Recherche Médicale, Paris), for helpful discussions about the results. E.A.G. had a fellowship from the Ministère de la Recherche, and G.G had a fellowship from the Fondation Jérôme Lejeune. This work was supported by the Fondation Jérôme Lejeune, European Economic Community grant T21 Targets, and AnEUpolidy.

Web Resources

Accession numbers and URLs for data presented herein are as follows:

Eleanor Roosevelt Institute: Chromosome 21 Gene Function and Pathway Database, http://chr21db.cudenver.edu/

GenBank, http://www.ncbi.nlm.nih.gov/Genbank/ (for accession numbers in tables 5–8)

Gene Expression Omnibus (GEO), http://www.ncbi.nlm.nih.gov/geo/ (for accession number GSE6408)

Max Planck Institute: Chromosome 21 Gene Catalog Based on the New AGP File July 2002, http://chr21.molgen.mpg.de/chr21_catalogs/chr21_mar_2002.html

NCBI Entrez, http://www.ncbi.nlm.nih.gov/gquery/gquery.fcgi (for accession numbers L13852 and AB000468)

NCBI Gene Database, http://www.ncbi.nlm.nih.gov/sites/entrez (for accession number 84221)

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for DS)

The R Project for Statistical Computing, http://www.r-project.org/

References

1. Epstein CJ, Korenberg JR, Anneren G, Antonarakis SE, Ayme S, Courchesne E, Epstein LB, Fowler A, Groner Y, Huret JL, et al (1991) Protocols to establish genotype-phenotype correlations in Down syndrome. Am J Hum Genet 49:207–235
2. Antonarakis SE, Lyle R, Dermitzakis ET, Reymond A, Deutsch S (2004) Chromosome 21 and Down syndrome: from genomics to pathophysiology. Nat Rev Genet 5:725–738
3. Lejeune J, Gautier M, Turpin R (1959) Etudes des chromosomes somatiques de neuf enfants mongoliens. C R Hebd Seances Acad Sci 248:1721–1722
4. Saran NG, Pletcher MT, Natale JE, Cheng Y, Reeves RH (2003) Global disruption of the cerebellar transcriptome in a Down syndrome mouse model. Hum Mol Genet 12:2013–2019
5. Shapiro BL (2001) Developmental instability of the cerebellum and its relevance to Down syndrome. J Neural Transm Suppl 61:11–34
6. Schinzel A (2001) Catalogue of unbalanced chromosome aberrations in man. Walter de Gruyter, Berlin
7. Olson LE, Roper RJ, Sengstaken CL, Peterson EA, Aquino V, Galdzicki Z, Siarey R, Pletnikov M, Moran TH, Reeves RH (2007) Trisomy for the Down syndrome "critical region" is necessary but not sufficient for brain phenotypes of trisomic mice. Hum Mol Genet 16:774–782
8. Dahmane N, Ghezala GA, Gosset P, Chamoun Z, Dufresne-Zacharia MC, Lopes C, Rabatel N, Gassanova-Maugenre S, Chettouh Z, Abramowski V, et al (1998) Transcriptional map of the 2.5-Mb CBR-ERG region of chromosome 21 involved in Down syndrome. Genomics 48:12–23
9. Delabar J, Theophile D, Rahmani Z, Chettouh Z, Blouin J, Prieur M, Noel B, Sinet P (1993) Molecular mapping of twenty-four features of Down syndrome on chromosome 21. Eur J Hum Genet 1:114–124
10. FitzPatrick DR, Ramsay J, McGill NI, Shade M, Carothers AD,

