European Journal of Human Genetics (2013) 21, 1253–1259

www.nature.com/ejhg

ARTICLE

The intellectual disability of trisomy 21: differences in gene expression in a case series of patients with lower and higher IQ

André Mégarbané*[,1,2], Florian Noguier[3], Samantha Stora[1], Laurent Manchon[3], Clotilde Mircher[1], Roman Bruno[3], Nathalie Dorison[1], Fabien Pierrat[3], Marie-Odile Rethoré[1], Bernadette Trentin[3], Aimé Ravel[1], Marine Morent[3], Gerard Lefranc[4] and David Piquemal[3]

Trisomy 21 (T21), or Down syndrome (DS), is the most frequent and recognizable cause of intellectual disabilities. The level of disability, as evaluated by the intelligence quotient (IQ) test, varies considerably between patients independent of other factors. To determine the genetic or molecular basis of this difference, a high throughput transcriptomic analysis was performed on twenty T21 patients with high and low IQ, and 10 healthy controls using Digital Gene Expression. More than 90 millions of tags were sequenced in the three libraries. A total of 80 genes of potential interest were selected for the qPCR experiment validation, and three housekeeping genes were used for normalizing purposes. *HLA DQA1* and *HLA DRB1* were significantly downregulated among the patients with a low IQ, the values found in the healthy controls being intermediate between those noted in the IQ+ and IQ− T21 patients. Interestingly, the intergenic region between these genes contains a binding sequence for the CCCTC-binding factor, or CTCF, and cohesin (a multisubunit complex), both of which are essential for expression of *HLA DQA1* and *HLA DRB1* and numerous other genes. Our results might lead to the discovery of genes, or genetic markers, that are directly involved in several phenotypes of DS and, eventually, to the identification of potential targets for therapeutic interventions.

European Journal of Human Genetics (2013) **21,** 1253–1259; doi:10.1038/ejhg.2013.24; published online 20 February 2013

Keywords: Down syndrome; Trisomy 21; intellectual disability; gene expression; HLA

INTRODUCTION

Trisomy 21 (T21) or Down syndrome (DS) is a chromosomal disorder resulting from the triplication of all or part of a chromosome 21. It is a common birth defect, the most frequent and recognizable form of intellectual disabilities (ID), appearing in about one out of every 700 newborns. The average intelligence quotient (IQ) of children with DS is around 50, ranging between 30 and 70. Remarkably, a small number of patients have a profound degree of ID, whereas others have a mild degree despite the absence of any genetic, cultural or familial favoring or disfavoring causes.[1]

Recent progress in studies of patients with partial T21 and mouse models of T21 suggest that it will be soon possible to link characteristic phenotype changes with differential gene expression of specific genes and help to decipher the molecular basis for these abnormalities, which may lead to treatment of the most distressing aspects of this disorder.[2] Such optimism is based on recent success with high-throughput genomic approaches in human medicine, gene expression signatures, and gene profiling studies that have linked specific gene regulation to specific phenotypic abnormalities, and aided the diagnosis, the treatment, or the prevention of diseases in DS.[3,4]

The purpose of this paper is to identify a case-series of patients with lower (IQ−) and higher IQ (IQ+) and to describe gene expression or transcriptome differences between them and healthy controls using the Digital Gene Expression (DGE) technique. This pilot description may suggest a genetic indicator of better intellectual prognosis in DS patients.

MATERIALS AND METHODS

Subjects

After the analysis of nearly 1000 clinical files of DS patients at the Jérôme Lejeune Institute, a series of 20 patients with a free and homogeneous T21, aged between 18 and 40 years were enrolled in this study. Patients were classified into two groups: those with a relatively lower IQ (IQ <20 or IQ−; four males and one female), and those with a higher one (IQ >70 or IQ+; six males and nine females). The IQ was measured using the Columbia test, a tool validated and largely used in similar settings in France.

Patients were not taking medications, had no neurological problems (epilepsy, seizures, west syndromes, and so on), no changes suggestive of early dementia, no autism, no endocrinological problems (hypo or hyperthyroidism, diabetes, and so on), no sleep-disordered breathing problems, no hearing impairment or vision impairment, no heart problems, no immune deficiency, and no cancer. In addition, in none of the families of the patients serious events were noted (death, frequent hospitalization, child abuse, and so on).

A third group of 10 individuals without T21 (four males and six females), aged between 26 and 39 years were added as controls.

[1]Institut Jérôme Lejeune, Paris, France; [2]Unité de Génétique Médicale et Laboratoire Associé INSERM UMR_S910, Beirut, Lebanon; [3]Skuldtech, Montpellier, France; [4]Université Montpellier 2 et CNRS UPR 1142, Institut de Génétique Humaine, Montpellier, France
*Correspondence: Professor A Mégarbané, Institut Jérôme Lejeune. 37, rue des Volontaires, Paris 75015, France. Tel: +33 1 56 58 63 00; Fax: +9611421632; E-mail: megarbane@usj.edu.lb
Received 21 August 2012; revised 27 December 2012; accepted 22 January 2013; published online 20 February 2013

This study was granted approval from the Institute Jérôme Lejeune Committee on Clinical Investigation and conformed to the tenets of the Declaration of Helsinki.

RNA samples

Blood samples were collected for genetic studies after informed written consent was obtained from all parents or guardians on behalf of the participants because of their inability to provide consent, and from the healthy controls. RNA samples were extracted from lymphoblastoid cell lines. RNA samples were obtained from 4×10^6 cells pellet with RNeasy Plus Mini kit (Qiagen Courtaboeuf, France) and QIAshredder (Qiagen). Control of RNA integrity was performed with the 2100 Bioanalyzer (Agilent Technologies, Lesulis, France) using Eukaryotic Total RNA 6000 Nano Chip (Agilent Technologies). RNA quantity was controlled using NanoDrop ND-1000 spectrophotometer (LABTECH, Palaiseau, France).

