Supplementary Information

Induced Chromosome Deletions Cause Hypersociability and Other Features of Williams-Beuren Syndrome in Mice

Hong Hua Li[1], Madhuri Roy[2,7], Unsal Kuscuoglu[1,7], Corinne M. Spencer[4,7], Birgit Halm[1], Katharine Carlsmith Harrison[1], Joseph Henry Bayle[3], Alessandra Splendore[1], Feng Ding[1], Leslie A. Meltzer[2], Elena Wright[1], Richard Paylor[4], Karl Deisseroth[2,5], and Uta Francke[1,6,*]

Department of [1]Genetics, [2]Bioengineering, [3]Pathology, [5]Psychiatry, [6]Pediatrics, Stanford University School of Medicine, Stanford, CA; [4]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX

Supplementary Methods and Results

1. Embryonic Stem Cell Culture

HPRT-deficient AB2.2' ES cells and mitotically inactivated mouse embryonic fibroblast feeders derived from STO cells were purchased from Stratagene (Mouse Kit, catalog # 745002). Feeder monolayers were prepared as recommended by the manufacturer. AB2.2 ES cell medium contained 10^3 U/ml ESGRO LIF (Chemicon, catalog # ESG1107) in addition to the components of feeder media. ES cell medium was conditioned by feeder monolayer for at least 3-4 days prior to plating ES cells. ES cell medium was replaced daily with fresh medium until cells reached 70-80% confluence.

2. Genotyping by PCR

To identify the product of Cre-mediated recombination for both DD and PD, we designed PCR primers unique to *5'Hprt* (F1: 5'-cgtcgtgattagtgatgatgaacc) and *3'Hprt* (R1: 5'-cactttgggaggctgggtaac) vector sequences. The 1loxP alleles produce a PCR fragment of 1374 bp while no product is generated from the WT allele. After the 1loxP genotypes of the founder mice had been confirmed, we designed primers that detect the junction of genomic sequence and *5'Hprt* vector sequence to identify the presence of each of the targeted chromosomes. Our protocol to identify WT, DD, PD and D/P genotypes uses a multiplex PCR with five primers: F2, 5'-agaaggtccaccagctcagt-3' (mouse *Limk1* genomic region), F3, 5'-aacagcctgccaacttcttt-3' (mouse *Gtf2i* genomic region) and R2, 5'-tttctgtggggcaaaatgta-3' (*5'Hprt* vector, human sequence). As control, primer pair (F4/R4) is specific to the *Hprt* minigene sequence and generates a product in 1loxP and 2loxP but not in WT; F4: 5'-cacatttcttctcaagcactggc-3' and R4: 5'-cactttgggaggctgggtaac-3'. Expected product sizes are 352 bp (for PD, F3/R2), 458 bp (for DD, F2/R2)

and 721 bp (for the *Hprt* control, F4/R4 primer pair). PCR to amplify the complete *Hprt* minigene was performed using primer pair F1/R1 at an annealing temperature of 60°C with 1.5mM $MgCl_2$. In a typical multiplex PCR for genotyping, we used 2.5 mM $MgCl_2$ in a 25*m*l reaction with five primers, F2, R2 (2x amount), F3, F4, and R4 at 60°C annealing temperature.

3. RNA isolation and Real-Time Quantitative RT-PCR (qRT-PCR)

Seventeen mice (4 D/P, 4 PD/+, 4 DD/+, and 5 WT) from three different litters were euthanized at 13 weeks of age using halothane and their brains were dissected. The right hemispheres were used for RNA extraction using TRIzol (Invitrogen). RNA from mice of the same genotype was pooled together in equal concentrations, and 2 mg of pooled RNA were used for first-strand cDNA synthesis with SuperScript II (Invitrogen). The resulting four cDNA pools, each representing 4 mice, were diluted 1:100, and 5ml were used as template in a qRT-PCR reaction. We performed qPCR by using SybrGreen Mix (Applied Biosystems). Typically, in a 25ml reaction, we used 5ml cDNA (1:100), 12.5ml SybrGreen master mix, 2ml primer (10 mM, F + R), and 5.5ml water. Each sample was analyzed in triplicate in an ABI5700 thermocycler (Applied Biosystems). Threshold cycle (Ct) values were set manually, exported and analyzed with qBASE software (v. 1.3.5; http://medgen.ugent.be/qbase). We designed transcript-specific primers for 15 genes in the WBS critical region that are expressed in brain by using Primer3 software v. 0.3.0. A combination of four reference primers was used for normalization (Suppl. Table S1)

