The genetic architecture of Down syndrome phenotypes revealed by high-resolution analysis of human segmental trisomies

Jan O. Korbel[a,b,c,1], Tal Tirosh-Wagner[d,1], Alexander Eckehart Urban[e,f,1], Xiao-Ning Chen[d], Maya Kasowski[e], Li Dai[d], Fabian Grubert[f], Chandra Erdman[g], Michael C. Gao[d], Ken Lange[h,i], Eric M. Sobel[h], Gillian M. Barlow[d], Arthur S. Aylsworth[j,k], Nancy J. Carpenter[l], Robin Dawn Clark[m], Monika Y. Cohen[n], Eric Doran[o], Tzipora Falik-Zaccai[p], Susan O. Lewin[q], Ira T. Lott[o], Barbara C. McGillivray[r], John B. Moeschler[s], Mark J. Pettenati[t], Siegfried M. Pueschel[u], Kathleen W. Rao[j,k,v], Lisa G. Shaffer[w], Mordechai Shohat[x], Alexander J. Van Riper[y], Dorothy Warburton[z,aa], Sherman Weissman[f], Mark B. Gerstein[a], Michael Snyder[a,e,2], and Julie R. Korenberg[d,h,bb,2]

Departments of [a]Molecular Biophysics and Biochemistry, [e]Molecular, Cellular, and Developmental Biology, and [f]Genetics, Yale University School of Medicine, New Haven, CT 06520; [b]European Molecular Biology Laboratory, 69117 Heidelberg, Germany; [c]European Molecular Biology Laboratory (EMBL) Outstation Hinxton, EMBL-European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, United Kingdom; [d]Medical Genetics Institute, Cedars–Sinai Medical Center, Los Angeles, CA 90048; [g]Department of Statistics, Yale University, New Haven, CT 06520; Departments of [h]Human Genetics, and [i]Biomathematics, University of California, Los Angeles, CA 90095; Departments of [j]Pediatrics and [k]Genetics, University of North Carolina, Chapel Hill, NC 27599; [l]Center for Genetic Testing, Tulsa, OK 74136; [m]Division of Medical Genetics, Department of Pediatrics, Loma Linda University School of Medicine, Loma Linda, CA 92354; [n]Department of Genetics, Kinderzentrum Munich, 81377 Munich, Germany; [o]Department of Pediatrics and Neurology, Irvine Medical Center, University of California, Orange, CA 92868; [p]Institute of Human Genetics, Western Galilee Hospital, Nahariya 22100, Israel; [q]Section of Clinical Genetics, Department of Pediatrics, University of Utah, Salt Lake City, UT 84132; [r]Department of Medical Genetics, Women's Health Centre, University of British Columbia, Vancouver, BC, Canada V6H 3V5; [s]Section of Clinical Genetics, Department of Pediatrics, Dartmouth–Hitchcock Medical Center, Lebanon, NH 03756; [t]Department of Pediatrics, Section on Medical Genetics, Wake Forest University School of Medicine, Winston-Salem, NC 27157; [u]Child Development Center, Rhode Island Hospital, Providence RI 02902; [v]Department of Pathology, University of North Carolina, Chapel Hill, NC 27514; [w]Signature Genomic Laboratories, Spokane, WA 99207; [x]Recanati Institute for Medical Genetics, Rabin Medical Center, Tel Aviv University, Petah Tikva, 49100 Israel; [y]Genetics Department, Kaiser Permanente Medical Offices, Sacramento, CA 95815; [z]Department of Genetics and Development, Columbia University, New York, NY 10032; [aa]Children's Hospital of New York, New York, NY 10032; and [bb]Department of Pediatrics, Center for Integrated Neurosciences and Human Behavior, The Brain Institute, University of Utah, Salt Lake City, UT 84108

Edited by Joseph R. Ecker, The Salk Institute for Biological Studies, La Jolla, CA, and approved May 29, 2009 (received for review December 30, 2008)

Down syndrome (DS), or trisomy 21, is a common disorder associated with several complex clinical phenotypes. Although several hypotheses have been put forward, it is unclear as to whether particular gene loci on chromosome 21 (HSA21) are sufficient to cause DS and its associated features. Here we present a high-resolution genetic map of DS phenotypes based on an analysis of 30 subjects carrying rare segmental trisomies of various regions of HSA21. By using state-of-the-art genomics technologies we mapped segmental trisomies at exon-level resolution and identified discrete regions of 1.8–16.3 Mb likely to be involved in the development of 8 DS phenotypes, 4 of which are congenital malformations, including acute megakaryocytic leukemia, transient myeloproliferative disorder, Hirschsprung disease, duodenal stenosis, imperforate anus, severe mental retardation, DS-Alzheimer Disease, and DS-specific congenital heart disease (DSCHD). Our DS-phenotypic maps located DSCHD to a <2-Mb interval. Furthermore, the map enabled us to present evidence against the necessary involvement of other loci as well as specific hypotheses that have been put forward in relation to the etiology of DS—i.e., the presence of a single DS consensus region and the sufficiency of *DSCR1* and *DYRK1A*, or *APP*, in causing several severe DS phenotypes. Our study demonstrates the value of combining advanced genomics with cohorts of rare patients for studying DS, a prototype for the role of copy-number variation in complex disease.

copy number variants | genomic structural variation | human genome | congenital heart disease | leukemia

For over two decades trisomy 21 has represented a prototype disorder for the study of human aneuploidy and copy-number variation (1, 2), but the genes responsible for most Down syndrome (DS) phenotypes are still unknown. The analysis of several overlapping segmental trisomies 21 has led to the suggestion that dosage alteration through duplication of an extended region on chromosome 21 (HSA21) is associated with DS features (2–5, 42). However, humans with segmental trisomy 21 are rare, and thus human-based DS-phenotypic maps have been of low resolution, far beyond the level of few or single genes, or even exons. Consequently, gene–disease links have often been based on indirect evidence from cellular or animal models (6, 7). Moreover, current hypotheses argue for the existence of a critical region, the DS consensus region (DSCR), responsible for most severe DS features (6, 8, 9), or presume the causative role of a small set of genes including *DSCR1* and *DYRK1A*, or *APP*, for these phenotypes (6, 7).

