

A Genome Model to Explain Major Features of Neurodevelopmental Disorders in Newborns

Bernard Friedenson 

Department of Biochemistry and Molecular Genetics, College of Medicine, The University of Illinois at Chicago, Chicago, IL, USA.

Biomedical Informatics Insights
Volume 11: 1–19
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1178222619863369



ABSTRACT: The purpose of this study was to test the hypothesis that infections are linked to chromosomal anomalies that cause neurodevelopmental disorders. In children with disorders in the development of their nervous systems, chromosome anomalies known to cause these disorders were compared with foreign DNAs, including known teratogens. Genes essential for neurons, lymphatic drainage, immunity, circulation, angiogenesis, cell barriers, structure, epigenetic and chromatin modifications were all found close together in polyfunctional clusters that were deleted or rearranged in neurodevelopmental disorders. In some patients, epigenetic driver mutations also changed access to large chromosome segments. These changes account for immune, circulatory, and structural deficits that accompany neurologic deficits. Specific and repetitive human DNA encompassing large deletions matched infections and passed rigorous artifact tests. Deletions of up to millions of bases accompanied infection-matching sequences and caused massive changes in human homologies to foreign DNAs. In data from 3 independent studies of private, familial, and recurrent chromosomal rearrangements, massive changes in homologous microbiomes were found and may drive rearrangements and encourage pathogens. At least 1 chromosomal anomaly was found to consist of human DNA fragments with a gap that corresponded to a piece of integrated foreign DNA. Microbial DNAs that match repetitive or specific human DNA segments are thus proposed to interfere with the epigenome and highly active recombination during meiosis, driven by massive changes in human DNA-foreign DNA homologies. Abnormal recombination in gametes produces zygotes containing rare chromosome anomalies that cause neurologic disorders and nonneurologic signs. Neurodevelopmental disorders may be examples of assault on the human genome by foreign DNAs at a critical stage. Some infections may be more likely tolerated because they resemble human DNA segments. Even rare developmental disorders can be screened for homology to infections within altered epigenomes and chromatin structures. Considering effects of foreign DNAs can assist prenatal and genetic counseling, diagnosis, prevention, and early intervention.

KEYWORDS: Genome, epigenetics, neurodevelopmental disorders, chromosome anomalies, retrotransposon, chromosome rearrangement, neurologic disease, birth defects, epigenome, infection

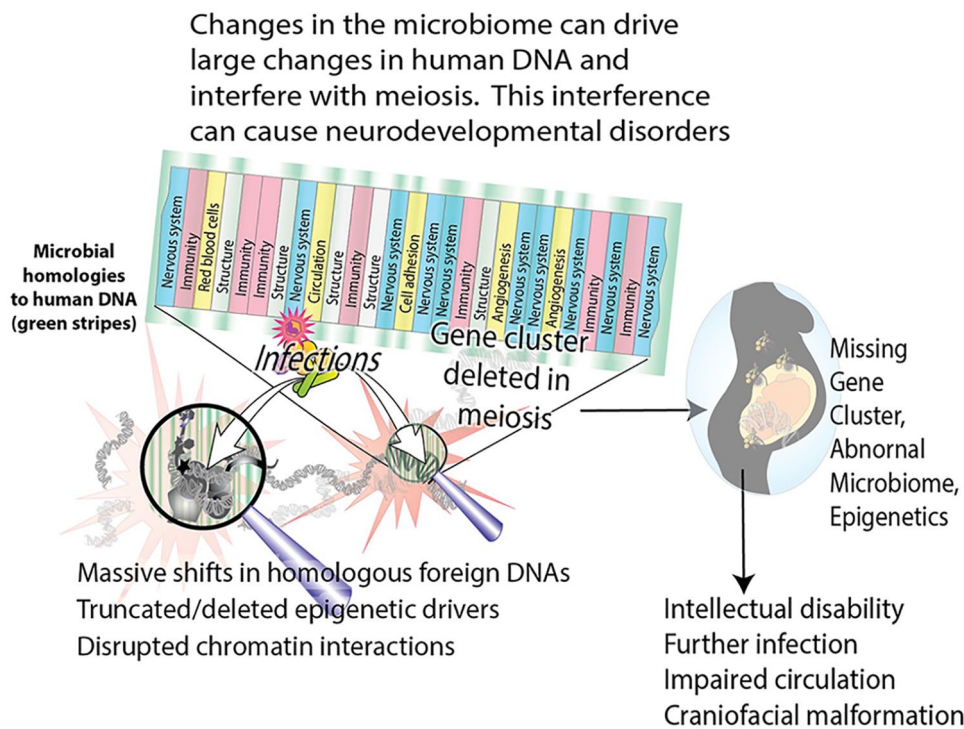
RECEIVED: June 21, 2019. **ACCEPTED:** June 21, 2019.

TYPE: Original Research

FUNDING: The author(s) received no financial support for the research, authorship, and/or publication of this article.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Bernard Friedenson, Department of Biochemistry and Molecular Genetics, College of Medicine, The University of Illinois at Chicago, Chicago, IL 60612, USA. Email: bernief@uic.edu



Introduction

An approach to preventing neurodevelopmental disorders is to gain better understanding of how neurodevelopment is coordinated and then to identify interference from environmental, genomic, and epigenomic factors. The development of the nervous system requires tight regulation and coordination of multiple functions essential to protect and nourish neurons. As the nervous system develops, the immune system, the circulatory system, and cranial and skeletal systems must all undergo synchronized and coordinated development. Neurodevelopmental disorders follow the disruption of this coordination.

A significant advance in genome sequence-level resolution of balanced cytogenetic abnormalities greatly improves the ability to document changes in regulation and dosage for genes critical for the neurologic system. Based on DNA sequence analyses, some chromosome rearrangements have been identified as causing individual congenital disorders because they disrupt genes essential for normal development.¹⁻³ There is poor understanding and no effective treatment for many of these overwhelming abnormalities. Signs and symptoms include autism, microcephaly, macrocephaly, behavioral problems, intellectual disability, tantrums, seizures, respiratory problems, spasticity, heart problems, hearing loss, and hallucinations.¹ Because the abnormalities do not correlate well with eventual outcome, genetic counseling is difficult and uncertain.³

In congenital neurologic disease, inheritance is usually autosomal dominant so the same chromosomal abnormalities occur in every cell. The genetic events that lead to most neurodevelopmental disorders are not understood⁴ but several maternal infections and other lifestyle factors are known to interfere.

DNA homology between non-humans and humans is a known fact, and DNA swapping between vertebrates and invertebrates has been reported. An early draft of the human genome found human genomes have undergone lateral gene transfer to incorporate microorganism genes.⁵ Lateral transfer from bacteria may have generated many candidate human genes.⁶ Genome-wide analyses in animals found up to hundreds of active genes generated by horizontal gene transfer. Fruit flies and nematodes have continually acquired foreign genes as they evolved. Although these transfers are thought to be rarer in primates and humans, at least 33 previously unreported examples of horizontally acquired genes were found.⁶ These findings argue that horizontal gene transfer continues to occur to a larger extent than previously thought. Transferred genes that survive have been largely concerned with metabolism and make important contributions to increasing biochemical diversity.⁷ A picture of humans as a large ecosystem of human and non-human DNAs is emerging.

The present work implicates foreign DNA largely from infections as a cause of the chromosome anomalies that cause birth defects. Infections replicate within the human central nervous system by taking advantage of immune deficiencies such as those traced back to deficient microRNA production⁸ or other gene damage. Disseminated infections can then interfere

with the highly active DNA break repair process required during meiosis. The generation of gametes by meiosis is the most active period of recombination, which occurs at chromatin positions enriched in epigenetic marks. Hundreds of double-strand breaks accompany meiotic recombination.⁹ Gametes with errors in how this recombination occurs cause chromosome anomalies in the zygote. In contrast to oocytes, meiotic recombination in sperm cells occurs continuously after puberty.

The exact DNA sequences of known pathogenic rearrangements in individual, familial, and recurrent congenital disorders¹⁻³ make it possible to test for association with foreign DNA. Even rare developmental disorders can be screened for homology to infections within altered epigenomes and chromatin structures. Considering effects of foreign DNAs can assist prenatal and genetic counseling, diagnosis, prevention, and early intervention.

The results showed that DNA abnormalities in some neurodevelopmental disorders closely match DNA in multiple infections that extend over long linear stretches of human DNA and often resemble repetitive human DNA sequences. Massive changes in the identity and distribution of sets of homologous alien DNAs that accompany chromosome abnormalities may drive and stabilize them. Removing competition by changing the sets of foreign DNAs may encourage pathogens.

The affected human sequences are shown to exist as linear clusters of genes closely spaced in 2 dimensions. Interference from infection can also delete or damage human gene clusters and change epigenetic functions that coordinate neurodevelopment. This microbial interference accounts for immune, circulatory, and structural deficits that accompany neurologic deficits.

Congenital neurodevelopmental disorders are thus viewed as resulting from an assault on human DNA by foreign DNA and perhaps subject to selection based on their similarity to host DNA. It is important to remember that effects in 2 dimensions can sometimes alter 3-dimensional topology as well.¹⁰

Testing and verifying predictions from a viable model may spur the development of methods for identifying contributions from infections in intractable rare disorders that are not now available. Convergent arguments from testing predictions based on any proposed model might lessen the effects of limitations on currently available technology.

Materials and Methods

Data sources

DNA sequences from acquired congenital disorders were from published whole genome sequences at chromosome breakpoints and rearrangement sites.¹⁻³ Comparison with multiple databases of microbial sequences determined whether there was significant human homology. Emphasis was on cases with strong evidence that a particular human chromosome rearrangement was pathologic for the congenital disorder. Patients in the 3 major studies¹⁻³ used in this analysis were 98 females and 144 males. Of the patients with background information available, most (46) were

younger than the age of 10, 7 were in the range of 10 to 20, 6 were in 20 to 40, and 1 patient was older than 40 years.

Testing for homology to microbial sequences

Hundreds of different private rearrangements in patients with different acquired congenital disorders were tested for homology¹¹ against nonhuman sequences from microorganisms known to infect humans as follows: Viruses (taxid: 10239), and retroviruses including HIV-1 (taxid: 11676), human endogenous retroviruses (taxids: 45617, 87786, 11745, 135201, 166122, 228277, and 35268); bacteria (taxid: 2); Mycobacteria (taxid: 85007); fungi (taxid: 4751), and chlamydias (taxid: 51291).

Because homologies represent interspecies similarities, “Discontinuous Megablast” was most frequently used, but long sequences were sometimes tested against highly similar microbial sequences. Significant homology (indicated by homology score) occurs when microbial and human DNA sequences have more similarity than expected by chance (E value $\leq e^{-10}$).¹² Confirmation of microorganism homologies was done by reverse testing multiple variants of complete microorganism genomes against human genomes and by extending the original analyses to 20 000 matches.

Various literature analyses have placed Alu repeats into 8 subfamilies having consensus sequences (GenBank; accession numbers U14567-U14574). Microbial sequences were independently compared with all 8 consensus Alu sequences and with 442 individual AluY sequences. Plots of chromosome locations of repetitive elements were from the UCSC genome browser (GrCH38)..