Hastie ND (2002) Transcriptome analysis of human autosomal trisomy. Hum Mol Genet 11:3249–3256
11. Gross SJ, Ferreira JC, Morrow B, Dar P, Funke B, Khabele D, Merkatz I (2002) Gene expression profile of trisomy 21 placentas: a potential approach for designing noninvasive techniques of prenatal diagnosis. Am J Obstet Gynecol 187:457–462
12. Mao R, Zielke CL, Zielke HR, Pevsner J (2003) Global up-regulation of chromosome 21 gene expression in the developing Down syndrome brain. Genomics 81:457–467
13. Giannone S, Strippoli P, Vitale L, Casadei R, Canaider S, Lenzi L, D'Addabbo P, Frabetti F, Facchin F, Farina A, et al (2004) Gene expression profile analysis in human T lymphocytes from patients with Down syndrome. Ann Hum Genet 68: 546–554
14. Tang Y, Schapiro MB, Franz DN, Patterson BJ, Hickey FJ, Schorry EK, Hopkin RJ, Wylie M, Narayan T, Glauser TA, et al (2004) Blood expression profiles for tuberous sclerosis complex 2, neurofibromatosis type 1, and Down's syndrome. Ann Neurol 56:808–814
15. Chung IH, Lee SH, Lee KW, Park SH, Cha KY, Kim YS, Lee S (2005) Gene expression analysis of cultured amniotic fluid cell with Down syndrome by DNA microarray. J Korean Med Sci 20:82–87
16. Mao R, Wang X, Spitznagel E, Frelin L, Ting J, Ding H, Kim JW, Ruczinski I, Downey T, Pevsner J (2005) Primary and secondary transcriptional effects in the developing human Down syndrome brain and heart. Genome Biol 6:R107
17. Li CM, Guo M, Salas M, Schupf N, Silverman W, Zigman W, Husain S, Warburton D, Thaker H, Tycko B (2006) Cell type-specific over-expression of chromosome 21 genes in fibroblasts and fetal hearts with trisomy 21. BMC Med Genet 7: 24
18. Amano K, Sago H, Uchikawa C, Suzuki T, Kotliarova SE, Nukina N, Epstein CJ, Yamakawa K (2004) Dosage-dependent over-expression of genes in the trisomic region of Ts1Cje mouse model for Down syndrome. Hum Mol Genet 13:1333–1340
19. Kahlem P, Sultan M, Herwig R, Steinfath M, Balzereit D, Eppens B, Saran NG, Pletcher MT, South ST, Stetten G, et al (2004) Transcript level alterations reflect gene dosage effects across multiple tissues in a mouse model of Down syndrome. Genome Res 14:1258–1267
20. Lyle R, Gehrig C, Neergaard-Henrichsen C, Deutsch S, Antonarakis SE (2004) Gene expression from the aneuploid chromosome in a trisomy mouse model of Down syndrome. Genome Res 14:1268–1274
21. Dauphinot L, Lyle R, Rivals I, Dang MT, Moldrich RX, Golfier G, Ettwiller L, Toyama K, Rossier J, Personnaz L, et al (2005) The cerebellar transcriptome during postnatal development of the Ts1Cje mouse, a segmental trisomy model for Down syndrome. Hum Mol Genet 14:373–384
22. Gardiner K (2006) 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 36:439–453
23. Wester U, Bondeson ML, Edeby C, Anneren G (2006) Clinical and molecular characterization of individuals with 18p deletion: a genotype-phenotype correlation. Am J Med Genet A 140:1164–1171
24. Hattori M, Fujiyama A, Taylor TD, Watanabe H, Yada T, Park HS, Toyoda A, Ishii K, Totoki Y, Choi DK, et al (2000) The DNA sequence of human chromosome 21. Nature 405:311–319
25. Kapranov P, Cawley SE, Drenkow J, Bekiranov S, Strausberg RL, Fodor SPA, Gingeras TR (2002) Large-scale transcriptional activity in chromosomes 21 and 22. Science 296:916–919
26. Reymond A, Camargo AA, Deutsch S, Stevenson BJ, Parmigiani RB, Ucla C, Bettoni F, Rossier C, Lyle R, Guipponi M (2002) Nineteen additional unpredicted transcripts from human chromosome 21. Genomics 79:824–832
27. Gardiner K, Fortna A, Bechtel L, Davisson MT (2003) Mouse models of Down syndrome: how useful can they be? Comparison of the gene content of human chromosome 21 with orthologous mouse genomic regions. Gene 318:137–147
28. Quackenbush J (2002) Microarray data normalization and transformation. Nat Genet Suppl 32:496–501
29. Wolfinger RD, Gibson G, Wolfinger ED, Bennett L, Hamadeh H, Bushel P, Afshari C, Paules RS (2001) Assessing gene significance from cDNA microarray expression data via mixed models. J Comput Biol 8:625–637
30. SAS Institute (2000) SAS/STAT software release 8. Cary, NC
31. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc B 57:289–300
32. Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. Proc Natl Acad Sci USA 100:9440–9445
33. Janel N, Christophe O, Ait Yahya-Graison E, Hamelet J, Paly E, Prieur M, Delezoide AL, Delabar JM (2006) Paraoxonase-1 expression is up-regulated in Down syndrome fetal liver. Biochem Biophys Res Commun 346:1303–1306
34. Brigand KL, Russell R, Moreilhon C, Rouillard JM, Jost B, Amiot F, Magnone V, Bole-Feysot C, Rostagno P, Virolle V, et al (2006) An open-access long oligonucleotide microarray resource for analysis of the human and mouse transcriptomes. Nucleic Acids Res 34:e87
35. Potier MC, Rivals I, Mercier G, Ettwiller L, Moldrich RX, Laffaire J, Personnaz L, Rossier J, Dauphinot L (2006) Transcriptional disruptions in Down syndrome: a case study in the Ts1Cje mouse cerebellum during post-natal development. J Neurochem Suppl 1 97:104–109
36. Pellegrini-Calace M, Tramontano A (2006) Identification of a novel putative mitogen-activated kinase cascade on human chromosome 21 by computational approaches. Bioinformatics 22:775–778
37. Birchler JA (1979) A study of enzyme activities in a dosage series of the long arm of chromosome one in maize. Genetics 92:1211–1229
38. Devlin RH, Holm DG, Grigliatti TA (1982) Autosomal dosage compensation in *Drosophila melanogaster* strains trisomic for the left arm of chromosome 2. Proc Natl Acad Sci USA 79: 1200–1204
39. Guo M, Birchler JA (1994) Trans-acting dosage effects on the expression of model gene systems in maize aneuploids. Science 266:1999–2002
40. Birchler JA, Riddle NC, Auger DL, Veitia RA (2005) Dosage balance in gene regulation: biological implications. Trends Genet 21:219–226
41. Yamada Y, Watanabe H, Miura F, Soejima H, Uchiyama M, Iwasaka T, Mukai T, Sakaki Y, Ito T (2004) A comprehensive analysis of allelic methylation status of CpG islands on human chromosome 21q. Genome Res 14:247–266
42. Sinet PM, Lavelle F, Michelson AM, Jerome H (1975) Super-