By using state-of-the-art genomics together with a large panel of partially trisomic individuals, we present the highest resolution DS phenotype map to date and identify distinct genomic regions that likely contribute to the manifestation of 8 DS features. Four of these phenotypes have never been associated with a particular region of HSA21. The map also enables us to rule out the necessary contribution of other HSA21 regions, thus providing strong evidence against the existence of a single DSCR, and a lack of support for the necessary synergistic roles of *DSCR1* and *DYRK1A*, or *APP*, as predominant contributors to many DS phenotypes.

Results and Discussion

To construct a high-resolution map of DS we assembled a panel of 30 individuals with rare, segmental trisomies 21 whose clinical features are summarized in Table 1. Nine patients are described for the first time, 16 were reassessed with respect to phenotype,

Author contributions: J.O.K., T.T.-W., A.E.U., S.W., M. Snyder, and J.R.K. designed research; J.O.K., T.T.-W., A.E.U., X.-N.C., M.K., L.D., F.G., M.C.G., K.L., E.M.S., G.M.B., A.S.A., N.J.C., R.D.C., M.Y.C., E.D., T.F.-Z., S.O.L., I.T.L., B.C.M., J.B.M., M.J.P., S.M.P., K.W.R., L.G.S., M. Shohat, A.J.V.R., D.W., M.B.G., M. Snyder, and J.R.K. performed research; C.E., K.L., and E.M.S. contributed new reagents/analytic tools; J.O.K., T.T.-W., A.E.U., M. Snyder, and J.R.K. analyzed data; and J.O.K., T.T.-W., A.E.U., M. Snyder, and J.R.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]J.O.K, T.T.-W., and A.E.U. contributed equally to this work.

[2]To whom correspondence may be addressed. E-mail: michael.snyder@yale.edu or julie.korenberg@hsc.utah.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0813248106/DCSupplemental.

Table 1. Major clinical features of patients

Feature	JL	JG	GY	IS	MI	WB	JS	KG	GP	KJ	BS	HOU	WS	MJF	WA	DS	SOS	SW	MB	ZSC	JJS	STO	SOL	JSB	NA	NO	BA	SM	HAD	PM
DSCHD	–	–	–	–	–	–	PDA†	–	–	–	AVSD*†	ASD*	–	–	AVSD§	–	+	TOF*†‡	–	TOF*†‡	MI†	VSD*	VSD*†‡	–	ASD	–	AVSD	PS†	AVSD	VSD*†‡
AMKL	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
TMD	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–
HSCR	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+*	+*	+*	–	–	–	–	–	–	–	–			
DST	–	–	–	–	–	–	+*	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–			
IA	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+*	–	–	–	–	–	–	–	–
MR	<30	P	59	Nor	M	M	M	M-P	85	70	52	+		47		P	M		M		69		P	M	42	P	P	52		

Numbers represent full scale IQ test. +, phenotype present; –, phenotype not present; AVSD, atrioventricular septal defect; PDA, patent ductus arteriosus; ASD, atrial septal defect; TOF, tetralogy of Fallot; MI, mitral insufficiency; VSD, ventricular septal defect; PS, pulmonic stenosis; blank, no information available; P, profound MR; M, moderate MR.

*Surgically repaired.

†Determined by echocardiogram.

‡Determined by cardiac catheterization.

§Determined by autopsy.

molecular cytogenetics, and breakpoints, and 5 were previously published (2, 10). We note that the ability to parse DS genes to phenotypes is determined by 3 factors: (*i*) phenotypic resolution, determined by the high-risk ratio of a given feature for DS versus normal and by the number of affected individuals; (*ii*) molecular resolution determined by technologies for identifying breakpoints and duplicated regions; and (*iii*) map resolution, determined by the density and locations of breakpoint positions in the patient panel. The high-risk ratio for a feature in DS suggests that a gene (or several genes) on HSA21 independently contribute(s) to the risk. The existence of segmental trisomies associated with the phenotype suggests that the gene(s) can be narrowed by analyzing cases displaying the phenotype, buttressed by cases without the region and without the phenotype. The candidate region for a DS phenotype is thereby determined by all 3 factors above, as formally described below and in the *SI Appendix*. With a larger number of cases [and as seen for DS-specific congenital heart disease (DSCHD) in our panel], the likelihood that a gene region contributes to the phenotypic variance is reflected by classical genetic concepts, namely, for all cases taken together, the duplication of the given region appears to generate the same penetrance (proportion of cases with the phenotype) and expressivity (variation of the phenotype, such as types of congenital heart disease) as is seen in full trisomy 21.

Specifically, we mapped HSA21 chromosomal rearrangements across patients by successively and systematically applying several technologies that interrogate HSA21 at increasingly high resolution (see e.g., Fig. 1; details in relation to the approaches used are in *Materials and Methods*). Initially, patients' karyotypes were ascertained through standard cytogenetic analyses. Then, breakpoint information and the orientation of segments were refined by using multicolor fluorescence in situ hybridization (FISH) with a panel of 350 bacterial artificial chromosomes and P1 artificial chromosomes (BAC/PACs). Furthermore, higher-resolution breakpoint mapping was achieved by using quantitative Southern blot analysis of single copy fragments in patient and parental control DNAs (3, 11, 12). Finally, to map chromosomal breakpoints at the level of single exons, we used high-density isothermal oligonucleotide DNA tiling arrays (13) with 380,000 probes that interrogate HSA21 at a median density <100 bp, i.e., in an unbiased fashion, providing an effective resolution <300 bp (13, 14). In general the tiling array results agreed well with the FISH, BAC/PAC, and Southern blot results (see Fig. 2), providing two or more independent determinations of copy number throughout the chromosome. However, the key significance of high-resolution breakpoint-mapping for linking phe-