Chromosome localizations

The positions of microbial homologies in human chromosomes were determined using BLAT or BLAST. Comparisons were also made to cDNAs based on 107 186 Reference Sequence (RefSeq) RNAs derived from the genome sequence with varying levels of transcript or protein homology support. Tests for contamination by vector sequences in these nontemplated sequences were also carried out with the BLAST program. Inserted sequences were also compared with *Mus musculus* GRCM38.p4 [GCF_000001635.24] chromosome plus unplaced and unlocalized scaffolds (reference assembly in Annotation Release 106). Homology of inserted sequences to each other was tested using the Needleman and Wunsch algorithm. Lists of total homology scores for microbes vs human chromosome rearrangements were compared by the Mann-Whitney U test using the StatsDirect Statistical analysis program. Fisher test was also used.

Results

Interdependent functions are clustered together on the same chromosome segment

The nervous system has a close relationship to structures essential for immunity, circulation, cell barriers, and protective enclosures. Genes essential for all these functions must develop in

concert so chromosome segments deleted in neurodevelopmental disorders may be critical for this coordination. Genes for these related functions are located close to each other on the same linear segment of a chromosome (Figure 1). For example, deletions at 4q34 in patient DGAP161 are shown in Figure 1. The genes within the 4q34 deletion in each of the 4 categories tested are color coded in Figure 1. Figure 1 is representative of 6 other chromosome bands that were also tested and gave similar results: 2q24.3, 6q13-6q14.1, 10p14-10p15.1, 13q14.2, 18p11.22-p11.32, and 19q12-q13.1. Deletion of these clustered arrangements has been correlated with serious neurodevelopmental disorders.¹ Alternatively spliced forms of the same gene may encode for pleiotropic functions that must be synchronized and coordinated among diverse cells. Multiple functions for the same gene in different cell types are commonly found. Hormonal signaling represents a major control mechanism.¹³

Many deleted gene clusters include long stretches of DNA strongly related to foreign DNAs

To investigate the chromosomal segment deletions that likely cause neurodevelopmental disorders, homologies to infection were tested in sequences within and flanking deleted clusters. Strong homologies to infections were interspersed. To demonstrate the extent of these relationships, a deleted 4q34 chromosome segment (Patient DGAP161)¹ was tested for homology to microbes and then compared to repetitive elements in the same region.

Figure 2 shows that stretches of homology to microorganisms are distributed along a 500,000 bp segment at the 5' end of chromosome 4q34. For comparison, the distribution of repetitive elements is shown below the plots of homologous foreign DNAs.

Enormous effects on microbial homologies accompany changes in junction sequences between chromosome bands caused by deletions.

Figure 3 represents a snapshot of how local homologies to foreign DNAs shift when a chromosome segment is deleted. Figure 3 (top) represents 200,000 bps on each side of the normal 4q33-4q34 junction and its change after the 4q34 deletion to the new junction 4q33-4q35. Around the breakpoint and 3' to it, microbial homologues become very different. Although the representations of microbial homologies as red rectangles appear to be small on Figure 3 (top), the matching sequences actually extend for hundreds of base pairs. Normally, there are multiple homologies to a variety of microorganisms at the breakpoint (yellow rectangle) which may help destabilize the area. After the deletion, the junction formed now has new and strong homology to pathogens *Neisseria gonorrhoeae*, *Klebsiella pneumoniae*, *Clostridium botulinum*, and other potential pathogens. Homology to some isolates of γ -herpesvirus 4 (Epstein-Barr virus) extends into the newly juxtaposed 4q35 region. The deletion significantly changes the sets of homologous microorganisms (Mann-Whitney test, $P = .0089$). Both maximum and total homology scores show that microbial distributions become very different after the rearrangement ($P = .0178$).

Chromosome 4q34 gene cluster, deleted in patient DGAP161	
Genes	Developmental Functions
HMGB2	Controls neural stem cell proliferation
	V(D)J recombination DNA RNA sensor in Innate immune response to nucleic acids.
	Chemokine. Anti-microbial
	Red blood cell differentiation
	Connective tissue differentiation
SAP30	Hematopoiesis, anti-inflammatory response to glucocorticoids. Hypoxia response
SCRG1	Autophagy, Anti-inflammatory, chemotaxis for immune system cells, prion response Bone differentiation from stem cells
HAND2	Noradrenergic properties of neurons, controls neuroblast multiplication
	Development of circulation, cardiac development
	Craniofacial skeleton
HPGD	Pro-inflammatory gene
	Bone structure and connective tissue
GLRA3	Tonic and agonist induced currents in forebrain, auditory nerve activity, brain development, vision
ADAM29	Cell adhesion and membrane structure
GPM6A	Signaling in neuron lipid rafts. Dendritic spine and synapse formation. Candidate causal gene for schizophrenia
WDR17	Retinal development
SPATA4	Apoptosis inhibitor
	Osteoblast differentiation
ASB5	Vascular remodeling in embryogenesis, artery development
SPCS3	Preprohormone convertase affects pituitary hormone ghrelin
VEGFC	Neural stem cell activation
	Formation of lymphatics, Endothelial cell migration, Circulation
NEIL3	Induces neurogenesis
	B-cell expansion in germinal center, prevention of autoimmunity
AGA	Thalamus structure
	Aspartylglucosaminidase, lysosomal hydrolase, joint inflammation
TENM	Development in nervous system Synapse organization

Figure 1. Chromosome 4q34 as a typical example of close relationships between nervous system genes and genes for other essential developmental functions.

Pleiotropic genes for multiple interdependent systems appear in clusters on 4q34. Nervous system genes are near genes essential for the immune system, connections to lymphatic circulation, ability to form tight junctions, structural enclosures, and chromatin control. Clusters of genes encoding these and other interdependent functions on chromosome segments are deleted in private neurodevelopmental disorders. These losses increase susceptibility to infections that have DNA homologous to long stretches of human DNA. In the example shown, genes are listed in the order they occur on 4q34. Developmental functions are to the right of the gene symbol. Blue genes are associated with the nervous system, pink with the immune system. Yellow genes have functions associated with angiogenesis, or lymphangiogenesis or cell barriers. Genes for development of essential bone structure or connective tissues needed to protect and house the nervous system are light gray. Isoforms of the same gene may encode different functions in different cellular locations. Consistent results were obtained from deletions involving 6 other chromosome bands (see text).

Extending these analyses to test for 20 000 homologues further supports this conclusion (data not shown). Depending on which foreign DNAs exist in the microenvironment, the changes could easily cause a change in free energies to drive, direct, and guide the chromosome rearrangement. Changes in topological chromosome interactions are also likely.

Deleted segments in familial chromosome anomalies point toward a general mechanism for infection as a cause of neurodevelopmental disorders

Genome sequencing of an entire family may be necessary³ because some family members carry balanced chromosomal translocations but do not have neurodevelopmental disease. In

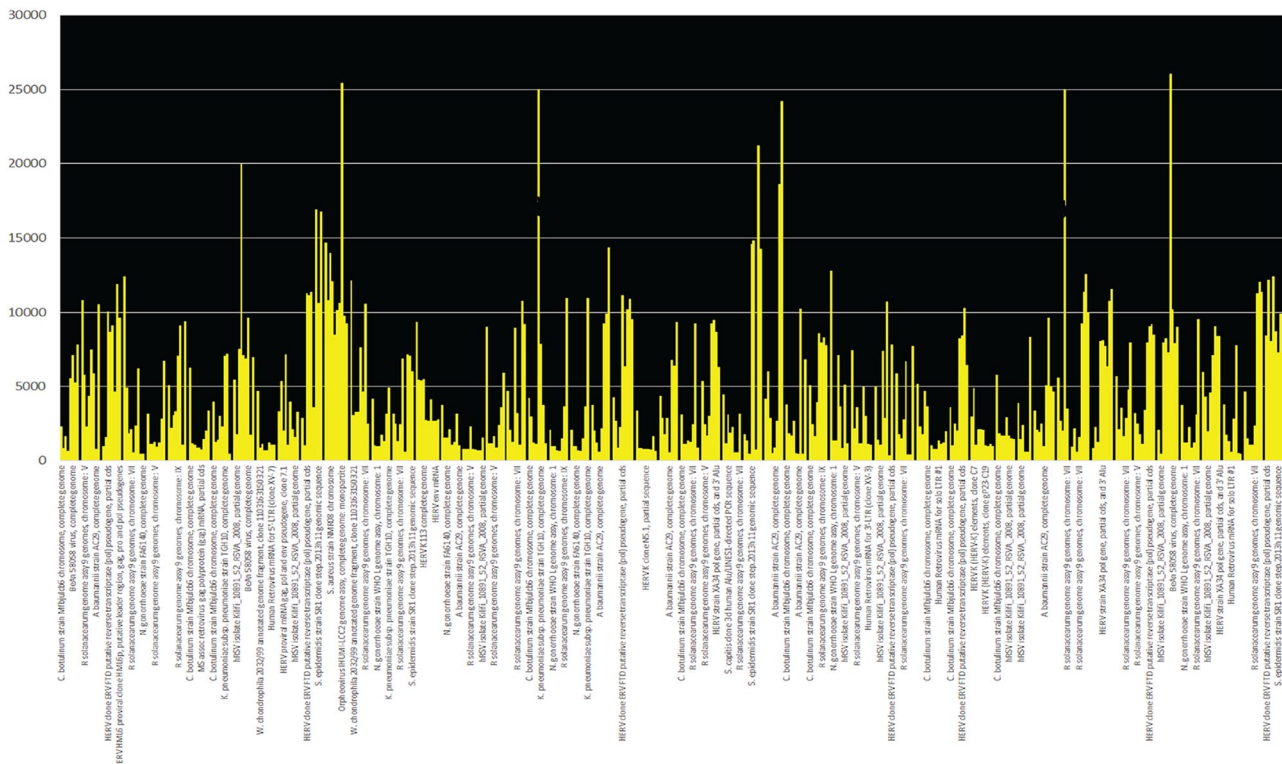


Figure 2. Example of total homologies to microbial sequences which are dispersed throughout the normal 4q34 chromosome segment deleted in patient DGAP161.

Homologies are shown for the 500,000bp segment at the 5' end of 4q34. Those with scores above 1000 are labeled. The y-axis plots the maximum homology scores ($E \leq 1 \times 10^{-10}$) and the blue dots indicate the 5' start point of significant homology scores. The horizontal axis gives the positions of homologies in bps along the 4q34 DNA sequence. Most of the highest homology scores involve human gammaherpesvirus 4 (EBV). Specific EBV isolates are indicated (HKNPC60 or HKHD40); the clostridium botulinum strain is also listed.

3 of the 4 families with familial balanced chromosomal translocations, patient-specific unbalanced deletions were found but the results did not overlap any database of human reference genomes.³ A disease-associated deletion in the study of Aristidou et al³ (family 2) was tested by comparing equivalent numbers of bps at the junction sequence created by the deletion vs the original chromosome junction sequence without the deletion (GRCh37: Chr16:49,741,265-49,760,865).

In Figure 4, the changes are enormous, involving new distributions (Mann-Whitney, $P < .0001$) and different homologous alien DNAs. Some microorganisms newly found in the rearranged sequence are known human pathogens. Their presence emphasizes how different the local homologous microbiome has become. This result was run many times and results were consistent if corrected or uncorrected (result shown) for low complexity human sequences and even if the test sets of sequences were varied somewhat. In agreement with results presented later in Table 1, critical genes linked to epigenetic modifications were among those interrupted by cryptic rearrangements.

