Natural copy number differences of tandemly repeated small nucleolar RNAs in the Prader-Willi syndrome genomic region regulate individual behavioral responses in mammals

Short title: Behavioral regulation through snoRNAs

Maryam Keshavarz, Rebecca Krebs-Wheaton, Peter Refki, Yoland Savriama, Yi Zhang, Anja Guenther[2], Tanja M. Brückl[3], Elisabeth B. Binder[3,4], Diethard Tautz#

Max-Planck Institute for Evolutionary Biology, Plön, Germany

2: previous address: Bielefeld University, Department of Animal Behavior, Bielefeld, Germany

3: Dept. of Translational Research in Psychiatry, Max-Planck Institute of Psychiatry, Munich, Germany

4: Dept. of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta GA

#corresponding author: tautz@evolbio.mpg.de

Abstract

The Prader-Willi Syndrome (PWS) gene region is an imprinted gene complex involved in behavioral, metabolic and osteogenic functions. We have analyzed here the variation of two families of regulatory small nucleolar RNAs (SNORD115 and SNORD116) that are coded within the PWS and are expressed from the paternal chromosome. They are organized in two tandemly repeated clusters which are naturally copy number variable between individuals. We find that the copy numbers at these loci correlate with repeatable individual test scores for anxiety that are considered to constitute a component of the "personality" of individuals. We show this for different populations and species of mice, cavies and for the anxiety component of personality tests in humans. This is also the case for an inbred mouse strain (C57Bl6) implying that copy number variation creates phenotypic variability even in an isogenic background. In transcriptome data from brain samples of this strain we find SNORD copy-number correlated regulation of target genes that are known to be involved in

influencing behavior. SNORD115 has previously been suggested to regulate splicing of the serotonin receptor *Htr2c* and we confirm this in our data. For SNORD116 we provide evidence that it regulates the expression level of the chromatin regulator *Ankrd11*, which itself regulates GABA receptors, metabolic pathways, cell differentiation and osteogenesis. Intriguingly, we find that craniofacial shapes in mice correlate also with SNORD116 copy numbers. New copy number variants are generated at very high rates in mice, possibly at every generation, explaining why conventional genetic mapping could not detect this association. Our results suggest that the variable dosage of two regulatory RNAs are major determinants of individual behavioral differences and correlated traits in mammals.

Introduction

The study of consistent individual differences in behavior has flourished over the last decades, because it has been recognised as a major contributor to differences in survival and fitness among individuals (Réale 2007; Wolf and Weissing 2012; Penke and Jokela 2016; Forkosh et al. 2019). The focus has been on traits showing relatively enduring patterns, both when measured across several occasions, as well as under different circumstances (Wilson 1998; Reale et al. 2007; Forkosh et al. 2019). This acknowledges that individuals of the same species, population, sex and even age or prior experiences show predictable differences in responding in certain ways under certain circumstances. Cornerstones of this emerging research field are to confirm that a) individuals vary from each other in behavioral responses, e.g. levels of anxiety in a specific test situation, b) this individual behavioral response shows temporal consistency and repeatability, e.g. an individual which reacts more scared then others to an Open Field Test at one occasion should react also more scared when tested at a later point in time and c) individual reactions should be consistent across different (test) situations, i.e. an individual that reacts more scared in an Open Field Test should also react more scared when confronted with an Elevated Plus Maze.

Such individual differences in behavioral patterns were found for many animal species and they resemble those found by psychologists studying human behavior , although there is an active discussion on whether they are really comparable (Uher 2011). Several studies have shown a heritable component for such behavioral differences implying some form of underlying genetic mechanisms (Dochtermann et al. 2015; Sanchez-Roige et al. 2018). Interestingly, however, such individual variation is also known to occur in genetically isogenic strains (Jakovcevski et al. 2008; Freund et al. 2013, 2015) which suggests that non-genetic mechanisms are also involved. On the other hand, a genetic locus with high rates of change between generations could potentially explain both observations.

A locus that has generally been implicated in behavioral traits in humans is the Prader–Willi syndrome (PWS) region. Mutations in this regions cause a neuro-developmental disorder which leads to several abnormalities in cognitive behaviors such as social communication, speech, anxiety, intellectual ability and decision making, but also to metabolic syndromes and craniofacial shape changes (Cassidy et al. 2012). The region is subject to imprinting, i.e. parentally biased expression of genes (Nicholls et al. 1998). The protein coding genes expressed in this region include *Ube3a*, *Snrpn* and *Magel2,* of which *Ube3a* is expressed from the maternally provided chromosome and *Snrpn* and *Magel2* from the paternally provided one. The PWS includes also several non-coding RNAs (see below). This arrangement has specifically evolved in mammals and is generally conserved among them, including humans (Zhang et al. 2014).

In mice we have previously identified the region between *Ube3a* and *Snrpn* in the PWS as one of two parentally imprinted regions that evolve particularly fast between natural populations and that could explain parentally biased mating preferences between them (Lorenc et al. 2014). Laboratory mice mutant for genes or gene regions from the PWS affect various behaviors including anxiety, activity, ultrasonic vocalization, social interaction, metabolism and foraging (Ding et al. 2008; Nakatani et al. 2009; Qi et al. 2016; Cavaille 2017; Homdli et al. 2019).

The PWS region between *Ube3a* and *Snrpn* includes two small nucleolar RNA (SNORD) gene families which are organized in large, tandemly repeated clusters known as SNORD115 and SNORD116 (Cavaille 2017) (Figure 1). The expression of both SNORD115 and SNORD116 is brain-specific and restricted to the alleles on the paternal chromosome (Ding et al. 2008).

SNORDs are part of a large group of small, metabolically stable snoRNAs which regulate post-transcriptional modification of their target genes (Kiss 2001; Falaleeva et al. 2017). One mechanism for snoRNA interaction is to bind via a complementary antisense region to a target RNA, whereby up to three mismatches between the antisense box and the target region can be tolerated (Kishore et al. 2010). The SNORD115 antisense element exhibits complementarity to the alternatively spliced exon V of the serotonin receptor *Htr2c* which is part of the serotonin regulatory pathway. It has been shown in cell-culture experiments and for PWS patients that there is a positive correlation between SNORD115 expression and exon part Vb usage (Kishore and Stamm 2006). However, ectopic expression of SNORD115 in non-target tissues of mice does not lead to a regulation of alternative splicing of this receptor (Raabe et al. 2019), implying that the interaction is more complex and possibly specific to the tissues of normal expression of SNORD115. Activation of the receptor by serotonin inhibits dopamine and norepinephrine release in certain areas of the brain which regulates mood, anxiety, feeding, and motoneuron functions (Stamm et al. 2017). Human SNORD116 shows a complementary sequence to an *Ankrd11* exon, allowing to suggest a direct regulation (Bazeley et al. 2008). ANKRD11 is a ankyrin repeat domain-containing protein that acts as a transcriptional co-factor with multiple possible target genes (Gallagher et al. 2015), as well as dendrite differentiation in

the cerebal cortex via the BDNF/TrkB signaling pathway (Ka and Kim 2018). But this interaction has only computationally been predicted and has not yet been experimentally confirmed.

Here we ask whether natural copy number variation in SNORD115 and SNORD116 would correlate with individual differences in behavioral traits in mammals. We used standardized tests for anxiety profiles for different populations and species of mice, as well as comparable tests for wild guinea pigs and a questionnaire-based test for humans. We find that there is indeed a strong correlation between copy numbers of the respective SNORD genes and behavioral measures. Intriguingly, this is not only found for wildtype strains, but also for the common laboratory mouse inbred strain C57BL/6J. Using transcriptomic analyses, we show that the predicted regulation of *Htr2c* and *Ankrd11* can indeed be observed and that the network of genes affected by the copy number variation of SNORD116 can explain the behavioral and osteogenic phenotypes. Our data suggest further that new alleles with different copy numbers are generated at an exceptionally high rate. Hence, we conclude that this variation could be a basis for the long sought molecular mechanism for the high variance of individual behavioral traits in families and populations of mammals.

Results

ddPCR to measure repeat numbers

Given the possible molecular links to the regulation of behavioral responses, we were interested to measure copy number variation within the SNORD115 and SNORD116 clusters. Both snoRNAs are part of a larger transcription unit, from which they are processed (Cavaille 2017). In the mouse reference genome, the repeat unit length for SNORD115 is 1.97kb including the RNA coding region with a length of 79nt. For SNORD116 the repeat unit length is 2.54kb and the RNA has a length of 94nt. Not all copies are annotated in the reference genome, but in the genome sequence itself, SNORD115 is represented with 143 copies and SNORD116 with 70 copies, with a 50kb sequence gap (S1 Figure).

Typing copy numbers in this range has long been a technological challenge, explaining also why natural copy number variation in this region has not received much attention so far. The classic quantitative PCR approaches do not have enough power to distinguish e.g. 150 from 160 copies. Also long read sequencing procedures (e.g. PacBio or Oxford Nanopore) come to their limits in this region, since they stretch across several hundred kb, which is beyond the standard limits of these technologies. S2 Figure shows an example of the non-suitable read coverage that is obtained with these technologies in the region spanning *Ube3a-Snrpn*. But a new technology, droplet digital PCR (ddPCR), has been developed to achieve the necessary resolution (Pinheiro et al. 2012; Quan et al. 2018). We have previously validated this technology by comparing predicted copy numbers based on sequencing read depth from whole mouse genome sequences with measured copy numbers based on

ddPCR assays (Pezer et al. 2015) S3 Figure shows that there is a very good correspondence between
145 ddPCR results and read depth calculation for different CNV loci with high copy numbers. A possibly confounding factor for ddPCR would be that primers may not bind equally in all repeat units, due to polymorphisms in the primer binding sequence. To reduce the chance of polymorphism in the primer binding sites, we placed them inside the SNORD RNA coding parts. Still, a survey of all described SNORD variants suggested that some polymorphism exists also in these regions. Hence, we have
150 tested two different sets of primers and found that both yield highly congruent results (see Methods). We have therefore conducted the bulk of the experiments with only one set of primers.

To control the technical replicability, we have run the ddPCR reactions for the individuals in triplicates. These show a very high correspondence, with average standard deviations of 3.54 for SNORD115 and 3.15 for SNORD116 in the *Mus* samples (S1 Table). We tested also biological
155 replicates by using ear samples and brain samples from the same individuals. We found also highly congruent results in these tests (S2 Table), implying also that the copy number in ear samples are a very good proxy for brain samples.

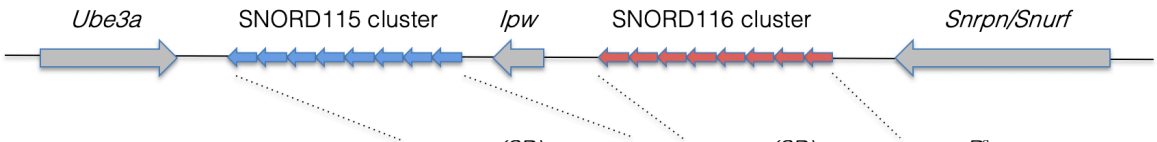
In contrast to the long-range sequencing approaches discussed above, the optical mapping procedure developed by Bionano has at least the chance to generate information for fragments longer than
160 500kb. But it has also its limitations. First, it is restricted to regions where the repeat units include at least one recognition site for the enzyme used for labelling. In the mouse this applies only to the SNORD116 repeat units, which have one recognition site for Nt.BspQI and two sites for Dle1 (these are the two alternative enzymes that are used for labelling). Second, the procedure requires to obtain a substantial amount of high-quality DNA and it is complex and expensive. Still, we have obtained
165 Bionano optical mapping data for seven individuals. Unfortunately, even with this technology and with 100x coverage of the whole genome, we found that reads spanning the whole SNORD116 cluster are very rare. In the end we could get enough suitable fragments to distinguish two alleles for one of the animals only (S4 Figure). For this animal we were able to confirm that the SNORD116 region is indeed a contiguous tandem repeat (i.e. the gap in the reference sequence could be filled with repeat
170 units). Further, we applied ddPCR to this animal and we found that the manually counted number of copies from the optical map (287) matched very closely the value that we obtained with ddPCR (293) (S4 Figure), providing another proof for the validity to use ddPCR. Because of the rareness of full length fragments spanning the full repeat regions and the high costs, optical mapping is not suitable for typing a large number of animals, especially not when the repeat units do not include the
175 recognition sites for labelling (as it is the case for the SNORD115 repeats in mouse and both types of SNORD repeats in humans). Hence, ddPCR is currently the only method that has shown a sufficient power and reproducibility to quantify the copy numbers in tandem repeat regions, which is the key to the results described below (see also discussion).

180 *SNORD copy number variation*

Using the ddPCR assay, we tested first individuals derived from two populations of *M. m. domesticus* that were originally caught in the wild in Germany (CB) and France (MC) and which were kept under outbreeding conditions (Harr et al. 2016). We find an average of 202 SNORD115 copies in CB animals and 170 in MC. A similar difference exists for SNORD116 copies, with 229 in CB and 188 in
185 MC (Figure 1). Although there is a large variation within the populations, these differences between the populations are significant (t-test; p=0.03 for SNORD115 and p=0.015 for SNORD116; all data normally distributed, Shapiro-Wilk normality test p>0.21). Intriguingly, we observe also a strong co-variation of copy number between the two SNORD clusters in both populations, with highly significant correlation coefficients *r* (Figure 1). Such a co-variation is unexpected, since the repeat
190 units are separated by a non-repeated region and since they share no sequence similarities. To test whether a possibly unrecognized technical factor leads to this correlation, we typed also the two other CNV loci in the mouse genome that show particularly high copy-number range (*Sfi1* and *Cwc22* – see S3 Figure) for the same set of individuals. *Sfi1* has between 18-98 copies in the CB population and 18-121 copies in the MC population, *Cwc22* shows high copy number variation in the CB population
195 only (18-99 copies) (Pezer et al. 2015). We find that neither of these two loci shows a correlation with SNORD115 or SNORD116 copy numbers (S3 Table). As a further control for the validity of the correlation between SNORD115 and SNORD116 copy numbers, we compared read coverage from the whole genome sequence data set of *Mus musculus* populations included in Harr et al. (Harr et al. 2016). We find indeed also a high correlation ($r^2 = 0.64$) between read coverage for SNORD115 and
200 SNORD116 in individuals (S5 Figure), supporting the observations from the ddPCR data.

Copy numbers and RNA expression

Given the proven efficiency of ddPCR for copy-number determination, we used it also to quantify RNA expression (Quan et al. 2018) in individuals from the wildtype population. The measured
205 SNORD gene copy numbers show a highly significant regression with the expression of the respective SNORD RNAs in the brain (Figure 2 A, B) (S4 Table), suggesting a direct relationship between copy number and expression level.



	average (SD)	range	average (SD)	range	R^2
M. m. domesticus CB (N=22)	202 (54)	95-290	229 (54)	132-325	0.94
M. m. domesticus MC (N=23)	168 (36)	108-246	188 (53)	86-290	0.87
mouse inbred strain C57BL/6J (N=60)	139 (48)	55-291	185 (51)	71-302	0.74
Apodemus uralensis (wood mouse) (N=32)	413 (170)	160-835	n.d.	n.d.	n.a.
Cavia aperea (wild guinea pig) (N=19)	n.d.	n.d.	10 (2)	6-14	n.a.
humans (N=96)	25 (7)	13-49	27 (4)	19-41	0.53

Figure 1: Structure of the PWS region and copy numbers of snoRNAs. The scheme depicts the genes in the cluster as arrows pointing in the direction of transcription. *Ube3a* is expressed from the maternal chromosome, the others from the paternal one. Note that the transcript structures are actually more complex and not yet fully resolved (Lorenc et al. 2014; Cavaille 2017). The tandem clusters coding for the SNORD115 and SNORD116 snoRNAs are depicted by smaller block arrows (not drawn to size). The table below provides the measures and ranges of copy numbers of the clusters in the different species analyzed in the present study. The correlation coefficients (*r*) between the numbers in the two clusters are listed to the right for the species where these are available.

Further, we asked whether the target RNA expressions correlate with their respective SNORD expression. The target of SNORD115 is the alternatively spliced exon V of *Htr2cr* (Kishore and Stamm 2006), which is split into two halves through the alternative splice – exonVa and exonVb (S6 Figure). We used primers that span specifically the exonVb to exonVI junction and we find a significant positive correlation for this variant with SNORD115 copy numbers in both wildtype populations (Figure 2C). The predicted target of SNORD116 is exon IX of *Ankrd11* (Bazeley et al. 2008) (S6 Figure). Since no alternative splice was described for this so far, we placed the primers close to the predicted SNORD116 binding site and we find a significant correlation with SNORD116 copy numbers (Figure 2 D) (all data in S4Table). Note that these patterns are confirmed and extended by the RNASeq analysis described below.