DGE library construction and tag-to-gene mapping

Three DGE libraries were constructed from pooled RNA samples of patients IQ+ and IQ−, and healthy controls. The libraries were constructed with Illumina's DGE Tag Profiling kit (ILLUMINA; SanDiego, CA, USA) according to the manufacturer's protocol (version 2.1B), using 5 μg of total RNA (mixing equal amounts of RNA from each individual). Sequencing analysis and base calling were performed using the Illumina Pipeline, and sequence tags were obtained after purity filtering. Data from each DGE library were analyzed with BIOTAG software (Skuldtech, Montpellier, France) for tag detection, tag counting and for assessing DGE library quality.[5] Raw and treated data are available on http://www.skuldtech.com/trisomie21.

Tag annotation and selection

A local database compiling *Homo sapiens* sequences and related information from well-annotated sequences of UniGene clusters (NCBI) was generated. For each sequence of this database, the expected DGE tag (canonical tag) located upstream of the 3′-nearest NlaIII restriction site (CATG) of the sequence (R1), as well as putative tags located in inner positions (labeled as R2, R3 and R4 starting from the 3′ end of the transcript) were extracted.[5] Experimental tags obtained from DGE libraries were matched and annotated (exact matches for the 17 bp) using this collection of virtual tags. First, a correspondence for each experimental tag with the virtual canonical tags (R1) was looked for. Then, unmatched experimental tags with the R2 tags, then with R3, and R4 were annotated. Targeted tags were selected using R package DESeq (Anders S, Bioconductor) for processing data without replicates. The analyzed genes were selected according to mathematic filters with the highest differential Fold Change (>1.5), FDR adjusted *P*-value criterion ($<10\%$, Benjamini-Hochberg Method) based on the type I ($\alpha \leq 5\%$) error.

cDNA synthesis and real-time PCR

Reverse transcriptions were performed for each of the 30 RNA samples in 20 μl final reaction volume with 1.5 μg of total RNA using 200 units of SuperScript II enzyme (M-MLV RT Type, Invitrogen, St Aubin, France) and 250 ng of random primers according to manufacturer's instructions (25 °C 10 min, 42 °C 50 min, 70 °C 15 min), the same day with the same pipettor set and the same manipulator. qPCR experiments were carried out using SYBR Green chemistry on LightCycler480 qPCR apparatus system I (Roche, Meylan, France). The reaction mix was prepared in a final volume of 10 μl as follows: 1 μl of cDNA matrix (1/30 diluted in H_2O) was added to 5 μl of SYBR Green I Master Mix (Roche Applied Science, Meylan, France) and 4 μl of Forward and Reverse primers mix (final concentration in PCR reaction of 0.5 μM each).

To discriminate specific from non-specific products and primer dimers, a melting curve was obtained by gradual increase in temperature from 65 to 95 °C. PCR efficiency was measured on standard curves performed for each primer pairs using a pool of all cDNA diluted as following: 1/10, 1/100, 1/1000, 1/10000 dilution factor. Primer pairs with PCR efficiency <1.8 were excluded of the analysis. The qPCR data were analyzed using the $2^{-(\Delta\Delta CT)}$ method.[6] Data were normalized using three housekeeping genes: *LIMK1* (NM_002314), *APEH* (BC000362) and *TUFM* (NM_003321) (Table 1). The Partial Least Squares Discriminant Analysis regression (PLSDA) analysis was performed with the R package mixOmics.[7] The PLSDA was used to select the most discriminant genes. According to the mixOmics vignette and from the variance importance plot, a score >1 represent a strong weight in the patient discrimination.[7]

RESULTS

By using Next-Generation Sequencing, we performed a transcriptomic study using the open method, DGE, to reveal genes differentially expressed in the two different selected phenotypes of DS. More than 90 million tags were sequenced in the two libraries. By taking into consideration the DGE-tags with a minimal occurrence of two times, the two libraries revealed 68 046 unique tags. Raw and treated data were integrated in a database with associated tools for annotation, *in silico* PCR, tag prediction and data visualization. This database is accessible via a user-friendly website on http://www.skuldtech.com/trisomie21 and can incorporate future data from the Next-Generation Sequencing.

After data were filtered and classified according to the statistical approach by DESeq and the fold induction among the well-annotated tags, 80 genes were selected (Table 2). To explore the individual variability, individual qPCR analysis was performed.

The qPCR data analysis showed that two genes were sufficient enough based on our selection criteria (fold change expression >2.5 and SD <0.1) to discriminate our entire population between IQ− and IQ+. The PLSDA run on these data with the R package mixOmics[7] showed that the genes *HLA-DQA1* (NM_002122, major histocompatibility complex (MHC), class II, DQ-α 1) and *HLA-DRB1* (NM_002124, MHC, class II, DR-β 1) obtained the highest values (*HLA-DQA1* got 2.84 and 2.22 for component 1 and 2, respectively, and *HLA-DRB1* got 2.56 and 1.90 for component 1 and 2, respectively, of the PLSDA) from the variance importance plot. Both expressions appeared to be less frequent in the IQ− group (Figure 1). Interestingly, the values found in the healthy controls were between both those noted in the IQ− and IQ+ T21 patients. A Monte–Carlo test on a factorial discriminate analysis gave a significant *P*-value of 0.02.

DISCUSSION

T21 is a direct consequence of either an additional copy of protein-coding genes that are dosage sensitive or an additional copy of non-protein-coding sequences that are regulatory or otherwise functional.