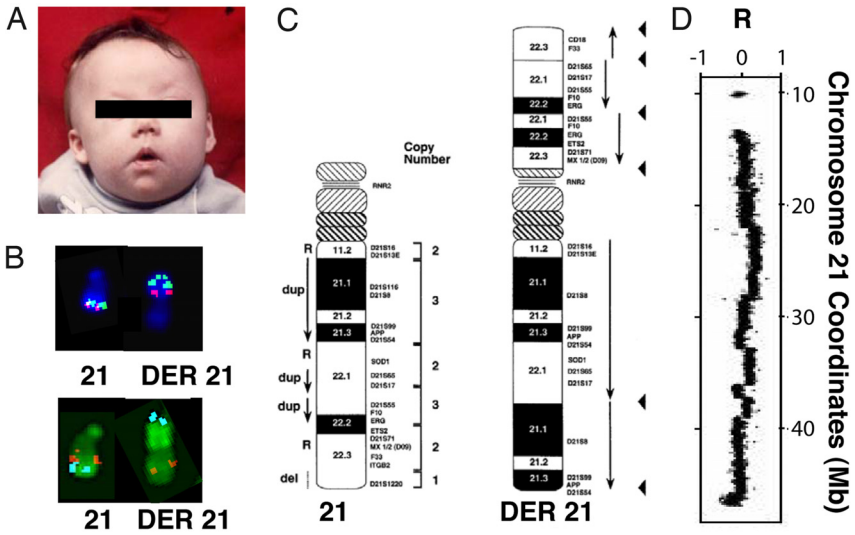

Fig. 1. High-resolution analysis of subject Dup21WB. (*A*) Photograph (age, 4 mos). (*B*) BAC-FISH analysis indicates duplication of *D21S55* (blue) but not *MX1/2* (fuchsia). 21, normal HSA21; DER21, rearranged chromosome. The orientation of nonduplicated segments is indicated in red (*SOD1*) and blue (*CD18/ITBG2*). (*C*) Summary of duplication and orientation of segments detected by using 19 FISH and 24 molecular markers. (*D*) HR-CGH analysis. Displayed are $\log_2$-ratios measured for Dup21WB relative to the control pool along HSA21.

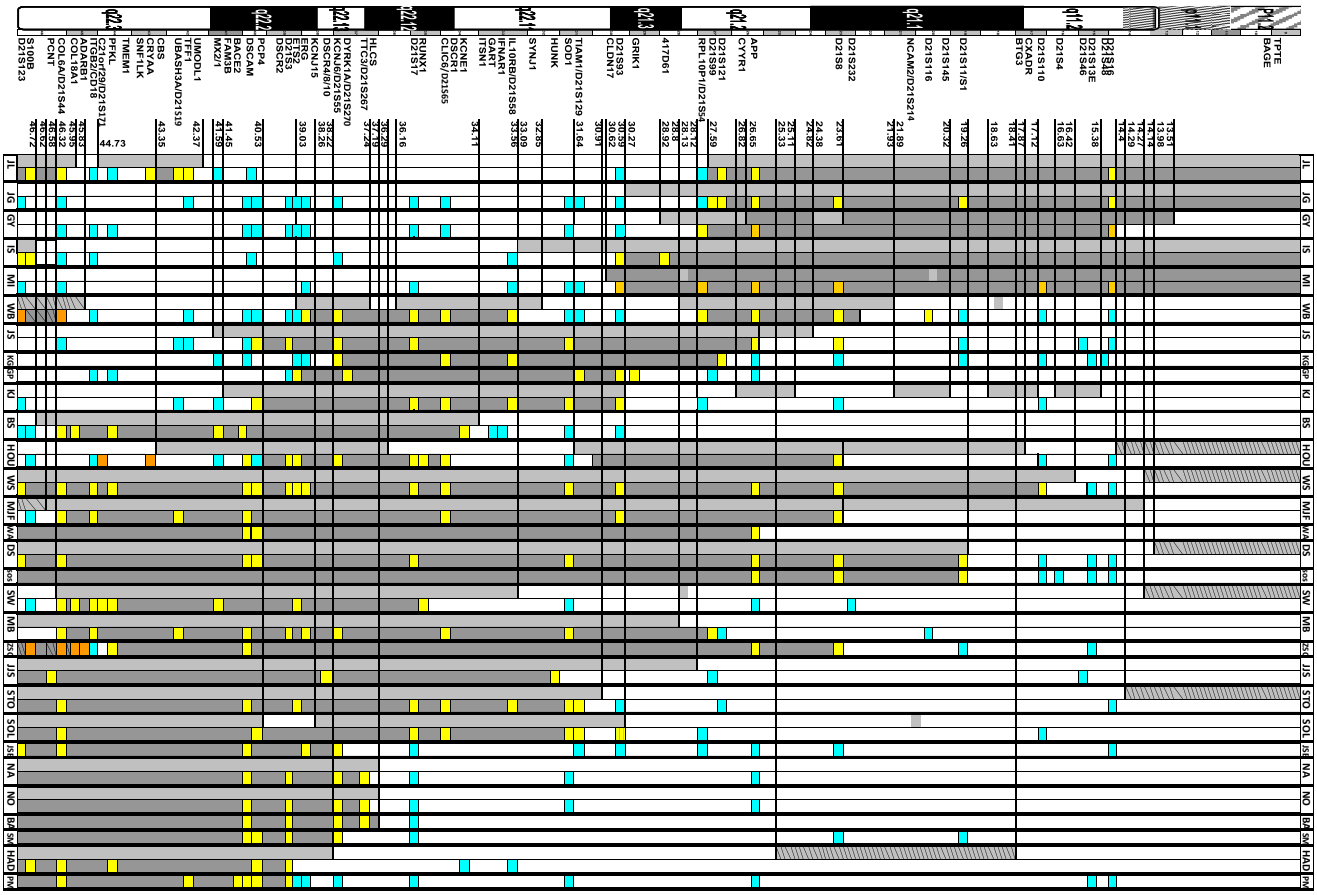

Fig. 2. Segmental 21 trisomy map for 30 individuals. Patient-IDs are displayed at the top and bottom. Columns to the right, or the only column indicated, represent FISH/Southern results (Table S6 in the *SI Appendix*, karyotypic features summarized in Table S7 in the *SI Appendix*). Copy-numbers are indicated by color-coding: orange, 1:2; blue, 2:2; Yellow, 3:2; Dark gray, 3:2. Dark gray with black lines, 1:2. The left columns depict oligonucleotide microarray results (see also Table S1 and Fig. S4 in the *SI Appendix*). Light gray, 3:2. Medium gray, 4:2. Light gray with black lines, 1:2.

notype and genotype is 2-fold, illustrated by the analysis of HSA21 in patient Dup21SOL, where the tiling microarrays detected a small gap in the duplicated region, unsuspected by the lower level of resolution afforded by the previous BAC/PAC analyses, thus eliminating the genes in the gap as candidates for his features. The microarrays thereby enabled substantially narrowing the region associated with DSCHD (see below). Furthermore, in patient Dup21NA, the tiling microarrays extend the Southern and BAC/PAC data to include *DYRK1A* but not *DSCR1* in the duplicated region and thereby argue against the necessary synergy of *DSCR1* and *DYRK1A* for producing the DS features found in this family.