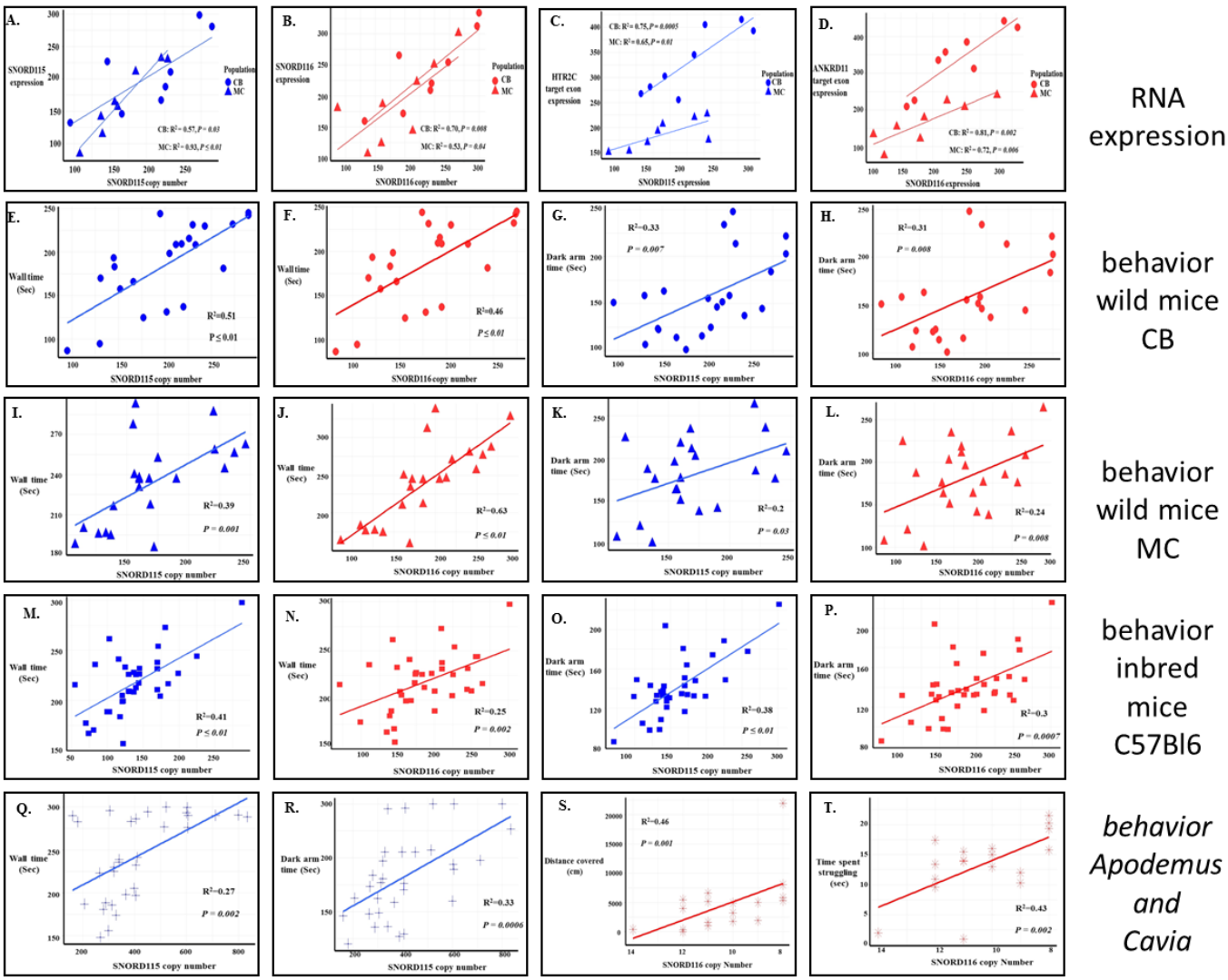


Figure 2: Regression between SNORD115 (blue) and SNORD116 (red) copy numbers and expression, anxiety-like behavior and skull shape. A) and B) Regression with the respective RNA expression values. C) and D) Regression between the respective SNORD expression and the splice-site regulated target RNAs. All values were determined by ddPCR in whole brain RNA preparations from eight outbred animals each of the CB and MC populations. The y-axis shows the RNA molecule copy number per µL input (see Methods) after normalization. β-catenin was used to normalize the data for HTR2C and ANKRD11 expression and SNORD66 was used to normalize the data from SNORD115/116 expression. The animals constitute a subset of the 45 animals used for the overall copy number variation reported in Figure 1 and were chosen to reflect the spread of the variation. E, F, I, J, M, N and Q show regression analysis between Time at wall in the Open Field test and SNORD copy number in wild mice CB (E, F), wild mice MC (I, J), inbred mice C57Bl6/J (M, N) and *Apodemus uralensis* (Q). G, H, K, L, O, P and R show regression analysis between Dark arm time in the Elevated Plus Maze test and SNORD copy number in CB (G, H), MC (K, L), inbred mice C57Bl6/J (O, P) and *Apodemus uralensis* (R). S,T: Regression analysis between SNORD116 copy number and distance covered (cm) from the Open Field (S) and time spent struggling from the Struggle Docility tests (T) in the wild guinea pig *Cavia aperea*. Note that the behavioral tests for cavies differ from the tests for the other species. Lower scores represent higher anxiety, i.e. the X-axis is reversed for better comparability. In all tested rodent species in this study, animals with higher copy numbers have a higher relative anxiety score.

250 *Behavioral tests*

To assess whether the SNORD gene copy number variation correlates also with behavioral scores of the mice, we used two standard tests in mice, the Open Field Test and the Elevated Plus Maze test. Both measure anxiety profiles and have been used before in studies of individual behavior (Jakovcevski et al. 2008; Freund et al. 2013). We determined which measurement scores generated in
255 these tests showed significant repeatability for individuals over the course of the experiment, using intra-class correlation coefficients (see details in Methods). These pre-tests for repeatability were carried out on individuals from the MC population, the C57Bl6/J strain and the wood mouse *Apodemus,* three times each with 4-week intervals between the tests (S5 Table, S6 Table). Among the seven measurements taken, only one was highly repeatable for each test, namely "Time at wall" for
260 the Open Field Test and "Dark arm time" for the Elevated Plus Maze test (Table 1). Hence, we used these two measurements for all mouse experiments, both as the first-time measurement for each individual, as well as in a mixed model analysis.

Table 1: Repeatability (**R**) for behavioral tests.

		MC (N=26)		*C57Bl6/J (N=37)*		*Apodemus (N=32)*	
Test	**Measurement**	**R**	*P-value*	**R**	*P-value*	**R**	*P-value*
Open Field	Time at wall	0.72	≤ 0.001	0.68	≤ 0.001	0.36	0.0002
	Distance travel at the wall	0.05	0.33	0.05	0.3	0.5	0.26
	Speed	0.16	0.09	0.035	0.357	0.11	0.11
	Total distance	0.07	0.26	0.08	0.2	0.09	0.11
Elevated Plus Maze	Dark arm time	0.62	≤ 0.001	0.55	≤ 0.001	0.64	≤ 0.001
	Speed	0.13	0.11	0.11	0.1	0.11	0.09
	Total distance	0.06	0.26	0.1	0.12	0.1	0.08

265

Behavioral scores and copy number

Using the first-time measurements for each individual, we find a highly significant regression between SNORD copy numbers and test scores in both the CB (Figure 2 E-H) and the MC (Figure 2 I-L) wild house mouse populations (data in S5 Table). The results from both populations show that
270 animals with higher copy numbers have higher relative anxiety scores.

For the repeated measurements per individual, we constructed univariate mixed models, including animal ID as random effect and triplicate measurements of SNORD copy number from ddPCR as fixed effects. This analysis was done to account for the potential differences over the course of the behavioral experiments (three times measurement) as well as technical error from triplicate
275 measurements from ddPCR for SNORDs copy number calculation. Regression analyses of SNORD copy number variation and behavioral variables confirmed the positive significant regression in both

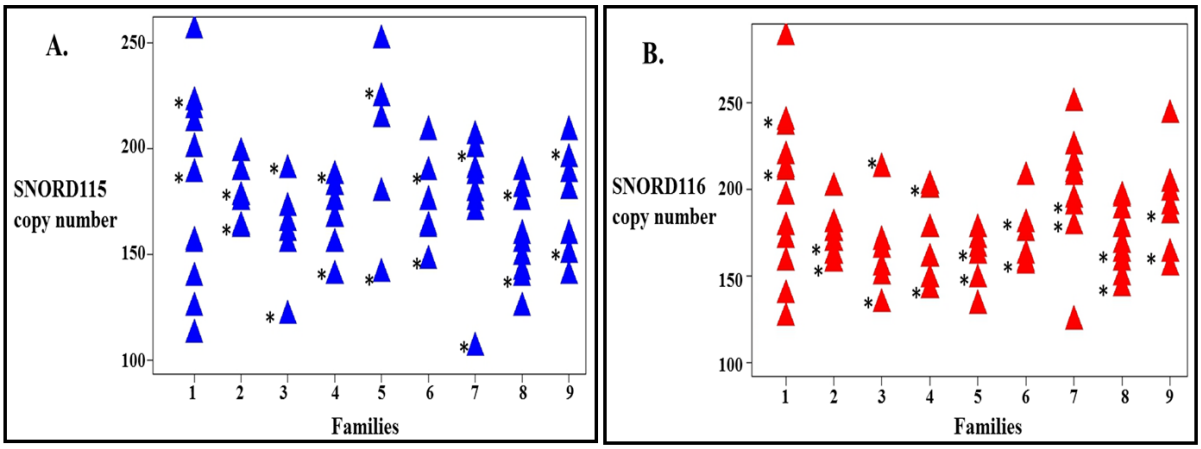
CB (in Open Field Test for SNORD115: R^2=0.3, P = 0.001; SNORD116: R^2 = 0.27, P = 0.001 and in Elevated Plus Maze for SNORD115: R^2=0.22, P = 0.004; SNORD116: R^2 = 0.20, P = 0.006) and MC (in Open Field Test for SNORD115: R^2=0.24, P = 0.001; SNORD116: R^2 = 0.41, P = 0.0001 and in Elevated Plus Maze for SNORD115: R^2=0.23, P = 0.003; SNORD116: R^2 = 0.22, P = 0.006).

To control whether the two other known high copy CNV loci would also show correlation with the behavioral scores, we analyzed the population samples for *Sfi1* and *Cwc22* copy numbers (see above). We find that both loci show no significant correlation with the behavioral scores in the populations tested (S3 Table), i.e. the correlations exist only for SNORD115 and SNORD116.

Inheritance pattern

Given the high variability of SNORD gene copy numbers in the mouse populations, one can ask whether they are subject to frequent changes, such as unequal cross-over events during meiosis. Given the total length of the repeat regions, it is difficult to test this directly, since the haplotypes cannot be distinguished by most conventional techniques (see above). However, we approached this question by comparing the inheritance of copy numbers in nine families of the MC population. We find that the offspring show a large variation, exceeding the spread of numbers measured for the parents in 7 out of the 9 families (Figure 3). Most importantly, if the alleles in the offspring would be only combinations of the parental alleles, one should expect at most four allele classes in the offspring. But two families with more than four offspring (families 1 and 7) show additional allele classes (even when one takes the standard deviation of measurement errors into account: S7 Table), suggesting that new length alleles have been generated in a single generation.

The experiment shows also that new large variation is generated in each generation, in line with the general observation that anxiety traits tend to differ between parents and offspring, as well as among the offspring. To confirm that this is also the case here, we subjected 26 individuals of the offspring to the behavioral tests described above. We find indeed a large variation in behavioral scores, but also the expected significant regression with copy numbers (Open Field Test for SNORD115: R^2=0.55, P << 0.001; SNORD116: R^2 = 0.40, P = 0.004; Elevated Plus Maze for SNORD115: R^2=0.33, P=0.002; SNORD116: R^2 = 0.31, P = 0.004) (data in S5 Table). The univariate mixed-model regression analysis also showed a positive significant regression between SNORD copy number and behavioral scores (Open Field Test for SNORD115: R^2=0.34, P << 0.001; SNORD116: R^2 = 0.3, P << 0.001; Elevated Plus Maze for SNORD115: R^2=0.24, P=0.003; SNORD116: R^2 = 0.22, P = 0.003).



310 **Figure 3: Inheritance patterns of SNORD gene copy numbers (CN) in families**. Nine families of the MC population were tested and the copy numbers were determined for each individual (blue triangles for SNORD115 and red for SNORD116). The parental copy numbers are marked with a star.

Inbred strain variation

315 It has been shown previously that anxiety scores can differ even between individuals in inbred strains (Jakovcevski et al. 2008; Freund et al. 2013). Hence, we asked whether inbred mice would also show copy number variation for SNORD115 and SNORD116. We used the C57BL/6J mouse strain for this purpose, which is known to harbor extremely little variation in the form of SNP polymorphisms (Zurita et al. 2011). We genotyped 60 individuals of both sexes (30 female and 30 male). We find that
320 there is indeed a large variation of copy numbers in these inbred mice, comparable to the spread seen in the outbred cohorts (Figure 1) (S1 Table). 20 mice of each sex were then subjected to the behavioral tests. We found no significant difference between males and females in these tests (ANOVA; p=0.62 for SNORD115 and p=0.28 for SNORD116) but we could confirm the strong correlation with copy numbers (Figure 2 M-P). Similarly, the univariate mixed-model regression
325 analysis also showed significant regression between SNORD copy number and behavioral scores (Open Field Test for SNORD115: R^2=0.40, P << 0.001; SNORD116: R^2 = 0.32, P = << 0.001; Elevated Plus Maze for SNORD115: R^2=0.33, P << 0.001; SNORD116: R^2 = 0.3, P = << 0.001).

Target gene analysis from transcriptome data

330 To investigate whole transcriptome changes in response to copy number variation of the SNORD genes, we choose ten C57BL/6J individuals from the previous experiment, representing a spread of copy numbers from low to high and sequenced RNA from their brains. First, we sought to confirm that the target genes that we had tested in the wildtype background by ddPCR (see above) would show the same patterns in the RNASeq data from the inbred strain.

335 The first analysis focused on *Htr2c* as target of SNORD115. The expected alternative splice is in exon V, which results in a 95nt long exon part (exon Vb) that is specific for one splice variants, while the rest of the mRNA is shared between both (S6 Figure). We have therefore determined RNASeq read coverage for the exon Vb part as well as exon VI and compared each to SNORD115 copy numbers from the respective animals. We find a significant positive correlation for exon Vb, but none for exon
340 VI (R^2 = 0.43 vs -0.12; P = 0.04 vs. 0.3) (S8 Table). This corroborates the original finding by (Kishore and Stamm 2006) that increasing the dose of SNORD115 transcripts in cell culture experiments leads to an increase in the relative concentration of exonVb into the full transcript. It confirms also our results from the ddPCR experiments in the wild type populations reported above. We have further also analyzed the edited sites in exon Vb in our RNASeq data. We find the same sites that were
345 previously described (Burns et al. 1997; Kishore and Stamm 2006) (S6 Figure), but we do not find the additional edited sites that were reported in the ectopic SNORD115 expression study (Raabe et al. 2019). Given that this study had also not found the alternative splicing signal, it would seem that it is not comparable with the normal expression of SNORD115.

Computational analysis has predicted *TAF1*, *RALGPS1*, *PBRM1, CRHR1,* and *DPM2* as five
350 additional possible targets for SNORD115 (Kishore et al. 2010). However, our RNAseq data analysis showed no significant correlation between SNORD115 copy number with the expression of these five genes, only a tendency for a negative correlation for *DPM2* (R^2 = 0.32, P = 0.09) (S8 Table).

For *Ankrd11* we correlated read coverage for the exon IX 5`- part including the PCR fragment that we had used for the ddPCR study above, as well as the SNORD116 binding site (S6 Figure). This shows
355 a significant positive correlation to SNORD116 copy number (R^2 = 0.5; P = 0.02) (S8 Table). We have also analyzed read coverage of the exons preceding exon IX to explore possible alternative splicing. Interestingly, exon VIII shows no significant correlation with SNORD116 copy number (R^2 = 0.009; P = 0.79), while exon VII shows a strongly negative one (R^2 = -0.53; P = 0.01) (S8 Table). This suggests that alternative splicing happens in this region, but the current data did not allow us to
360 resolve this unequivocally. However, in addition to the several annotated splice variants (compare S6 Figure), we find further ones in the RNASeq data that will require a much deeper analysis. Also, in contrast to *Htr2C*, we find no evidence for RNA editing around the predicted SNORD116 binding site in *Ankrd11*. Hence, we conclude that the mechanism for regulation may be different, but this is not unexpected given the diversity of regulatory interactions that are known for C/D-box snoRNAs
365 (Falaleeva et al. 2017).

Whole transcriptome analysis

Given that we test a range of copy number variants, rather than two states, it is not possible to do a conventional analysis of significantly differentially expressed genes across the whole genome, where

370 only two states are compared. Hence, we used correlation coefficients (R^2) as a basis for identifying possible target genes. Since SNORD116 appears to regulate *Ankrd11* and since ANKRD11 is a chromatin regulator which influences the expression of downstream genes by binding chromatin modifying enzymes like histone deacetylases, we started with a list of previously identified target genes of ANKRD11 (Gallagher et al. 2015). We were able to match the data for 635 predicted target
375 genes from the (Gallagher et al. 2015) study with our data from the whole transcriptome analysis. We used these to correlate the expression data of these genes from our whole transcriptome analysis to both the *Ankrd11*-exon IX 5`-part ddPCR and the SNORD116 copy numbers. When applying a cutoff of R^2=0.4 for both correlations, we find 72 positively correlated and 32 negatively correlated genes (S9 Table). A subunit of the GABA A receptor (*Gabrg1*) shows the strongest positive correlation
380 among the known target genes. *Gabrg1* is known to play a crucial role in the modulation of the anxiety response in both humans and mice (Nuss 2015).

In a second step, we applied the same correlation analysis to all genes in the transcriptome, using a correlation with either SNORD115 or SNORD116 copy numbers. We found an additional 67 candidate genes for SNORD115 and 89 for SNORD116 (S9 Table). This suggests that the
385 transcriptome effects of the two SNORD genes go beyond the known target pathways, although this could of course be a secondary effect. Note that the *Ankrd11* level correlates as a whole with SNORD116 copy number, as predicted from the ddPCR on the *Ankrd11*-exon IX 5`-part results above, i.e. the exon which represents most of the full-length RNA (S6 Figure). In contrast, the level of the whole *Htr2c* transcript does not correlate with SNORD115 copy number, since the alternative
390 splice changes are only a minor part of it (S6 Figure).

SNORD116 copy number and craniofacial features

Given that a previous study of a particular mutation in the *Ankrd11* gene has shown craniofacial abnormalities (Barbaric et al. 2008), we tested whether SNORD116 copy number variation would also
395 correlate with craniofacial features in the C57BL/6J mouse strain. We used a 3D landmarking approach to quantify differences in skull shapes based on CT scans, following the procedures described in (Pallares et al. 2015). A multivariate regression of the shape data (Procrustes coordinates) on SNORD116 copy number indicated a strong relationship (percentage predicted = 7.97%; p-value < 0.0001; 10,000 rounds of permutation) (Figure 4A).

