Computational Detection of Homologous Recombination Hotspots in X-Chromosome Autism-Associated Genes

A. Ard[a], S. Bwabye[c], K. Koyama[c], N. Doan[b], M. A. LaMadrid[c], T. A. Deisher[c]

[a]University of Portland, Portland, OR; [b]Seattle University, Seattle, WA; [c]Sound Choice Pharmaceutical Institute, Seattle, WA

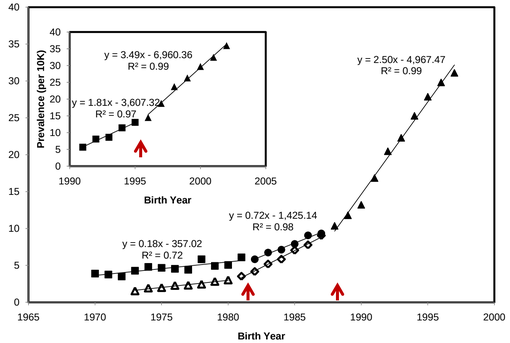
Introduction

Recently, changepoint analyses by Environmental Protection Agency (EPA) scientists (McDonald & Paul 2010) identified changepoint birth years in autistic disorder data from US and Denmark. We have carried out further analyses and found additional changepoints, shown in Figure 1 for the US, summarized in Table 1. In the countries studied, the only universal environmental cause correlated with the changepoint years that we have identified is the introduction of vaccines containing human DNA residuals.

The safety of human DNA residuals has been debated for 50 years (Sheng et al. 2009). Potential dangers of the residuals include auto-immune reactions to the non-host human DNA or improper integration of DNA fragments into the host genome or host mitochondrial genome during base lesion repair by homologous recombination.

This study focuses on improper integration of the residual DNA as a possible contributor to autism, particularly in genetically susceptible infants. It is known from gene therapy studies that injected naked DNA can be transported to the brain (Wang et al. 2001); that improperly integrated therapeutic DNA has caused cancer in young children(Hacein-Bey-Abina et al. 2008); and that shorter DNA fragments have a higher probability of entering the nucleus (Lechardeur et al. 2002). To investigate whether improper DNA integration can contribute to autism, we are undertaking the following: (1) measure the amount and length distribution of residual human DNA in vaccines; (2) predict sites of DNA insertion via homologous recombination (HR) and measure insertion rates; (3) model how brain cell function might be affected, either via loss of the ability to make proper connections or via selective growth of cells with improperly integrated DNA at the expense of healthy cells; (4) conduct epidemiology studies comparing autism rates in children injected with vaccines containing human DNA residuals.

Fig. 1: Changepoint analysis for US(DOE) and CA(DDS) Autistic Disorder



Arrows point to changepoint years obtained using hockey-stick analysis. Filled points are from California (CA) Department of Developmental Services (DDS), open points are from Department of Education (DOE), 19 year olds. Inset is adapted from CA DDS data in Schechter & Grether (2008).

Table 1: Events related to vaccines with human DNA residuals

Country	Changepoint (Birth Year)	Vaccine-related event
Denmark	1988	2 dose MMR introduced 1987
USA (CA DDS)	1988	MMR-II booster recommendation (1989); increased MCV compliance
USA (CA DDS)	1995	Varivax (chicken pox vaccine) licensed in 1995
USA (DOE, CA DDS)	1981, 1982	Meruvax-II, MMR-II licensed in 1979; market exclusivity by early 1980s
Japan	1988*	Introduction of chicken pox vaccine in 1988

*Changepoint algorithm could not detect Japan changepoint due to scatter. However, the data (McDonald & Paul 2010; Honda et al. 2005; Ohtaki et al. 1992; Tanoue et al. 1988; Matsuishi et al. 1987; Ishii et al. 1988;Hoshino et al. 1982) are suggestive of a slope change starting near 1988.

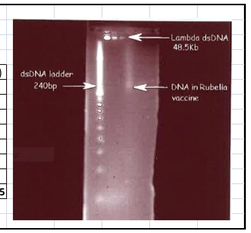
Methods and Results

DNA levels and lengths
Vials of Meruvax II (rubella, Merck&Co. Inc) and Havrix (hepatitis A, Glaxo Smith Kline Biologicals) were heat inactivated by placement in a 60degC water bath for 2 hours. Meruvax contents were reconstituted in Tris-EDTA (TE), ph8, then loaded onto 4% agarose gel. Havrix came as a suspension. Human DNA was isolated using ethanol precipitation, then resuspended in TE. DNA was loaded onto 4% agarose gel. After electrophoresis, gels were stained with SYBR Gold dye (Invitrogen). Human DNA was quantified by labeling double stranded DNA (dsDNA) with picogreen (Invitrogen) and single-stranded (ssDNA) with oligreen (Invitrogen), then reading with a spectrofluorometer.