Some chromosome regions with microbial homologies are only deleted in affected family members in families that share a recurrent translocation

Recurrent de novo translocations between chromosomes 11 and 22 have so far only been detected during spermatogenesis

and have been attributed to palindromic structures that induce genomic instability.² The recurrent breakpoint t(8;22)(q24.13;q11.21)² was tested to determine whether palindromic rearrangements might arise because infection interferes with normal chromatin structures.

Figure 5 shows strong homology to bacterial and viral sequences in a family with a recurrent translocation DNA sequences from an unaffected mother carry a balanced translocation rearrangement² with different homologies than are present in affected cases (Figure 5). The distributions of homologous microorganisms are clearly different for the unaffected mother vs affected Case 12 Der(8) (Mann-Whitney, $P = .0049$) and vs Case 13 (not shown). The rearrangement accompanies profound local changes in the homologous microbiome ($P = .0022$). Some prominent nonhuman DNAs that show large differences, eg, fungi and yeasts have been isolated from oral cavities of children.⁴¹

Microbial DNA homologies in areas around a mutated epigenetic driver gene

In some patients with neurodevelopmental disorders, a chromosomal anomaly disrupts a critical driver gene with strong evidence that the disrupted driver contributes to the disease.¹ The genes identified as underlying phenotypic drivers of congenital neurologic diseases include chromatin modifiers.¹ In

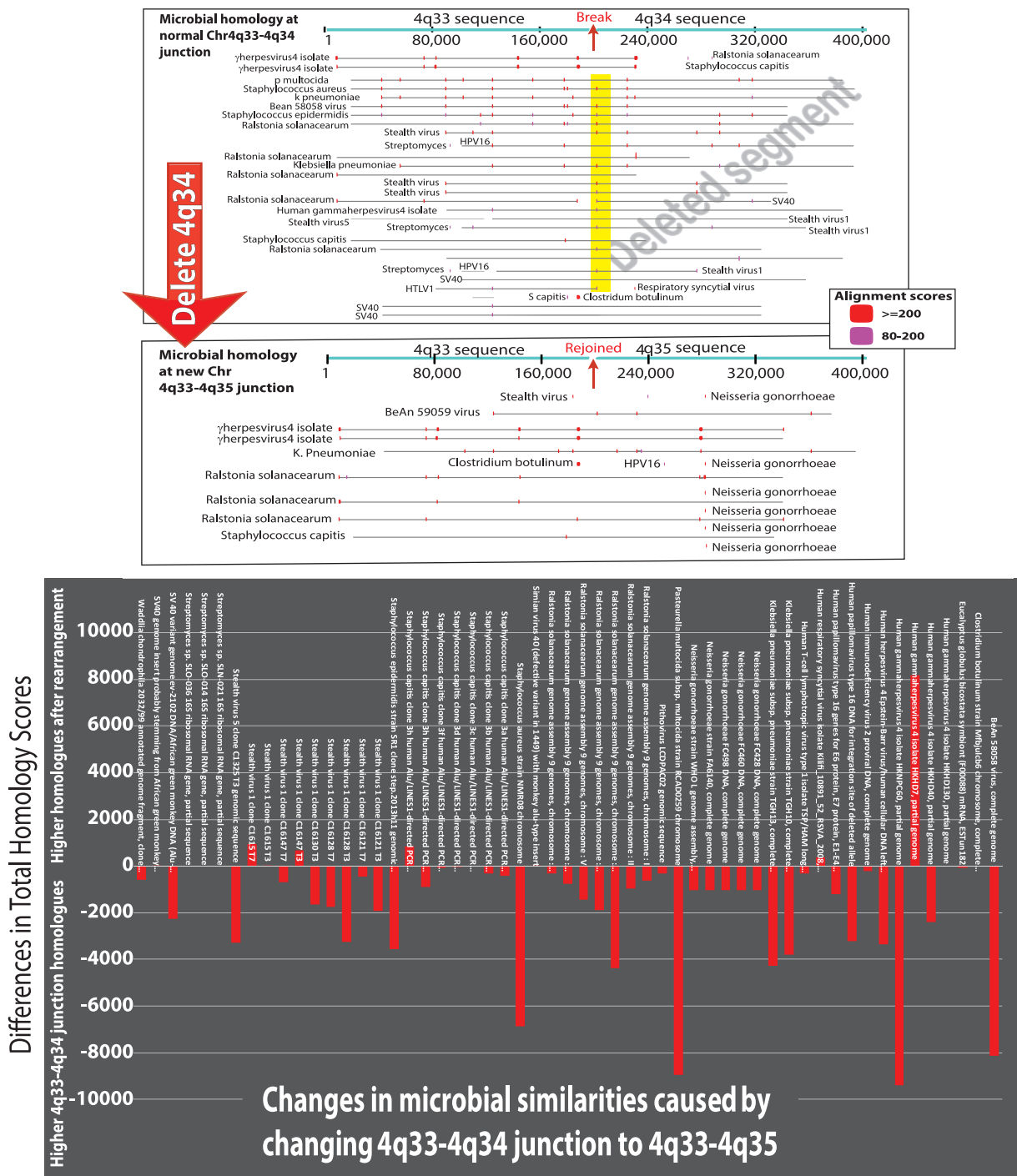


Figure 3. Snapshot of very large differences in local microbial homologies in one 200 kb section of chromosome 4q34. The same 200 kbp sections after vs before deletion of 4q34 are shown. Top: The 4q34 deletion changes the local microbiome homologous to human sequences in this region. The top panel shows the breakpoint of the deletion. There is homology to stealth viruses and bacteria near the breakpoint (yellow box). After the deletion occurs, homology to human gammaherpesvirus 4 (Epstein-Barr virus “EBV”) extends into 4q35 and there is now homology to pathogens: *Clostridium botulinum*, *Klebsiella pneumoniae*, and *Neisseria gonorrhoeae*. Bottom: In the rearranged 200,000 bp segment, very different microorganism DNAs have homology to the human chromosome segment and are shown above the x-axis. Including 20 000 foreign DNA sequences in the test gave the same conclusion.

agreement with this designation, Table 1 shows that most pathogenic driver genes are more specifically epigenetic factors (at least 45 of the 66 patients listed in Table 1). Using a value of 815 as a rough estimate of the total number of epigenetic factors in the human genome⁴² containing 20 000 genes, the probability that association between neurodevelopmental patients and epigenetic modifications occurs by chance is $P < .0001$.

Pathogenic driver gene mutations caused by large chromosome deletions amplify their effects because of epigenetics

Because most identified driver genes of neurodevelopmental disorders¹ are epigenetic factors (Table 1), the functions they control in individual patients and in families with members

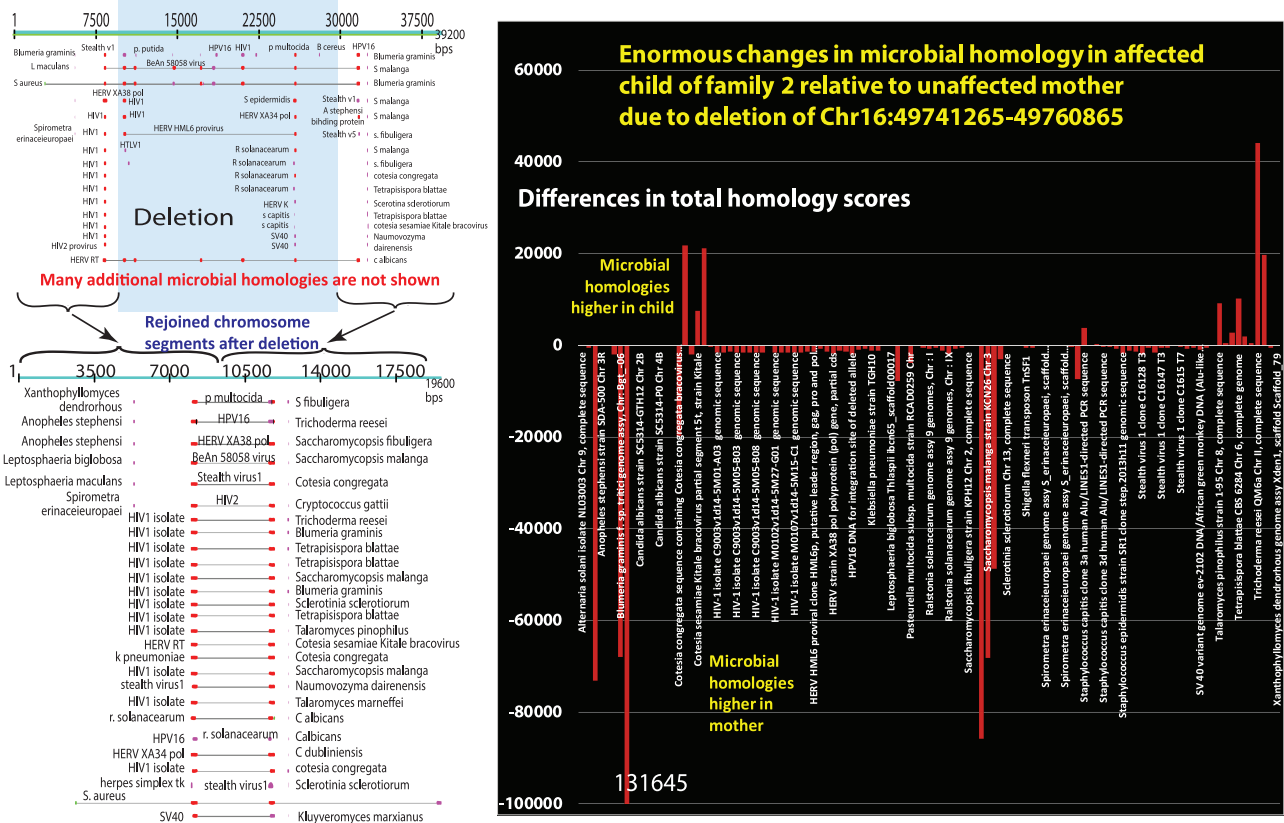


Figure 4. A structural variant unique to an affected member of family 2 in Aristidou et al³ has massive changes in the distribution and identities of homologous microorganisms when compared with the unaffected mother. At the left (top) are representations of the homologies to normal human chromosome 16:49,741,265-49,760,865 (GrCh37) with 9600 bps on both the 5' and 3' ends. After the deletion (lower left), the two 9600bp additions are juxtaposed. The rearrangement creates multiple differences in the distributions of foreign DNA homologies. The graph at the right shows the quantitative differences created by the rearrangement. Foreign DNAs are arranged alphabetically. Bars above the line are microbes more strongly represented in the affected child's sequence and those below the line are stronger in the unaffected mother's sequence. This result was repeated many times with a variety of different assumptions both for the regions and for the homology criteria.

affected by neurodevelopmental disorders³ were compared with genes in pathogenic chromosomal deletions. Parts of pathogenic chromosome deletions affected these kinds of critical neurodevelopmental driver genes.¹ Like clustered chromosomal deletions, virtually all pathogenic driver genes have strong effects on the immune system, angiogenesis, circulation, and craniofacial development. Figure 6 summarizes how the functions of damaged epigenetic drivers are distributed and shows that all 46 gene drivers of neurodevelopment have pleiotropic effects. By comparison, pleiotropy has been documented for 44% of 14459 genes in the GWAS catalog.⁴³ By this standard, neurodevelopmental driver genes are disproportionately pleiotropic ($P < .0001$).