We combined the results from all experimental analyses into the breakpoint map shown in Fig. 2. For many individuals the map displays subtle translocations and internal rearrangements causing duplications and deletions of HSA21 regions. Interestingly, for several patients, additional copies of several noncontiguous regions were detected by the tiling arrays, indicating that complex rearrangements involving multiple events had occurred (Table S1 in *SI Appendix*). For instance, Dup21WB displayed 5 obligatory breaks resulting in a mosaic of inverted and direct duplications and deletions (Fig. 1). For 3 individuals (Dup21GY, Dup21KJ, Dup21SOL), the segmental trisomic, tetrasomic, and monosomic regions as well as chromosomal breakpoints were validated by quantitative PCR (qPCR) (Fig. S1 in the *SI Appendix*). Furthermore, as we described in the *SI Appendix*, for Dup21JG and Dup21NA, sequencing the breakpoint junctions of partially trisomic regions (Fig. S2 in the *SI Appendix*) revealed that these rearrangements were likely caused by a microhomology-associated process, such as nonhomologous end-joining.

In addition to the large chromosomal variations, smaller regions [i.e., copy-number variants (CNVs) <500 Kb] of dosage alteration were observed (Table S2 in the *SI Appendix*). CNVs at this size-range are common in the normal population (15–18) and a substantial fraction of such events detected in this study overlapped with those found previously in unaffected individuals (Table S2 in the *SI Appendix*). Thus, their relationship to DS phenotypes is unknown. Unexpectedly, the array data for 4 individuals (Dup21JL, Dup21JG, Dup21IS, Tetra21MI) detected a copy-number increase of the HSA21 short arm region containing the gene TPTE and the pericentromeric long arm, whereas for 5 patients (Dup21HOU, Dup21WS, Dup21DS, Dup21SW, and Dup21STO; see Table S1 in the *SI Appendix*) the copy-number appears to be decreased. Homologous copies of the TPTE region are found on the acrocentric arms of 4 other chromosomes (12) and it remains unclear which chromosome(s) are affected. Nonetheless, our results indicate variation in copy-number of an expressed gene in acrocentric short arms, variability that may contribute to both normal and DS phenotypes.

Our breakpoint assignments also enable us to classify each segmental trisomy with respect to the genes/exons affected. As we describe in the *SI Appendix*, 20 of 59 breakpoints occur within annotated genes, leading to possible gene fragments/fusions with potentially aberrant function (Table S1 in the *SI Appendix*). Note that the relatively high portion of breakpoints within genes may suggest a role for transcription in the mechanisms for genome structural rearrangement.