Table 1 PCR primers list for the two targeted genes and the three housekeeping genes (*)

Gene	*Gb id.*	*Primer forward*	*Primer reverse*
HLA-DQA1	NM_002122	5′-GACCACGTCGCCTCTTATGG-3′	5′-ACGTAGAACTGCTCATCTCCA-3′
HLA-DRB1	AK290388	5′-TGTTCTCCAGCATGGTGTGT-3′	5′-AGAAACGTGGTCTGGTGTCC-3′
*LIMK1**	NM_002314	5′-ATGAGGTTGACGCTACTTTGTTG-3′	5′-CCTCTCCCATACGTTCTTCCC-3′
*APEH**	BC000362	5′-CTGGAACGCATGGAGAACATT-3′	5′-CCGTCATGGAACACCAGGTA-3′
*TUFM**	NM_003321	5′-GGGGCTAAGTTCAAGAAGTACG-3′	5′-CACATGAGCCGCATTGATGG-3′

Table 2 Expression data from DGE analysis of the 80 selected transcripts in whole-blood cells in IQ+ *versus* IQ− patients

DGE-tag	*IQ+*	*IQ−*	*P-value*	*Sequence i.d.*	*Cyt loc*	*Hugo name*	*Description*	*GO annotation*
5′-CAAATGAGGAGTGCCAG-3′	0	541	1E−160	NM_001130523	1p13.2	CSDE1	Neuroblastoma RAS viral (v-ras) oncogene homolog	P:regulation of transcription, DNA-dependent
5′-TCCAAAGTAATGGAGAT-3′	0	247	1E−72	BC000301	1p34	HDAC1	Histone deacetylase 1	F:histone deacetylase activity
5′-ATCAGTGGCTTTGAATG-3′	0	239	3E−70	BC000331	1q21	PSMB4	Proteasome (prosome, macropain) subunit, β-type 4	P:small molecule metabolic process
5′-GAAGCCCCAGCTCAGCT-3′	842	3	1E−124	BC066343	2p12	IGKC	Immunoglobulin κ-constant	P:regulation of immune response
5′-ATATTTTAAATGTTAAG-3′	0	183	2E−53	BC047933	2p15	B3GNT2	UDP-GlcNAc: β-Gal-β-1,3-N-acetylglucosaminyltransferase 2	F:transferase activity, transferring glycosyl groups
5′-TTCATACACCTATCCCC-3′	432	27556	3E−118	BC001797	2p16.1	CCDC104	Coiled-coil domain-containing 104	
5′-TTAACAACATTAAAAAC-3′	22	64	2E−06	NM_001139488	2p25.1-p24.1	RASGRP3	RAS guanyl releasing protein 3 (calcium and DAG-regulated)	F:Ras GTPase binding
5′-TCATTGTAATGATTTCA-3′	0	237	1.46E−69	BC014274	2q11.2	STARD7	StAR-related lipid transfer (START) domain containing 7	
5′-GAAAAATGGTTGATGGA-3′	0	4692	3E−174	BC010054	3p22.2	RPSA	Ribosomal protein SA	P:mRNA metabolic process
5′-CGTGTTAATGGCTGTTC-3′	0	171	7E−50	BC000288	3q21	CNBP	CCHC-type zinc finger, nucleic acid-binding protein	P:positive regulation of transcription from RNA polymerase II promoter
5′-TCTTAATGAAGTTTGAA-3′	0	286	3E−84	CR612348	3q28	EIF4A2	Eukaryotic translation initiation factor 4A2	F:helicase activity
5′-ATGTCATCAAATGGGTG-3′	0	217	1E−63	NM_004068	3q28	AP2M1	Adaptor-related protein complex 2, mu 1 subunit	P:intracellular protein transport
5′-TGTGCTAAATGTGTTCG-3′	0	673	2E−200	BC070208	4q25	RPL34	Ribosomal protein L34	F:structural constituent of ribosome
5′-ATACTTTAATCAGAAGC-3′	0	186	2E−54	NM_001154	4q26-q28l4q27	ANXA5	Annexin A5	F:receptor tyrosine kinase binding
5′-TCTGCAATGAAGAGATT-3′	0	372	1E−109	BC005288	5q13.3	NSA2	NSA2 ribosome biogenesis homolog (S. cerevisiae)	P:rRNA processing
5′-TTACTAAATGGTGTTAC-3′	0	428	8E−127	NM_001746	5q35	CANX	Calnexin	P:protein secretion
5′-ACACTAAAATGGCAGAG-3′	0	191	8E−56	BC011676	6p21.1	NFKBIE	Nuclear factor of κ-light polypeptide gene enhancer in B-cells inhibitor, epsilon	P:D-serine transport