Fig. 2: Levels and residual size (SYBR gold) of human dsDNA (picogreen assay) and ssDNA (oligreen assay) in Havrix (HepA) and Meruvax II (Rubella)

2. Rubella
SS DNA and DsDNA contents per vial
Name of Vaccine: Meruvax II
Manufactured by Merk & Co. Inc.

	ssDNA (ng/vial)	dsDNA (ng/vial)	size (bp)
Vial 1	135.66	59.34	
Vial 2	145.82	24.54	
Vial 3	-	36.052	240
Vial 4	-	20.064	190
Vial 5	118.54	-	
vial 6	164.40	-	
vial 7	145.82	-	
Average	**142.05**	**35.00**	**215**

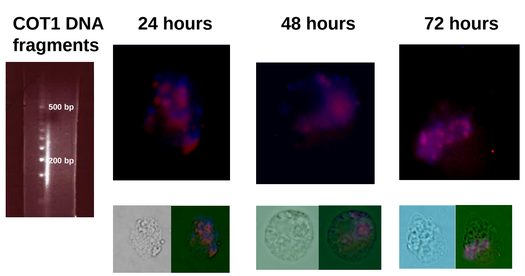


1. Hepatitis A
SS DNA and DsDNA contents per vial
Name of Vaccine: HAVRIX
Manufactured by GlaxoSmithKline Biologicals

	ssDNA (ng/vial)	dsDNA (ng/vial)	size (bp)
Vial 1	844.58	112.69	
Vial 2	104.88	10.69	
Vial 3	213.6	15.32	
Vial 4	213.6	15.32	
Vial 5	213.6	15.32	
Vial 6	165.43	73.33	
Vial 7	176.31	7.49	>48.5K
Average	**276.00**	**35.74**	

Varivax dsDNA levels are >2 ugs (Merck Summary Basis for Approval). An entire human genome consists of approximately 6-7pg of nuclear DNA. Varivax human dsDNA levels are 300,000X the amount of one human genome.

Fig. 3:Human DNA accumulation in nucleus of human U937 cells



On day (-1), U937 cells were permeabilized with 0.2% saponin and then treated with DAPI to inhibit cell proliferation and to label endogenous cellular DNA blue. On day 0, 2 ugs Cy3-labeled red COT1 DNA fragments were added to the culture. Nuclear COT1 DNA accumulation (red) is evident after 24 hours and persists out to 72 hours.

Recombination Hotspots
Chromosomal coordinates of hotspots (Myers et al. 2004) were overlaid with coordinates (transcription starts and ends) of autism-associated genes downloaded from the ACGMAP website. This procedure finds kb-length hotspot regions in the genes. More localized searches for hotspot motifs were done using BLAST. Gene coordinates are from build 36. Initial focus is on X-chromosome genes due to the >3:1 male:female ASD ratio.

Fig. 4: X-chromosome autism-associated genes with recombination hotspots.

Gene	Function or Involvement with Disease	Tissue Specificity
GRPR: Gastrin-releasing peptide receptor	Receptor for gastrin releasing peptide (GRP). This receptor mediates its action by association with G proteins that activate a phosphatidylinositol-calcium second messenger system.	Highly expressed in pancreas. Also expressed in stomach, adrenal cortex and brain.
NLGN3: neuroligin-3	Neuronal cell surface protein thought to be involved in cell-cell-interactions by forming intercellular junctions through binding to beta-neurexins. May play a role in formation or maintenance of synaptic junctions. May also play a role in glia-glia or glia-neuron interactions in the developing peripheral nervous system.	Brain
NLGN4X: neuroligin-4, X-linked	Putative neuronal cell surface protein involved in cell-cell-interactions.	Expressed at highest levels in heart. Expressed at lower levels in liver, skeletal muscle and pancreas and at very low levels in brain.
IL1RAPL1: X-linked interleukin-1 receptor accessory protein-like 1	Defects in IL1RAPL1 are the cause of mental retardation X-linked type 21 (MRX21) [MIM:300143]. Mental retardation is a mental disorder characterized by significantly sub-average general intellectual functioning associated with impairments in adaptative behavior and manifested during the developmental period. Non-syndromic mental retardation patients do not manifest other clinical signs.	Detected at low levels in heart, skeletal muscle, ovary, skin, and in amygdala, caudate nucleus, corpus callosum, hippocampus, substantia nigra and thalamus. Detected at very low levels in tonsil, prostate, testis, small intestine, placenta, colon and fetal liver.
AFF2	Defects in AFF2 are the cause of FRAXE [MIM:309548]. FRAXE is an X-linked form of mental retardation. Loss of FMR2 expression is correlated with FRAXE CCG_n expansion. Normal individuals have 6-35 copies of the repeat, whereas cytogenetically positive, developmentally delayed males have more than 200 copies and show methylation of the associated CPG island.	Brain (most abundant in hippocampus and amygdala), placenta and lung.