400 The variation in SNORD116 copy number is associated with skull shape changes mostly localized at the back of the skull (i.e. braincase, squamosal, and occipital) with a low copy number being linked to more slender and straight skulls while a high copy number is related to more rounded and bent skulls with an enlarged braincase (Figure 4B, C).

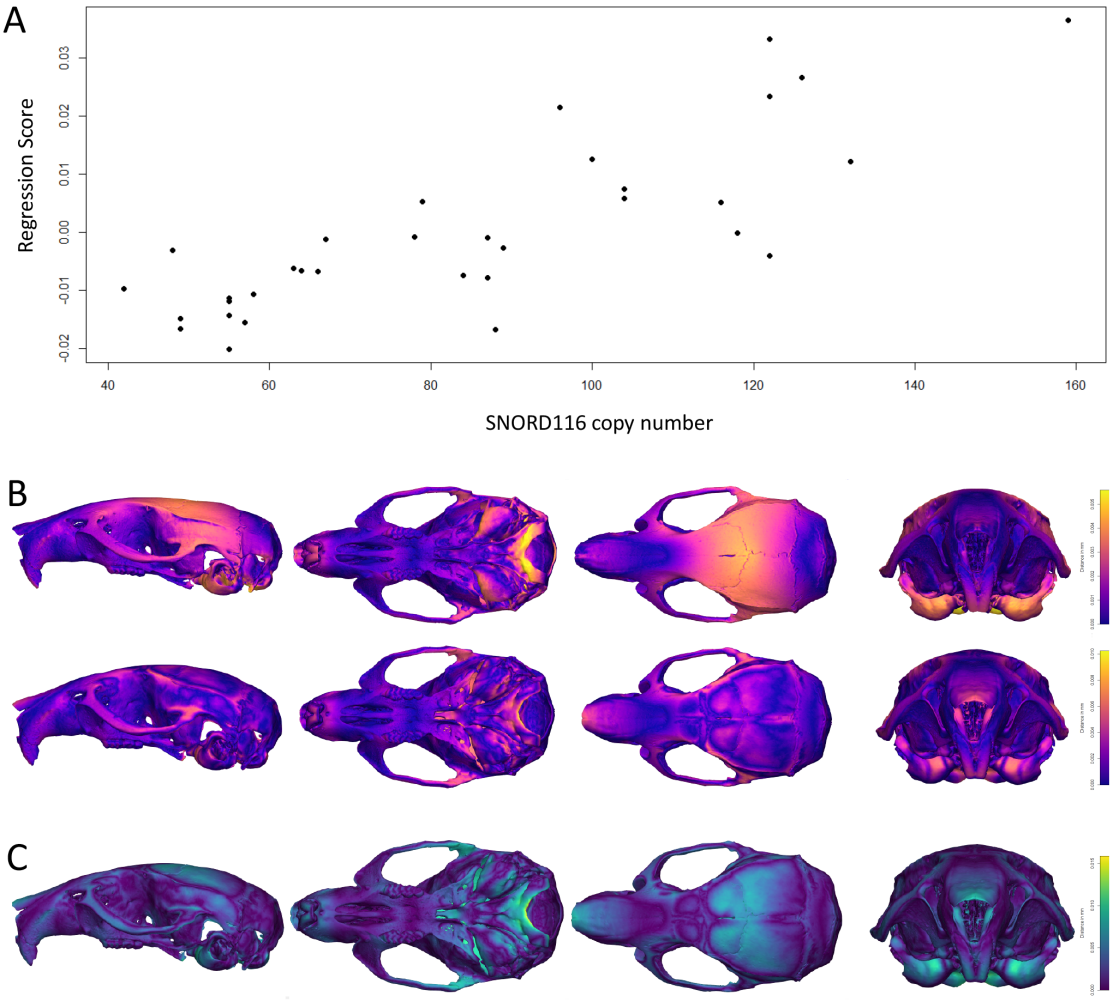


Figure 4: Morphometric analysis of mouse (C57Bl6) crania in dependence to SNORD116 copy number variation. A) Shape scores obtained from the multivariate regression plotted against SNORD116 copy number. B) Shape change associated with a low SNORD116 copy number (top) and high copy number (bottom) visualized by warped 3D surfaces of mouse skulls representing the deviation of each predicted shape and the overall mean shape (warm colors indicate high deviation while cold colors stand for low deviation measured in Procrustes distance). C) Same as B, but differences between predicted shape change for a low and high SNORD116 copy number.

Other rodents

Given that the arrangement of the PWS region with the two clusters of snoRNAs is conserved throughout mammals (Sato 2017), we were interested to assess in how far one can trace the association between copy number variation and anxiety scores also in other species.

We have first analyzed a second mouse species, the wood mouse *Apodemus uralensis*, which is separated from *Mus musculus* since about 10 million years and which has different ecological adaptations. Still, given its general similarity to house mice, we applied the same phenotyping scheme and verified that the phenotype scores are repeatable for individuals (Table 1). From the currently

available genomic data, we could only retrieve sufficient information for the SNORD115 cluster, but this has even higher copy numbers than in *Mus musculus* (Figure 1). We find indeed a significant regression with SNORD115 copy numbers and behavioral scores (Time at wall from Open Field and Dark arm time from Elevated Plus Maze) in *Apodemus* as well (Figure 2 Q, R). Mixed-model regression analysis also showed a significant regression between SNORD115 copy number and behavioral measurements (in Open Field Test for SNORD115: R^2=0.35, P << 0.001; and in Elevated Plus Maze for SNORD115: R^2=0.24, P = 0.001) (all data in S6 Table).

As a second, even more distant rodent, we used a cavy species, *Cavia aparea,* the wild congener of domesticated guinea pigs. For this species we used two previously established behavioral tests that can be considered to reflect anxiety behavior and were shown to be repeatable (Guenther and Trillmich 2015), but differ in their execution and their scores from the mouse tests (see Methods). The current genome sequence of a close relative, *Cavia aperea* f. *porcellus,* provides only the annotation for SNORD116 repeats, hence we focused our analysis on this cluster. We find a much smaller range of copy numbers than in mouse (Figure 1), but still a significant regression with the specific cavy anxiety scores from the two tests (Figure 2 S,T) (all data in S10 Table). As with the other rodents in this study, individuals with higher copy numbers showed higher anxiety.

Humans

To assess a possible association between SNORD copy numbers and personality traits in humans, we tested a subset from a cohort of 541 healthy individuals that had taken part in a study based on the Tridimensional Personality Questionnaire (TPQ) which is designed to explore novelty seeking, harm avoidance and reward dependence (Cloninger et al. 1991; Weyers et al. 1995). From these individuals, we chose the top 48 each for low and high anxiety scores (see Methods) and typed them for SNORD copy numbers. SNORD115 in humans is represented by a single class only, while the SNORD116 family is split in three subclasses, for which we designed different ddPCR assays. We find significant differences of copy numbers between the two groups (Table 2). Particularly significant are SNORD115 and SNORD116_2. The latter is the variant that is predicted to bind to *Ankrd11* exon IX, while the possible target genes for the other two SNORD116 variants are not yet clear. These latter ones show also generally only little copy number variation (Table 2) (all data in S11 Table). However, we note that the direction of the correlation is different between rodents and humans. In humans, the relatively higher anxiety group has the smaller number of copies, while it is the other way around in the three tested rodent species (see above).

Table 2: SNORD copy numbers (mean) for humans scoring low or high on anxiety and activity

ddPCR assay	SNORD115	SNORD116 _1	SNORD116_2	SNORD116_3	all SNORD
low anxiety and activity group, N=48 (SD)	28 (8.0)	6 (1.1)	16 (4.4)	8 (1.4)	57 (12.6)
high anxiety and activity group, N=48 (SD)	21 (4.5)	6 (0.8)	12 (1.7)	7 (1.1)	47 (6.8)
p-value#	7.2×10^{-6}	0.29	1.1×10^{-5}	0.06	4.4×10^{-6}

unpaired t-test, two-tailed

460

Discussion

The PWS region has been extensively studied in the past and it is clear that it is involved in a multitude of phenotypes. Given its implication in disease phenotypes in humans, a strong focus has
465 been on loss of function studies, i.e. mostly deletions in the region, since these mimic the disease phenotypes. Our focus here was to look at the effects of natural variation caused by different copy numbers of two SNORD RNA clusters. These form two large tandemly repeated arrays that provide a technical challenge for their analysis. We show here that ddPCR is currently the only technique by which this can be solved. By comparing copy numbers with behavioral scores in a variety of
470 populations and species, we were able to establish a link between these scores and copy number variation at the SNORD115 and SNORD116. This includes also an isogenic mouse strain, which is not expected to carry many other genetic variants (Zurita et al. 2011). Furthermore, RNA expression analysis reveals target genes that have previously been implicated in behavioral and other phenotypes known from PWS patients and knockout mice. Intriguingly, our data establish also a possible direct
475 link to the skull shape phenotypes known in humans.

Further, our data show that new copy number alleles are generated every generation, which provides an explanation for the high individual variances observed between parents and offspring, as well as the low general heritability. At the same time, it can explain the observation that the behavioral scores tested for "personality" are repeatable within the lifetime of an individual. The new length variants are
480 most likely generated in the germline, possibly only during meiosis, through unequal crossing over. Hence, a given individual carries a specific genetic variant, even though it would not necessarily inherit it to its offspring.

Technical considerations

485 Our conclusions rely on the validity of the ddPCR data for measuring the repeat numbers. As discussed above, there are already a number of studies that show the reliability of ddPCR, including a validation by comparing ddPCR scores with sequencing depth data. In fact, if the ddPCR would be unreliable or too noisy, it would be very difficult to explain the observed correlations with SNORD copy numbers. Also, given that all copy number determination experiments were done blinded (i.e. 490 the experimenter did not know the scores for the behavior tests), we can exclude observer bias. Finally, a technical bias is highly unlikely, since SNORD115 and SNORD116 were amplified with different primer sets and since primer sets for two other high copy number CNV regions (*Sfi1* and *Cwc22*) did not yield any correlation.

There are also two further observations that make it unlikely that a technical error could cause the 495 observed results. 1) We found an a priori unexpected correlation of SNORD115 and SNORD116 copy number in the ddPCR experiments. We have then re-checked previous data on genome read coverage for these regions (Pezer et al. 2015) and find the same correlation in the comparisons among the mouse populations. Hence, this can be considered as an additional validation of the ddPCR results. 2) The ddPCR results had suggested for the target gene expression data that the expression level of target 500 exons correlated with copy number. Our subsequent RNASeq data allowed to test this prediction and we found it confirmed.

The choice of the behavioral assays could be another reason for concern. While these tests seem noisy, they constitute standard tests of behavioral biologists studying "personality", i.e. they were found to be readily repeatable and consistent. We have paid extra attention that the tests were 505 repeatable for given individuals over time. Forkosh et al. (Forkosh et al. 2019) have recently shown that the paradigm of screening for repeatable behavioral patterns among complex behaviors can be a proxy for "personality". Our tests address one particular dimension of "personality", namely anxiety. Interestingly, we find that the anxiety dimension of human personality tests via questionnaires shows also a correlation with SNORD copy numbers.

510 Finally, one can raise the concern that all our results are based on correlation only, rather than direct genetic interference. For example, we have not attempted to modify the number of active copies by a dedicated CRISPR/Cas mutagenesis approach. Rather, we looked at the effects of natural variation. Any attempt to do genetic manipulation would actually run into the problem of high natural variability between generations, i.e. it would not be possible to distinguish effects generated by CRISPR/Cas 515 mutagenesis from the already existing normal variation, even in inbred strains. Hence, we opted to solve this verification problem by analyzing multiple different natural systems (i.e. animals form different populations and species). When one gets consistent results across these multiple different gnomic backgrounds, one can consider this as an independent confirmation that it is indeed the

SNORD copy number variation that creates the effect rather than a consequence of something else in the genomic background of a given strain.

Finally, the PWS region has not shown an association in large scale GWA studies in mice that have included anxiety tests (Nicod et al. 2016; Parker et al. 2016). This would seem in contrast to the rather strong effects that we trace here. However, GWA studies rely on SNP typing and linkage to the typed SNPs. Such a linkage is not given when new length variants are generated at a high rate. Hence, the effects we trace here cannot be properly addressed via linkage analysis, but only via direct testing of copy number polymorphism, as we have done.

Origin of variation

The high level of copy number variation in the SNORD115 and SNORD116 clusters can be ascribed to their tandem repeat organization, which should facilitate unequal crossover and thus length variation of the cluster. Both the family study in the wildtype mice, as well as the observation of many alleles in the otherwise isogenic strain C57BL/6J suggest an exceptional high rate of generation of new repeat number variants. C57BL/6J mice were originally generated from a single pair of mice in 1921 and have been kept as a well identifiable strain with very little SNP variation (Zurita et al. 2011). Still, we find at least 25 copy number variants for each SNORD cluster (when binned in distances of five) among the 57 animals that were provided by a single supplier (Jackson laboratories). While the exact number of generations that have passed since the founding of the strain is difficult to ascertain, it would not be more than a few hundred. Given that we find at least 25 alleles among them (and this can be considered to be a conservative estimate), this suggests at minimum the generation of one new variant every few generations. Imprinted regions are known to include hotspots of meiotic recombination (Paigen and Petkov 2010). However, such an extreme rate of change has not been described so far.

Another very unexpected finding is the covariation of repeat numbers between the SNORD115 and SNORD116 clusters. The repeat units of these clusters do not show any sequence similarity and they are separated by a non-repetitive region. We are not aware of any mechanism that could explain this covariation, but it is consistently found in all tested species. Given the overlapping regulation of the same behavioral pathways by the two clusters, it makes sense that they should show this covariation, but how this is achieved will require further studies.

Target genes

Our transcript analysis data in response to SNORD115 copy number variation confirm the previously described interaction and regulation of the alternatively spliced exon Vb of the serotonin receptor

Htr2c through this snoRNA (Kishore and Stamm 2006). We suggest that the corresponding predicted interaction of human SNORD116 with *Ankrd11* may also be a direct interaction, but we have so far not been able to clarify the mechanism. In contrast to the SNORD115 interaction, we do not find signs of RNA editing around the binding site in the transcriptome sequences. Instead, we find an influence on the expression level of the whole *Ankrd11* RNA. While this could be a consequence of alternative splicing, other regulatory mechanisms are not excluded. Note that C/D-box snoRNAS can function in many different ways (Falaleeva et al. 2017).

There are a number of mouse knockout lines that delete more or less large parts of the PWS region. However, only one knockout line exists that has removed specifically one SNORD gene cluster, namely SNORD116. The mice from this line show an early-onset postnatal growth deficiency, are deficient in motor learning and show slightly increased anxiety in one of the tests that were performed (Elevated Plus Maze) (Ding et al. 2008). The latter finding would be somewhat at odds with our results that the anxiety scores increase with the number of SNORD115 and SNORD116 copies. However, we expect that the anxiety phenotypes are primarily regulated via the serotonin pathway that is affected by SNORD115 copy number and this was not controlled for in the (Ding et al. 2008) study. The SNORD116 copy number variation would be more relevant for the other associated phenotypes. In fact, a later analysis of a full deletion variant of this line has revealed further effects on feeding related pathways, as well as bone mineral density (Qi et al. 2016). These findings are in line with our findings of a disturbance of the multiple pathways that are regulated by *Ankrd11* as a direct target gene of SNORD116.

Interestingly, we find that not only behavior, but also craniofacial shape correlates with the SNORD copy number variation and the above reasoning would suggest that this is conveyed by SNORD116. For humans it has been suggested that there is indeed a link between personality and facial characteristics (Kramer and Ward 2010), but a possible causality of these observations was left open. Our data suggest such causality for mice. But given that mice are essentially nocturnal animals that communicate mostly via scents and ultrasonic vocalization (von Merten et al. 2014), it is unclear whether they would even recognize different craniofacial shapes among their conspecifics. However, the cranial shape changes could reflect a general osteogenic effect on the whole bone system that could be of relevance for being combined with the behavioral tendency. For example, bold animals might profit from stronger bone structures, in case they get more involved in fights. However, this connection will need further study, especially since the SNORD116 expression is confined to the brain. Qi et al. (Qi et al. 2016) have suggested that a general osteogenic effect may be mediated via the metabolic pathways that are regulated by *Ankrd11*.

Implications for the genetics of anxiety traits

Our findings can resolve a long-standing controversy about the genetics and plasticity of anxiety or "personality" traits. Our data suggest that these traits are indeed genetically specified, but through a hypervariable locus that would escape conventional genetic mapping approaches. Hence, it is not necessary to invoke developmental or secondary epigenetic effects in explaining variance in such traits, given that there is a mechanistic pathway of modulation of target RNA expression through different concentrations of snoRNAs. However, such alternative mechanisms may have additional contributions. In future studies on such traits it would seem advisable to treat SNORD copy number variation as a covariate to assess whether other mechanisms exist as well.

The fact that we find significant regressions between behavioral scores and SNORD copy numbers also in other rodents and even in humans, implies that the system is conserved throughout mammals. In humans, we observed this within a psychiatrically healthy control group, where we find that extremes of personality traits for anxiety and activity are associated with significant differences in copy numbers. The effects are somewhat weaker than for the rodents, but humans are evidently also more subjected to environmental influences than the animals that were bred under controlled conditions. Most intriguingly, however, the effect in humans is in the opposite direction, i.e. more anxious individuals harbor lower copy numbers. This would suggest that the copy number variation acts only as a general regulator, but the actual behavioral consequences are modified by downstream pathways.

Evolutionary implications

One can raise the question of how and why such a hypervariable system could have evolved and how it is maintained. This question has in fact been posed since a long time and there have been a number of attempts to propose evolutionary models that could explain the large variability in behavioral traits. These include drift effects in a neutral context (Tooby and Cosmides 1990), mutation-selection balance (Zhang and Hill 2005) or balancing selection processes (Dingemanse and Wolf 2010; Penke and Jokela 2016). But none of these models has taken the possibility into account that a hypervariable locus could control this variation. Hence, our results establish a new paradigm in the regulation of behavioral variance, namely the involvement of dosage changes of a regulatory RNA through recurring copy number mutations.