5′-TTGATTTCTTAGCTGAC-3′	425	1	6E−126	AF533900	6p21.3	HLA-DQA1	Major histocompatibility complex, class II, DQ-α 1	P:T-cell receptor-signaling pathway
5′-GCAGTTCTGACAGTGAC-3′	1616	1	1E−119	AK290388	6p21.3	HLA-DRB1	Major histocompatibility complex, class II, DR-β 1	F:peptide antigen binding
5′-TAGATAATGGCCATCAT-3′	0	372	5E−110	AK098772	6p21.33	TUBB	Tubulin-β	F:structural constituent of cytoskeleton
5′-GTCTTAAAGTGAGATTT-3′	4	16	4E−03	NM_000636	6q25.3	SOD2	Superoxide dismutase 2, mitochondrial	P:negative regulation of fat cell differentiation
5′-TACATCCGAATGCTAAA-3′	0	212	4E−62	NM_001128619	7q33	LUZP6	Myotrophin	P:positive regulation of macromolecule biosynthetic process
5′-AACAGAAGCAAATGATG-3′	0	396	3E−117	BG480804	8q13l8q13	SNHG6	Small nucleolar RNA host gene 6 (non-protein coding)	
5′-ATCAAATGCAACCTCAC-3′	0	364	1E−107	NM_002467	8q24.21	MYC	V-myc myelocytomatosis viral oncogene homolog (avian)	F:sequence-specific DNA-binding transcription factor activity
5′-GACAATGCCAGCAAGAA-3′	0	165	4E−48	NM_001001973	10p15.1	ATP5C1	ATP synthase, H+ transporting, mitochondrial F1 complex, gamma polypeptide 1	P:small molecule metabolic process
5′-GCCATAAAATGGCTTTA-3′	0	745	5E−222	NM_002727	10q22.1	SRGN	Serglycin	P:mast cell secretory granule organization
5′-CTAAATATTCTTTCCTA-3′	22	55	6E−05	NM_005445	10q25	SMC3	Structural maintenance of chromosomes 3	P:chromosome organization
5′-GGGACGGCGCGGAGGAG-3′	29	51	4E−03	NM_030930	11q13	UNC93B1	Unc-93 homolog B1 (C. elegans)	P:toll-like receptor-3 signaling pathway
5′-GTACAGGCTTGGAGCTT-3′	85	32	3E−07	NM_012296	11q14.1	GAB2	GRB2-associated binding protein 2	P:phosphatidylinositol-mediated signaling
5′-GACTTTTCTGGGAAATG-3′	0	179	3E−52	Z49194	11q23.1	POU2AF1	POU class 2 associating factor 1	P:regulation of transcription, DNA-dependent
5′-AATAGGTCCAACCAGCT-3′	9081	577	5E−123	NM_001028	11q23.3	RPS25	Ribosomal protein S25	P:mRNA metabolic process
5′-CCAGGAGGAATGCCTGG-3′	0	243	2E−71	BC042163	11q24.1	HSPA8	Heat shock 70 kDa protein 8	F:ATPase activity, coupled
5′-TTTCTGTATGTTAATGA-3′	0	415	6E−123	NM_001013699	12p11.21	H3F3C	H3 histone, family 3C	F:DNA binding
5′-GAAATGATGAGTCAGAA-3′	0	250	3E−73	BM562920	12q12	PFDN5	Prefoldin subunit 5	P:cellular protein metabolic process
5′-GAAATGTAAGAGTGGAA-3′	0	357	1E−105	AK123458	12q13.12-q13.13	PCBP2	Poly(rC)-binding protein 2	F:RNA binding
5′-GTGCTGAATGGCTGAGG-3′	0	1162	9E−173	BC017455	12q13.2	MYL6	Myosin, light chain 6, alkali, smooth muscle and non-muscle	P:skeletal muscle tissue development
5′-TCACAAGCAAATGTGTC-3′	0	5532	9E−206	NM_001113203	12q23-q24.1	NACA	Nascent polypeptide-associated complex-α subunit	P:translation
5′-TTTCTGTGAAAATGTAT-3′	0	232	4E−68	NM_004592	12q24.33	SFSWAP	Splicing factor, suppressor of white-apricot homolog (Drosophila)	P:regulation of transcription, DNA-dependent
5′-CTGTTGATTGCTAAATG-3′	0	334	1E−98	AK126454	13q14.3	HNRNPA1L2	Heterogeneous nuclear ribonucleoprotein A1-like 2	F:RNA binding