oxide dismutase activities of blood platelets in trisomy 21. Biochem Biophys Res Commun 67:904–909
43. Sherman L, Levanon D, Lieman-Hurwitz J, Dafni N, Groner Y (1984) Human Cu/Zn superoxide dismutase gene: molecular characterization of its two mRNA species. Nucleic Acids Res 12:9349–9365
44. Deutsch S, Lyle R, Dermitzakis ET, Attar H, Subrahmanyan L, Gehrig C, Parand L, Gagnebin M, Rougemont J, Jongeneel CV, et al (2005) Gene expression variation and expression quantitative trait mapping of human chromosome 21 genes. Hum Mol Genet 14:3741–3749
45. Cheung VG, Conlin LK, Weber TM, Arcaro M, Jen KY, Morley M, Spielman RS (2003) Natural variation in human gene expression assessed in lymphoblastoid cells. Nat Genet 33:422–425
46. Spielman RS, Bastone LA, Burdick JT, Morley M, Ewens WJ, Cheung VG (2007) Common genetic variants account for differences in gene expression among ethnic groups. Nat Genet 39:226–231
47. Storey JD, Madeoy J, Strout JL, Wurfel M, Ronald J, Key JM (2007) Gene-expression variation within and among human populations. Am J Hum Genet 80:502–509
48. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, et al (2006) Global variation in copy number in the human genome. Nature 444:444–454
49. O'Leary DA, Pritchard MA, Xu D, Kola I, Hertzog PJ, Ristevski S (2004) Tissue-specific overexpression of the HSA21 gene *GABPα*: implications for DS. Biochim Biophys Acta 1739:81–87
50. Nikolaienko O, Nguyen C, Crinc LS, Cios KJ, Gardiner K (2005) Human chromosome 21/Down syndrome gene function and pathway database. Gene 364:90–98
51. Chen H, Chrast R, Rossier C, Morris MA, Lalioti MD, Antonarakis SE (1996) Cloning of 559 potential exons of genes of human chromosome 21 by exon trapping. Genome Res 6: 747–760
52. Ohira M, Seki N, Nagase T, Suzuki E, Nomura N, Ohara O, Hattori M, Sakaki Y, Eki T, Murakami Y, et al (1997) Gene identification in 1.6-Mb region of the Down syndrome region on chromosome 21. Genome Res 7:47–58