Acknowledgements

We thank Chen Xie for his valuable advice with transcriptome analysis, Stefan Dennenmoser, Cemalettin Bekpen and Guy Reeves for their advice on SNORDs copy number analysis, Sven Künzel

for sequencing, Ellen McConnell, Nicole Thomsen and Cornelia Burghardt for their great lab-related support, Elke Blohm-Sievers for her help in CT scanning, the entire mouse team at MPI for evolutionary biology, in particular Christine Pfeifle, Susanne Holz and Camilo Medina for taking care of the mice in this study and Heike Harre and Fikri Tuğberk Kara for their help with behavioral tests and brain dissection. Animals were kept according to FELASA (Federation of European Laboratory Animal Science Association) guidelines, with the permit from the Veterinäramt Kreis Plön: 1401-144/PLÖ-004697and the Veterinäramt Bielefeld for cavies, respectively. The experiments involving the behavioral analyses was approved and registered under V244-71173/2015 through the local authorities. The animal welfare officer was informed about the sacrifice of the animals for the molecular analyses. The study on humans was approved by the ethics committee of the Ludwig-Maximilians-University (LMU) in Munich and written informed consent was obtained from all subjects.

Methods

Mouse strains

Wildtype mice (*Mus musculus domesticus* and *Apodemus uralensis*) used in this study were offspring of mice that originated from wild populations sampled in the Massif Central region of France (MC), the Cologne/Bonn region of Germany (CB) and Kazhakstan, Almaty (*Apodemus*) and then held under outbreeding conditions at the Max-Planck-Institute for Evolutionary Biology in Plön. The C57BL/6J inbred strain (30 male and 30 female) was purchased at the age of 3 weeks from Charles River Laboratories, Research Models and Services, Germany GmbH (Sandhofer Weg 7, 97633 Sulzfeld). These inbred mice were not siblings.

All mice were bred in type III cages (Bioscape, Germany) and were weaned at the age of 3 weeks. At weaning, the sexes were separated. Males were housed together with brothers or in individual cages. Females were housed in sister groups to a maximum of 5 mice per cage. Enrichment, including wood wool, toilet paper, egg cartons and a spinning wheel (Plexx, Netherland), was provided in each cage. Mice were fed standard diet 1324 (Altromin, Germany) and provided water ad libitum. Housing was at 20–24°C, 50–65% humidity and on a 12:12 light-dark schedule with lights on at 7 am. While the C57BL/6J and the *Apodemus* mice and MC mice used in family study were kept under identical conditions throughout the experiment, the CB and MC mice in the first part of study were released into semi-natural conditions at an age of 3-4 months. Details of this experiment are provided in the PhD thesis of Rebecca Krebs which is available at https://macau.uni-kiel.de/receive/dissertation_diss_00022305.

Wild cavies:

Details on the behavioral work with wild cavies are described in previous papers (Guenther et al. 2014; Guenther and Trillmich 2015). Wild cavies were descendants from a population of wild animals caught in Uruguay in 2003 and 2007. Experimental animals were bred in outdoor enclosures consisting of a roofed hut ($2m^2$) and an outdoor compartment ($14m^2$) in the Department of Animal Behaviour at Bielefeld University, Germany. They experienced natural temperature and photoperiodic conditions. Food (hay and guinea pig chow, Höveler, Germany) and water was provided ad libitum and supplemented with fresh green (e.g. carrots, bellpepper or apples) three times a week.

At weaning (21 days of age), animals were individually marked and transferred into indoor standard housing conditions. Indoors, animals were kept in unrelated same-sex groups of four individuals each and housed in $1.6m^2$ enclosures. Each enclosure was equipped with three shelters, three feeding dispensers and three water bottles. Animals were kept under natural light conditions with additional

670 artificial light from 6 am to 7 pm, at constant temperatures of 20 ± 2 °C. The feeding regime was the same as under outdoor conditions. Animals used for this study were all unrelated. In total, 10 males and 9 females were used for the experiments. Since wild cavies do not show sex-differences in the expression of anxiety-related behaviors (Guenther and Trillmich 2015), we did not include "sex" as a factor in statistical analyses.

675

Human cohort and evaluation of personality traits.

541 healthy controls were randomly selected from the Munich registry of residents and interviewed for the presence of DSM-IV anxiety, affective, somatoform, eating, alcohol dependence, drug abuse or dependence, disorders using a modified version of the Munich Composite International Diagnostic
680 Interview (Wittchen and Pfister 1997) at the MPIP. Only individuals negative for the above-named disorders were included in the context of a large study on the genetics of major depression using EDTA blood as the source for DNA (see (Heck et al. 2009) for more detail).

DNA and RNA extraction and cDNA synthesis

685 All dissections were done following standardized protocols and personal instructions. Prepared tissues were immediately frozen and kept at −70 ° until DNA/RNA preparation.

DNA extraction was performed according to a standard salt extraction protocol. Briefly, samples were lysed by using HOM buffer (80 mM EDTA, 100 mM Tris and 0.5 % SDS) with Proteinase K (0.2 mg/mL) for 16 hours in Thermomixer (Eppendorf, Germany) at 55°C. 500 µL sodium chloride (4.5
690 M) was added to each sample and was incubated on ice for 10 minutes. Then chloroform was added, mixed and spun for 10 minutes at 10,000 rpm. The upper aqueous phase was separated, mixed well with Isopropanol (0.7 volume) and spun for 10 minutes at 13,000 rpm. The pellet was washed with Ethanol (70 %), air dried and dissolved in TE-buffer (10 mM Tris, 0.1 mM EDTA). DNA concentration was measured on the Nano Drop 3300 Fluorospectrometer using Quant-iT dsDNA BR
695 Assay kit (Invitrogen) reagent.

RNA extraction (from whole brain) was done by using Trizol reagent. 1mL Trizol was added to each sample. Then the samples were lysed by Tissue lyser II (QIAGEN, Germany) at 30 Hertz for 5 minutes. Homogenized samples were incubated at room temperature for 5 minutes. 200µL chloroform (per 1 mL TRIzol) was added to each sample, shook vigorously by hand 15 seconds, followed by 3
700 minutes incubation at room temperature and spun at 12,000g for 15 minutes at 4°C. The aqueous phase was transferred to a new tube and 0.5 volumes Isopropanol was added, incubated at room temperature for 10 minutes and spun at 12,000g at 4°C. The supernatant was removed and the pellet was washed with 75% EtOH (made with RNAse-free water). Samples were mixed by hand several times and then spun at 7,500g for 5 minutes at 4°C. The supernatant was removed and the pellet dried

705 shortly at room temperature, dissolved in 200µl RNAse free water and stored at -20°C for overnight. An equal volume of LiCL (5M) was added to the crude RNA extract, mixed by hand and incubated for one hour at -20°C. Samples were spun at 16,000g for 30 minutes. The supernatant was removed; samples were washed twice with EtOH 70% and spun at 10,000 at 4°C. The pellet was dried at room temperature, dissolved in RNAse free water and kept for long-term storage at -70°C.

710 The quality of the RNA samples was measured with Bio-Analyzer chips and samples with RIN values below 7.5 were discarded. cDNA was synthesized using the MMLV High Performance Reverse Transcriptase kit according to the instructions of the supplier (epicenter, an Illumina company).

Small RNA extraction and cDNA synthesis

715 Total RNA was extracted from the whole brain by using the mirVana miRNA Isolation Kit, which enriches small RNAs. The quality of the RNA was measured with BioAnalyzer chips and samples with RIN values below 8 were discarded. Illumina® TruSeq® Small RNA Library Prep kit was used for small RNA cDNA synthesis. The protocol in the Illumina kit takes advantage of the common natural structure in most known small RNA molecules. Most mature small RNAs have a 5'-phosphate
720 and a 3'-hydroxyl group. So, the Illumina adapters in this kit are directly ligated to these small RNAs. Then Reverse Transcriptase with a primer for this adaptor was used to synthesize cDNA from small RNA. 1 µg of total RNA which was enriched in small RNA was used as input. To control for DNA contamination, we included a tube with all ingredients apart of the reverse transcriptase. The results obtained from this tube were used as gDNA contamination to normalize the gene expression data for
725 each sample. The data provided in Suppl. file 12 show that gDNA contamination is expected to be very small for each sample.

RNAseq analysis

Poly-A^+ RNA was used for cDNA synthesis and Illumina library preparation by using the Truseq
730 stranded RNA HT kit. The libraries passing quality control were subjected to sequencing on an Illumina NextSeq 500 sequencing system and paired-end reads were 150 bases long. Raw sequence reads were quality trimmed using Trimmomatic (Bolger et al. 2014). The quality trimming was performed base wise, removing bases below quality score of 20 (Q20), and keeping reads whose average quality was of at least Q60. Reads were mapped to the mouse mm10 reference genome by
735 using Hisat2 (Kim et al. 2015). Between 20.1 to 28.7 million mapped reads were obtained per sample (average 23.8 million). Differential expression analysis was performed with the DESeq2 package (Love et al. 2014) in the R environment. To perform the alternatively spliced isoforms analysis, BAM

files from Hisat2 were used as input for SAMtools (Li et al. 2009) by using option –c for total read from each exon and option –q 60 for total read of each sample.

740

Droplet digital PCR

Digital PCR is a method enabling absolute quantification of DNA targets without the need to construct a calibration curve as used in qPCR (Zhao et al. 2016). It requires a reference to calculate gene copy number and to normalize gene expression level. For the latter we used β-catenin for
745 mRNAs and SNORD66 for standardizing SNORD RNAs. SNORD 66 is a single copy gene located in an intron of the eukaryotic translation initiation factor 4 and was therefore used also as reference gene for copy number calculations. All reactions were done as technical triplicates which yielded highly congruent results (S1 Table). Averages of the triplicates were used to calculate copy numbers. Values for copy number represent the total number of copies per diploid genome. The values for expression
750 represent the number of RNA copies per µL of input material.

Primers (Table 3) were designed aiming for a balanced GC content <50% and with low potential for primer dimer structure. All primers were verified that they amplify only the expected length fragment.

The SNORD115/116 primers were placed within the coding sequences of the respective snoRNAs. Two pairs were tested for each for an initial set of 19 animals and found to yield highly congruent
755 results (R^2=0.99, S3 Table). Hence, we used only the first primer set (labelled with -a) for the rest of the analyses. To test for biological replicability, we obtained ear clip samples from 41 four-week old animals and later (at 14 weeks) brain samples from the same animals. Copy number determination was highly congruent between these samples (R^2=0.99, S3 Table).

The QX100™ Droplet Digital™ PCR System (Bio-Rad, Hercules, CA, USA) was used in this study
760 according to the manufacturer's instructions. Briefly, fluorescent PCR reactions for each sample were prepared in a 23 µL volume containing 12µL 2 X EvaGreen supermixes, 200nM of each forward and reverse primers, DNA or cDNA (1ng in mice study,10ng in Guinea pig and 5ng in human study) and water. For tandem copy separation, sample viscosity reduction and improved template accessibility, DNA digestion was done by different restriction enzymes (Table 3, Thermo Fisher Scientific,
765 Germany). 5 units of restriction enzymes were added to each sample. The samples were kept 20 minutes at room temperature for complete digestion.

Droplets were generated using a Droplet Generator (DG) with an 8-channel DG8 cartridge and cartridge holder with 70 µL of DG oil/well, 20 µL of fluorescent PCR reaction mixture and a DG8 gasket. The prepared droplets were transferred to corresponding wells of a 96-well PCR plate
770 (Eppendorf, Germany). The PCR plate was subsequently heat-sealed with pierceable foil using a PX1™ PCR plate sealer (Bio-Rad) and then amplified in a LifeEco thermal cycler (Bioer, China). The

thermocycling protocol was: initial denaturation at 95°C for 5 min, then 40 cycles of denaturation at 95°C for 30 s, annealing at 60°C for 45 s and, finally, incubation first at 4°C for 5 min and then at 90°C for 5 min. After cycling, the 96-well plate was fixed into a plate holder and placed into the Droplet Reader. Droplets of each sample were analyzed sequentially and fluorescent signals of each droplet were measured individually by a detector. Copy numbers were calculated according to the procedures suggested by BioRad using QuantaSoft 1.7.

Table 3. Primer sequences and enzymes for pre-digestion

Locus	Forward sequence 5´→3´	Reverse sequence 5´→3´	length (bp)	enzyme
Mouse samples				
SNORD 66	GTGTCTGGGCCACTGAGAC	TTCCTCAGGTCCTCAATCCCA	47	EcoR1
SNORD 115-a	GGGCCTCAGCGTAATCCTAT	ACCCAATGTCATGAAGAAAGGTG	64	EcoR1
SNORD 115-b	GACTCAGCGTAATCCTATTGAGC	TGACAACCCAATGTCATGAAGAA	67	EcoR1
SNORD 116-a	ACCTCAGTTCCGATGAGAGTG	TTCCCAGTCAAACATTCCTTGG	79	BamHI
SNORD 116-b	TCAGTTCCGATGAGAGTGGC	ACATTCCTTGGAAAAGCTGAACA	65	BamHI
β-catenin	GGAAAAGAGCCTCAGGGCAT	CTGCCTGACGGCCAGG	64	EcoR1
HTR2C*	ATTCGCGGACTAAGGCCATC	GAAGTTCGGGTCATTGAGCAC	146*	
Ankrd11	ACGGGGGTCACAGTCTTCTG	AAGAGGACGCCCCGTCATTC	124	EcoR1
Cwc22	TAAAGACTGCTCGCAAACACC	CTAATGCTCTTGGTGGCACTT	78	ClaI
Sfi1	ATAGGTTTGGGCTGAGGATATGTA	AACCCTTCTTACTTGGTTCTTCAC	112	EcoRI
Cavie samples				
SNORD116	GCTCGGATCGATGATGATGC	CAGTGTGCTGGACCTCAGTC	107	BamHI
SNORD66	AGGGACTGGGATTTTGCTTTCC	GCTTCCTTTCCTCAGGTCCTC	101	EcoR1
Human samples				
SNORD66	CTGATGACTTCCTGTTAGTGCC	TCCTCAGATCCTCAGTTCCATC	70	EcoR1
SNORD115	CTGGGTTGGGTCGATGATGAG	GGGCCTCAGCGTAATCCTATT	89	AvaI
SNORD116- I	GTGCTTGGATCGATGATGAGT	CAATGTGCTGGACCTCAGT	110	AvaI+HaeIII
SNORD116-II	TGCGTGGATCGATGATGAC	GACCTCAGTTCGACGAGGATG	96	AvaI
SNORD116- III	TGTGCGTGGATCGATGATGAC	GACCTCAGTTCGACGAGGATG	98	HindIII+XbaI

* amplifies only from the spliced RNA

Behavioral Tests

Mouse strains

Behavioral tests were performed on the *M. m. domesticus* mice starting at an age of 24 weeks, on theC57BL/6J inbred strain at the age of 16 weeks and on the *A. uralensis* mice at the age of 18 weeks. The testing included an Elevated Plus Maze and an Open Field, all adapted towards the use of wild mice to avoid escape. Test setups were cleaned with 30% ethanol between individual tests. All tests were done during the day (between 8 am to 1pm) with constant artificial light of about 300 lux.

The Elevated Plus Maze consisted of four arms 1m above ground, each 50 cm long and a 10x10 cm neutral area in the middle. Two of the arms were made of clear Plexiglas, indicating the unsafe zone and two were made of grey PVC, indicating a safe zone. The floor was made of white PVC. Mice were placed in the center and the behavior of each mouse was monitored for 5 min. During this

experiment, the time spent in the dark and light arms were measured, as well as the speed and distance travelled.

For the Open Field test, mice were placed in a 60x60 cm arena with 60 cm high white PVC walls and they were allowed to explore it for 5 min. The speed of the mouse, the distance travelled and time spent within 10 centimeters off the wall vs in the central area were measured.

The behavioral tests were filmed using a TSE camera (TSE system, Germany). To score the videos from each test, all the videos were transferred to Videomot2 system (TSE system, Germany). Mice were detected by the software in 3 points (head/center/tail base tracking) and then the software automatically generates the numerical data for the above-mentioned measurements.

Wild cavies

The behavioral tests on the wild cavies (*Cavia aperea*) were conducted as described in (Guenther and Trillmich 2015). The tests started when individuals were 55 ± 3 days old. In short, to measure struggle docility, the animal was turned on its back and held in the hand of an observer. For 30 s, the time the animal struggled to actively escape that situation was recorded. To measure open field anxiety, an Open Field test was used. This consisted of a $1m^2$ arena and behaviour was scored based on the videos from a camera on top of the field. Light conditions were artificial and constant around 250 lux. Experiments were conducted between 9-12 am or 3-5 pm, as this was shown to have no systematic influence on the behaviour (Guenther et al. 2014). The distance moved (cm) when individuals were exposed to an open field for 20 min, was scored. The first 10 min, a semi-transparent shelter (20cm x20cm with four exits) was present in the arena under which animals could hide. For the second 10 min, this shelter was lifted out of the arena. Individuals which are anxious in this situation freeze and move about only little while individuals which are not anxious run around and explore the situation (Guenther and Trillmich 2015). Hence, for wild cavies, high values equal low anxiety.

Humans

Five hundred and forty-one individuals also completed the Tridimensional Personality Questionnaire (TPQ), a validated personality measure (Cloninger et al. 1991; Weyers et al. 1995). To match the behavioral assessments in mice, single items of the TPQ reflecting anxiety and activity were used to select individuals on the extremes of this distribution. Individuals can score from 0 to 6 on the activity and anxiety scale, respectively and their sum was used to identify the extreme groups, while matching for age and sex. The high anxiety and activity group (N = 48) had a mean additive score of 7.35 (SD 1.04). The low anxiety and activity group (N = 48) had a mean additive score of 0.73 (SD 0.49). The

mean age of the high anxiety group was 42.3 year and 41.6 years for the low anxiety group with 33.3% and 62.5% females, respectively.