Table S1. Primer sequences for qRT-PCR

Gene	Forward	Reverse	PCR product
Primer sequences for tested genes			**Size (bp)**
Fzd9	AACATAGGCTAGGCTGTGAGTTATG	GTGTCTCACTTGTCTGACCTTGAC	111
Baz1b	CAGAAAACGCTAGGTGCAAAG	TCAGACAGAACAGGTGGAAGG	99
Bcl7b	ATTCCTCCCTCCTCCTTGAAT	GACTTGAGTTGGTGCTGCTGT	99
Tbl2	GGTTCGTAAGGAGAAGCCTCA	CTGCTGAAGTCCATGCAAGAT	102
Wbscr18	GTACTACCGGCAGAGCTTCCT	AGACCAGGTAAGCTTCGGAGA	99
Wbscr22	TGGTAGACTTCCCCAACAGTG	CTGGTCTGCATCCTGACTTTC	111
Stx1a	AACCCCGATGAGAAGACAAAG	GATGCTCTGCTCAATGCTCTT	105
Abhd11	AAGGCAATGGTTCAGAGGACT	TGACTCATGGCCTCATAGCTT	104
Limk1	GGACAAGCGGCTAAACTTCA	AAGCTGACCCTCTGACTCCA	105
Eif4h	GGACTTCGATACCTACGACGA	GCTCTGTGGGTAGCTCCTTCT	114
Rfc2	ACTGAACAACCTGCAGTCCAC	GATCATCTCCTTCACCAGCAG	106
Cyln2	GAAGGTGGATCTGTCCAATCA	TCCAAGTCTCCTTTGGTGATG	99
Gtf2i	TCACCAAACTGAGGAAGATGG	TTTTGGTACGGTACAGCCTTG	94
Gtf2ird1	GGGACTGCTTGGTCAGACAT	TCCTCTGGCTCGATTTTGAT	101
Hip1	ACCTGGTGGACAAGGAGATG	CTCCTGCTCGGGACTTACTG	93
Primer sequences for reference (normalization) genes			
Atp10a	GGCGACTTTGTCCGTCTTTG	CAGGTTAGCAGTCTCGATGTG	105
Rps28	TAGGGTAACCAAAGTGCTGGGCAG	GACATTTCGGATGATAGAGCGG	103
Reln	CTGTGTCATACGCCACGAACA	GGGGAGGTACAGGATGTGGAT	101
Snrpn	TGCTACGTGGGGAGAACTTG	CCTGGGGAATAGGTACACCTG	157

4. Behavior testing at BCM

A test battery strategy developed by Paylor and others (Crawley and Paylor, 1997, Paylor et al., 2006) was utilized to assess multiple domains of CNS function. Behavioral testing was performed between 8 AM and 3 PM. At the start of testing the mice were approximately 8 weeks of age. The experimenter was blind to the genotypes. Tests were performed in the following order with 1-3 days between tests: open-field, light-dark, marble-burying, rotarod, acoustic startle/ PPI, GAP detection, conditioned fear, hotplate, tube test, partition test, and direct social interaction test.

Marble burying. This paradigm is used to assess anxiety-related responses and obsessive-compulsive like responses (Broekkamp et al., 1986; Njung'e and Handley, 1991). A standard mouse cage is filled with 10 cm of corn-cob bedding and 20 small (1.5-2 cm) black marbles are placed equidistant on top of the bedding. A mouse is placed in this cage and allowed to explore and bury the marbles. After 30 minutes the mouse is removed and we measure the number of marbles that were buried (a marble is said to be 'buried' if more than 50% of it is under the bedding). No significant differences in marble-burying behavior ($p > 0.05$) were found among the four genotypes (data not shown).