Clear evidence of a non-human insertion

In 48 patients,¹ multiple infection-matching sequences were included in chromosomal anomalies generated by balanced chromosomal translocations (data not shown). Sequences around individual breakpoints were tested for microbial insertions by first comparing the sequences with human and then with microbial DNA. For example, chromosome breakpoint 2

in patient DGAP154 matched human DNA X-chromosome in 2 segments with a gap in the sequence (Figure 7). The gap did not match human sequences but did correspond to nematodes and yeast-like fungi, suggesting one or more of these microorganisms had inserted foreign DNA into patient DGAP154 (Figure 7). More frequently, however, other breakpoints in DGAP154 chromosomes matched many microbial sequences. A simple example of one of these alignments around DGAP154 breakpoint 3 shows that many microorganisms align with human DNA. The similarity between critical human DNA epigenetic factors and microbial DNA (which is more abundant) can set up competitions during recombination and break repair (Figure 7 bottom). Large changes in the sets of homologous foreign DNAs also accompany chromosome gene rearrangements that affect driver genes.

Multiple infections identified by homology match signs and symptoms of neurodevelopmental disorders

These kinds of alignments suggest candidates that can contribute to the signs and symptoms in each individual (Table 2). Eight patients have growth retardation, and 20 of 48 patients

Table 1. Epigenetic functions of mutated and deleted genes in neurodevelopmental disorders relate neurologic deficits to deficits in the immune system, the circulatory system, and structural genes.

PATIENT	PROPOSED PRIMARY PHENOTYPIC DRIVERS OF ANOMALY OR DISRUPTED GENES	GENES DELETED OR MUTATED LINKED TO NERVOUS SYSTEM FUNCTION ¹⁴	GENES DELETED OR MUTATED LINKED TO IMMUNITY, INFECTION	GENES DELETED OR MUTATED WITH LINKED TO CIRCULATION AND BLOOD BRAIN BARRIER	GENES DELETED OR MUTATED LINKED TO BONE, STRUCTURAL REQUIREMENTS
DGAP002	14q12-q21.1 deleted. FOXA1 functions in epigenetic reprogramming as part of a DNA repair complex ¹⁵	NOVA1, AKAP6	FoxG1, NFKB1A, NKX2, BAZ1A, FoxA1	FoxA1	FoxA1
DGAP011	FGFR1 integrates many epigenomic signals that control development ¹⁶	✓	✓	✓	✓
DGAP012	PHF21A , ¹⁷ Histone methyl-reader ¹⁷ within the histone deacetylase complex PHF21A encodes BHC80, a component of a BRAF35 histone deacetylase complex.	✓	✓	✓	Not found
DGAP093	CDKL5 gene disrupted at breakpoint, phosphorylates HDAC4 histone deacetylase causing cytoplasmic retention. ¹⁸	✓	✓	✓	Not found
DGAP096	WAC at breakpoint, disrupted by translocation regulates histone ubiquitination	✓	✓	✓	Not found
DGAP099	ZBTB20 controlled by epigenetic processes, important in the development of the hippocampus. Hypermethylation is linked to major depressive disorder. ¹⁹	✓	✓	✓	✓
DGAP100	KDM6 Histone Demethylase and Epigenetic switch to specify stem cell fate ²⁰	✓	✓	✓	✓
DGAP112	12p12.1-p11.22 deleted. SOX5 , a histone demethylase at breakpoint is disrupted by rearrangement. MED21 at breakpoint may be required to displace histones to enable heat-shock response. ²¹	✓	✓, KRAS, MED21	✓, KRAS	✓
DGAP124	NRXN1 , repressive histone marker at rearrangement breakpoint, controls choice of NRXN splicing isoform. Repressed by histone methyltransferase ²²	✓	✓	✓	Not found
DGAP127	PAK3 , compensates for PAK1 which interacts with histone H3	✓	✓	✓	Not found
DGAP133	6q13-q14.1 deleted includes the epigenetic controller HMGN3 which normally stimulates acetylation of histone H3. ²³	SMAP1, B3GAT2, OGFRL1, RIMS1, SLC17A5, FILIP1, LCA5, SENP6, etc (see Figure 1)	KHDC1, EEF1A1, MYO6, IRAK1BP1, B3GAT2, etc (see Figure 1)	Collagen genes 19A, 12A, COX7A2	Col9A
DGAP139	13q14.2 deleted includes SETDB2 a gene that codes for the epigenetic regulator methyltransferase ²⁴ (see Figure 1)	SUCLA2, ITM2B, RCBTB2, CYSLTR2, KPNA3	RB1, NUDT15, DLEU1, DLEU2, PHF11, SETDB2, KCNRG	ITM2B, CYSLTR2, LPAR6, DLEU2	DLeu2

(Continued)

Table 1. (Continued)

PATIENT	PROPOSED PRIMARY PHENOTYPIC DRIVERS OF ANOMALY OR DISRUPTED GENES	GENES DELETED OR MUTATED LINKED TO NERVOUS SYSTEM FUNCTION ¹⁴	GENES DELETED OR MUTATED LINKED TO IMMUNITY, INFECTION AND BLOOD BRAIN BARRIER	GENES DELETED OR MUTATED LINKED TO BONE, STRUCTURAL REQUIREMENTS
DGAP142	MBD5 Epigenetic regulator Histone acetylation disrupted by rearrangement.	✓	✓	Not found
DGAP145	EFTUD2 [RNA splicing regulator in spliceosome] at rearrangement breakpoint. An epigenetic effector that connects methylated histone to RNA splicing ²⁵	✓	✓	✓
DGAP147	NALCN encodes sodium leak channel	✓	Not found	Not found
DGAP154	Xq25 (Duplication) XIAP, STAG2 . ²⁶ STAG2 inhibition increases the ability to reprogram cardiac fibroblasts implicating STAG2 in epigenetics.	XIAP, THOC2, GRIA3	XIAP	SMAD5
DGAP155	EHMT1 /GLP encodes a histone methyltransferase, an epigenetic regulator.	✓	✓	✓
DGAP157	FOXP1 "PRMT5 recruitment to the FOXP1 promoter facilitates H3R2me2s, SET1 recruitment, H3K4me3, and gene expression," ²⁷ disrupted by rearrangement.	✓	✓	✓
DGAP159	10p15.3-p14 deletion includes GATA3 , see Figure 1. GATA3 controls enhancer accessibility ²⁸	✓	✓	✓
DGAP161	4q34 (see Figure 1) HMGB2 , effects epigenetic status ²⁹	AGA, SPCS3,TENM	HMGB2, HPGD, VEGFC, ADAM29	HMGB2, VEGFC
DGAP164	NODAL/activin a targets KDM6B (epigenetic regulator of neuronal plasticity) signaling in mouse brain. Rearrangement also disrupts TET1 (a methylation eraser (demethylase))	✓	✓	✓
DGAP166	SCN1A Sodium voltage gated channel	✓	Not found	Not found
DGAP169	NR2F1 disrupted by rearrangement, target of the epigenetic regulator complex PRC2	✓	Not found	✓
DGAP173	11p14.2 FIBIN, BBOX1, SLC5A12, ANO3	ANO3, BBOX1	Mucin15	Not found
DGAP186	NR5A1 disrupted, participates in resetting pluripotency. Differentially methylated in endometriosis	✓	✓	Not found
DGAP189	SOX5 participates in epigenetics as a histone demethylase	✓	✓	✓
DGAP190	SMS Spermine synthetase gene. Polyamines control the epigenetic regulators histone demethylases	✓	✓	✓

(Continued)

Table 1. (Continued)

PATIENT	PROPOSED PRIMARY PHENOTYPIC DRIVERS OF ANOMALY OR DISRUPTED GENES	GENES DELETED OR MUTATED LINKED TO NERVOUS SYSTEM FUNCTION ¹⁴	GENES DELETED OR MUTATED LINKED TO IMMUNITY, INFECTION	GENES DELETED OR MUTATED WITH LINKED TO CIRCULATION AND BLOOD BRAIN BARRIER	GENES DELETED OR MUTATED LINKED TO BONE, STRUCTURAL REQUIREMENTS
DGAP193	SPAST	✓	✓	✓	✓
DGAP199	NOTCH2, disrupted at breakpoint by rearrangement.	✓	✓	✓	✓
DGAP201	AUTS2 , associated with H3K4me3 and epigenetic control	✓	✓	Not found	Not found
DGAP202	KDM6A Histone di/tri demethylase disrupted at breakpoint by rearrangement.	✓	✓	✓	
DGAP211	SATB2 Epigenetic functions, interacts with chromatin remodelers.	✓	✓	Not found	✓
DGAP219	CUL3 Ubiquitin ligase recruited to epigenetically regulate lymphoid effector programs, disrupted at breakpoint of rearrangement.	✓	✓	✓	✓
DGAP232	SNRPN-SNURF/SMN ; PWCR; SM-D; sm-N; RT-LI; HCERN3; SNRNP-N; required for epigenetic imprinting.	✓	✓	Not found	✓
DGAP235	MBD5 Epigenetic regulator involved in histone acetylation disrupted at breakpoint of rearrangement.	✓	✓	✓	Not found
DGAP239	CHD7 Chromodomain helicase collaborates with the epigenetic regulator SOX2 ³⁰ and is the gene that encodes functions essential for chromatin remodeling	✓	✓	Not found	✓
DGAP244	CTNND2	✓	✓	Not found	✓
DGAP278	SNRPN-SNURF/SMN ; PWCR; SM-D; sm-N; RT-LI; HCERN3; SNRNP-N; required for epigenetic imprinting	✓	✓	Not found	✓
DGAP301	MEF2C associated with histone hypermethylation in disease, ³¹ disrupted by rearrangement	✓	✓	✓	✓
DGAP316	18p11.32-p11.22 deletion: PHF21A "Histone methyl-reader" ¹⁷ within the histone deacetylase complex PHF21A Encodes BHC80, a component of a BRAF35 histone deacetylase complex	✓	✓	YES1, MYOM1, EPB41L3, ARHGAP28, LAMA1, PTPRM, MTCL1, TWSG1, RALBP1, RAB31	SMCHD1, TGIF1, LAMA1, PTPRM, TWSGH1
DGAP317	6q14:1 TBX18: IRAK1BP1; IBTK	LCA5, ELOV4, HTR1B, DOPEY	IRAK1BP1, T-BX18, IBTK, NFKB	COX7A	Not found

(Continued)

Table 1. (Continued)

PATIENT	PROPOSED PRIMARY PHENOTYPIC DRIVERS OF ANOMALY OR DISRUPTED GENES	GENES DELETED OR MUTATED LINKED TO NERVOUS SYSTEM FUNCTION ¹⁴	GENES DELETED OR MUTATED LINKED TO IMMUNITY, INFECTION AND BLOOD BRAIN BARRIER	GENES DELETED OR MUTATED LINKED TO BONE, STRUCTURAL REQUIREMENTS
MGH7	GRIN2B reflects changes in methylation levels ³²	✓	Not found	Not found
MGH8	CHD8 binds methylated histone H3K4 at active promoters, protein is also an ATP dependent chromatin remodeler. ³³	✓	Not found	Not found
MGH9	TCF4 regulated by histone deacetylases. Binding sites overlaps histone acetylation of enhancers ³⁴	✓	✓	✓
NIJ2	PHIP , linked to histone methylation ³⁵ Disrupted at breakpoint by rearrangement.	PHIP, MYO6	PHIP, MYO6	PHIP, MYO6
NIJ5	IL1RAPL1	✓	Not found	Skeletal growth
NIJ6	KAT6B/MORF/MYST4 Lysine acetyl transferase.	✓	✓	✓
NIJ14	NFIA in response to NOTCH1, Nfia displaces DNMT1 to demethylate astrocyte specific gene promoters ³⁶	✓	✓	✓
NIJ15, NIJ6	MYT1L	✓	Not found	Not found
ROC4	CAMTA1	✓	✓	✓
ROC17	AUTS2 , associated with H3K4me3 and epigenetic control	✓	Not found	Not found
ROC23	TCF12	✓	✓	✓
ROC43	SOX5 is a histone demethylase	✓	✓	✓
ROC62	SNRPN-SNURF , required for epigenetic imprinting	✓	Not found	✓
UTR7	NFIX , controlled by methylation	✓	✓	✓
UTR12	DYRK1A , histone phosphorylation	✓	✓	✓
UTR13	MBD5 , Epigenetic regulator disrupted in topology associated domain	✓	✓	Not found
UTR17	ZBTB20 , binds to genes that control chromatin architecture including MEF2c and SAT2b	✓	✓	✓
UTR20	FOXP2	✓	✓	✓