Table 2 (Continued)

DGE-tag	*IQ+*	*IQ−*	P-*value*	*Sequence i.d.*	*Cyt loc*	*Hugo name*	*Description*	*GO annotation*
5′-TCAAATGCATCCTCTAG-3′	0	702	4E−209	AK126950	14q11.2	HNRNPC	Heterogeneous nuclear ribonucleoprotein C (C1/C2)	F:nucleic acid binding
5′-GAGATCCGCAATGCTTA-3′	0	745	5E−222	NM_006263	14q11.2	PSME1	Proteasome (prosome, macropain) activator subunit 1 (PA28-α)	P:small molecule metabolic process
5′-GAGATGCCTTTGGTGCT-3′	17	3	1E−03	NM_005132	14q11.2-q12	REC8	REC8 homolog (yeast)	P:meiosis
5′-GAAATAAAGCACCCACC-3′	8003	739	1E−96	BC016381	14q32.33	IGHM	Immunoglobulin heavy constant gamma 1 (G1m marker)	F:antigen binding
5′-TTTAATACATAGGTGAT-3′	0	189	3E−55	NM_001014812	15q22.31	FAM96A	Family with sequence similarity 96, member A	F:molecular_function
5′-TAATGGTAACTTGGACT-3′	0	177	1E−51	NM_004255	15q24.1	COX5A	Cytochrome c oxidase subunit Va	F:metal ion binding
5′-TTTGTATTGCACAGATC-3′	29	49	6E−03	NM_144572	15q24.3-q25.1	TBC1D2B	TBC1 domain family, member 2B	F:phospholipid binding
5′-AGACCTGTAATAAAATA-3′	6	20	3E−03	NM_006565	16q21-q22.3	CTCF	CCCTC-binding factor (zinc finger protein)	P:regulation of histone methylation
5′-AAAAAGATAATGGAGAC-3′	0	215	5E−63	AK295346	18q21.32	SEC11C	SEC11 homolog C (S. cerevisiae)	P:cellular protein metabolic process
5′-GCCTTATCTGTTCCAGT-3′	33	9	1E−04	NM_002918	19p13.1	RFX1	Regulatory factor X, 1 (influences HLA class II expression)	F:RNA polymerase II distal enhancer sequence-specific DNA-binding transcription factor activity
5′-TGTTAATGTTACGATGT-3′	0	243	2E−71	AB209074	19p13.3	MKNK2	MAP kinase-interacting serine/threonine kinase 2	F:protein serine/threonine kinase activity
5′-CACGCAATGCTAGCTGC-3′	0	654	1E−194	AK095154	19p13.3	AES	Amino-terminal enhancer of split	P:response to interleukin-1
5′-CCCCCAATGCTGAGGCC-3′	0	684	1E−203	NM_007165	19p13.3-p13.2	SF3A2	Splicing factor 3a, subunit 2, 66 kDa	F:metal ion binding
5′-GCCCAGCTCAAATGCTA-3′	0	189	3E−55	AK308809	19q13.2	PSMD8	Proteasome (prosome, macropain) 26S subunit, non-ATPase, 8	P:proteolysis
5′-AAATGTTTGGCCTTAGT-3′	0	368	8E−109	NM_012068	19q13.3	ATF5	Activating transcription factor 5	F:sequence-specific DNA binding
5′-CCTTCGAGATCATACAC-3′	147	5856	2E−189	BG165682	19q13.4	RPS5	Ribosomal protein S5	P:mRNA metabolic process
5′-CCAGGAACAATGTCTCC-3′	0	309	4E−91	AK054634	20p11	SNX5	Sorting nexin 5	P:cell communication
5′-ACAAATCCTTGAATGTT-3′	0	290	1E−85	NM_000801	20p13	FKBP1A	FK506-binding protein 1A, 12 kDa	F:transforming growth factor-β-activated receptor activity
5′-AAGGAAGCAATGGTTCA-3′	0	163	3E−47	NM_006392	20p13	NOP56	NOP56 ribonucleoprotein homolog (yeast)	F:RNA binding
5′-TAAATAAATACATTCTT-3′	0	6	5E−03	NM_000484	21q21.2	APP	Amyloid-β (A4) precursor protein	P:axon cargo transport
5′-GACTTGGCCCAAAAGAA-3′	64	5	1E−14	NM_000819	21q22.11	GART	Phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole synthetase	P:small molecule metabolic process
5′-GCTTCCTAAATGGCCCT-3′	0	6	500E−03	NM_000071	21q22.3	CBS	Cystathionine-β-synthase	P:cartilage development involved in endochondral bone morphogenesis
5′-CTCAGGAAATAAATGTG-3′	0	192	4E−56	BC005402	22cen-q12.3	UQCR10	Ubiquinol-cytochrome c reductase, complex III subunit X	P:small molecule metabolic process
5′-AACGCGGCCAATGTGGG-3′	0	623	4E−185	BC036909	22q11.23	LOC284889	Hypothetical LOC284889	F:phenylpyruvate tautomerase activity
5′-AAACTAGAAATGTCATC-3′	0	433	2E−128	AF230904	Xp22.1-p21.3	SH3KBP1	SH3-domain kinase-binding protein 1	P:regulation of cell shape
5′-CAATGTGTTATGTAGTG-3′	0	197	1E−57	NM_004541	Xq24	NDUFA1	NADH dehydrogenase (ubiquinone) 1-α subcomplex, 1, 7.5 kDa	F:NADH dehydrogenase (ubiquinone) activity
5′-GTTAATAACAAATGAAT-3′	0	35	2E−11	NM_006603	Xq25	STAG2	Stromal antigen 2	P:negative regulation of DNA endoreduplication
5′-TGAAGGATGCCAATGGC-3′	0	220	1E−64	NM_001008	Yp11.3	RPS4Y1	Ribosomal protein S4, Y-linked 1	P:mRNA metabolic process
5′-GAATCCAACTGCTTCGA-3′	3597	176	5E−105	AJ712551	—	—	Transcribed locus, moderately similar to NP_061929.2 NADH dehydrogenase (ubiquinone) 1-β subcomplex subunit 11, mitochondrial isoform 1 (*Homo sapiens*)	P:small molecule metabolic process
5′-CTGCTATACGAGAGAAT-3′	22	2699	2E−192	AV755966	—	—	Transcribed locus, strongly similar to XP_003086526.1 PREDICTED: 60S ribosomal protein L5-like isoform 3 (*Mus musculus*)	P:mRNA metabolic process
5′-AGGGCAGGGAATGTGTC-3′	0	175	4E−51	AW967735	—	—	Transcribed locus	
5′-ACTTTTTCAAAAAAAAA-3′	5286	213	3E−159	BE877281	—	—	Transcribed locus, strongly similar to NP_001460.1 X-ray repair cross-complementing protein 6 (*Homo sapiens*)	F:helicase activity
5′-GGGACGGCGCGGAGGAG-3′	29	51	4E−03	BG469659	—	—	Transcribed locus, moderately similar to XP_001721426.3 PREDICTED: protein unc-93 homolog B1-like (*Homo sapiens*)	P:toll-like receptor 3-signaling pathway

Table 2 (Continued)