Light-dark exploration. This test revealed no significant differences in the total number of transitions between the light and dark side ($p > 0.3$) in the light-dark box among the four genotypes (data not shown). We do not report latency to enter the dark side and the time in the dark (or light) because these measures are extremely sensitive to outliers that display extreme levels of heightened anxiety. It is true that the number of transitions is sensitive to activity differences, but it is not sensitive to outliers for 'anxiety' (Bouwknecht and Paylor (2002). We did not notice differences in activity levels between the genotypes during these tests.

GAP detection. The GAP detection test is very similar to the prepulse (PPI) test with the exception that instead of low-level prepulse sound presented before the startle, a 'gap' in sound of different duration is presented. Previous research (Clark et al., 2000; Friedman et al., 2004; Peiffer et al., 2004; Threlkeld et al., 2006) has shown that the absence of a sound immediately preceding a loud startle noise also attenuates the startle response. In the GAP detection assay, mice are exposed to trials that either have a loud sound, or a loud sound preceded by a gap of either 20, 40, or 100 ms. Percent inhibition of the startle response by the GAP is calculated by using the same formula used for the PPI assay.

References

Bouwknecht JA and Paylor R (2002) Behavioral and physiological mouse assays for anxiety: a survey in nine mouse strains, *Behav Brain Res* **136:** 489-501.

Broekkamp CL, Rijk HW, Joly-Gelouin D, Lloyd KL (1986) Major tranquillizers can be distinguished from minor tranquillizers on the basis of effects on marble burying and swim-induced grooming in mice. *Eur J Pharmacol* **126:** 223-229

Clark MG, Sherman GF, Bimonte HA, Fitch RH (2000) Perceptual auditory gap detection deficits in male BXSB mice with cerebrocortical ectopias. *Neuroreport* **11:** 693-696

Crawley JN, Paylor R (1997) A proposed test battery and constellations of specific behavioral paradigms to investigate the behavioral phenotypes of transgenic and knockout mice. *Horm Behav* **31:** 197-211

Friedman JT, Peiffer AM, Clark MG, Benasich AA, Fitch RH (2004) Age and experience-related improvements in gap detection in the rat. *Brain Res Dev Brain Res* **152:** 83-91

Njung'e K, Handley SL (1991) Evaluation of marble-burying behavior as a model of anxiety. *Pharmacol Biochem Behav* **38:** 63-67

Paylor R, Spencer CM, Yuva-Paylor LA, Pieke-Dahl S (2006) The use of behavioral test batteries, II: effect of test interval. *Physiol Behav* **87:** 95-102

Peiffer AM, Friedman JT, Rosen GD, Fitch RH (2004) Impaired gap detection in juvenile microgyric rats. *Brain Res Dev Brain Res* **152:** 93-98

Threlkeld SW, McClure MM, Rosen GD, Fitch RH (2006) Developmental timeframes for induction of microgyria and rapid auditory processing deficits in the rat. *Brain Res* **1109:** 22-31

5. Growth and developmental phenotypes

Dental abnormalities. Individuals with WBS may have small, malformed or absent teeth (Axelsson et al., 2003). In the *Wbscr* deletion mice, we observed no differences in number, length and structure of teeth.

Reference

Axelsson S, Bjornland T, Kjaer I, Heiberg A, Storhaug K (2003) Dental characteristics in Williams syndrome: a clinical and radiographic evaluation. Acta Odontol Scand **61:** 129-136

Postnatal growth retardation in deletion mice.

As shown in Figure S1, all female deletion mice were significantly smaller relative to WT controls at three weeks of age (ANOVA, $p < 0.05$ for DD and PD, and $p < 0.0001$ for D/P). D/P females also weighed significantly less compared to WT during the 2-8 week period. In males, weight differences were more obvious. D/P males were smaller from five weeks onward, with $p < 0.05$ up to 12 weeks and $p < 0.01$ from 13 to 24 weeks. For PD, the differences were significant ($p < 0.05$) from 1-7 weeks and from 21-24 weeks. In contrast, DD did not differ significantly from WT in their body weight.