(Continued)

Table 1. (Continued)

PATIENT	PROPOSED PRIMARY PHENOTYPIC DRIVERS OF ANOMALY OR DISRUPTED GENES	GENES DELETED OR MUTATED LINKED TO NERVOUS SYSTEM FUNCTION ¹⁴	GENES DELETED OR MUTATED WITH LINKED IMMUNITY, INFECTION AND BLOOD BRAIN BARRIER	GENES DELETED OR MUTATED LINKED TO BONE, STRUCTURAL REQUIREMENTS
UTR21	NSD1 lysine methylation	✓	✓	✓
UTR22	2q24.3 SCN9A	SCN9A, TTC21B, STK39	NOSTRIN, CERS6	TTC21B
Affected member of family 1	NCALD1 exists in myristoyl and non myristoyl forms. Histone deacetylase 11 is a myristoyl hydrolase. ³⁷	✓	Not found	✓
Affected member of family 1	NPL—N-acetylneuraminase lyases regulate cellular sialic acid concentrations by mediating its reversible conversion into N-acetylmannosamine and pyruvate	✓	✓	✓
Affected member of family 1	BCAR3 . A candidate epigenetic mediator for increased adult body mass index in socially disadvantaged females ³⁸	Not found	Not found	✓
Affected member of family 2	ZNF423 —epigenetic modifications associated with obesity, ³⁹ methylation marker for age prediction.	✓	Not found	✓
Affected member of family 3	RAP1GDS1 — RAP1 GTP-GDP dissociation stimulator 1. ⁴⁰ Rap1 occupancy antagonizes some histones.	✓	✓	✓

Genes in bold type are related to epigenetic control; checkmark indicates identical gene with mainly epigenetic functions is listed in column 2. Other genes that do not completely match column 2 are listed individually.

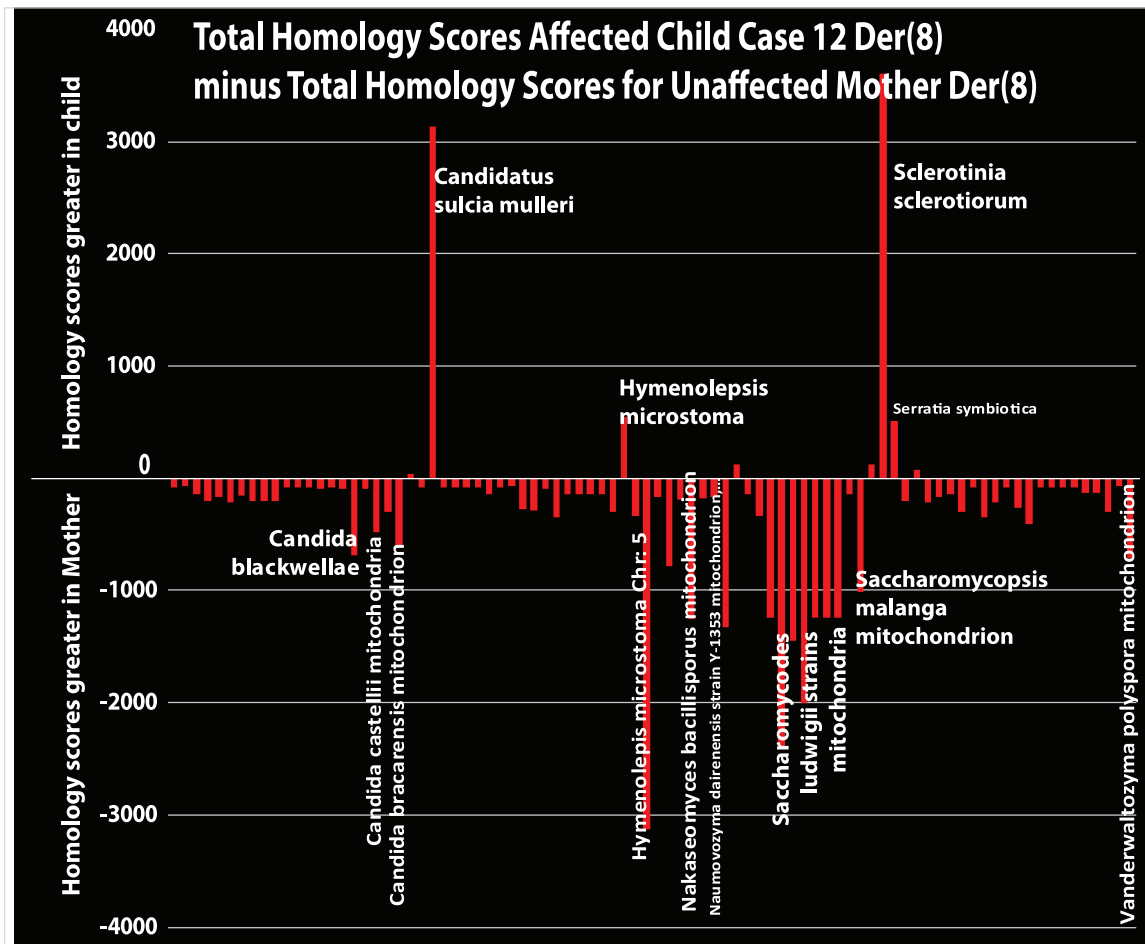
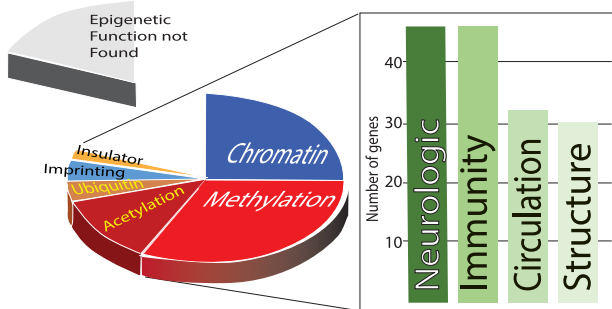


Figure 5. Changes in alien DNA homologies in an affected child born to a mother with a recurrent translocation. Parental translocation t(8;22) in Family FHU13-027 from Mishra et al² is homologous to microorganisms as shown for der(8). The graph represents the difference in total homology scores for an affected child case 12 subtracted from those of the normal healthy mother with balanced translocation t(8;22). A very different set of foreign DNAs is present in affected case 12. The homologous foreign to DNAs are in alphabetical order but only those that show the greatest differences are labeled.



Driver genes are epigenetic regulators of coordinated neurodevelopment functions

Figure 6. As a result of chromosome anomalies, driver genes truncated or deleted in congenital neurodevelopmental disorders are mainly epigenetic regulators or effectors. The pie chart shows the percentages of 46 driver genes that have the epigenetic functions indicated. Loss of these driver gene functions then impacts a group of functions that must be synchronized during the complex process of neurodevelopment. These are the same general functions lost in deleted gene clusters.

had impaired speech.¹ Multiple infections can cause these problems. For instance, HIV-1 causes white matter lesions

associated with language impairments and also harms fetal growth. There are nearly 50 matches to HIV-1 DNA in the chromosome anomalies of 35 patients.

Within chromosome anomalies, stealth viruses have about 35 matching sequences. Stealth viruses are mostly herpes derivatives that emerge in immunosuppressed patients such as cytomegalovirus (CMV). Stealth virus 1 (Table 2) is Simian CMV with up to 95% sequence identity to isolates from human patients. First trimester CMV infection can cause severe cerebral abnormalities followed by neurologic symptoms.⁴⁹ CMV is also a common cause of congenital deafness and visual abnormalities. Twenty-seven of 48 neurodevelopmental patients had hearing loss. Herpes simplex virus is another stealth virus that directly infects the central nervous system and can cause seizures (reported for 9 patients).

Chromosome anomalies in patient DGAP159 have strong homology to *Neisseria meningitidis*.

Signs and symptoms in patient DGAP159 are consistent with known neurodevelopmental effects of bacterial meningitis including hearing loss, developmental delay, speech failure, and visual problems.

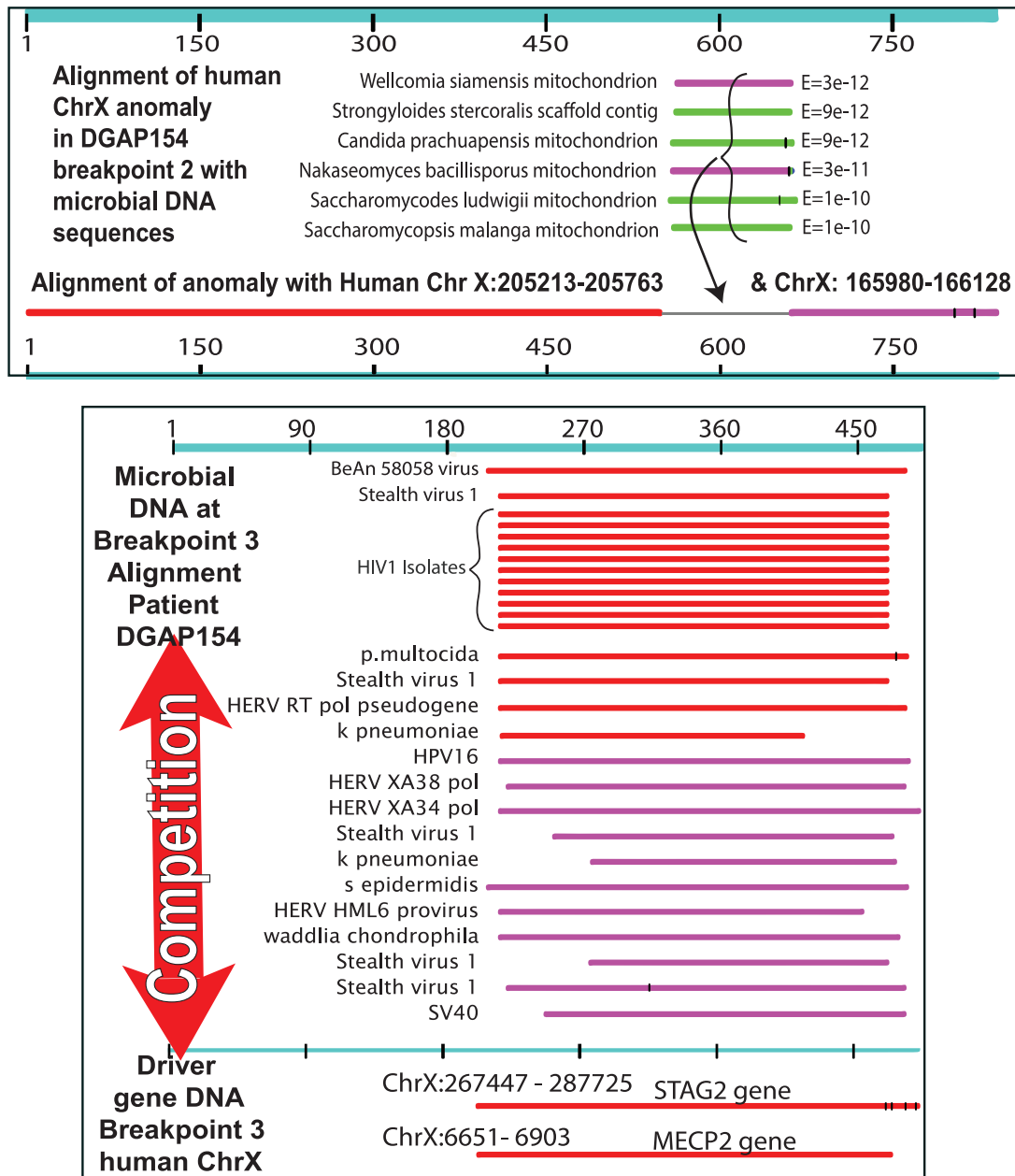


Figure 7. Foreign DNA sequences can compete with human DNA at epigenetic regulators around breakpoints.