Statistical analyses

Repeatability of behavioral measures

As pointed out in the introduction, one of the cornerstones when aiming to describe *consistent* individual differences in behavior, is the validation of the temporal consistency of the behavioral trait. Commonly, an index called "repeatability" is used to describe the temporal consistency. Repeatability reflects the amount of variation between individuals compared to the amount of all phenotypic variation in the population (including also within-individual variation and measurement error).

To estimate temporal consistency for our animals, we conducted repeatability analyses for each of our three mouse model systems separately (i.e.: C57BL/6J, the wild mice MC population used in the family study and *Apodemus*). Each model included three measurements per behavioral trait in total, with 4 weeks interval between them. The CB and MC mice in the first part of study were part of an experiment in semi-natural environments. They were tested 8 weeks apart of each other.

Statistical analyses were carried out using R 3.3.3 and R 3.3.2. *Cavia aperea* and the CB and MC mice in the first part of study were a subset of mice that have been used in other projects and had already been analyzed for repeatability, so we did not include them for repeatability assessment in this study. For C57BL/6J, the wild mice MC population used in the family study mice and *Apodemus*, repeatability was calculated using the "rptR" package (Stoffel et al. 2017). Normally distributed data were calculated using rpt (ANOVA based) and for count data rpt.poisGLMM was used.

Linear Regression analyses SNORD copy number variation and behavioral variables

To create a regression model for SNORD copy number and behavioral variables, the lm function in R. 3.3.3 and R 3.3.2 was used and population origin, sex and age were modelled as main factors. The average from triplicate measurements in ddPCR was used as SNORD copy number for each individual and the direct scores from first time measurement of each behavioral test were used as behavioral score in this study. p-values ≤ 0.05 were considered to be significant.

Mixed-model regression SNORD copy number variation and behavioral variables

Univariate mixed models for each repeatable behavioral measurement in the behavioral experiments (for Open Field: time at the wall; for Elevated Plus Maze: time in the dark arm) were constructed

using lmer function in R. 3.3.3 and R 3.3.2, including animal ID as random effect and triplicate measurements of SNORDs copy number from ddPCR as fixed effects. This analysis has been performed to account for the potential differences over the course of the behavioral experiments (three times behavioral measurement) as well as technical error from triplicate measurements from ddPCR for SNORDs copy number calculation. R^2 was calculated to assess the proportion of variance explained by the model using r.squaredGLMM function from MuMIn package (Barton K., 2019 - https://CRAN.R-project.org/package=MuMIn). The P-value was calculated using sjstat package (Lüdecke 2019). p-values ≤ 0.05 were considered to be significant.

Craniofacial shape analysis

Landmarks for the morphometric analysis were obtained as described previously (Pallares et al. 2015). Briefly, mouse heads were scanned using a computer tomograph (micro-CT—vivaCT 40; Scanco, Bruettisellen, Switzerland) at a resolution of 48 cross-sections per millimeter. 44 three-dimensional landmarks were positioned in the skull (the landmarks are described in (Pallares et al. 2015)). The raw 3D landmark coordinates obtained in TINA tool were exported to MorphoJ (Klingenberg 2011) for further morphometric analyses. The symmetric component of the skull was obtained through a mirror image of the landmark configuration of each individual was generated, and a full GPA was performed with the original and mirror configurations. Again, the resulting configurations were averaged to obtain the symmetric component of shape variation. The new landmark coordinates generated by the GPA are called "Procrustes coordinates". The multivariate regression of the shape data (Procrustes coordinates) on SNORD116 copy number was carried out via function procD.lm from the geomorph R package v3.0.7 (Adams and Otarola-Castillo 2013).

To visualize the patterns of predicted shape differences for minimum and maximum values of SNORD116 copy number, the closest specimen to the overall mean shape in multidimensional space was first identified via the function 'findMeanSpec' in the geomorph package v3.0.7 (Adams and Otarola-Castillo 2013) and a 3D CT-scan of the specimen closest to the mean shape was warped to this overall mean shape using the Thin Plate Spline (TPS) method via the function 'tps3d' in R-package Morpho v2.6 and using warping and heatmap functions from the R-package Rvcg v0.18 (Schlager 2017) as well as the color palette from R-package viridis v0.5.1 (https://CRAN.R-project.org/package=viridis).

890 **References**

Adams, D. C. and E. Otarola-Castillo. 2013. geomorph: an r package for the collection and analysis of geometric morphometric shape data. Methods in Ecology and Evolution 4:393-399.

Barbaric, I., M. J. Perry, T. N. Dear, A. R. Da Costa, D. Salopek, A. Marusic, T. Hough, S. Wells, A. J. Hunter, M. Cheeseman, and S. D. M. Brown. 2008. An ENU-induced mutation in the Ankrd11
895 gene results in an osteopenia-like phenotype in the mouse mutant Yoda. Physiological Genomics 32:311-321.

Bazeley, P. S., V. Shepelev, Z. Talebizadeh, M. G. Butler, L. Fedorova, V. Filatov, and A. Fedorov. 2008. snoTARGET shows that human orphan snoRNA targets locate close to alternative splice junctions. Gene 408:172-179.

900 Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-2120.

Burns, C. M., H. Chu, S. M. Rueter, L. K. Hutchinson, H. Canton, E. SandersBush, and R. B. Emeson. 1997. Regulation of serotonin-2C receptor G-protein coupling by RNA editing. Nature 387:303-308.

905 Cassidy, S. B., S. Schwartz, J. L. Miller, and D. J. Driscoll. 2012. Prader-Willi syndrome. Genetics in Medicine 14:10-26.

Cavaille, J. 2017. Box C/D small nucleolar RNA genes and the Prader-Willi syndrome: a complex interplay. Wiley Interdisciplinary Reviews-Rna 8.

Cloninger, C. R., T. R. Przybeck, and D. M. Svrakic. 1991. THE TRIDIMENSIONAL PERSONALITY
910 QUESTIONNAIRE - UNITED-STATES NORMATIVE DATA. Psychological Reports 69:1047-1057.

Ding, F., H. H. Li, S. W. Zhang, N. M. Solomon, S. A. Camper, P. Cohen, and U. Francke. 2008. SnoRNA Snord116 (Pwcr1/MBII-85) Deletion Causes Growth Deficiency and Hyperphagia in Mice. Plos One 3.

Dingemanse, N. J. and M. Wolf. 2010. Recent models for adaptive personality differences: a review.
915 Philosophical Transactions of the Royal Society B-Biological Sciences 365:3947-3958.

Dochtermann, N. A., T. Schwab, and A. Sih. 2015. The contribution of additive genetic variation to personality variation: heritability of personality. Proceedings of the Royal Society B-Biological Sciences 282.

Falaleeva, M., J. R. Welden, M. J. Duncan, and S. Stamm. 2017. C/D-box snoRNAs form methylating
920 and non-methylating ribonucleoprotein complexes: Old dogs show new tricks. Bioessays 39.

Forkosh, O., S. Karamihalev, S. Roeh, U. Alon, S. Anpilov, C. Touma, M. Nussbaumer, C. Flachskamm, P. M. Kaplick, Y. Shemesh, and A. Chen. 2019. Identity domains capture individual differences from across the behavioral repertoire. Nature Neuroscience 22:2023-+.

Freund, J., A. M. Brandmaier, L. Lewejohann, I. Kirste, M. Kritzler, A. Kruger, N. Sachser, U.
925 Lindenberger, and G. Kempermann. 2013. Emergence of Individuality in Genetically Identical Mice. Science 340:756-759.

Freund, J., A. M. Brandmaier, L. Lewejohann, I. Kirste, M. Kritzler, A. Kruger, N. Sachser, U. Lindenberger, and G. Kempermann. 2015. ASSOCIATION BETWEEN EXPLORATORY ACTIVITY AND SOCIAL INDIVIDUALITY IN GENETICALLY IDENTICAL MICE LIVING IN THE SAME
930 ENRICHED ENVIRONMENT. Neuroscience 309:140-152.

Gallagher, D., A. Voronova, M. A. Zander, G. I. Cancino, A. Bramall, M. P. Krause, C. Abad, M. Tekin, P. M. Neilsen, D. F. Callen, S. W. Scherer, G. M. Keller, D. R. Kaplan, K. Walz, and F. D. Miller. 2015. Ankrd11 Is a Chromatin Regulator Involved in Autism that Is Essential for Neural Development. Developmental Cell 32:31-42.

935 Guenther, A., M. A. Finkemeier, and F. Trillmich. 2014. The ontogeny of personality in the wild guinea pig. Animal Behaviour 90:131-139.

Guenther, A. and F. Trillmich. 2015. Within-litter differences in personality and physiology relate to size differences among siblings in cavies. Physiology & Behavior 145:22-28.

Harr, B., E. Karakoc, R. Neme, M. Teschke, C. Pfeifle, Z. Pezer, H. Babiker, M. Linnenbrink, I. Montero,
940 R. Scavetta, M. R. Abai, M. P. Molins, M. Schlegel, R. G. Ulrich, J. Altmuller, M. Franitza, A.

Buntge, S. Kunzel, and D. Tautz. 2016. Genomic resources for wild populations of the house mouse, Mus musculus and its close relative Mus spretus. Scientific Data 3.
Heck, A., R. Lieb, A. Ellgas, H. Pfister, S. Lucae, A. Erhardt, H. Himmerich, S. Horstmann, S. Kloiber, S. Ripke, B. Muller-Myhsok, T. Bettecken, M. Uhr, F. Holsboer, and M. Ising. 2009. Polymorphisms in the Angiotensin-Converting Enzyme Gene Region Predict Coping Styles in Healthy Adults and Depressed Patients. American Journal of Medical Genetics Part B-Neuropsychiatric Genetics 150B:104-114.
Homdli, C. N. S., E. Wong, E. Ferris, K. Bennett, S. Steinwand, A. N. Rhodes, P. T. Fletcher, and C. Gregg. 2019. Complex Economic Behavior Patterns Are Constructed from Finite, Genetically Controlled Modules of Behavior. Cell Reports 28:1814-+.
Jakovcevski, M., M. Schachner, and F. Morellini. 2008. Individual variability in the stress response of C57BL/6J male mice correlates with trait anxiety. Genes Brain and Behavior 7:235-243.
Ka, M. H. and W. Y. Kim. 2018. ANKRD11 associated with intellectual disability and autism regulates dendrite differentiation via the BDNF/TrkB signaling pathway. Neurobiology of Disease 111:138-152.
Kim, D., B. Landmead, and S. L. Salzberg. 2015. HISAT: a fast spliced aligner with low memory requirements. Nature Methods 12:357-U121.
Kishore, S., A. Khanna, Z. Y. Zhang, J. Y. Hui, P. J. Balwierz, M. Stefan, C. Beach, R. D. Nicholls, M. Zavolan, and S. Stamm. 2010. The snoRNA MBII-52 (SNORD 115) is processed into smaller RNAs and regulates alternative splicing. Human Molecular Genetics 19:1153-1164.
Kishore, S. and S. Stamm. 2006. The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. Science 311:230-232.
Kiss, T. 2001. Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. Embo Journal 20:3617-3622.
Klingenberg, C. P. 2011. MorphoJ: an integrated software package for geometric morphometrics. Molecular Ecology Resources 11:353-357.
Kramer, R. S. S. and R. Ward. 2010. Internal facial features are signals of personality and health. Quarterly Journal of Experimental Psychology 63:2273-2287.
Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and P. Genome Project Data. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078-2079.
Lorenc, A., M. Linnenbrink, I. Montero, M. B. Schilhabel, and D. Tautz. 2014. Genetic Differentiation of Hypothalamus Parentally Biased Transcripts in Populations of the House Mouse Implicate the Prader-Willi Syndrome Imprinted Region as a Possible Source of Behavioral Divergence. Molecular Biology and Evolution 31:3240-3249.
Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15.
Lüdecke, D. 2019. sjstats: Statistical Functions for Regression Models. Zenodo 1284473.
Nakatani, J., K. Tamada, F. Hatanaka, S. Ise, H. Ohta, K. Inoue, S. Tomonaga, Y. Watanabe, Y. J. Chung, R. Banerjee, K. Iwamoto, T. Kato, M. Okazawa, K. Yamauchi, K. Tanda, K. Takao, T. Miyakawa, A. Bradley, and T. Takumi. 2009. Abnormal Behavior in a Chromosome-Engineered Mouse Model for Human 15q11-13 Duplication Seen in Autism. Cell 137:1235-1246.
Nicholls, R. D., S. Saitoh, and B. Horsthemke. 1998. Imprinting in Prader-Willi and Angelman syndromes. Trends in Genetics 14:194-200.
Nicod, J., R. W. Davies, N. Cai, C. Hassett, L. Goodstadt, C. Cosgrove, B. K. Yee, V. Lionikaite, R. E. McIntyre, C. A. Remme, E. M. Lodder, J. S. Gregory, T. Hough, R. Joynson, H. Phelps, B. Nell, C. Rowe, J. Wood, A. Walling, N. Bopp, A. Bhomra, P. Hernandez-Pliego, J. Callebert, R. M. Aspden, N. P. Talbot, P. A. Robbins, M. Harrison, M. Fray, J. M. Launay, Y. M. Pinto, D. A. Blizard, C. R. Bezzina, D. J. Adams, P. Franken, T. Weaver, S. Wells, S. D. M. Brown, P. K. Potter, P. Klenerman, A. Lionikas, R. Mott, and J. Flint. 2016. Genome-wide association of

multiple complex traits in outbred mice by ultra-low-coverage sequencing. Nature Genetics 48:912-+.
Nuss, P. 2015. Anxiety disorders and GABA neurotransmission: a disturbance of modulation. Neuropsychiatric Disease and Treatment 11:165-175.
Paigen, K. and P. Petkov. 2010. Mammalian recombination hot spots: properties, control and evolution. Nature Reviews Genetics 11:221-233.
Pallares, L. F., P. Carbonetto, S. Gopalakrishnan, C. C. Parker, C. L. Ackert-Bicknell, A. A. Palmer, and D. Tautz. 2015. Mapping of Craniofacial Traits in Outbred Mice Identifies Major Developmental Genes Involved in Shape Determination. Plos Genetics 11.
Parker, C. C., S. Gopalakrishnan, P. Carbonetto, N. M. Gonzales, E. Leung, Y. J. Park, E. Aryee, J. Davis, D. A. Blizard, C. L. Ackert-Bicknell, A. Lionikas, J. K. Pritchard, and A. A. Palmer. 2016. Genome-wide association study of behavioral, physiological and gene expression traits in outbred CFW mice. Nature Genetics 48:919-+.
Penke, L. and M. Jokela. 2016. The evolutionary genetics of personality revisited. Current Opinion in Psychology 7:104-109.
Pezer, Z., B. Harr, M. Teschke, H. Babiker, and D. Tautz. 2015. Divergence patterns of genic copy number variation in natural populations of the house mouse (Mus musculus domesticus) reveal three conserved genes with major population-specific expansions. Genome Research 25:1114-1124.
Pinheiro, L. B., V. A. Coleman, C. M. Hindson, J. Herrmann, B. J. Hindson, S. Bhat, and K. R. Emslie. 2012. Evaluation of a Droplet Digital Polymerase Chain Reaction Format for DNA Copy Number Quantification. Analytical Chemistry 84:1003-1011.
Qi, Y., L. Purtell, M. Fu, N. J. Lee, J. Aepler, L. Zhang, K. Loh, R. F. Enriquez, P. A. Baldock, S. Zolotukhin, L. V. Campbell, and H. Herzog. 2016. Snord116 is critical in the regulation of food intake and body weight. Scientific Reports 6.
Quan, P. L., M. Sauzade, and E. Brouzes. 2018. dPCR: A Technology Review. Sensors 18.
Raabe, C. A., R. Voss, D. M. Kummerfeld, J. Brosius, C. R. Galiveti, A. Wolters, J. Seggewiss, A. Huge, B. V. Skryabin, and T. S. Rozhdestvensky. 2019. Ectopic expression of Snord115 in choroid plexus interferes with editing but not splicing of 5-Ht2c receptor pre-mRNA in mice. Scientific Reports 9.
Reale, D., S. M. Reader, D. Sol, P. T. McDougall, and N. J. Dingemanse. 2007. Integrating animal temperament within ecology and evolution. Biological Reviews 82:291-318.
Réale, D. 2007. What do we really know about selection on personality? European Journal of Personality 21:618-620.
Sanchez-Roige, S., J. C. Gray, J. MacKillop, C. H. Chen, and A. A. Palmer. 2018. The genetics of human personality. Genes Brain and Behavior 17.
Sato, M. 2017. Early Origin and Evolution of the Angelman Syndrome Ubiquitin Ligase Gene Ube3a. Frontiers in Cellular Neuroscience 11.
Schlager, S. 2017. Morpho and Rvcg - Shape Analysis in R . Pp. 217-256 *in* G. Zheng, S. Li, and G. Szekely, eds. Statistical Shape and Deformation Analysis . Academic Press.
Stamm, S., S. B. Gruber, A. G. Rabchevsky, and R. B. Emeson. 2017. The activity of the serotonin receptor 2C is regulated by alternative splicing. Human Genetics 136:1079-1091.
Stoffel, M. A., S. Nakagawa, and H. Schielzeth. 2017. rptR: repeatability estimation and variance decomposition by generalized linear mixed-effects models. Methods in Ecology and Evolution 8:1639-1644.
Tooby, J. and L. Cosmides. 1990. ON THE UNIVERSALITY OF HUMAN-NATURE AND THE UNIQUENESS OF THE INDIVIDUAL - THE ROLE OF GENETICS AND ADAPTATION. Journal of Personality 58:17-67.
Uher, J. 2011. Individual Behavioral Phenotypes: An Integrative Meta-Theoretical Framework. Why "Behavioral Syndromes" Are Not Analogs of "Personality". Developmental Psychobiology 53:521-548.

von Merten, S., S. Hoier, C. Pfeifle, and D. Tautz. 2014. A role for ultrasonic vocalisation in social communication and divergence of natural populations of the house mouse (Mus musculus domesticus). PloS one 9:e97244.