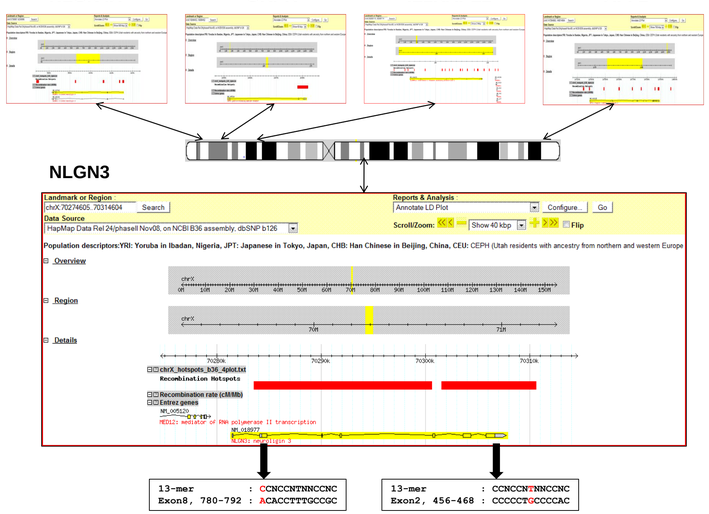
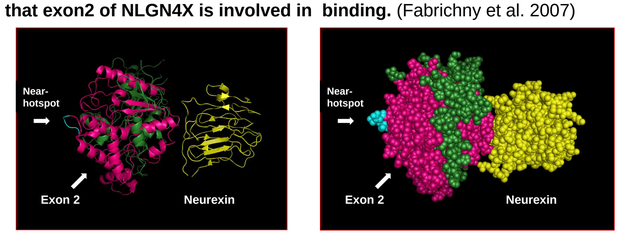


Fig. 5:3-dimensional models for NLGN 4X and NRXN1 (neurexin) demonstrate that exon2 of NLGN4X is involved in binding. (Fabrichny et al. 2007)



Discussion

Changepoint analysis of autism disorder demonstrates a temporal correlation with events associated with human DNA residuals in vaccines. The levels of residual DNA are well over FDA-recommended limits. To reduce the dangers of residual DNA, recommendations were made to fragment the DNA. Unfortunately, in vitro studies in model organisms have shown that shorter fragments have a higher chance of entering the nucleus. Cell culture experiments are in progress to determine the rate and sites at which these residual DNA fragments integrate into the genome.

Our preliminary bioinformatic analysis has identified sites at which these DNA residuals might integrate into the genome and predicted that disruption of exon 2 of NLGN4X could alter binding to neurexin. Neuroligin binding to neurexin is critical for synapse maturation and function in the brain. Across the entire genome, the vast majority of recombination hotspots are located outside the transcribed regions of a gene (Myers et al. 2004). In contrast, we find 5 of 15 autism-associated X-chromosome genes contain hotspots within the transcribed regions. Among all 238 published autism-associated genes, 119 genes have a combined total of 536 hotspots within transcribed regions. Moreover, we find almost-perfect-matches to the most common hotspot motif (Myers et al. 2008) inside exons of two X-chromosome neuroligin genes. Mouse models have demonstrated that loss of binding of NLGN4X to neurexin leads to deficits in social interactions and communication that are similar to autism spectrum disorder (Jamain et al. 2008).

Summary

1. Meruvax-II contains >140ng/vial ssDNA and >30ng/vial dsDNA, with average lengths of 215bp. Havrix contains >270ng/vial ssDNA and >30ng/vial dsDNA. The FDA-recommended amounts are 10ng/dose.
2. There are 5/15 autism-associated genes in the X-chromosome with recombination hotspots inside the transcribed regions.
3. NLGN3 (exons 2,8) and NLGN4X (exons 2,3) contain near-matches to the most common recombination hotspot motif in humans. Structural modeling shows that exon 2 is involved in the binding to neurexin (NRXN1), which is important for synapse formation.