GENETICS

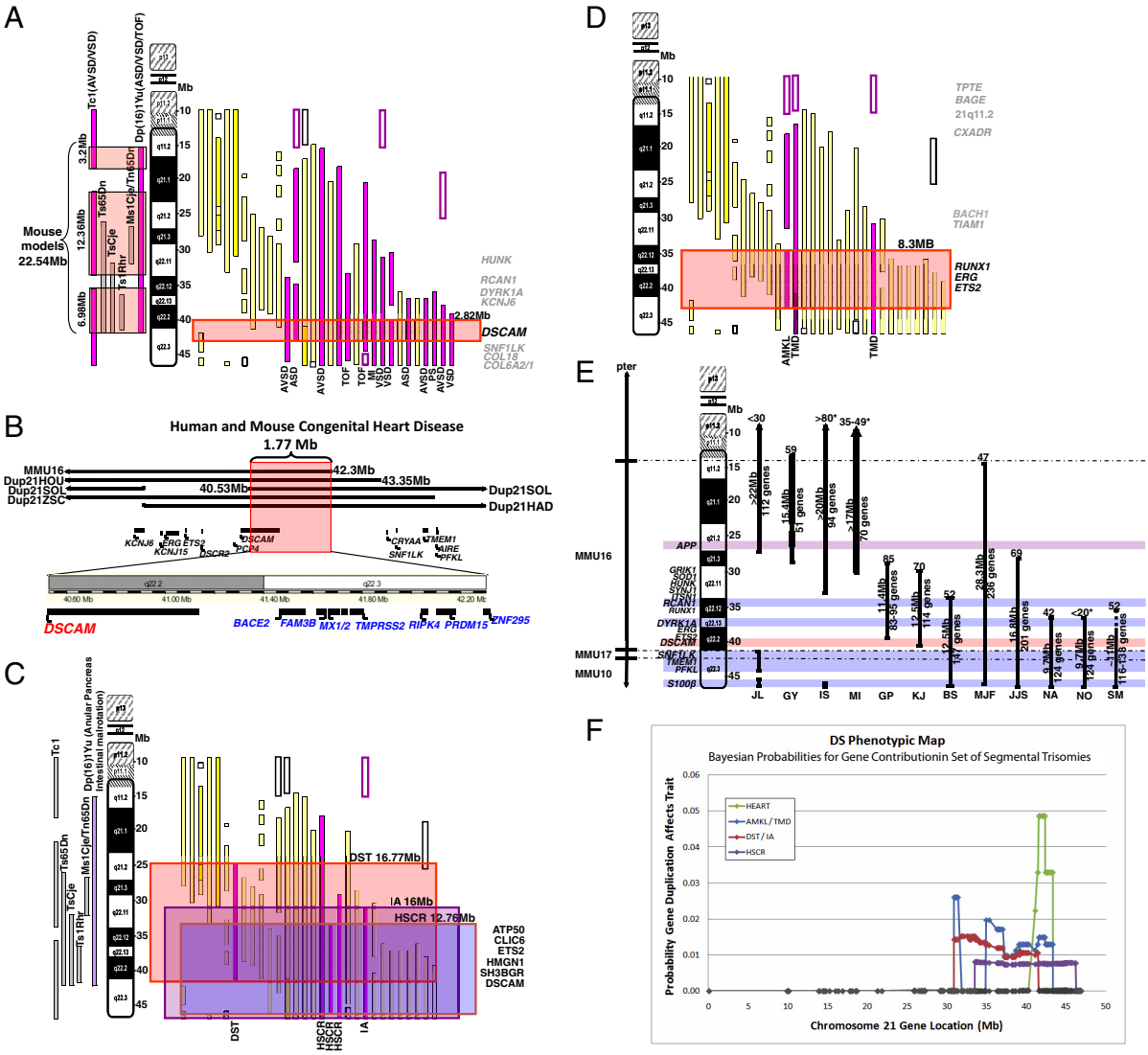

Fig. 3. Metaanalysis of DS phenotypes. A panel of 30 patients with segmental trisomy 21 defines DS phenotype candidate regions. Purple boxes, presence of phenotype; yellow boxes, no phenotype; solid boxes, increased copy-number; open boxes, 1:2 (monosomies). (*A*) DSCHD region. Red box, DSCHD candidate region. Twenty-three subjects have duplications including the DSCHD region, 14 thereof have DSCHD. No subject lacking a segmental trisomy involving the DSCHD critical regions was diagnosed with DSCHD. Corresponding regions for 6 mouse models are indicated to the left (21, 22, 39–41). TOF, tetralogy of Fallot; PS, pulmonic stenosis; PDA, patent ductus arteriosus; VSD, ventricular septal defect; ASD, atrial septal defect; MI, mitral insufficiency. (*B*) Proposed DSCHD critical region (red box) determined by combining human and mouse data from *A*. MMU16 indicates the extent of the duplication in the mouse model Dp (16)1Yu with DSCHD. (*C*) Gut region. Red box, DST candidate region; purple box, IA; orange box, HSCR. (*D*) AMKL and TMD. Red box, AMKL/TMD candidate region. (*E*) IQ in segmental trisomy 21. Vertical black bars at the right indicate the regions of segmental trisomy in individuals with IQs indicated above (thicker bars indicate regions of 4 copies). (*F*) Bayesian probability map of DS phenotype. The probabilities plotted are per gene or partial gene and therefore each data point is the posterior probability that 3 copies of the given gene cause the relevant phenotype. Thus, the posterior probability that a gene influencing a trait is within a larger chromosomal interval is the area under the curve for that interval. For example, for the DSCHD trait, the area under the curve from 41 Mb to 43.35 Mb is 0.9878, which is calculated by adding the values in Table S4C in the *SI Appendix* for each gene or partial gene located within the interval (see Table S3 in the *SI Appendix* for further examples). For AMKL/TMD, the first high-probability interval (30.9–31.4 Mb) corresponds to *KRTAP* genes and the 3′-end of *TIAM1* and is thus unlikely to be relevant to the phenotype.

To generate a DS phenotypic map, we used a Bayesian model and correlated the breakpoint information on segmental trisomies with the entire set of phenotypic features recorded for the patients. This approach provides evidence for the involvement of discrete genomic regions in the development of several DS phenotypes. Note that although some individuals contained translocations involving other chromosomes, map assignments were based on 3–14 individuals carrying aneuploidies of HSA21 only [as an exception, for duodenal stenosis (DST), acute megakaryocytic leukemia (AMKL), and imperforate anus (IA), only one individual was available, respectively]. Furthermore, Southern and FISH comparisons to parental DNAs confirmed that all detected duplications correspond to de novo events. Once a phenotype and the trisomic regions are well defined, an application of Bayes' theorem provides the probability that a given gene influences the trait; this approach exploits general penetrance rates provided by Torfs and Christianson (19) for full trisomy cases and normal controls. Fig. 3 and Table S3 in the *SI Appendix* depict the resulting phenotypic map of DS. This Bayesian approach converts prior knowledge of gene location into posterior probabilities via the likelihood of the observed data. The likelihood depends on the pattern of dosage alteration across patients. Here we treat each gene as a single locus on HSA21 that can influence each trait and assign a uniform prior probability to all HSA21 genes. The complete derivation of probabilities for the DS map is given in the *SI Appendix*.

We first analyzed DSCHD, a major phenotype associated with DS that is thought to derive from the abnormal development of the endocardial cushions and that results in a spectrum of defects involving the atrioventricular septum and valves. In DS, the risk of atrioventricular septal defect (AVSD) is ≈1,000-fold increased relative to that of the non-DS population (Table S4 in the *SI Appendix*) and DS accounts for 70% of all AVSDs (19). The overall risk of DSCHD in DS is 40–60%, of which approximately half are AVSDs (19) (*SI Appendix*). Although some candidate genes have been implicated for DSCHD (see below), conclusive evidence for their involvement is lacking. We have previously mapped the DSCHD region in humans to a 5.27-Mb chromosomal segment containing 82 genes (10). By using an expanded panel with 14 subjects with DSCHD, we have narrowed this segment to a 2.82-Mb critical region likely involved in DSCHD endocardial cushion defects (Fig. 3*A*). Interestingly, our results exclude a necessary role for a number of genes previously suggested to be critical for DSCHD, including *D21S55*/*KCNJ6* (5), *RCAN1* (20), Collagens 6A1/2 and 18 (10), and *DYRK1A* (6) (see *SI Appendix* and Fig. 3*A*).

By integrating our map with several other lines of evidence, in particular, information from segmental trisomic mouse models with DSCHD (21, 22), we further limit the region (Fig. 3*B*). Specifically, the model Dp (16)1Yu/+ displaying DSCHD involves only HSA21 regions orthologous to MMU16 (located at 14.4 Mb-42.3 Mb of HSA21), thereby defining the telomeric DSCHD border and suggesting lack of a necessary role of the adjacent telomeric region for DSCHD. Thus, we propose a 1.77-Mb DSCHD critical region, which contains 10 genes including the promoter and a portion of the Down syndrome cell adhesion molecule (*DSCAM*) gene (up to intron 11–12) that may result in an alternative transcript and defines the centromeric border. Of the genes in the region, only *DSCAM* is known to be highly expressed in the developing heart (Fig. S3 in the *SI Appendix*), implicating it as a likely candidate for causing DSCHD/AVSD.