Weyers, P., H. Krebs, and W. Janke. 1995. RELIABILITY AND CONSTRUCT-VALIDITY OF THE GERMAN VERSION OF CLONINGERS TRIDIMENSIONAL PERSONALITY QUESTIONNAIRE. Personality and Individual Differences 19:853-861.

Wilson, D. S. 1998. Adaptive individual differences within single populations. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 353:199-205.

Wittchen, H. U. and H. Pfister. 1997. DIA-X-Interview. Instruktionsmanual zur Durchführung von DIA-X-Interviews. Swets & Zeitlinger, Frankfurt.

Wolf, M. and F. J. Weissing. 2012. Animal personalities: consequences for ecology and evolution. Trends in Ecology & Evolution 27:452-461.

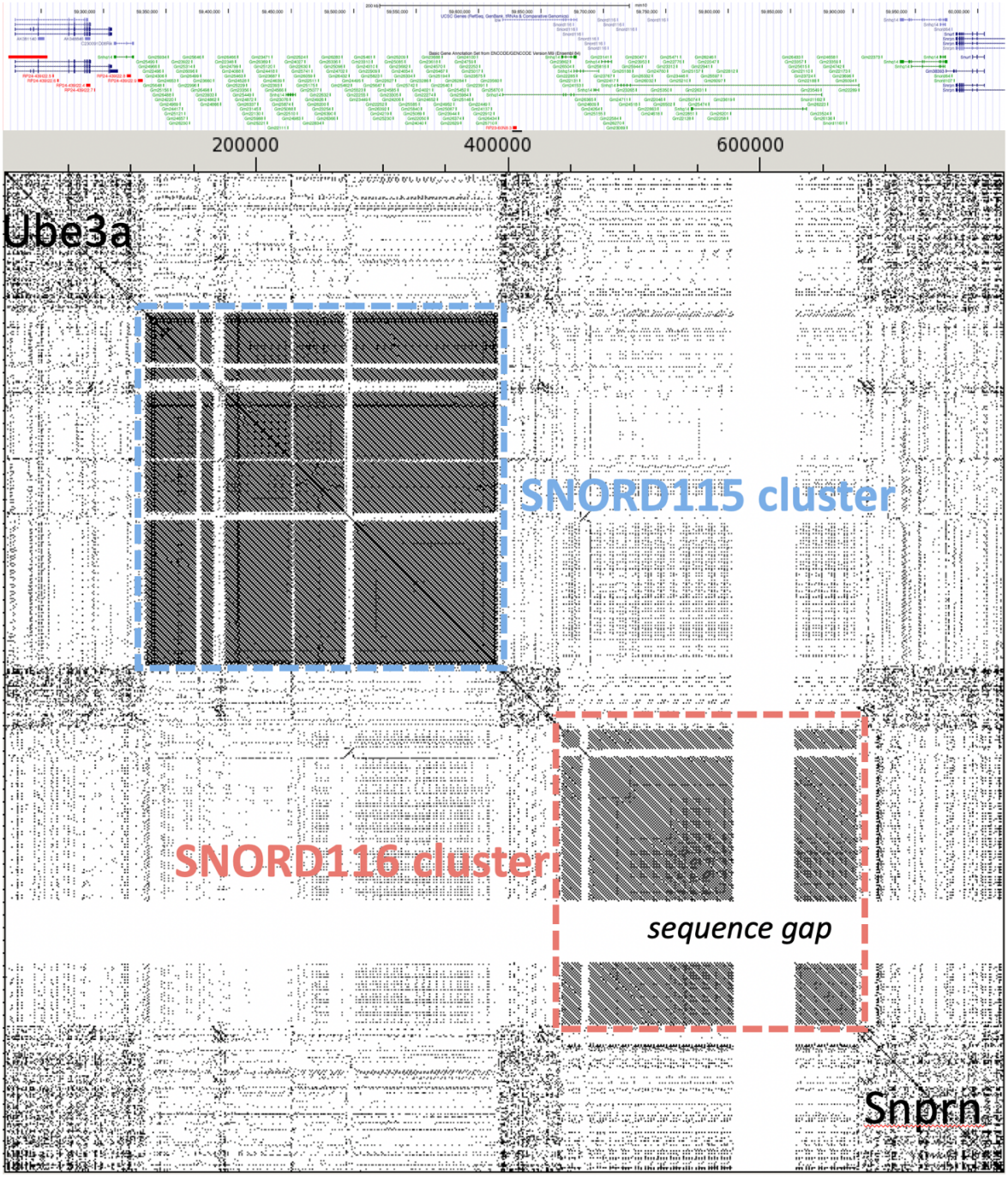
Zhang, X. S. and W. G. Hill. 2005. Genetic variability under mutation selection balance. Trends in Ecology & Evolution 20:468-470.

Zhang, Y. J., J. H. Yang, Q. S. Shi, L. L. Zheng, J. Liu, H. Zhou, H. Zhang, and L. H. Qu. 2014. Rapid Birth-and-Death Evolution of Imprinted snoRNAs in the Prader-Willi Syndrome Locus: Implications for Neural Development in Euarchontoglires. Plos One 9.

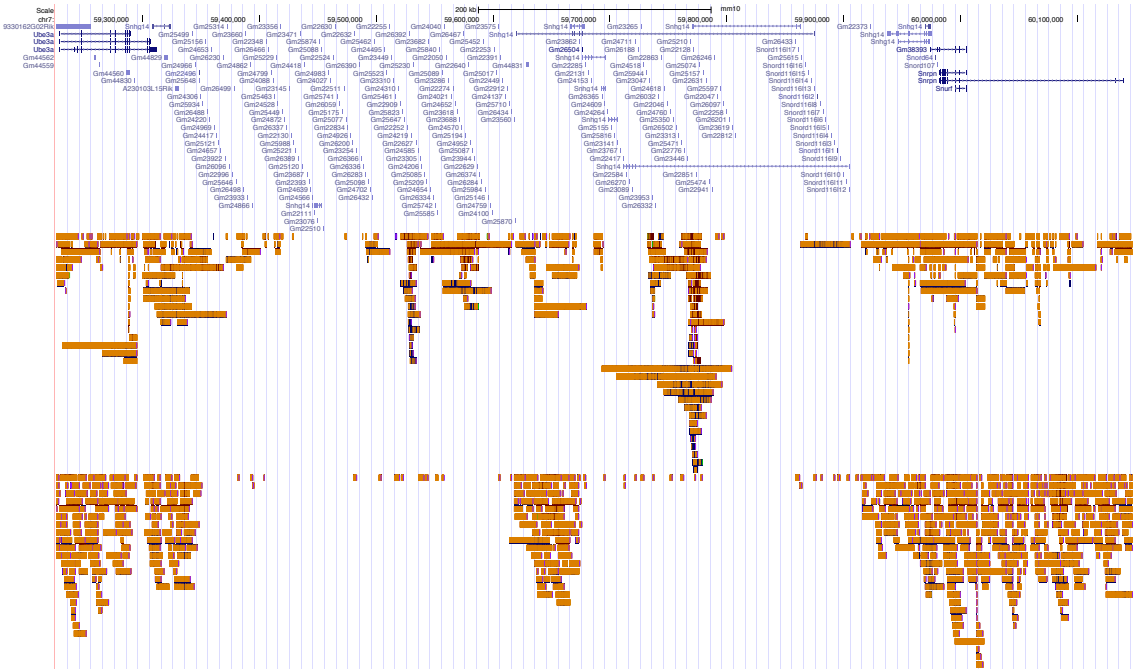
Zhao, Y., Q. Y. Xia, Y. P. Yin, and Z. K. Wang. 2016. Comparison of Droplet Digital PCR and Quantitative PCR Assays for Quantitative Detection of Xanthomonas citri Subsp citri. Plos One 11.

Zurita, E., M. Chagoyen, M. Cantero, R. Alonso, A. Gonzalez-Neira, A. Lopez-Jimenez, J. A. Lopez-Moreno, C. P. Landel, J. Benitez, F. Pazos, and L. Montoliu. 2011. Genetic polymorphisms among C57BL/6 mouse inbred strains. Transgenic Research 20:481-489.

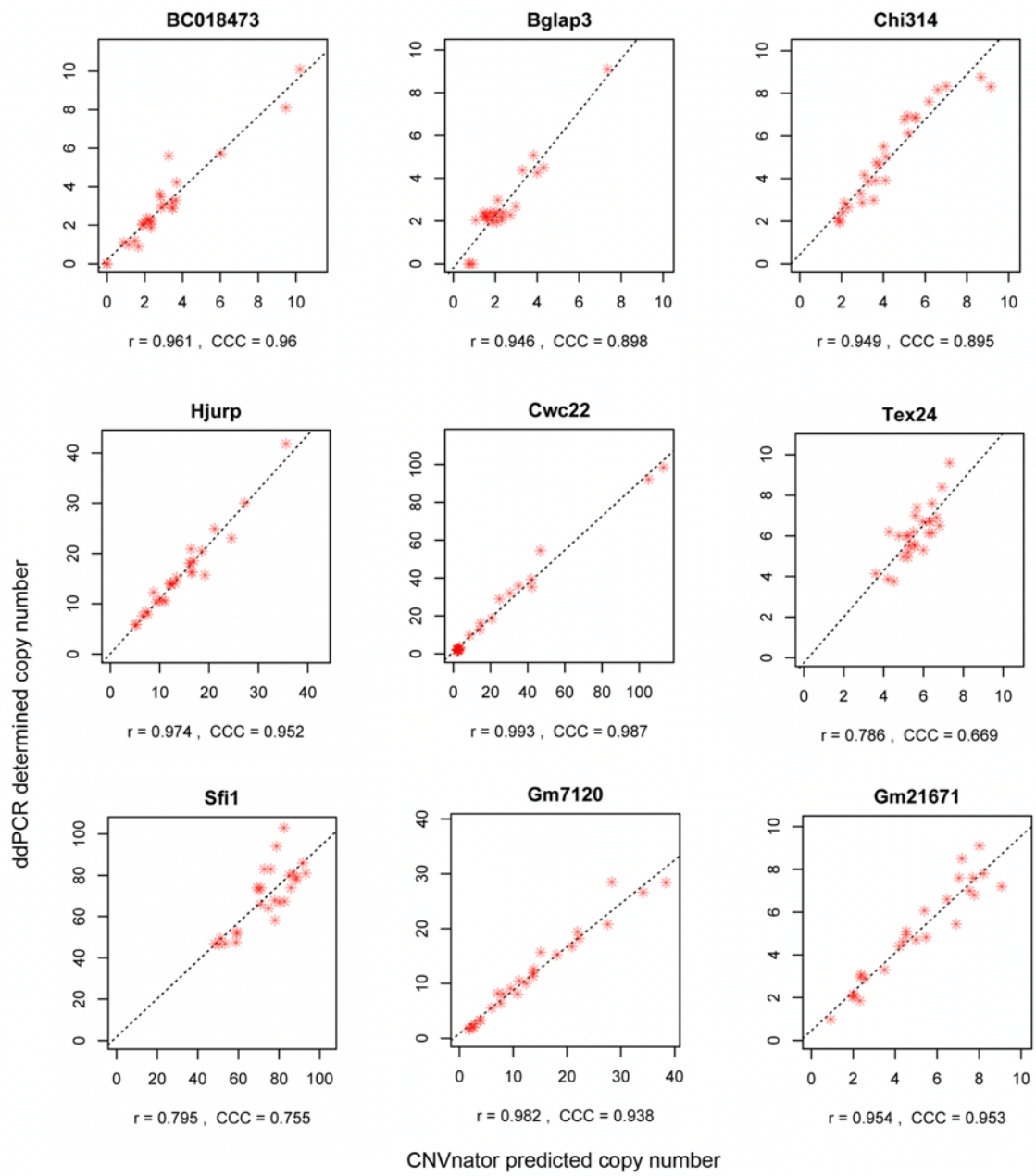
Supplementary Figures



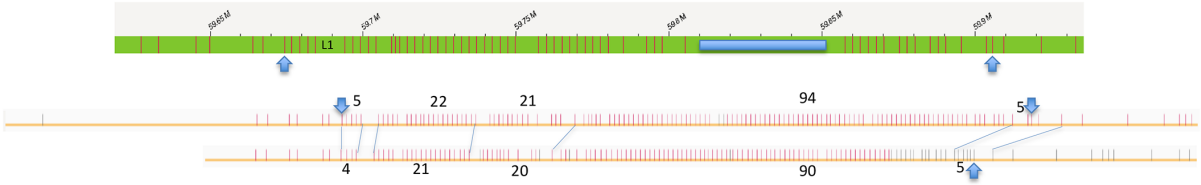
S1 Figure: SNORD115 and SNORD116 tandem repeat clusters. Dotplot representation of the PWS region between *Ube3a* and *Snrpn* matched against itself. Sequence regions with matches appear as diagonals, tandem repeats as clusters of diagonals. The corresponding annotations from the UCSC browser are shown at the top. The Snord115 and SNORD116 clusters are boxed, the sequence gap (i.e. a 50kb region filled with N in the reference sequence) is indicated.



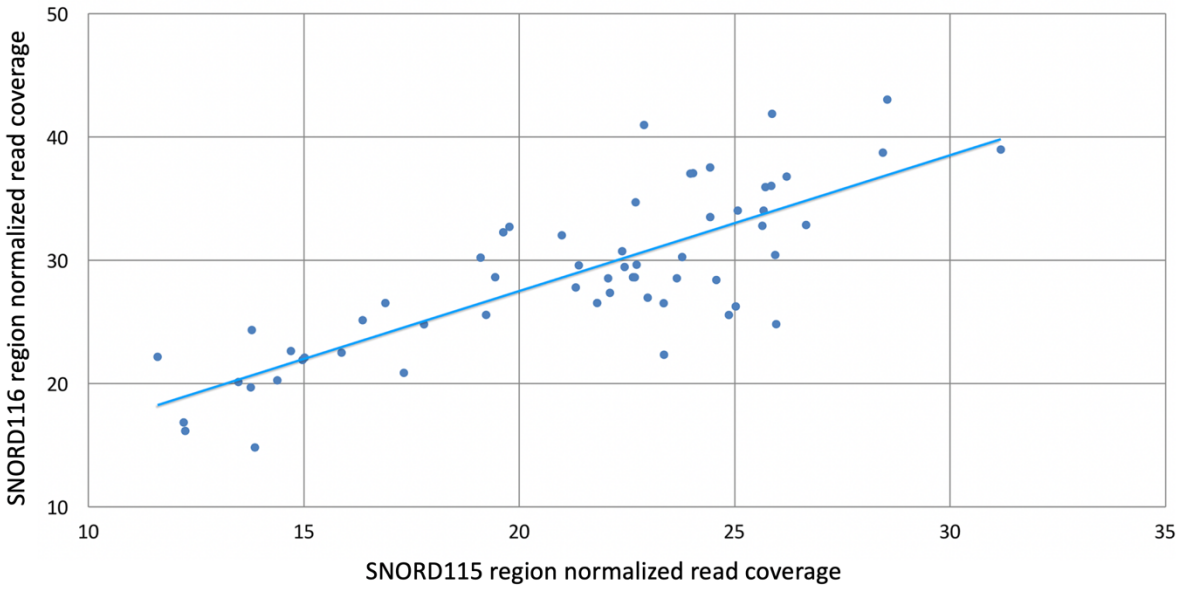
S2 Figure: Long read coverage comparisons. UCSC browser tracks showing the region between *Ube3a* and *Snrpn* with the corresponding reads obtained from Oxford Nanopore (top track) and PacBio (bottom track) sequencing of whole genomes (compiled from several runs). Each read is represented by an orange bar showing the extent of its coverage. Note that the mapping algorithm does not place reads in regions of internal repeats unless there are specific additional SNPs that allow unequivocal placements. This explains the relative paucity of reads in the SNORD115 and SNORD116 tandem repeat clusters.



S3 Figure. Comparison of copy numbers at tandem repeat loci determined by sequencing read depth versus digital PCR. Assays for 9 loci with copy numbers of 10 or higher were designed and used to screen 27 individuals (red asterisks) for which read depth information was available from
1090 whole genome sequencing; data taken from Pezer et al. (Genome Res 25, 1114-1124, 2015). The results of ddPCR runs (y axis) are plotted against copy numbers predicted with CNVnator based on sequencing read depth (x axis) for the amplicon regions. Fitted regression lines are shown dashed; *r* is Pearson's correlation coefficient; CCC is Lin's Concordance Correlation Coefficient.