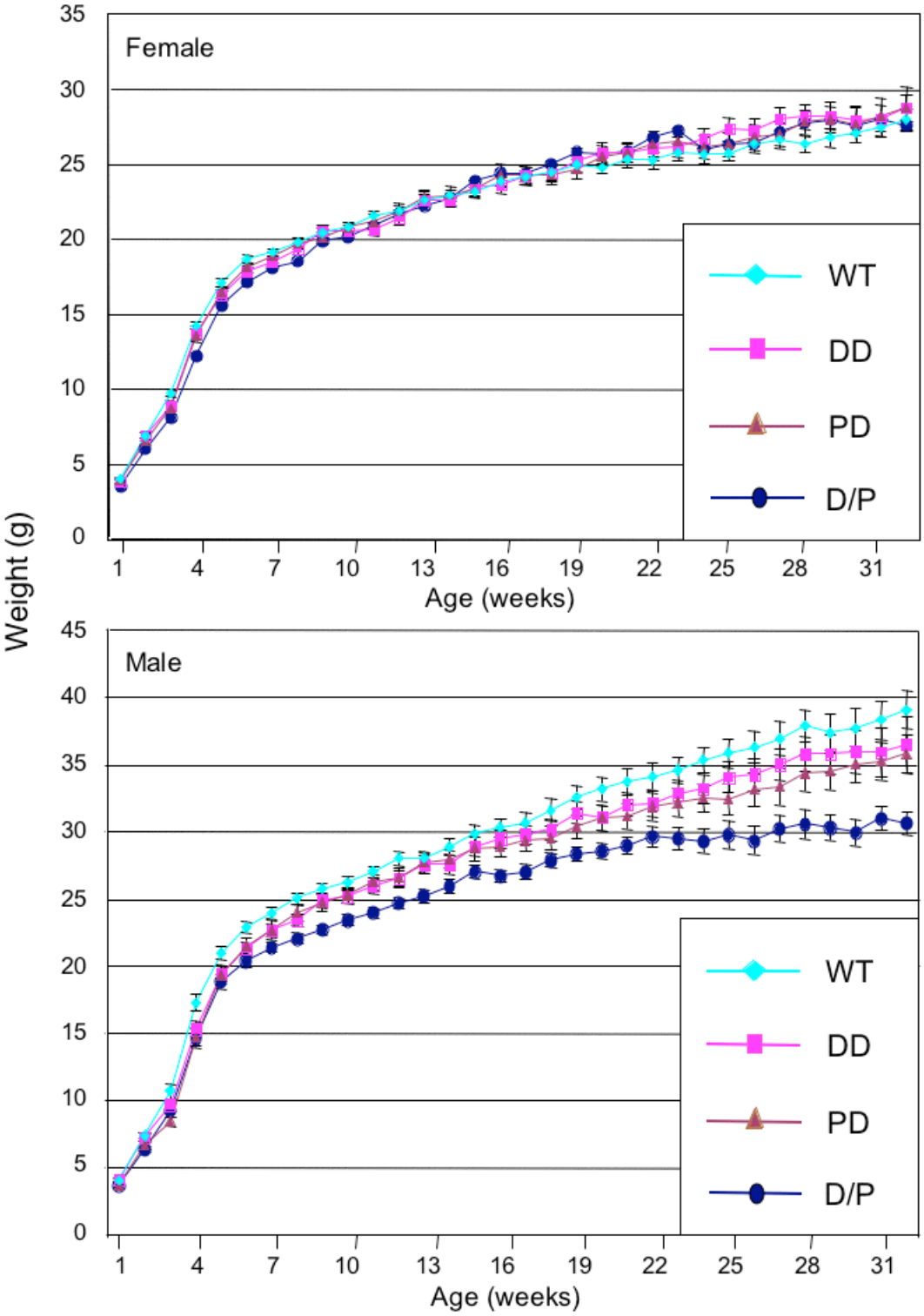

Figure S1.
Weight curves for females (top, N = 8-29) and males (bottom, N = 8-25). Error bars represent SEM.