In the top panel (breakpoint 2 in patient DGAP154), mitochondrial DNA from potential pathogens, such as nematodes and yeast-like fungi, aligns well with an anomaly in a damaged epigenetic factor. The sequence at breakpoint 2 on chromosome X indicated by the black line does not match any human sequences. Any of the nematodes or yeast-like fungi could insert their DNA at this position. The lower panel (breakpoint 3 in the same patient) shows that microbial DNA can compete with human DNA around damaged epigenetic regulators such as MECP2. Microbial homology to this major epigenetic regulator was confirmed by testing human MECP2 against microbes. STAG2 is also implicated in epigenetics (Table 1).

Tests for artifacts in matches to human-microbial DNA

Genome rearrangements for patient data¹ produced 1986 matches to microbial sequences (range=66%-100%) with $E \leq e-10$ and a mean value of 83% identity. About 190 Alu sequences resembled microbial sequences, supporting the idea that homologies among repetitive human sequences and microbes are real. Correspondence between microbial sequences and multiple human repetitive sequences increases possibilities that microbial sequences can interfere with essential human processes. Contamination of microbial DNA sequences by

human Alu elements⁵⁰ was ruled out by comparing about 450 AluJ, AluS, and AluY sequences with all viruses and bacteria in the National Center for Biotechnology Information (NCBI) database.

Tests for DNA sequence artifacts

To further test the possibility that some versions of these microbial sequences were sequencing artifacts or contaminated by human genomes, microbial genomes were (reverse) tested for homology to human genomes. Similarities to human

Table 2. Recurrent infections found to have homology to chromosomal abnormalities in neurologic birth defects can cause developmental defects.

RECURRENT INFECTIONS MATCHING DNA REARRANGEMENTS IN CONGENITAL NEUROLOGIC DISORDERS	CNS, PHYSIOLOGIC, AND OR TERATOGENIC EFFECTS IF KNOWN	NUMBER OF PATIENTS WITH CHROMOSOMAL ABNORMALITIES WITH SIGNIFICANT HOMOLOGIES TO THIS INFECTION
HIV-1, HIV-2	Impaired fetal growth, premature delivery, chorioamnionitis, deciduitis, immunodeficiency. HIV-1 causes white matter lesions associated with language impairments. HIV-1 infects CNS by targeting microglial cells.	600 matches in 36 patients
HPV16	One of a group of infections that coexist with other viral infections, bacterial infections, and chemicals in autism syndrome disorders. ⁴⁴	35 patients
Staphylococcal infection (<i>Staphylococcus aureus</i> , <i>Staphylococcus capitis</i> , <i>Staphylococcus epidermidis</i>)	May infect brain and interfere with normal nerve transmission. Causes meningitis, brain abscess. Major hospital pathogens	206 matches in 40 patients
Stealth viruses, Stealth virus 1, CMV, HHV-4/EBV, herpes simplex, SV40	Associated with neurologic impairments, hearing loss, ophthalmic problems.	425 matches in 39 patients
<i>Ralstonia solanacearum</i>	Opportunistic bacterial pathogen, associated with pneumonia and neonatal sepsis, especially in patients with compromised immunity on ventilator. Has some resemblance to other human pathogens including <i>Borrelia</i> , <i>Bordetella</i> , and <i>Burkholderia</i> .	158 matches in 44 patients
HTLV-1	Associated with neuropathy. Immune-mediated disease of the nervous system affects spinal cord and peripheral nerves. Fetal neural cells are highly susceptible to infection	24 patients
BeAn 58058 virus, Bandra megavirus	BeAn 58058 is almost identical (97%) to Cotia virus which can infect human cells. Represents a distinct branch of poxviruses ^{45,46}	50 matches in 36 patients
<i>Klebsiella pneumoniae</i> , <i>Escherichia coli</i>	Neonatal pneumonia and sepsis.	35 patients
Human respiratory syncytial virus Killifi	Neonatal pneumonia, trouble breathing	7 patients: DGAP099, DGAP127, DGAP137, DGHAP159, DGAP167, DGAP172, DGAP225
<i>Waddlia chondrophila</i>	Recently discovered. Causes preterm birth, miscarriage, spontaneous abortion. Chlamydia like microorganism. Survives in macrophages, multiplies in endometrial cells	26 different patients
<i>Neisseria meningitidis</i> , <i>Neisseria gonorrhoeae</i>	Bacteria that can cause damage to human DNA. Premature birth, miscarriage, severe eye infection in infant. Antibodies to <i>N gonorrhoeae</i> impair outgrowth of neurites, and cross react with specific brain proteins.	Patient DGAP159 shows 25 different homologies to <i>N meningitidis</i> . DGAP017 and in DGAP101 have homology to <i>N gonorrhoeae</i> .
<i>Pseudomonas putida</i>	Gram-negative bacterium found in cases of acute bacterial meningitis. ⁴⁷ Reported in contaminated water and heparinized flush solutions ⁴⁸	13 patients: DGAP012, DGAP93, DGAP100, DGAP125, DGAP133, DGAP134, DGAP154, DGAP159, DGAP167, DGAP169, DGAP170, DGAP220, ROC17
<i>Exiguobacterium oxidotolerans</i>	High catalase activity may disable immune defenses depending on peroxide.	DGAP127, DGAP225

Abbreviations: CMV, cytomegalovirus; CNS, central nervous system; EBV, Epstein-Barr virus; HTLV, Human T-cell lymphotropic virus.

sequences were found across multiple strains of the same microorganism (Table 3). For example, an Alu homologous region of the HIV-1 genome (bps = 7300-9000) in 28 different HIV-1 isolates was compared with human DNA. All 28 HIV sequences matched the same region of human DNA, at up to 98% identity. In contrast, only 1 of 20 zika virus sequences matched humans, and zika virus was not considered further.

A model for infection interference in neurodevelopment

Autosomal dominant inheritance of neurodevelopmental disorders containing microbial DNA suggests interference with gamete generation in 1 parent. The mechanism proposed in Figure 8 is based on significant changes in microbial homologies on multiple human chromosomes. Large amounts of

Table 3. Independent evidence that microbial genomes have regions of homology to human DNA as predicted by results.

MICROBE	HUMAN CHROMOSOME HOMOLOGIES	LENGTH OF HOMOMOLOGY (BP)	% HOMOMOLOGY, E VALUE
<i>Neisseria gonorrhoeae</i> strain 1090	NotI sites	233-345	100%, E=5e-89
<i>N gonorrhoeae</i> WHO-L genome (LT591901.1)	SPARC, SPOCK3, IMP2L, LAMA2, FRMD4, parts of reference sequences for most human chromosomes.	685	99.4%-100%, E=0.0
<i>N gonorrhoeae</i> strain 1090 (Accession AE004969)	EEFA1L14	865	68%, E=3e-77
<i>Staphylococcus aureus</i>	ATPase	1198	98%, E=0.0
<i>S aureus subsp. aureus</i> NC_007795.1	P143 mRNA	642	70%, E=9e-58
<i>S aureus</i> MRSA strain UTSW (Accession CP01323.1)	Succinate CoA dehydrogenase flavoprotein	1131	98%, E=0.0
<i>Staphylococcus capitis</i>	ABC8	218	71%, E=5e-15
<i>Ralstonia solanacearum</i>	s-adenosyl-homocysteine hydrolase, NM_001242673.1 Homo sapiens adenosylhomocysteinase like 1 (AHCYL1), transcript variant 2, mRNA (>100 significant matches to other microbial sequences)	791	69%, E=2e-73
<i>Pseudomonas putida</i> strain IEC33019 complete genome NZ_CP016634.1 vs strain T1E (CP003734.1)	Many hundreds of homologies, eg, homo sapiens sequence around Not1 site clone H5J-DM24RS	674 bp	92%, E=0.0
<i>Clostridium botulinum</i>	HSP70 member 9	1066	65%, E=5e-54
<i>C botulinum BrDura</i>	Human cDNA	692	83%, E=0.0
<i>Waddlia chondrophila</i> (CP001928.1) whole genome 2.1 million bp	Phosphoglycerate dehydrogenase	734	66%, E=9e-71
	Fumarate hydratase	1034	67%, E=3e-63
	Elongation factor alpha	992	71%, E=1e-112
	Hsp70 family member 9	1548	65%, E=1e-68
	Aldehyde dehydrogenase	1168	64%, E=1e-30
67% BeAn 58058 virus NM_001165931.1 9 ribonucleotide reductase > 100 significant matches to cotia virus (90% homology), Volepox, monkeypox, cowpox, and vaccinia virus, eg, 675/939 bp (72%) E=4e-136	Ribonucleotide regulatory subunit (Homo sapiens ribonucleotide reductase regulatory subunit M2 [RRM2], transcript variant 1, mRNA)	870	68%, E=8e-69
Cowpox, KY369926.1 Cowpox virus strain Kostroma_2015, complete genome	At least 100 homologies. Homologies to ribonucleotide reductase, z-protein mRNA, transmembrane BAX inhibitor motif, EF hand domain containing EFHC2	68%-78% homology	2279 bp at 69% homology (E=0.0)
HIV-1 28 different isolates	Alu homology bps 7300 to 8000		Up to 98% identity for all 28 sequences
HTLV1 J02029 vs HTLV1/HAM Long terminal repeat	Homologous to hydroxysteroid dehydrogenase like 1 variant, mRNA for hGLI2	97% homology	2162 bp at 97% homology (E=0.0)

foreign DNA present during human meiosis with its many double-strand breaks during the most active period of recombination produce defective gametes. Errors in spermatogenesis underlie a prevalent and recurrent gene rearrangement that causes intellectual disability, and dysmorphism (Emanuel syndrome).² In contrast, recombination in ova occurs in fetal life and then meiosis is arrested until puberty.⁵¹

The resemblance of foreign DNA to host background DNA may be a major factor in selecting infection and in human ability to clear the infection. Only 1 rare defective gamete is modeled in Figure 8 but the male generates 4 gametes during meiosis beginning at puberty. Only 1 gamete survives in the female because 3 polar bodies are generated. In neurodevelopmental disorders, massive changes in similarities