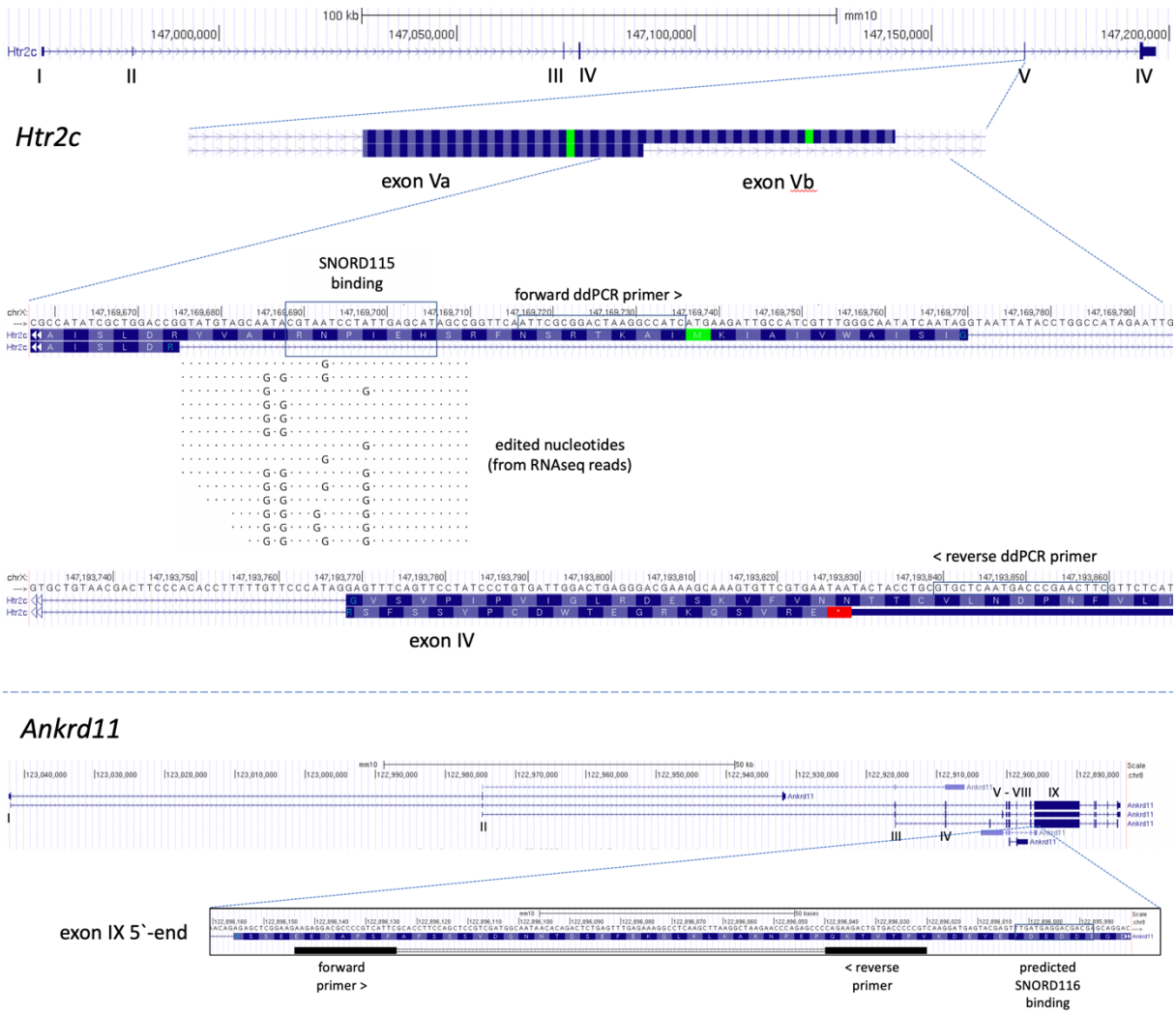


S4 Figure. Optical mapping analysis of the SNORD116 region. Optical mapping was done following the standard procedures on a Bionano Saphyr instrument using Nt.BspQI, which labels GCTCTTC strings in the DNA. This string occurs once in the SNORD116 repeat consensus sequence. Fragments that span the whole region of the SNORD116 cluster plus the adjacent regions were manually retrieved from the fragment display function of the instrument software. The green bar at the top represents the predicted labelling positions for the restriction enzyme from the mouse reference genome. The tandem repeat region (marked by the arrow heads) becomes apparent by the regular spacing of sites, which are, however, not perfect. The string includes at one end an inserted L1 element that interrupts this spacing and which is marked in the bar for the reference sequence. We find this same interruption in the recorded fragments and this serves as a suitable marker to align the fragments with the reference sequence. There are further interruption of the regular pattern that are due to repeat copies that lack the Nt.BspQI site. The reference sequence includes also a 50kb gap filled with N (marked in blue). The lines below the reference bar represent two fragments that are interpreted to constitute the two alleles of this individual. Visible gaps are used to align them (blue line connectors) and count the number of repeats (annotated numbers). Note that the repeat counts consider the gaps in the regular string as variants that have no Nt.BspQI site or labelling drop outs, i.e. they are added to the total numbers. Allele 1 of this mouse shows 147 copies, allele 2 shows 140 copies. The ddPCR measurement for this mouse was 293 copies, which is well within the range of the total of 287 predicted copies from the optical mapping.



1120

S5 Figure. Correlation of read coverage between the SNORD115 and SNORD116 regions based on whole genome sequencing data. Genome sequencing reads from individuals of Mus musculus populations, including the three subspecies (Harr et al. 2016) were mapped against the house mouse
1125 reference genome mm10 with 'ngm' (Sedlazeck et al. 2013), followed by sorting, marking and removing duplicates with the picard software suite (https://broadinstitute.github.io/picard/). The resulting bam files can be obtained via http://wwwuser.gwdg.de/~evolbio/evolgen/wildmouse/introgression/ngm/bam/. We used the 'genomeCoverageBed' tool from the bedtools software suite v2.26.0 (Quinlan and Hall 2010) with the
1130 option 'bga' to obtain site-specific coverage across the snord115 (chr7:59,340,775-59,620,255) and snord116 (chr7:59,671,705-59,905,815) region. Subsequently, the coverage files were used to calculate the mean coverage per site in the respective genome regions. These mean coverages were then normalized according to the total reads obtained from each genome.



S6 Figure. Structures of the *Htr2C* and Ankrd11 genes. The structure details with exon arrangements are taken from the UCSC browser. Exon numbers are indicated in roman numerals. For *Htr2c*, exon V and VI are enlarged, for *Ankrd11* exon IX is enlarged. SNORD binding sites and primer sites for the ddPCR experiment for determining copy numbers are marked with boxes. Further, we added a read alignment from RNASeq reads to show the edited sites in *Htr2c*.

Supplementary Tables

S1 Table

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**SNORD115 and SNORD116 copy numbers for *Mus* strains and stocks**												
2													
3				**SNORD115 copy numbers**					**SNORD116 copy numbers**				
4	MouseID	Populations / strain	sex	SNORD115 (replicate1)	SNORD115 (replicate2)	SNORD115 (replicate3)	average	SD	SNORD116 (replicate1)	SNORD116 (replicate2)	SNORD116 (replicate3)	average	SD
77	50046761	*Mouse inbred strain C57BL6/J*	Female	103	105	107	98	1.84	143	146	144	140	1.39
78	50046787	*Mouse inbred strain C57BL6/J*	Female	125	122	128	126	3.01	149	147	155	144	4.22
79	50046785	*Mouse inbred strain C57BL6/J*	Female	144	148	143	141	2.65	173	177	169	174	4.00
80	50046773	*Mouse inbred strain C57BL6/J*	Female	55	58	53	55	2.50	71	73	69	72	2.01
81	50046777	*Mouse inbred strain C57BL6/J*	Female	74	71	77	73	3.01	135	138	129	139	4.62
82	50046786	*Mouse inbred strain C57BL6/J*	Female	225	223	221	232	2.17	258	259	256	260	1.58
83	50046778	*Mouse inbred strain C57BL6/J*	Female	116	119	115	113	2.14	166	164	168	165	2.01
84	50046770	*Mouse inbred strain C57BL6/J*	Female	118	120	114	121	3.10	139	137	135	144	1.84
85	50046762	*Mouse inbred strain C57BL6/J*	Female	122	123	119	124	2.08	168	167	164	173	2.08
86	50046788	*Mouse inbred strain C57BL6/J*	Female	100	101	104	96	1.95	141	139	143	140	2.01
87	50046775	*Mouse inbred strain C57BL6/J*	Female	181	176	186	182	5.00	209	206	212	208	3.01
88	50046781	*Mouse inbred strain C57BL6/J*	Female	288	293	281	291	6.05	302	305	306	296	1.90
89	50046792	*Mouse inbred strain C57BL6/J*	Male	54	51	59	55	4.04	85	89	86	87	2.08
90	50046797	*Mouse inbred strain C57BL6/J*	Male	83	89	90	87	3.79	134	138	132	135	3.06
91	50046801	*Mouse inbred strain C57BL6/J*	Male	99	91	90	93	4.93	137	130	139	135	4.73
92	50046806	*Mouse inbred strain C57BL6/J*	Male	177	182	185	181	4.04	228	226	222	225	3.06
93	50046807	*Mouse inbred strain C57BL6/J*	Male	208	193	195	199	8.14	223	229	230	227	3.79
94	50046809	*Mouse inbred strain C57BL6/J*	Male	161	163	157	160	3.06	219	225	218	221	3.79
95	50046814	*Mouse inbred strain C57BL6/J*	Male	172	173	169	171	2.08	220	227	216	221	5.57
96	50046815	*Mouse inbred strain C57BL6/J*	Male	176	179	168	174	5.69	221	222	228	224	3.79
97	50046816	*Mouse inbred strain C57BL6/J*	Male	117	120	116	118	2.08	161	160	164	162	2.08
98	50046820	*Mouse inbred strain C57BL6/J*	Male	108	103	107	106	2.65	142	148	146	145	3.06
99	50046763	*Mouse inbred strain C57BL6/J*	Male	204	201	196	200	4.04	229	229	224	227	2.89
100	50046764	*Mouse inbred strain C57BL6/J*	Male	193	206	200	200	6.51	235	229	230	231	3.21
101	50046765	*Mouse inbred strain C57BL6/J*	Male	227	229	220	225	4.73	261	259	253	258	4.16
102	50046766	*Mouse inbred strain C57BL6/J*	Male	150	153	154	152	2.08	203	207	206	205	2.08
103	50046767	*Mouse inbred strain C57BL6/J*	Male	158	152	156	155	3.06	211	209	207	209	2.00
104	50046769	*Mouse inbred strain C57BL6/J*	Male	158	157	155	157	1.53	213	220	214	216	3.79
105	50046772	*Mouse inbred strain C57BL6/J*	Female	165	158	157	160	4.36	219	223	222	221	2.08
106	50046774	*Mouse inbred strain C57BL6/J*	Female	80	83	81	81	1.53	134	128	135	132	3.79
107	50046776	*Mouse inbred strain C57BL6/J*	Female	101	98	102	100	2.08	145	143	136	141	4.73
108	50046779	*Mouse inbred strain C57BL6/J*	Female	97	103	99	100	3.06	139	147	142	143	4.04
109	50046780	*Mouse inbred strain C57BL6/J*	Female	117	121	115	118	3.06	167	163	157	162	5.03
110	50046784	*Mouse inbred strain C57BL6/J*	Female	118	123	115	119	4.04	163	160	171	165	5.69
111	50046790	*Mouse inbred strain C57BL6/J*	Female	128	124	129	127	2.65	172	169	177	173	4.04
112							**average SD:**	**3.54**				**average SD:**	**3.15**
113													

S2 Table

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	**ddPCR results for biologcal replicates**																
2	mouse samples from *M. m. domesticus populations* CB and MC																
3																	
4		**ear samples**								**brain samples**							
5																	
6	Animal ID	SNORD115 (replicate 1)	SNORD115 (replicate 2)	SNORD115 (replicate 3)	**SNORD115 average**	SNORD116 (replicate 1)	SNORD116 (replicate 2)	SNORD116 (replicate 3)	**SNORD116 average**	SNORD115 (replicate 1)	SNORD115 (replicate 2)	SNORD115 (replicate 3)	**SNORD115 average**	SNORD116 (replicate 1)	SNORD116 (replicate 2)	SNORD116 (replicate 3)	**SNORD116 average**
7	1789_1	56	58	52	**55.3**	68	62	64	**64.7**	52	60	55	**55.7**	69	60	64	**64.3**
8	1286_3	69	67	67	**67.7**	69	67	73	**69.7**	61	70	65	**65.3**	78	72	74	**74.7**
9	1326_10	69	73	74	**72.0**	116	112	114	**114.0**	68	66	72	**68.7**	118	114	112	**114.7**
10	1789_11	69	71	72	**70.7**	101	99	95	**98.3**	66	74	70	**70.0**	104	92	98	**98.0**
11	1429_4	85	81	77	**81.0**	88	85	84	**85.7**	78	84	82	**81.3**	82	85	90	**85.7**
12	1429_30	78	69	73	**73.3**	78	70		**74.0**	73	72	74	**73.0**	78	70	74	**74.0**
13	1403_30	115	110	104	**109.7**	127	123	121	**123.7**	109	113	107	**109.7**	121	127	123	**123.7**
14	1403_1	87	81	83	**83.7**	95	90	87	**90.7**	84	79	91	**84.7**	91	94	86	**90.3**
15	1789_3	146	149	142	**145.7**	99	96	91	**95.3**	144	148	140	**144.0**	95	93	99	**95.7**
16	1286_1	104	96	99	**99.7**	71	77	74	**74.0**	97	103	99	**99.7**	74	80	76	**76.7**
17	1454_10	105	107	102	**104.7**	113	103	107	**107.7**	106	111	99	**105.3**	107	112	104	**107.7**
18	1444_4	154	151	159	**154.7**	165	163	167	**165.0**	157	153	155	**155.0**	166	160	162	**162.7**
19	1311_1	70	67	62	**66.3**	82	79	80	**80.3**	60	72	65	**65.7**	79	85	75	**79.7**
20	1892_3	65	62	68	**65.0**	74	68	70	**70.7**	64	68	62	**64.7**	70	73	69	**70.7**
21	1454_1	160	165	155	**160.0**	166	169	163	**166.0**	157	160	165	**160.7**	164	169	157	**163.3**
22	1789_10	103	106	100	**103.0**	107	110	113	**110.0**	101	94	106	**100.3**	110	112	108	**110.0**
23	1429_10	170	174	166	**170.0**	179	182	180	**180.3**	169	164	178	**170.3**	179	183	180	**180.7**
24	1429_30	128	118	121	**122.3**	128	124	126	**126.0**	118	128	122	**122.7**	131	125	127	**127.7**
25	1429_31	153	147	149	**149.7**	147	153	149	**149.7**	150	154	146	**150.0**	151	154	146	**150.3**
26	1429_3	97	103	99	**99.7**	98	102	99	**99.7**	98	106	101	**101.7**	99	103	97	**99.7**
27	1349_30	72	68	71	**70.3**	80	86	74	**80.0**	71	67	73	**70.3**	79	83	77	**79.7**
28	1311_30	91	99	94	**94.7**	100	106	103	**103.0**	90	94	100	**94.7**	105	108	104	**105.7**
29	1326_1	71	61	67	**66.3**	81	85	84	**83.3**	65	63	69	**65.7**	77	81	75	**77.7**
30	1286_10	73	65	70	**69.3**	90	82	85	**85.7**	69	73	65	**69.0**	86	83	89	**86.0**
31	1349_3	126	123	122	**123.7**	120	123	126	**123.0**	120	126	124	**123.3**	125	128	124	**125.7**
32	1444_10	38	41	33	**37.3**	44	50	46	**46.7**	41	33	37	**37.0**	49	51	47	**49.0**
33	1892_11	86	89	94	**89.7**	101	99	100	**100.0**	93	87	91	**90.3**	100	96	104	**100.0**
34	1892_40	107	111	113	**110.3**	123	120	119	**120.7**	105	115	110	**110.0**	121	122	123	**122.0**
35	1326_3	155	152	149	**152.0**	161	158	166	**161.7**	148	156	152	**152.0**	160	162	168	**163.3**
36	1311_3	84	81	76	**80.3**	90	85	95	**90.0**	79	84	76	**79.7**	89	84	96	**89.7**
37	1429_11	83	78	73	**78.0**	94	96	100	**96.7**	79	81	77	**79.0**	96	95	99	**96.7**
38	1326_30	203	201	197	**200.3**	201	199	205	**201.7**	201	198	204	**201.0**	204	205	203	**204.0**
39	1349_1	122	114	117	**117.7**	130	133	127	**130.0**	115	114	118	**115.7**	129	129	127	**128.3**
40	1789_4	91	87	93	**90.3**	79	83	81	**81.0**	87	89	93	**89.7**	75	87	80	**80.7**
41	1892_10	90	94	100	**94.7**	109	103	105	**105.7**	93	95	97	**95.0**	103	100	110	**104.3**
42	1403_11	91	85	95	**90.3**	97	93	103	**97.7**	87	93	90	**90.0**	98	96	100	**98.0**
43	1349_10	52	47	54	**51.0**	51	53	55	**53.0**	51	47	54	**50.7**	53	55	51	**53.0**
44	1403_10	81	79	82	**80.7**	90	86	94	**90.0**	79	74	86	**79.7**	96	84	89	**89.7**
45	1311_11	140	142	148	**143.3**	163	167	170	**166.7**	140	142	149	**143.7**	168	172	162	**167.3**
46	1311_10	150	154	159	**154.3**	160	163	168	**163.7**	150	147	155	**150.7**	165	168	160	**164.3**
47								correlation between ear and brain samples					R^2=0.99				R^2=0.99
48																	
49																	
50	**primer comparison for SNORD loci**																
51																	
52		primer															
53	Animal ID	SNORD115	SNORD115-b		SNORD116	SNORD116-b											
54	50046802	103	104		200	200											
55	50046795	137	137		153	155											
56	50046796	121	121		161	161											
57	50046803	174	170		224	222											
58	50046818	140	140		186	186											
59	50046805	145	140		174	177											
60	50046819	199	199		196	196											
61	50046813	137	136		175	174											
62	50046817	122	123		146	145											
63	50046804	121	121		155	155											
64	50046811	131	131		200	202											
65	50046812	130	130		250	250											
66	50046799	225	224		256	256											
67	50046800	140	141		225	225											
68	50046793	145	146		245	246											
69	50046794	185	185		265	265											
70	50046808	170	171		244	243											
71	50046810	170	170		210	210											
72	50046798	170	170		210	210											
73	correlation between primers		R^2=0.99			R^2=0.99											

S3 Table

Lack of correlation with SNORD copy numbers to two other CNV loci (Sfi1 and CWC22)

Population CB

Animal ID	Sfi copy number	SNORD115 copy number	SNORD116 copy number
252980	29	95	132
254434	33	195	224
251505	69	177	206
253612	53	130	155
251736	44	166	197
253911	98	145	171
252388	25	131	167
250737	18	152	180
254089	19	226	243
253366	47	205	193
253721	42	212	245
254154	46	218	241
252264	27	220	245
254321	28	243	255
250148	42	233	273
252408	37	263	294
252909	85	202	228
252396	71	273	322
254175	50	290	324
250230	56	290	325
253670	46	230	231
251786	36	146	190
R^2		**0.007**	**0.01**
P-value		**0.71**	**0.65**