We next focused on the congenital gut diseases associated with DS, i.e., Hirschsprung disease (HSCR), DST, and IA, which occur with 100-, 270-, and 30-fold increased risks, respectively, in DS relative to the general population (Table S4*A* in the *SI Appendix*) (19). The risk-ratio of HSCR in DS is known to be greater than the risk conferred by any of the single gene mutations for HSCR (23), none of which localize to HSA21. Our map (see Fig. 3*C*) suggests a discrete critical region <13 Mb that may be involved in this DS feature. The region contains a candidate gene, *DSCAM*, which was previously implicated in HSCR (24) through its predominant expression in brain, peripheral ganglia, and the developing neural crest of the gut (see also *SI Appendix*). Our results provide independent support for the involvement of *DSCAM* in HSCR. For the remaining 2 gut disorders, DST and IA, candidate genes have not been reported yet. Using single individuals from our panel for DST and IA respectively, we suggest regions that are likely involved in both features (see Fig. 3*C* and *SI Appendix*).

We then evaluated candidate genes for DS-associated leukemia. DS is associated with a 500-fold increased risk of transient myeloproliferative disorder (TMD) and a rare form of leukemia, AMKL. Previous studies have found that mutation/overexpression of *RUNX1*, *ERG*, *ETS2* (25, 26), and *TIAM1* (27) are associated with AMKL. Other indirect studies, using gene profiling of DS vs. non-DS AMKL, suggested that *CXADR* and *BACH1*, the downstream targets of GATA1, are associated with AMKL (28). We found a critical region of 8.35 Mb (35–43.35) that is likely contributing to the risk-increase for both TMD and AMKL (Fig. 3*D*). This region includes *RUNX1*, *ERG*, and *ETS*, but does not include *CXDR* and *BACH1*. Furthermore, it likely excludes a necessary role of *TIAM1* in AMKL because the breakpoint in the only DS patient with AMKL that was available to us deletes the promoter (see *SI Appendix*). Finally, an unexpected deletion of the HSA21 short arm observed in individuals with AMKL and TMD may also contribute to the phenotype (see *SI Appendix*).

Another DS phenotype addressed by our map is Alzheimer's disease (AD). Increased AD risk in DS has been linked to increased copy-number of the *APP* gene and genes nearby (29). Furthermore, genes in the vicinity of *APP* may act together to increase the risk of plaques and tangles in DS, and alternatively, genes elsewhere on HSA21 may protect against the neuronal damage incurred by increased *APP*. Although definitive conclusions will require neuropathology or other evidence of APP processing not found in DS, the trisomic region in Dup21JJS, a 65-year-old subject with DS who does not have dementia and has no amyloid accumulation by functional brain imaging, suggests the involvement in AD of a 1.95 Mb interval including *APP*. Further, as we point out in the *SI Appendix*, the duplicated region of subject JJS argues against essential roles of genes located distal to 28.12 Mb.

Finally, we suggest regions that may be involved in mental retardation (MR) (see Fig. 3*E*), a DS phenotype based on IQ, for which ascertainment of genes in DS patients remains challenging. Nevertheless, our map suggests that more than one MR critical region exists (see Fig. 3*E* and additional details presented as *SI Appendix*) and argues against an essential role of *APP* in DS-associated MR (7). Further studies of this panel, with neurocognitive and neural imaging tests focused on DS features, are needed to parse the contributions of HSA21 regions to brain development and function in DS.

The resolution of our map, which is equal to or better than that obtained by most standard linkage analysis studies, enabled us to evaluate specific hypotheses that have been put forward concerning the etiology of DS (Table S5 in the *SI Appendix*). In particular, our map rules out an essential role for several genes in specific DS phenotypes because the phenotype is observed in the absence of trisomy of the relevant region. For example, although both may contribute, our data do not support a necessary synergistic contribution to MR or DSCHD of the genes *DSCR1* and *DYRK1A* that were proposed to destabilize NFATC pathways (6). In a similar fashion, our study significantly limits the hypothesis of a coordinated role for the protein kinase genes *DYRK1A*, *HUNK*, and *SNF1LK*. Although these pathways may still contribute to DS cardiac valve defects, they are unlikely to play a central role in defects of the atrioventricular septae. Nonetheless, the current data may facilitate identifying unknown members of these pathways located in genomic regions telomeric to *DYRK1A*.

A second hypothesis that is challenged by our map relates to the essential role for *APP* in MR, although its proposed contribution to AD (7) is supported (Table S5 in the *SI Appendix*). Finally, our data are inconsistent with the hypothesis that the *DSCR1* and *DYRK1A* genes (6) form a critical region causing most DS features—indeed, several patients with severe DS features display segmental trisomies that do not include *DSCR1*. Thus, our results indicate that there is no DSCR, i.e., no single region of HSA21 responsible for all or most severe DS features.

The construction of our DS phenotype map required a large panel of well-characterized individuals with diverse duplicated regions spanning HSA21. Our map can be combined with other data, such as gene expression (43) or phenotypic data gathered in humans or mouse models, to refine candidate genes and define mechanisms involved in the etiology of DS.

Materials and Methods

Patient Examination. Procedures for human subjects and confidentiality were followed as approved by the Cedars–Sinai Medical Center Institutional Review Board. All data were taken from the original medical records and were confirmed by follow-up examinations or discussions with the family and patient.