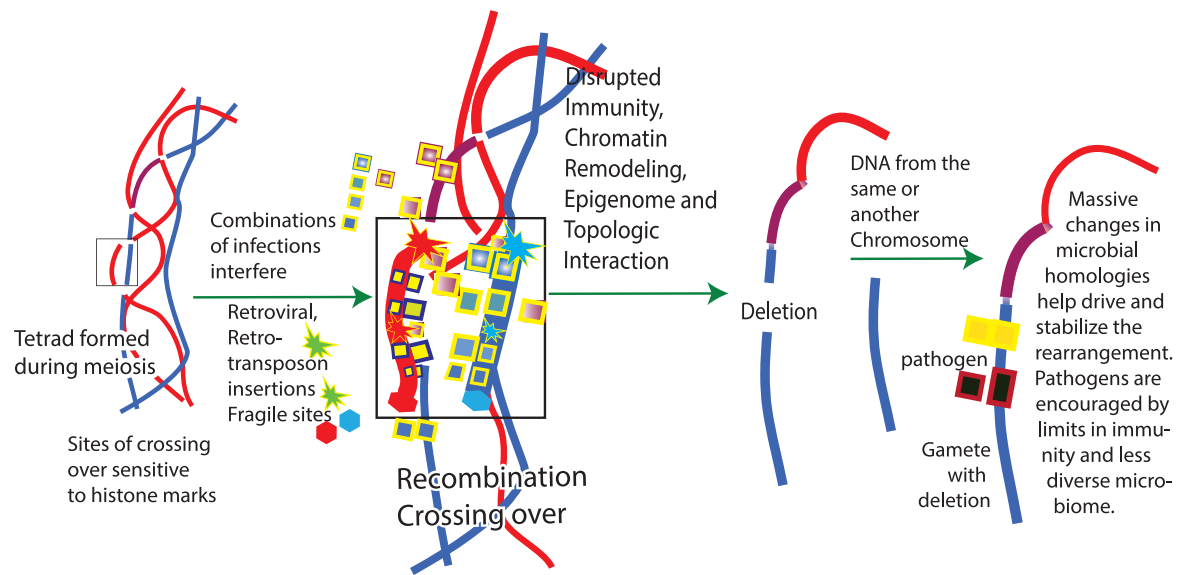


Figure 8. Soon after conception, erasure of epigenetic marks generates pluripotent stem cells and then the epigenome is reprogrammed. Foreign DNA can interfere with both processes leading to neurodevelopmental disorders. The model above shows interference with meiosis by DNA from infections. After duplication of parental chromosomes prior to generation of gametes, foreign DNA matches and associates with strands of DNA in many places, including repetitive sequences of human DNA. In addition, retroviruses and retrotransposons may integrate their DNAs and fragile sites may further destabilize the area. These events interfere with reductive cell division, topological relationships among chromosomes, epigenetic regulation, and high-fidelity break repair. Interference from foreign DNAs may favor the illegitimate combinations due to palindromes reported by Mishra et al.² Large numbers of DNA breaks occur during meiosis. Initiation sites for recombination are enriched in histone methylation and acetylation marks. Incorrect repair of recombination breaks is known to occur,¹ causing chromosome anomalies such as deletions (shown). Massive changes in distributions of homologous foreign DNAs occur to drive and stabilize the rearrangements. Clustered genes responsible for linked functions and epigenetic regulation in neurodevelopment are lost or displaced. Other chromosome segments with microbial homology do not contain identified genes but may be essential control regions, insulators, or essential for chromatin structures.

to foreign DNAs accompany chromosome anomalies such as deletions and insertions. Foreign DNAs can insert itself, interfere with epigenomic marking or with break repairs during meiosis. A preexisting balanced chromosomal translocation in the family³ increases the chances of generating a defective gamete during meiosis.

Discussion

Long stretches of DNA in many foreign DNAs match millions of repetitive human DNA sequences. Individual microorganisms also match nonrepetitive sequences. Human infections may be selected for and initially tolerated because of these matches. It is almost impossible to completely exclude the possibility of sequencing artifacts or contamination of microbial sequences with human sequences. However, rather than reflecting widespread, wholesale error due to human DNA contamination in many laboratories over many years, microbial homologies more likely suggest that DNA sequences in the microbiome have been selected because they are homologous to regions of human DNA. This may be a driving force behind the much slower evolution of human repetitive DNAs.

Infections such as exogenous or endogenous retroviruses are known to insert into DNA hotspots.¹⁰ Foreign DNAs are proposed to drive neurodevelopmental anomalies because humans harbor large numbers of foreign DNAs. Changes in the composition of foreign DNAs can stabilize rearrangements and favor pathogens. (Figure 8). Thus, the genetic background of an individual may be a key factor in determining the susceptibility to infection and to the effects of infection. At the genetic level,

this suggests selective pressures for infections to develop and use genes that are similar to human versions and to silence or mutate genes that are immunogenic. Infection genomes evolve rapidly on transfer to a new host.⁵² The presence of genes in infections that have long stretches of identity with human genes makes the infection more difficult to recognize as non-self. For example, there is no state of immunity to *N gonorrhoeae*. Long stretches of *N gonorrhoeae* DNA are almost identical to human DNA. Alu sequences and other repetitive elements are thought to underlie some diseases by interfering with correct homologous recombination as in hereditary colon cancer⁵³ or abnormal splicing. Why this occurs is not well understood. The presence of infection DNA that is homologous to multiple, long stretches of human DNA may mask proper recombination sites and encourage this abnormal behavior.

Neurons interact with cells in the immune system, sensing and adapting to their common environment. These interactions prevent multiple pathological changes.⁵⁴ Many genes implicated in neurodevelopmental diseases reflect strong relationships between the immune system and the nervous system. It was always possible to find functions within the immune system for genes involved in neurodevelopmental disorders (Figures 1, 4, and 6). Damage to genes essential to prevent infection leads to more global developmental neurologic defects including intellectual disability. These homologies include known microbes known to produce teratogens. Analysis of mutations within clusters of genes deleted in neurodevelopmental disorders predicts loss of brain-circulatory barriers,

facilitating infections. Damage to cellular genes essential for autophagy may lead to abnormal pruning of neural connections during postnatal development.

Aggregated gene damage accounts for immune, circulatory, and structural deficits that accompany neurologic deficits. Other gene losses listed in Table 1 and in deleted chromosome segments (Figure 1) account for deficits in cardiac function, cell barriers, bone structure, skull size, muscle tone, and many other nonneurologic signs of neurodevelopmental disorders.

The arrangement of genes in clusters converging on the same biological process may simplify the regulation and coordination between neurons and other genes during neurodevelopment and neuroplasticity. Genes that are required for related functions, requiring coordinated regulation have been shown to be organized into individual topologically associated domains.⁵⁵ Neurons are intimately connected to chromatin architecture and epigenetic controls.⁵⁶ A disadvantage of the clustered arrangement of co-regulated accessory genes is that homology to microbial infections or other causes of chromosome anomalies anywhere in the cluster can then ruin complex coordinated neurological processes.

The results in Table 1 and Figure 6 emphasize the role of epigenetic factors in neurodevelopmental diseases. Chromatin modifier genes are disproportionately affected in patients with neurodevelopmental disorders¹ and include 2 types of modifiers. Epigenetic factors signal chromatin remodelers, which are large multi-protein complexes. Epigenetic factors are responsible for differentiation from pluripotent states; chromatin remodeling also has major roles in developmental stage transitions. There are 5 families of chromatin remodelers that all control access to DNA within nucleosomes, exchanging and repositioning them. Chromatin remodeling arrays contain an ATPase subunit resembling motor proteins⁵⁷ and are distinct from epigenetic factors.

Epigenetic regulators that affect multiple functions required by the same process make their mutation especially critical. Longer range developmental interactions in chromosome regions exacerbate the effects of infection. Mutations or deletions (Figure 1 and Table 1) show that this effect can occur in neurodevelopmental disorders. Functions that must be synchronized are grouped together on the same chromosome region and can be lost together.

Microbial DNA sequences are unlikely to be contaminants or sequencing artifacts. They are all found connected to human DNA in disease chromosomes; for example, multiple microbial sequences from different laboratories are all homologous to the same Alu sequence. Alu element-containing RNA polymerase II transcripts (AluRNAs) determine nucleolar structure and rRNA synthesis and may regulate nucleolar assembly as the cell cycle progresses and as the cell adapts to external signals.⁵⁸ HIV-1 integration occurs with some preference near or within Alu repeats.⁵⁹ Alu sequences are largely inactive retrotransposons, but some human-microbial homologies detected may be

due to insertions from Alu or other repetitive sequences. Neuronal progenitors may support de novo retrotransposition in response to the environment or maternal factors.⁶⁰

Their variability and rarity make neurologic disorders difficult to study by conventional approaches. The techniques used here can improve prenatal and genetic counseling. However, a limitation is the inability to unequivocally identify 1 infection and to absolutely distinguish infection by 1 foreign DNA from multiple infections.

Conclusions

1. DNAs in some congenital neurodevelopmental disorders closely match multiple infections that extend over long linear stretches of human DNA and often involve repetitive human DNA sequences. The affected human sequences are shown to exist as linear clusters of genes closely spaced in 2 dimensions.
2. Interference from infection and foreign DNAs can delete or damage human gene clusters and alter the epigenome. This interference accounts for immune, circulatory, and structural deficits that accompany neurologic deficits.
3. Neurodevelopmental disorders are proposed to begin when parental infections cause insertions or interfere with epigenetic markings and meiosis. Shifts in homologous sets of foreign DNAs can be massive and may drive chromosomal rearrangements.
4. Congenital neurodevelopmental disorders are thus viewed as resulting from an assault on human DNA by foreign organisms. Recognizing and considering these effects can improve prenatal and genetic counseling.

Acknowledgements

This work is based on the outstanding DNA sequence information and publications from investigators listed in the references. These results emphasize the importance of their work. It is a pleasure to acknowledge the inspiration, challenge, and stimulation provided by the Department of Biochemistry and Molecular Genetics and the College of Medicine at University of Illinois at Chicago.

Author Contributions

Bernard Friedenson is entirely responsible for all content and research in this article.

ORCID iD

Bernard Friedenson  <https://orcid.org/0000-0003-0590-2081>

REFERENCES

1. Redin C, Brand H, Collins RL, et al. The genomic landscape of balanced cytogenetic abnormalities associated with human congenital anomalies. *Nat Genet.* 2017;49:36–45.
2. Mishra D, Kato T, Inagaki H, et al. Breakpoint analysis of the recurrent constitutional t(8;22)(q24.13;q11.21) translocation. *Mol Cytogenet.* 2014;7:55.