References

1. *Autism Candidate Gene Map (ACGMAP) Database*. http://www.meddean.luc.edu/node/375 (accessed July 2009).
2. Dept. of Education, Office of Special Education and Rehabilitative Services. *www.ed.gov*. http://www2.ed.gov/about/offices/list/osers/osep/research.html (accessed 2009).
3. Fabrichny IP, Leone P, Sulzenbacher G, Comoletti D, Miller MT, Taylor P, Bourne Y, Marchot P. "Structural analysis of the synaptic protein neuroligin and its beta-neurexin complex: determinants for folding and cell adhesion." *Neuron* 56, no. 6 (2007): 979-91.
4. Hacein-Bey-Abina S, Garrigue A, Wang GP, Soulier J, Lim A, Morillon E, Clappier E, Caccavelli L, Delabesse E, Beldjord K, Asnafi V, MacIntyre E, Dal Cortivo L, Radford I, Brousse N, Sigaux F, Moshous D, Hauer J, Borkhardt A, Belohradsky BH, Wintergerst U,. "Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1." *J Clin Invest* 118, no. 9 (2008): 3132-42.
5. Honda H, Shimizu Y, Rutter M. "No effect of MMR withdrawal on the incidence of autism: a total population study." *Journal of Child Psychology and Psychiatry* 46, no. 6 (2005): 572-79.
6. Hoshino Y, Kumashiro H, Yashima Y, Tachibana R, Watanabe M. "The epidemiological study of autism in Fukushima-ken." *Folia Psychiatr Neurol Jpn* 36, no. 2 (1982): 115-24.
7. Ishii, T, and O Takahashi. "The epidemiology of autistic children in Toyota, Japan." *Jpn. J. Child Adolesc. Psychiat.* 24 (1983): 311-321.
8. Jamain S, Radyushkin K, Hammerschmidt K, Granon S, Boretius S, Varoqueaux F, Ramanantsoa N, Gallego J, Ronnenberg A, Winter D, Frahm J, Fischer J, Bourgeron T, Ehrenreich H, Brose N. "Reduced social interaction and ultrasonic communication in a mouse model of monogenic heritable autism." *Proc Natl Acad Sci* 105, no. 5 (2008): 1710-5.
9. Lechardeur D, Lukacs GL. "Intracellular barriers to non-viral gene transfer." *Curr Gene Ther* 2, no. 2 (2002): 183-94.
10. Matsuishi T, Shiotsuki Y, Yoshimura K, Shoji H, Imuta F, Yamashita F. "High prevalence of infantile autism in Kurume City, Japan." *J Child Neurol* 2, no. 4 (1987): 268-71.
11. McDonald ME, Paul JF. "Timing of increased autistic disorder cumulative incidence." *Environ Sci Technol* 44, no. 6 (2010): 2112-8; Supporting info: http://pubs.acs.org/doi/suppl/10.1021/es902057k.
12. Myers S, Bottolo L, Freeman C, McVean G, Donnelly P. "A fine-scale map of recombination rates and hotspots across the human genome." *Science* 310, no. 5746 (Oct 2005): 321-4.
13. Myers S, Freeman C, Auton A, Donnelly P, McVean G. "A common sequence motif associated with recombination hot spots and genome instability in humans." *Nat Genet* 40, no. 9 (2008): 1124-9.
14. Ohtaki E, Kawano Y, Urabe F, Komori H, Horikawa M, Yamashita Y, Katfuchi Y, Kuriya N, Matsuishi T, Yamashita F. "The prevalence of Rett syndrome and infantile autism in Chikugo District, the southwestern area of Fukuoka prefecture, Japan." *J Autism Dev Disord. 1992 Sep;22(3):452-4* 22, no. 3 (1992): 452-4.
15. Schechter R, Grether JK. "Continuing increases in autism reported to California's developmental services system: mercury in retrograde." *Arch Gen Psychiatry* 65, no. 1 (Jan 2008): 19-24; http://archpsyc.ama-assn.org/cgi/reprint/65/1/19.
16. Services, California Health and Human Services Agency Department of Developmental. "Autistic Spectrum Disorders. Changes in California Caseload.AnUpdate: 1999 through 2002." Sacramento, CA, 2003.
17. Sheng-Fowler L, Lewis AM Jr, Peden K. "Issues associated with residual cell-substrate DNA in viral vaccines." *Biologicals* 37, no. 3 (2009): 190-5.
18. Tanoue Y, Oda S, Asano F, Kawashima K. "Epidemiology of infantile autism in southern Ibaraki, Japan: differences in prevalence in birth cohorts." *J Autism Dev Disord* 18, no. 2 (1988): 155-66.
19. Wang S, Ma N, Gao SJ, Yu H, Leong KW. "Transgene expression in the brain stem effected by intramuscular injection of polyethylenimine/DNA complexes." *Mol Ther* 3, no. 5 Pt 1 (2001): 658-64.

Conflict of Interest: None