Overview—Mapping Chromosomal Rearrangements. HSA21 genetic abnormalities were mapped with several approaches by using DNA and chromosomes generated from blood cells and cell lines. Patients' karyotypes were first ascertained through standard cytogenetic analyses. To determine breakpoints and orientations of duplicated chromosomal segments, all patient genomes were analyzed by FISH with subsets of a panel of 350 BAC/PACs, each of which was

GENETICS

validated for copy-number normals. This analysis was followed by a higher resolution mapping by using quantitative Southern blot analysis of single copy fragments in patient and parental control DNAs (3, 12, 30). To determine each breakpoint region, FISH used multiple BAC/PAC DNAs, labeled with combinations of FITC, Texas Red, Cy3, Cy5, and simultaneous hybridization to chromosome and interphase preparations. The chromosomal constitutions are available in Table S7 in the *SI Appendix*, and the raw data in Table S6 in the *SI Appendix*. The 30 individuals either had subtle translocations (Dup21JG, Dup21SOS, Dup21DS, Dup21JSB, Dup21NA, Dup21NO, and Dup21BA) or internal rearrangements, duplications, and deletions of HSA21 regions (Dup21JL, Dup21GY, Dup21IS, Dup21JS, Dup21KG, Dup21GP, Dup21KJ, Dup21BS, Dup21HOU, Dup21WA, Dup21SOL, Dup21SM, Dup21ZSC, Dup21WB). Further, of 2 individuals lacking the HSA21 telomeric 100-kb region, one carried an HSA21 region translocated to chromosome 1 (Dup21JG) validated by PCR and subsequent DNA sequencing. In some individuals (Dup21GY and Tetra21MI) 2 extra copies (segmental tetrasomy) were detected, and in others (Dup21WB, Dup21HOU, Dup21WS, Dup21MJF, Dup21DS, Dup21SW, Dup21STO, and Dup21HAD) only one copy of HSA21 regions (deletion) was detected. Tiling arrays were used for further breakpoint fine mapping.

Cytogenetic Analysis. Extended chromosome preparations were prepared from cultured peripheral lymphocytes by methotrexate synchronization (11). Chromosomes were stained by G-banding (GTG) and R-banding (RHG) techniques (11).

FISH and Quantitative Southern Blot Dosage Analysis. Large-fragment physical maps of human HSA21 were generated and integrated to define the copy-numbers and/or structural rearrangements (30–35). Sources and references for each DNA probe are in Table S6 in the *SI Appendix*. Probes were extracted from BAC, PAC, or cosmid clones and labeled by indirect or direct methods using a Nick Translation Kit (Invitrogen). Biotin-11-dUTP or Dig-11-dUTP (Sigma) were used for indirect labeling, and Alexa Fluor 488–5-dUTP, Alexa Fluor 568–5-dUTP, Alexa Fluor 594–5-dUTP, and Alexa Fluor 647–5-dCTP (Molecular Probes) were used for direct labeling (36). To define breakpoints we performed an initial screen followed by more specific testing with BACs progressively closer to the breakpoints. In each case, 10–30 copy-number determinations were performed, 2–4 BAC DNAs were hybridized simultaneously to target chromosomes, and 20–50 cells were analyzed. FISH and posthybridization detection were performed as in ref. 36. Multicolor images were captured with a Zeiss Axioplan 2 microscope equipped with an AxiocamMRMcamera and conjugated to Metasystem software (Microsoft). Procedures for DNA isolation and digestion, agarose gel construction, Southern blotting, probe labeling, hybridization, and autoradiogram development were conducted as in ref. 37. Southern blots used 8–12 paired lanes (16 to 24 totals) of patient and control DNAs. Densitometric analyses used the log transformation of density measurements from autoradiograms. All probes were isolated as DNA fragments for Southern blot procedures or as plasmids or cosmids for FISH studies. DNAs were obtained from peripheral blood, fibroblasts, or lymphoblastoid cell lines, with confirmed karyotype.

High-Density Oligonucleotide Tiling Array Analysis. DNA from cell lines or blood was used to probe a custom tiling array platform (Nimblegen Technology) by using high-resolution comparative genome hybridization [HR-CGH; (13)]; the isothermal array (13) contained 45–85-bp oligonucleotide probes. The array covered all regions of HSA21 to which probes could be uniquely mapped, including regions of the HSA21 short arm. The median probe distance was 90 bp, a density enabling breakpoint-mapping at 200–300-bp resolution (13, 14). DNA from the patients was labeled with Cy3 and hybridized to the array along with Cy5-labeled DNA from a reference pool of 7 individuals (healthy male individuals, Promega). Array normalization was performed with the Qspline algorithm (38). Copy-number changes were called by using the *BreakPtr* algorithm (with the ''core parameterization''; default parameters) as described in ref. 14. In the majority of cases a single array was used; in one case averaged signals from 2 arrays were used.

ACKNOWLEDGMENTS. We thank the individuals and families described in this article for their support and participation. Funding was provided by the European Union Sixth Framework Program (J.O.K.), the Save a Heart Foundation (T.T.W.), and grants from the National Institutes of Health (A.E.U., K.L., E.M.S., S.W., M.B.G., M.S.), the National Heart, Lung and Blood Institute (J.R.K.), and the Department of Energy (J.R.K.).