3. Aristidou C, Koufaris C, Theodosiou A, et al. Accurate breakpoint mapping in apparently balanced translocation families with discordant phenotypes using whole genome mate-pair sequencing. *PLoS ONE*. 2017;12:e0169935.
4. Banik A, Kandilya D, Ramya S, Stünkel W, Chong YS, Dheen ST. Maternal factors that induce epigenetic changes contribute to neurological disorders in offspring [published online ahead of print March 24, 2017]. *Genes (Basel)*. doi:10.3390/genes8060150.
5. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409:860–921.
6. Salzberg SL, White O, Peterson J, Eisen JA. Microbial genes in the human genome: lateral transfer or gene loss? *Science*. 2001;292:1903–1906.
7. Crisp A, Boschetti C, Perry M, Tunnacliffe A, Micklem G. Expression of multiple horizontally acquired genes is a hallmark of both vertebrate and invertebrate genomes. *Genome Biol*. 2015;16:50.
8. Bhela S, Mulik S, Reddy PB, et al. Critical role of microRNA-155 in herpes simplex encephalitis. *J Immunol*. 2014;192:2734–2743.
9. Vrielynck N, Chambon A, Vezon D, et al. A DNA topoisomerase VI-like complex initiates meiotic recombination. *Science*. 2016;351:939–943.
10. Babaei S, Akhtar W, de Jong J, Reinders M, de Ridder J. 3D hotspots of recurrent retroviral insertions reveal long-range interactions with cancer genes. *Nat Commun*. 2015;6:6381.
11. Zhang Z, Schwartz S, Wagner L, Miller W. A greedy algorithm for aligning DNA sequences. *J Comput Biol*. 2000;7:203–214.
12. Pearson WR. An introduction to sequence similarity (“homology”) searching. *Curr Protoc Bioinformatics*. 2013;42:3.1.1–3.1.8.
13. Di Donato M, Bilancio A, D’Amato L, et al. Cross-talk between androgen receptor/filamin A and TrkA regulates neurite outgrowth in PC12 cells. *Mol Biol Cell*. 2015;26:2858–2872.
14. Wright CF, Fitzgerald TW, Jones WD, et al. Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet*. 2015;385:1305–1314.
15. Zhang Y, Zhang D, Li Q, et al. Nucleation of DNA repair factors by FOXA1 links DNA demethylation to transcriptional pioneering. *Nat Genet*. 2016;48:1003–1013.
16. Stachowiak MK, Stachowiak EK. Evidence-based theory for integrated genome regulation of ontogeny: an unprecedented role of nuclear FGFR1 signaling. *J Cell Physiol*. 2016;231:1199–1218.
17. Garay PM, Wallner MA, Iwase S. Yin-yang actions of histone methylation regulatory complexes in the brain. *Epigenomics*. 2016;8:1689–1708.
18. Trazzi S, Fuchs C, Viggiano R, et al. HDAC4: a key factor underlying brain developmental alterations in CDKL5 disorder. *Hum Mol Genet*. 2016;25:3887–3907.
19. Davies MN, Krause L, Bell JT, et al. Hypermethylation in the ZBTB20 gene is associated with major depressive disorder. *Genome Biol*. 2014;15:R56.
20. Hemming S, Cakouros D, Isenmann S, et al. EZH2 and KDM6A act as an epigenetic switch to regulate mesenchymal stem cell lineage specification. *Stem Cells*. 2014;32:802–815.
21. Kremer SB, Kim S, Jeon JO, et al. Role of Mediator in regulating Pol II elongation and nucleosome displacement in *Saccharomyces cerevisiae*. *Genetics*. 2012;191:95–106.
22. Ding X, Liu S, Tian M, et al. Activity-induced histone modifications govern Neurexin-1 mRNA splicing and memory preservation. *Nat Neurosci*. 2017;20:690–699.
23. Barkess G, Postnikov Y, Campos CD, et al. The chromatin-binding protein HMGN3 stimulates histone acetylation and transcription across the Glyt1 gene. *Biochem J*. 2012;442:495–505.
24. Parfett CL, Desaulniers D. A Tox21 approach to altered epigenetic landscapes: assessing epigenetic toxicity pathways leading to altered gene expression and oncogenic transformation in vitro [published online ahead of print June 1, 2017]. *Int J Mol Sci*. doi:10.3390/ijms18061179.
25. Guo R, Zheng L, Park JW, et al. BS69/ZMYND11 reads and connects histone H3.3 lysine 36 trimethylation-decorated chromatin to regulated pre-mRNA processing. *Mol Cell*. 2014;56:298–310.
26. Zhou Y, Alimohamadi S, Wang L, et al. A loss of function screen of epigenetic modifiers and splicing factors during early stage of cardiac reprogramming [published online ahead of print March 18, 2018]. *Stem Cells Int*. doi:10.1155/2018/3814747.
27. Chiang K, Zielinska AE, Shaaban AM, et al. PRMT5 is a critical regulator of breast cancer stem cell function via histone methylation and FOXF1 expression. *Cell Rep*. 2017;21:3498–3513.
28. Krendl C, Shaposhnikov D, Rishko V, et al. GATA2/3-TFAP2A/C transcription factor network couples human pluripotent stem cell differentiation to trophoblast with repression of pluripotency. *Proc Natl Acad Sci U S A*. 2017;114:E9579–E9588.
29. Bronstein R, Kyle J, Abraham AB, Tsirka SE. Neurogenic to gliogenic fate transition perturbed by loss of HMG2B. *Front Mol Neurosci*. 2017;10:153.
30. Amador-Arjona A, Cimadamore F, Huang CT, et al. SOX2 primes the epigenetic landscape in neural precursors enabling proper gene activation during hippocampal neurogenesis. *Proc Natl Acad Sci U S A*. 2015;112:E1936–E1945.
31. Mitchell AC, Javidfar B, Pothula V, et al. MEF2C transcription factor is associated with the genetic and epigenetic risk architecture of schizophrenia and improves cognition in mice. *Mol Psychiatry*. 2018;23:123–132.
32. Ryley Parrish R, Albertson AJ, Buckingham SC, et al. Status epilepticus triggers early and late alterations in brain-derived neurotrophic factor and NMDA glutamate receptor Grin2b DNA methylation levels in the hippocampus. *Neuroscience*. 2013;248:602–619.
33. Barnard RA, Pomaville MB, O’Roak BJ. Mutations and modeling of the chromatin remodeler CHD8 define an emerging autism etiology. *Front Neurosci*. 2015;9:477.
34. Forrest MP, Hill MJ, Kavanagh DH, Tansey KE, Waite AJ, Blake DJ. The psychiatric risk gene transcription factor 4 (TCF4) regulates neurodevelopmental pathways associated with schizophrenia, autism, and intellectual disability. *Schizophr Bull*. 2018;44:1100–1110.
35. Morgan MAJ, Rickels RA, Collings CK, et al. A cryptic Tudor domain links BRWD2/PHIP to COMPASS-mediated histone H3K4 methylation. *Genes Dev*. 2017;31:2003–2014.
36. Namihira M, Kohyama J, Semi K, Sanosaka T, Deneen B, Taga T, Nakashima K. Committed neuronal precursors confer astrocytic potential on residual neural precursor cells. *Dev Cell*. 2009;16:245–255.
37. Moreno-Yruela C, Galleano I, Madsen AS, Olsen CA. Histone deacetylase 11 is an epsilon-N-myristoyllysine hydrolase. *Cell Chem Biol*. 2018;25:849–856.e8.
38. Chu SH, Loucks EB, Kelsey KT, et al. Sex-specific epigenetic mediators between early life social disadvantage and adulthood BMI. *Epigenomics*. 2018;10:707–722.
39. Lin X, Lim IY, Wu Y, et al. Developmental pathways to adiposity begin before birth and are influenced by genotype, prenatal environment and epigenome. *BMC Med*. 2017;15:50.
40. Chen M, Licon K, Otsuka R, Pillus L, Ideker T. Decoupling epigenetic and genetic effects through systematic analysis of gene position. *Cell Rep*. 2013;3:128–137.
41. Ejdy E. Environment of school as potential place of interindividual transmissions. *Wiad Parazytol*. 2001;47:353–358.
42. Medvedeva YA, Lennartsson A, Ehsani R, et al. EpiFactors: a comprehensive database of human epigenetic factors and complexes [published online ahead of print July 7, 2015]. *Database (Oxford)*. doi:10.1093/database/bav067.
43. Chesmore K, Bartlett J, Williams SM. The ubiquity of pleiotropy in human disease. *Hum Genet*. 2018;137:39–44.
44. Omura Y, Lu D, Jones MK, et al. Early detection of autism (ASD) by a non-invasive quick measurement of markedly reduced acetylcholine & DHEA and increased beta-amyloid (1-42), asbestos (chrysotile), titanium dioxide, Al, Hg & often coexisting virus infections (CMV, HPV 16 and 18), bacterial infections etc. in the brain and corresponding safe individualized effective treatment. *Acupunct Electrother Res*. 2015;40:157–187.
45. Haller SL, Peng C, McFadden G, Rothenburg S. Poxviruses and the evolution of host range and virulence. *Infect Genet Evol*. 2014;21:15–40.
46. Afonso PP, Silva PM, Schnellrath LC, et al. Biological characterization and next-generation genome sequencing of the unclassified Cotia virus SPAN232 (Poxviridae). *J Virol*. 2012;86:5039–5054.
47. Huang CR, Lien CY, Tsai WC, et al. The clinical characteristics of adult bacterial meningitis caused by non-*Pseudomonas* (Ps.) *aeruginosa* *Pseudomonas* species: a clinical comparison with *Ps. aeruginosa* meningitis. *Kaohsiung J Med Sci*. 2018;34:49–55.
48. Perz JF, Craig AS, Stratton CW, Bodner SJ, Phillips WE Jr, Schaffner W. *Pseudomonas putida* septicemia in a special care nursery due to contaminated flush solutions prepared in a hospital pharmacy. *J Clin Microbiol*. 2005;43:5316–5318.
49. Oosterom N, Nijman J, Gunkel J, et al. Neuro-imaging findings in infants with congenital cytomegalovirus infection: relation to trimester of infection. *Neonatology*. 2015;107:289–296.
50. Claverie JM, Makalowski W. Alu alert. *Nature*. 1994;371:752.
51. Hassold T, Hunt P. To err (meiotically) is human: the genesis of human aneuploidy. *Nat Rev Genet*. 2001;2:280–291.
52. Woolfit M, Iturbe-Ormaetxe I, Brownlie JC, et al. Genomic evolution of the pathogenic *Wolbachia* strain, wMelPop. *Genome Biol Evol*. 2013;5:2189–2204.
53. Nystrom-Lahti M, Kristo P, Nicolaides NC, et al. Founding mutations and Alu-mediated recombination in hereditary colon cancer. *Nat Med*. 1995;1:1203–1206.
54. Veiga-Fernandes H, Mucida D. Neuro-immune interactions at barrier surfaces. *Cell*. 2016;165:801–811.
55. Dixon JR, Gorkin DU, Ren B. Chromatin domains: the unit of chromosome organization. *Mol Cell*. 2016;62:668–680.
56. Fasolino M, Zhou Z. The crucial role of DNA methylation and MeCP2 in neuronal function [published online ahead of print May 13, 2017]. *Genes (Basel)*. doi:10.3390/genes8050141.
57. Saha A, Wittmeyer J, Cairns BR. Chromatin remodelling: the industrial revolution of DNA around histones. *Nat Rev Mol Cell Biol*. 2006;7:437–447.
58. Caudron-Herger M, Pankert T, Rippe K. Regulation of nucleolus assembly by non-coding RNA polymerase II transcripts. *Nucleus*. 2016;7:308–318.
59. Cohn LB, Silva IT, Oliveira TY, et al. HIV-1 integration landscape during latent and active infection. *Cell*. 2015;160:420–432.
60. Muotri AR, Zhao C, Marchetto MC, Gage FH. Environmental influence on L1 retrotransposons in the adult hippocampus. *Hippocampus*. 2009;19:1002–1007.