DGE-tag	IQ+	IQ−	P-value	Sequence i.d.	Cyt loc	Hugo name	Description	GO annotation
5′-AATAAATGGATCTGTGA-3′	0	224	1E−65	BG498681	—	—	Transcribed locus, moderately similar to NP_004885.1 6.8 kDa mitochondrial proteolipid isoform 1 (*Homo sapiens*)	
5′-CAAATGAGGAGTGCCAG-3′	0	541	1E−160	BG680089	—	—	Transcribed locus, strongly similar to XP_002729159.1 PREDICTED: cold shock domain-containing E1, RNA binding (*Rattus norvegicus*)	
5′-TTTATTTAAGAAATGGA-3′	0	239	3E−70	BM541359	—	—	Transcribed locus, strongly similar to NP_001013208.1 UV excision repair protein RAD23 homolog A (*Rattus norvegicus*)	F:metal ion binding
5′-CAATAAATGTTCTGGTT-3′	0	1496	6E−223	BM553029	—	—	CDNA clone IMAGE:6385453	
5′-TTTTTAATGTTGTCTGT-3′	0	233	2E−68	BU566862	—	—	Transcribed locus	
5′-CTTCGGATGTCTTGGAG-3′	16	73	2E−10	NM_001007226	—	SPOP	Transcribed locus	F:ubiquitin protein ligase binding
5′-CCACCCCGAATGGCTCA-3′	0	771	8E−230	NM_001098576	—	—	Transcribed locus	P:apoptotic process
5′-AATACTTAAATGCCAAC-3′	0	399	4E−118	NM_006990	—	WASF2	MRNA adjacent to 3- end of integrated HPV16 (INT290)	P:lamellipodium assembly

The 17 bp DGE-tags are indicated in the first column. The 2nd and 3rd columns indicate the exact occurrences of the DGE-tags sequenced. The '0' value indicate that a DGE-tag was not identified in one of the 2 groups. Some GO (http://www.geneontology.org/) annotations are listed in the last column with 'P:*' for Biological process and 'F:*' for molecular function.

	HLA-DQA1	HLA-DRB1
IQ −	$3.70\ E^{-3} \pm 4.4\ E^{-3}$	0.384 ± 0.09
IQ +	1	1
Control	0.129 ± 0.19	0.641 ± 0.18

Figure 1 HLA-DRB1 and HLA-DQA1 genes expression profiles in the two conditions of IQ and one in Control patients for DS patients using the $2^{-(\Delta\Delta CT)}$ method from qPCR data. The arbitrary value of 1 is given to the IQ+ patient values. ± indicates the SD values.

The effect of some dosage-sensitive genes on the phenotypes might be allele-specific, and could also depend on the polymorphism of coding and non-coding (regulatory, mRNAs, and so on) nucleotide sequences and on the combination and interaction of all these variants coding for proteins alleles with qualitative (alleles with amino-acid variation) or quantitative (alleles with variation in gene expression level) traits. Dosage sensitive genes could act directly to induce pathogenesis or indirectly by interacting with genes or gene products of either aneuploid or non-aneuploid genes or gene products.[8] It is a reasonable speculation that the genetic background of individuals has an important role in the variability of phenotypic severity that is seen in DS.[2]

In order to understand why some T21 patients present a wide difference in their intelligence despite the absence of any known favoring or disfavoring factors, we initiate the construction of three transcriptomic libraries, in T21 IQ− and IQ+ patients and in healthy controls using the SAGE technique. Indeed, searching for transcriptional alterations is an easier (or more effective) means to find relevant loci than would be complete sequencing. Such approach was already tested in DS and different pathologies.[9–12]

The low number of patients, especially in the IQ− group, was secondary to the fact that we wanted to restrict this study only to the T21 patients that we could discriminate by IQ test. Interestingly, the group IQ− was predominantly males, whereas the IQ+ group was frequently females with a sex ratio of 2/3. Few reports already pointed the fact that boys with DS are more affected than girls but without any reason.[13]

Two genes with major expression differences were found: *HLA-DQA1* and *HLA-DRB1*. Both genes were less expressed in the T21 IQ− population (Figure 1). HLA loci were already reported as leading to significant risk factor for celiac disease in DS patients,[14] but never for ID. Nevertheless, HLA has been associated with cognitive ability. For example, Cohen *et al.*[15] showed that a proportion of patients with primary neuronal degeneration of the Alzheimer type have the HLA-B7 antigen, and that patients with HLA-B7 antigens had selective attention span significantly lower than Alzheimer patients without the antigen. Recently, HLA-DRB1 has been associated with cognitive ability in both demented and non-

demented individuals,[16] and a positive association of HLA-DR4 with attention deficit and hyperactivity disorder and the role of the HLA-DRB1 in the etiology of some types of childhood neuropsychiatric illnesses were reported.[17] Future qualitative and quantitative studies at the level of the HLA-DQA1 and HLA-DRB1 proteins on the PBLs of more IQ+ and IQ− DS patients might further validate our results found at the transcriptional level. Likewise, the comparison between the SNPs located within the DS critical region on chromosome 21 in the two groups of DS patients might be informative.

Interestingly, the intergenic region between *HLA-DQA1* and *HLA-DRB1*, is characterized by the presence of a CTCF-binding CCCTC sequence, *XL9*, of high histone acetylation and of particular importance for the control of the expression of both HLA genes and for the chromatin architecture of the MHC class II locus.[18–21] The fact that CTCF expression is increased and transcription of *HLA DQA1* and *HLA DRB1* is decreased in the IQ− group could be in favor to a repressor role for *CTCF*. On the other hand, the *RFX1* gene, coding for a protein that binds to the X-boxes of MHC class II genes and which is essential in their expression, shows a three-fold decreased level of expression in the IQ− group in comparison to the IQ+ group. In contrast, in the IQ+ group, the downregulated *CTCF* gene expression and the increased *RFX1* gene expression could explain the upregulation of the *HLA-DQA1* and *HLA-DRB1* genes. Furthermore, it was found recently that cohesin and CTCF-binding sites have a high degree of overlap, and interacts with each other.[20,22–24] Moreover, often cohesin and CTCF have been found to colocalize at several thousand sites in non-repetitive sequences in the human genome.[22,24,25] In addition to the regulation of gene expression, cohesin has also other important functions: it forms a huge tripartite ring, mediates the sister chromatid cohesion, and facilitates the repair of damaged DNA.[26,27] From our analysis, we can postulate that *HLA-DQA1* and *HLA-DRB1* may represent a genetic biomarker for predicting differences in ID conditions, but also that polymorphisms or mutations of the cohesin subunits might have an important role for the non-disjunction of the chromosomes 21 and/or for the dysregulation of the expression of many genes.