1. Epstein CJ (1986) *The Consequences Of Chromosome Imbalance Principles, Mechanisms, and Models*, eds Barlow PW, Green PB, Wylie CC (Cambridge Univ Press, New York), pp 253–323.
2. Korenberg JR, et al. (1994) Down syndrome phenotypes: The consequences of chromosomal imbalance. *Proc Natl Acad Sci USA* 91:4997–5001.
3. Korenberg JR, et al. (1990) Down syndrome: Toward a molecular definition of the phenotype. *Am J Med Genet Suppl* 7:91–97.
4. Niebuhr E (1974) Down's syndrome. The possibility of a pathogenetic segment on chromosome no. 21. *Humangenetik* 21:99–101.
5. Rahmani Z, et al. (1989) Critical role of the D21S55 region on chromosome 21 in the pathogenesis of Down syndrome. *Proc Natl Acad Sci USA* 86:5958–5962.
6. Arron JR, et al. (2006) NFAT dysregulation by increased dosage of DSCR1 and DYRK1A on chromosome 21. *Nature* 441:595–600.
7. Belichenko PV, Kleschevnikov AM, Salehi A, Epstein CJ, Mobley WC (2007) Synaptic and cognitive abnormalities in mouse models of Down syndrome: Exploring genotype-phenotype relationships. *J Comp Neurol* 504:329–345.
8. Delabar JM, et al. (1993) Molecular mapping of twenty-four features of Down syndrome on chromosome 21. *Eur J Hum Genet* 1:114–124.
9. Ronan A, et al. (2007) Familial 4.3 Mb duplication of 21q22 sheds new light on the Down syndrome critical region. *J Med Genet* 44:448–451.
10. Barlow GM, et al. (2001) Down syndrome congenital heart disease: A narrowed region and a candidate gene. *Genet Med* 3:91–101.
11. Korenberg JR, Chen XN (1995) Human cDNA mapping using a high-resolution R-banding technique and fluorescence in situ hybridization. *Cytogenet Cell Genet* 69:196–200.
12. Korenberg JR, et al. (1999) Human genome anatomy: BACs integrating the genetic and cytogenetic maps for bridging genome and biomedicine. *Genome Res* 9:994–1001.
13. Urban AE, et al. (2006) High-resolution mapping of DNA copy alterations in human chromosome 22 using high-density tiling oligonucleotide arrays. *Proc Natl Acad Sci USA* 103:4534–4539.
14. Korbel JO, et al. (2007) Systematic prediction and validation of breakpoints associated with copy-number variants in the human genome. *Proc Natl Acad Sci USA* 104:10110–10115.
15. Iafrate AJ, et al. (2004) Detection of large-scale variation in the human genome. *Nat Genet* 36:949–951.
16. Korbel JO, et al. (2007) Paired-end mapping reveals extensive structural variation in the human genome. *Science* 318:420–426.
17. Redon R, et al. (2006) Global variation in copy number in the human genome. *Nature* 444:444–454.
18. Sebat J, et al. (2004) Large-scale copy number polymorphism in the human genome. *Science* 305:525–528.
19. Torfs CP, Christianson RE (1998) Anomalies in Down syndrome individuals in a large population-based registry. *Am J Med Genet* 77:431–438.
20. Lange AW, Molkentin JD, Yutzey KE (2004) DSCR1 gene expression is dependent on NFATc1 during cardiac valve formation and colocalizes with anomalous organ development in trisomy 16 mice. *Dev Biol* 266:346–360.
21. Li Z, et al. (2007) Duplication of the entire 22.9 Mb human chromosome 21 syntenic region on mouse chromosome 16 causes cardiovascular and gastrointestinal abnormalities. *Hum Mol Genet* 16:1359–1366.
22. O'Doherty A, et al. (2005) An aneuploid mouse strain carrying human chromosome 21 with Down syndrome phenotypes. *Science* 309:2033–2037.
23. Gabriel SB,et al. (2002) Segregation at three loci explains familial and population risk in Hirschsprung disease. *Nat Genet* 31:89–93.
24. Yamakawa K, et al. (1998) DSCAM: A novel member of the immunoglobulin superfamily maps in a Down syndrome region and is involved in the development of the nervous system. *Hum Mol Genet* 7:227–237.
25. Levanon D, et al. (2001) Architecture and anatomy of the genomic locus encoding the human leukemia-associated transcription factor RUNX1/AML1. *Gene* 262:23–33.
26. Rainis L, et al. (2005) The proto-oncogene ERG in megakaryoblastic leukemias. *Cancer Res* 65:7596–7602.
27. Ives JH, et al. (1998) Increased levels of a chromosome 21-encoded tumour invasion and metastasis factor (TIAM1) mRNA in bone marrow of Down syndrome children during the acute phase of AML(M7). *Genes Chromosomes Cancer* 23:61–66.
28. Bourquin JP, et al. (2006) Identification of distinct molecular phenotypes in acute megakaryoblastic leukemia by gene expression profiling. *Proc Natl Acad Sci USA* 103:3339–3344.
29. Rovelet-Lecrux A, et al. (2006) APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. *Nat Genet* 38:24–26.
30. Korenberg JR, et al. (1995) A high-fidelity physical map of human chromosome 21q in yeast artificial chromosomes. *Genome Res* 5:427–443.
31. Cox DR, Shimizu N (1990) Report of the committee on the genetic constitution of chromosome 21. *Cytogenet Cell Genet* 55:235–244.
32. Gardiner K, Davisson M (2000) The sequence of human chromosome 21 and implications for research into Down syndrome. *Genome Biol* 1:REVIEWS0002.
33. Korenberg JR, Pulst SM, Neve RL, West R (1989) The Alzheimer amyloid precursor protein maps to human chromosome 21 bands q21.105-q21.05. *Genomics* 5:124–127.
34. Owen MJ, James LA, Hardy JA, Williamson R, Goate AM (1990) Physical mapping around the Alzheimer disease locus on the proximal long arm of chromosome 21. *Am J Hum Genet* 46:316–322.
35. Hubert RS, et al. (1997) BAC and PAC contigs covering 3.5 Mb of the Down syndrome congenital heart disease region between D21S55 and MX1 on chromosome 21. *Genomics* 41:218–226.
36. Korenberg JR, Yang-Feng T, Schreck R, Chen XN (1992) Using fluorescence in situ hybridization (FISH) in genome mapping. *Trends Biotechnol* 10:27–32.
37. Korenberg JR, Bradley C, Disteche CM (1992) Down syndrome: Molecular mapping of the congenital heart disease and duodenal stenosis. *Am J Hum Genet* 50:294–302.
38. Workman C, et al. (2002) A new non-linear normalization method for reducing variability in DNA microarray experiments. *Genome Biol* 3:research0048.
39. Davisson MT, Schmidt C, Akeson EC (1990) Segmental trisomy of murine chromosome 16: A new model system for studying Down syndrome. *Prog Clin Biol Res* 360:263–280.
40. Olson LE, Richtsmeier JT, Leszl J, Reeves RH (2004) A chromosome 21 critical region does not cause specific Down syndrome phenotypes. *Science* 306:687–690.
41. Sago H, et al. (1998) Ts1Cje, a partial trisomy 16 mouse model for Down syndrome, exhibits learning and behavioral abnormalities. *Proc Natl Acad Sci USA* 95:6256–6261.
42. Lyle R, et al. (2009) Genotype-phenotype correlations in Down syndrome indentified by array CGH in 30 cases of partial trisomy and partial monosomy chromosome 21. *Eur J Hum Genet* 17:454–466.
43. Sultan M, et al. (2007) Gene expression variation in Down's syndrome mice allows prioritization of candidate genes. *Genome Biol* 8:research0091.