Other genes present a different transcriptional pattern between the IQ− and IQ+ groups. Although this difference is not as obvious as with HLA, it might be interesting to be carefully looked at (Table 2). For instance, *APP* (Amyloid-β (A4) precursor protein), located on chromosome 21q21.3, was overexpressed in the IQ− group. Duplication of the *APP* gene was found to lead to early-onset Alzheimer disease (AD) and prominent cerebral amyloid angiopathy.[28] Triplication of the *APP* gene accelerates the APP expression, leading to cerebral accumulation of APP-derived amyloid-β peptides, early-onset AD neuropathology, and age-dependent cognitive sequelae.[29] At relatively early ages, DS patients develop progressive formation and extracellular aggregation of amyloid-β peptide, considered as one of the causal factors in the pathogenesis of AD.[30] The cystathionine β-synthase (*CBS*) gene, located on human chromosome 21q22.3 encodes a key enzyme of sulfur-containing amino acid metabolism, a pathway involved in several brain physiological processes. It is overexpressed in the brain of individuals with DS,[31] and thus was considered as a good candidate in having a role in the DS cognitive profile.[32] Recently, Régnier *et al.*[33] studied the neural consequences of CBS overexpression in a transgenic mouse line expressing the human *CBS* gene. They observed that the transgenic mice showed normal behavior, and that hippocampal synaptic plasticity was facilitated. Thus, they raised the possibility that *CBS* overexpression might have an advantageous effect on some cognitive functions in DS. Our results do not confirm the latter hypothesis as we found that *CBS* was overexpressed in the IQ− group v/s IQ+ group. The glycinamide phosphoribosyltransferase (*GART)* gene, located on 21q22.11 was found overexpressed in the IQ+ group. GART is an essential enzyme in *de novo* purine biosynthesis (OMIM 138440). In 1993, Peeters *et al.*[34] studied the variations in mitotic index of lymphocyte cultures to which various metabolites of purine synthesis (inosine, adenosine and guanosine) were added. They unexpectedly found a significant decrease in mitotic index in DS patients without psychiatric complications when compared with normal controls, and opposite reactions in DS patients presenting psychotic features. They concluded that T21 patients may have different purine metabolism depending on whether or not they have associated psychiatric complications. Purine-rich diet or the prescription of exogenous inosine, and serotonine-rich diet were used to treat psychotic T21 patients and some results in reduction of self-injurious behavior were reported.[34,35] Our results tend to give a beginning of explanation to the latter observations. The overexpression of *GART* in IQ+ patients might prevent the apparition of any abnormal behavior in T21 patients leading to a better IQ. Supplementing such patients with a purine-rich diet might not be beneficial (as seen by the decrease mitotic index) on the contrary of patients with a lower expression of *GART*.

Interestingly, *DYRK1A* (Dual-specificity tyrosine-(Y)-phosphorylation-regulated kinase 1A), located on chromosome 21q22.2 and known to have a significant role in developmental brain defects and in early onset neurodegeneration, neuronal loss and dementia in DS[36] did not show a significantly different expression between the IQ− and IQ+ groups.

In conclusion, we established a transcriptome of DS patients with IQ− and IQ+ and found that *HLA-DQA1* and *HLA-DRB1* may discriminate between the two populations. However, the pools of eligible patients available for such studies in particular centers are usually small, and therefore findings remain limited in their power to significantly rule in or out a marker that discriminates early between DS who will be IQ− and who will be IQ+. Larger multicenter series will allow to determine in a valid way the presence of such markers and whether they are sex-specific. Beyond providing evidence to support the hypothesis that has directed much of the work discussed, the importance of determining valid markers would have major consequences for informing pathogenesis and for providing defined targets to combat pathogenesis. This may lead to the discovery of genes that are 'directly' involved in several phenotypes of DS and eventually to the identification of potential targets for therapeutic interventions. For example, the genetic association with HLA supports the involvement of the immune system in ID and offers new targets for drug development.

Continued and increasing investments in research on the genetic and molecular basis of T21 promise to transform the lives of these individuals and the communities in which they live.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by the Jérôme Lejeune Foundation.

1 Carr J: Six weeks to 21 Years: A longitudinal study of children with Down's syndrome and their families. *J Child Psychol Psychiatry* 1991; **29**: 407–431.
2 Mégarbané A, Ravel A, Mircher C *et al*: The 50th anniversary of the discovery of trisomy 21: the past, present, and future of research and treatment of Down syndrome. *Genet Med* 2009; **11**: 611–616.

3 Lima FA, Moreira-Filho CA, Ramos PL *et al*: Decreased AIRE expression and global thymic hypofunction in Down syndrome. *J Immunol* 2011; **187**: 3422–3430.
4 Loudin MG, Wang J, Eastwood Leung HC *et al*: Genomic profiling in Down syndrome acute lymphoblastic leukemia identifies histone gene deletions associated with altered methylation profiles. *Leukemia* 2011; **25**: 1555–1563.
5 Piquemal D, Commes T, Manchon L *et al*: Transcriptome analysis of monocytic leukemia cell differentiation. *Genomics* 2002; **80**: 361–371.
6 Livak KJ, Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 2001; **25**: 402–408.
7 Lê Cao KA, Boitard S, Besse P, Sparse PLS: discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems. *BMC Bioinformatics* 2011; **12**: 253.
8 Antonarakis SE, Lyle R, Dermitzakis ET, Reymond A, Deutsch S: Chromosome 21 and down syndrome: from genomics to pathophysiology. *Nat Rev Genet* 2004; **5**: 725–738.
9 Booij BB, Lindahl T, Wetterberg P *et al*: A gene expression pattern in blood for the early detection of Alzheimer's disease. *J Alzheimers Dis* 2011; **23**: 109–119.
10 Chrast R, Scott HS, Papasavvas MP *et al*: The mouse brain transcriptome by SAGE: differences in gene expression between P30 brains of the partial trisomy 16 mouse model of Down syndrome (Ts65Dn) and normals. *Genome Res* 2000; **10**: 2006–2021.
11 Costa V, Angelini C, D'Apice L *et al*: Massive-scale RNA-Seq analysis of non ribosomal transcriptome in human trisomy 21. *PLoS ONE* 2011; **6**: e18493.
12 Sommer CA, Pavarino-Bertelli EC, Goloni-Bertollo EM, Henrique-Silva F: Identification of dysregulated genes in lymphocytes from children with Down syndrome. *Genome* 2008; **51**: 19–29.
13 Kittler P, Krinsky-McHale SJ, Devenny DA: Sex differences in performance over 7 years on the Wechsler Intelligence Scale for Children-Revised among adults with intellectual disability. *J Intellect Disabil Res* 2004; **48**: 114–122.
14 Book L, Hart A, Black J, Feolo M, Zone JJ, Neuhausen SL: Prevalence and clinical characteristics of celiac disease in Down's syndrome in a US study. *Am J Med Genet* 2001; **98**: 70–74.
15 Cohen D, Eisdorfer C, Walford RL: Histocompatibility antigens (HLA) and patterns of cognitive loss in dementia of the Alzheimer type. *Neurobiol Aging* 1981; **2**: 277–280.
16 Payton A, van den Boogerd E, Davidson Y *et al*: Influence and interactions of cathepsin D, HLA-DRB1 and APOE on cognitive abilities in an older non-demented population. *Genes Brain Behav* 2006; **5**: 23–31.
17 Aureli A, Sebastiani P, Del Beato T *et al*: Investigation on the possible relationship existing between the HLA-DR gene and attention deficit hyperactivity disorder and/or mental retardation. *Int J Immunopathol Pharmacol* 2008; **21**: 985–991.
18 Majumder P, Gomez JA, Boss JM: The human major histocompatibility complex class II HLA-DRB1 and HLA-DQA1 genes are separated by a CTCF-binding enhancer-blocking element. *J Biol Chem* 2006; **281**: 18435–18443.
19 Majumder P, Gomez JA, Chadwick BP, Boss JM: The insulator factor CTCF controls MHC class II gene expression and is required for the formation of long-distance chromatin interactions. *J Exp Med* 2008; **205**: 785–798.
20 Majumder P, Boss JM: CTCF controls expression and chromatin architecture of the human major histocompatibility complex class II locus. *Mol Cell Biol* 2010; **30**: 4211–4223.
21 Majumder P, Boss JM: DNA methylation dysregulates and silences the HLA-DQ locus by altering chromatin architecture. *Genes Immun* 2011a; **12**: 291–299.
22 Parelho V, Hadjur S, Spivakov M *et al*: Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* 2008; **132**: 422–433.
23 Majumder P, Boss JM: Cohesin regulates MHC class II genes through interactions with MHC class II insulators. *J Immunol* 2011; **187**: 4236–4244.
24 Wendt KS, Yoshida K, Itoh T *et al*: Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* 2008; **451**: 796–801.
25 Wendt KS, Peters JM: How cohesin and CTCF cooperate in regulating gene expression. *Chromosome Res* 2009; **17**: 201–214.
26 Kudo NR, Wassmann K, Anger M *et al*: Resolution of chiasmata in oocytes requires separase-mediated proteolysis. *Cell* 2006; **126**: 135–146.
27 Kudo NR, Anger M, Peters AH *et al*: Role of cleavage by separase of the Rec8 kleisin subunit of cohesin during mammalian meiosis I. *J Cell Sci* 2009; **122**: 2686–2698.
28 Rovelet-Lecrux A, Hannequin D, Raux G *et al*: APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. *Nat Genet* 2006; **38**: 24–26.
29 Moncaster JA, Pineda R, Moir RD *et al*: Alzheimer's disease amyloid-beta links lens and brain pathology in Down syndrome. *PLoS ONE* 2010; **5**: e10659.
30 Azkona G, Amador-Arjona A, Obradors-Tarragó C *et al*: Characterization of a mouse model overexpressing beta-site APP-cleaving enzyme 2 reveals a new role for BACE2. *Genes Brain Behav* 2010; **9**: 160–172.
31 Lockstone HE, Harris LW, Swatton JE *et al*: Gene expression profiling in the adult Down syndrome brain. *Genomics* 2007; **90**: 647–660.
32 Lejeune J: Pathogenesis of mental deficiency in trisomy 21. *Am J Med Genet* 2000; **7**: 20–30.
33 Régnier V, Billard JM, Gupta S *et al*: Brain phenotype of transgenic mice over-expressing cystathionine *β*-synthase. *PLoS ONE* 2012; **7**: e29056.
34 Peeters MA, Megarbane A, Cattaneo F, Rethore MO, Lejeune J: Differences in purine metabolism in patients with Down's syndrome. *J Intellect Disabil Res* 1993; **37**: 491–505.
35 Gedye A: Dietary increase in serotonine reduces self-injurious behavior in a Down's syndrome adult. *J Ment Defic Res* 1990; **34**: 247–258.
36 Wegiel J, Gong CX, Hwang YW: The role of DYRK1A in neurodegenerative diseases. *FEBS J* 2011; **278**: 236–245.