Animal ID	Cwc22 copy number	SNORD115 copy number	SNORD116 copy number
252980	18	95	132
254434	54	195	224
251505	69	177	206
253612	89	130	155
251736	31	166	197
253911	25	145	171
252388	49	131	167
250737	67	152	180
254089	55	226	243
253366	48	205	193
253721	47	212	245
254154	33	218	241
252264	22	220	245
254321	99	243	255
250148	39	233	273
252408	69	263	294
252909	71	202	228
252396	74	273	322
254175	39	290	324
250230	64	290	325
253670	35	230	231
251786	52	146	190
R^2		**0.04**	**0.04**
P-value		**0.35**	**0.36**

Population MC

Animal ID	Sfi copy number	SNORD115 copy number	SNORD116 copy number
250687	85	137	137
254940	121	108	86
250556	49	115	110
250622	63	127	116
253429	75	133	127
254298	58	139	160
251209	79	155	190
254488	39	157	200
253218	61	160	213
250481	52	169	185
250400	112	175	220
250061	55	160	170
252446	49	156	162
254418	76	160	185
251023	69	190	205
250053	36	221	241
254918	95	229	249
251016	81	168	172
251249	18	237	257
253816	67	246	267
250545	69	220	290
252169	114	172	169
251664	99	185	220
R^2		**0.04**	**0.05**
P-value		**0.34**	**0.32**

Lack of behavioral correlation to two other CNV loci (Sfi and Cwc2

Population CB

Animal ID	Sfi copy number	Wall time	Dark arm time
252980	29	86.68	150
254434	33	243.84	115.32
251505	69	124.48	101
253612	53	94.64	158.4
251736	44	165.76	113.64
253911	98	193.2	123.16
252388	25	169.6	106.28
250737	18	157.4	163.08
254089	19	215.28	158.48
253366	47	198.12	124.56
253721	42	208.24	146.4
254154	46	209.08	151.84
252264	27	136.92	165.1
254321	28	229.44	137
250148	42	208.12	213.68
252408	37	181.04	144.56
252909	85	131	155.44
252396	71	231.64	183.76
254175	50	242	221.96
250230	56	245.08	202.96
253670	46	230.88	248.1
251786	36	183.04	122.2
R^2		**0.0006**	**0.0004**
P-value		**0.92**	**0.93**

Animal ID	Cwc22 copy number	Wall time	Dark arm time
252980	18	86.68	150
254434	54	243.84	115.32
251505	69	124.48	101
253612	89	94.64	158.4
251736	31	165.76	113.64
253911	25	193.2	123.16
252388	49	169.6	106.28
250737	67	157.4	163.08
254089	55	215.28	158.48
253366	48	198.12	124.56
253721	47	208.24	146.4
254154	33	209.08	151.84
252264	22	136.92	165.1
254321	99	229.44	137
250148	39	208.12	213.68
252408	69	181.04	144.56
252909	71	131	155.44
252396	74	231.64	183.76
254175	39	242	221.96
250230	64	245.08	202.96
253670	35	230.88	248.1
251786	52	183.04	122.2
R^2		**0.001**	**0.01**
P-value		**0.88**	**0.65**

Population MC

Animal ID	Sfi copy number	Wall time	Dark arm time
250687	85	193.92	100.04
254940	121	187.6	106.8
250556	49	199.44	223.84
250622	63	195.24	119.68
253429	75	195.76	186.36
254298	58	215.64	174.88
251209	79	276.84	195
254488	39	292.52	162.88
253218	61	237.24	176.12
250481	52	216.92	210.44
250400	112	251.88	136.8
250061	55	230.04	150.04
252446	49	239.64	162.32
254418	76	236	217.36
251023	69	236.36	140.76
250053	36	257.96	184.44
254918	95	244	235
251016	81	236.36	233.76
251249	18	255.64	175.12
253816	67	261.96	207
250545	69	286.52	263.3
252169	114	185.08	202
251664	99	262.88	260.1
R^2		**0.09**	**0.0009**
P-value		**0.16**	**0.89**

S4 Table

	A	B	C	D	E	F	G	H	I
1	**relative RNA expression in the brain, based on ddPCR scores**								
2									
3	MouseID	Population	sex	SNORD 115 copy number	SNORD 115 expression	Htr2c ExonV expression	SNORD116 copy number	SNORD 116 expression	ANKRD11 exon 10 expression
22	251664	MC	Male	185	211	220	220	138	140
23		correlation of copy number with SNORD expression			R^2=0.93			R^2=0.53	
24		correlation of SNORD expression with target exon expression				R^2=0.65			R^2=0.72
25									
26	MouseID	Population	sex	SNORD 115 copy number	SNORD 115 expression	Htr2c ExonV expression	SNORD116 copy number	SNORD 116 expression	ANKRD11 exon 10 expression
27	252980	CB	Male	95	132	267	131	173	227
28	251736	CB	Male	166	146	281	197	186	244
29	254089	CB	Male	226	188	255	243	227	363
30	252264	CB	Male	220	167	302	245	239	387
31	250148	CB	Male	233	211	345	270	275	417
32	252396	CB	Male	275	298	393	320	339	479
33	250230	CB	Male	291	281	416	325	362	460
34	251786	CB	Male	146	227	405	190	287	339
35	250687	MC	Male	137	142	171	137	66	83
36	254940	MC	Male	108	84	151	86	112	75
37	254298	MC	Male	139	115	153	160	77	47
38	253218	MC	Male	160	157	193	213	89	96
39	252446	MC	Male	156	164	207	160	116	111
40	254918	MC	Male	221	230	227	249	156	129
41	250545	MC	Male	220	232	235	290	187	149
42	251664	MC	Male	185	211	220	220	138	140
43		correlation of copy number with SNORD expression			R^2=0.71			R^2=0.53	
44		correlation of SNORD expression with target exon expression				R^2=0.56			R^2=0.91
45									

1160 **S5 Table**

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	**Behavioral score data for *Mus* strains and stocks**																									
2						**Open field test**												**Elevated Plus Maze tes**								
3	MouseID	Populations / strain	sex	SNORD 115 copy number	SNORD 116 copy number	WallTime (replicate1)	WallTime (replicate2)	WallTime (replicate3)	Travel Distance at the Wall (replicate 1)	Travel Distance at the Wall (replicate 2)	Travel Distance at the Wall (replicate 3)	Speed (replicate 1)	Speed (replicate 2)	Speed (replicate 3)	Total distance (replicate 1)	Total distance (replicate 2)	Total distance (replicate 3)	DarkArm (replicate1)	DarkArm (replicate2)	DarkArm (replicate3)	Speed (replicate 1)	Speed (replicate 2)	Speed (replicate 3)	Total distance (replicate 1)	Total distance (replicate 2)	Total distance (replicate 3)
94	50046804	Mouse inbred strain C57BL6/J	Male	117	153	205	169	146	3711	4145	6412	28	13	19	6353	4582	4573	98	132	110	9	40	29	2735	9996	7783
95	50046811	Mouse inbred strain C57BL6/J	Male	127	199	209	212	212	2888	3336	3208	12	15	20	3921	4462	4924	133	116	161	15	38	15	3339	5556	4245
96	50046812	Mouse inbred strain C57BL6/J	Male	134	253	209	202	215	3914	2307	5434	12	10	26	4372	3475	7852	157	144	165	34	13	53	8387	2742	4866
97	50046799	Mouse inbred strain C57BL6/J	Male	231	251	244	213	236	4098	1587	2534	19	5	10	5545	2596	3670	150	214	224	50	16	14	7253	4563	3980
98	50046800	Mouse inbred strain C57BL6/J	Male	139	228	226	231	199	1063	2050	3641	4	11	17	1584	3100	4527	135	123	131	34	34	23	5421	9064	3920
99	50046794	Mouse inbred strain C57BL6/J	Male	181	263	217	211	218	3874	2101	3218	23	11	14	6759	4700	3448	148	196	198	23	55	30	6880	5532	5559
100	50046798	Mouse inbred strain C57BL6/J	Male	166	209	238	220	250	4504	2138	3162	20	11	14	6452	3537	4261	143	190	137	50	44	10	4910	10824	2900
101	50046768	Mouse inbred strain C57BL6/J	Male	86	112	136	112	132	1692	1850	1454	18	9	4	3411	2745	2347	104	129	134	64	24	19	14578	6844	5609
102	50046771	Mouse inbred strain C57BL6/J	Female	125	184	226	236	219	1661	2544	4674	8	12	26	2361	3457	6325	132	120	179	4	44	82	1086	5513	4626
103	50046761	Mouse inbred strain C57BL6/J	Female	98	140	161	152	160	477	3694	3473	7	15	17	1331	4000	5012	133	215	115	20	79	45	4574	13991	11039
104	50046787	Mouse inbred strain C57BL6/J	Female	126	144	233	215	232	2329	2051	3127	11	11	19	3157	3139	4152	130	178	133	29	102	14	3556	16039	3775
105	50046785	Mouse inbred strain C57BL6/J	Female	141	174	217	252	215	1423	1886	1976	4	8	10	1992	2489	2901	164	180	192	25	38	50	4520	8600	9377
106	50046773	Mouse inbred strain C57BL6/J	Female	55	72	215	138	122	2101	2900	1744	10	15	7	3173	3670	2130	86	47	67	31	24	15	6978	4548	4136
107	50046777	Mouse inbred strain C57BL6/J	Female	73	139	167	180	159	2001	1523	2013	6	7	10	2943	2013	2924	98	190	109	64	87	26	16963	15006	7387
108	50046786	Mouse inbred strain C57BL6/J	Female	232	260	244	249	241	1104	1480	1477	6	7	12	1820	2191	2789	178	121	100	88	74	104	16322	14899	17895
109	50046778	Mouse inbred strain C57BL6/J	Female	113	165	241	241	236	2055	2925	1581	15	10	8	3965	2925	2456	135	131	135	11	30	20	3678	7460	5030
110	50046770	Mouse inbred strain C57BL6/J	Female	121	144	184	187	198	3473	1284	1995	17	5	8	5012	1640	2483	181	198	158	17	32	52	5552	8179	9888
111	50046762	Mouse inbred strain C57BL6/J	Female	124	173	199	205	211	3896	1871	2018	18	9	10	4897	2622	2858	132	153	151	10	22	25	3050	2822	5782
112	50046788	Mouse inbred strain C57BL6/J	Female	96	140	189	128	163	3333	1360	2197	21	6	9	4589	1796	2700	149	105	154	27	23	37	6898	3522	8089
113	50046775	Mouse inbred strain C57BL6/J	Female	182	208	273	282	270	2900	1744	2023	15	7	8	3670	2130	2491	174	197	179	69	50	34	1558	6107	9600
114	50046781	Mouse inbred strain C57BL6/J	Female	291	296	298	288	278	590	2032	1451	2	10	7	616	2836	2583	226	273	300	14	43	25	4050	11442	5050

S6 Table

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC
1	**Behavioral score data for Apodemus**				**Behavioral score data for *Apodemus***																								
2									**Open field test**												**Elevated Plus Maze tes**								
3	MouseID	Strain	sex	SNORD115 (replicate1)	SNORD115 (replicate2)	SNORD115 (replicate3)	average	SD	WallTime (replicate1)	WallTime (replicate2)	WallTime (replicate3)	Travel Distance at the Wall (replicate 1)	Travel Distance at the Wall (replicate 2)	Travel Distance at the Wall (replicate 3)	Speed (replicate 1)	Speed (replicate 2)	Speed (replicate 3)	Total distance (replicate 1)	Total distance (replicate 2)	Total distance (replicate 3)	DarkArm (replicate1)	DarkArm (replicate2)	DarkArm (replicate3)	Speed (replicate 1)	Speed (replicate 2)	Speed (replicate 3)	Total distance (replicate 1)	Total distance (replicate 2)	Total distance (replicate 3)
4	50042188	Apodemus uralensis	Male	414	412	389	405	13.9	233	282	243	3214	1998	1893	6	18	10	514	2495	1050	59	108	89	2224	2650	478	14	16	5
5	50052192	Apodemus uralensis	Male	601	621	595	606	13.6	290	252	277	2134	3001	1147	16	29	13	2548	3410	254	292	233	269	751	2506	3124	6	12	27
6	50042198	Apodemus uralensis	Male	608	597	594	600	7.4	293	221	279	459	2134	322	8	26	6	603	3203	544	120	212	130	4206	1842	6308	32	12	55
7	50051183	Apodemus uralensis	Male	183	184	174	180	5.5	283	111	155	2578	5214	2146	17	32	6	2105	7126	764	41	111	165	2428	3578	1181	18	31	11
8	50041177	Apodemus uralensis	Male	392	409	411	404	10.4	197	197	177	3120	4456	5543	44	55	39	8963	10430	7895	141	121	139	4100	7698	3125	35	54	28
9	50040724	Apodemus uralensis	Male	157	159	163	160	3.1	290	106	119	1123	4454	3077	16	7	24	1603	733	3563	92	152	79	2575	3172	5448	14	24	32
10	50042894	Apodemus uralensis	Male	606	593	615	605	11.1	276	296	259	5211	3457	4322	55	42	17	8193	5484	1479	179	124	152	7414	1281	3767	48	10	29
11	50040725	Apodemus uralensis	Male	332	346	343	340	7.4	239	195	174	1236	5479	2984	18	14	39	2983	7328	35476	290	150	277	326	5038	2278	9	44	16
12	50042597	Apodemus uralensis	Female	415	405	409	410	5.0	241	194	254	988	4504	1875	25	40	16	1706	5548	2574	292	233	269	751	2506	3219	8	15	21
13	50042596	Apodemus uralensis	Female	452	445	454	450	4.7	294	200	182	3988	2876	5574	47	39	58	4771	3458	7086	211	169	117	2789	3671	2568	17	27	17
14	50042196	Apodemus uralensis	Female	532	518	515	522	9.1	291	265	300	987	2875	421	10	36	4	1133	3775	596	300	299	300	1571	690	123	9	8	3
15	50041182	Apodemus uralensis	Female	269	265	271	268	3.1	148	188	198	3785	1746	1888	32	14	6	4575	2248	2349	147	178	197	5156	4041	2488	55	30	15
16	50042193	Apodemus uralensis	Female	811	787	803	800	12.2	291	269	280	2987	2222	355	47	30	8	4213	3025	555	300	300	300	5306	567	2272	44	6	11
17	50041181	Apodemus uralensis	Female	713	722	700	712	11.1	290	197	290	1698	1896	1556	13	7	22	1072	576	2434	195	188	190	3270	1578	380	31	12	4
18	50041176	Apodemus uralensis	Female	849	836	821	835	14.0	289	249	286	2456	3246	4334	25	31	46	3761	4978	7391	253	207	212	6838	4015	3214	47	32	27
19	50040725	Apodemus uralensis	Female	331	332	322	328	5.5	239	295	174	1786	1454	2988	18	14	19	2983	2328	35476	210	173	181	2612	7763	4691	17	58	31
20	50041175	Apodemus uralensis	Female	205	209	203	206	3.1	187	247	193	954	1646	1237	17	34	6	2974	4824	2407	126	114	101	1972	2616	4561	15	17	31
21	50041179	Apodemus uralensis	Female	404	402	410	405	4.2	286	245	224	1236	1087	2746	13	19	25	1857	2090	3979	153	259	148	3707	1754	4127	25	11	38
22	50042601	Apodemus uralensis	Female	609	606	595	603	7.4	299	268	200	1787	553	3678	16	8	32	2139	780	4785	188	88	198	3478	9746	3487	19	64	17
23	50042197	Apodemus uralensis	Female	317	319	323	320	3.1	225	212	191	2987	367	286	16	6	4	3002	477	354	194	254	158	5612	1926	7128	41	12	54
24	50041180	Apodemus uralensis	Female	277	282	278	279	2.6	181	190	197	988	2999	1999	9	25	17	1456	3351	2356	168	228	158	3876	542	3757	31	5	28
25	50041174	Apodemus uralensis	Female	522	514	507	514	7.5	277	229	277	955	3678	746	11	35	9	1258	5470	1098	214	228	264	3651	2809	6793	29	26	51
26	50042199	Apodemus uralensis	Female	406	396	393	398	6.8	205	195	199	3347	987	698	23	10	6	4078	1411	935	211	200	118	3222	5789	3026	24	41	29
27	50042600	Apodemus uralensis	Female	366	370	361	366	4.5	198	188	198	2747	1010	1688	24	12	8	3541	1947	2021	117	168	45	2250	7621	2443	16	47	15
28	50041178	Apodemus uralensis	Female	318	319	323	320	2.6	230	210	234	987	7845	4453	9	55	44	1365	9793	7317	161	112	155	6278	4124	2578	44	32	15
29	50040647	Apodemus uralensis	Female	293	285	286	288	4.4	188	177	126	2741	7746	1759	22	36	16	3580	9660	2987	71	89	67	4459	1068	2541	45	71	16
30	50042200	Apodemus uralensis	Female	296	298	302	299	3.1	156	199	130	4786	2479	325	39	18	5	5210	3471	574	125	185	157	12314	1985	7138	88	13	72
31	50042195	Apodemus uralensis	Female	307	302	304	304	2.5	296	208	212	2746	1987	3999	19	18	27	3191	2781	4001	98	103	130	9427	2574	4325	57	16	31
32	50041185	Apodemus uralensis	Female	393	380	385	386	6.6	283	237	231	2779	2001	974	19	4	11	3606	573	1548	55	117	77	9045	4226	2145	81	38	27
33	50041184	Apodemus uralensis	Female	269	264	261	265	4.0	223	203	216	3452	1454	1999	25	14	7	4358	1742	497	96	118	95	2257	5356	21345	22	48	21
34	50042198	Apodemus uralensis	Female	314	312	309	312	2.5	186	170	185	2785	773	1987	27	9	14	3763	955	2367	154	144	213	5454	1235	3763	44	16	30
35	50042189	Apodemus uralensis	Female	336	346	338	340	5.3	236	225	216	745	3333	2988	10	7	19	1494	579	3697	73	62	76	5028	7612	3075	45	67	32

1165

S7 Table

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1		**SNORD copy numbers family analysis**																
2		copy number classes were assigned in bins with at least 10 copies difference (aprrox. 2SD); lettering starts with the largest numbers																
3		yellow marked fields: cases with more than 4 classes detected in offspring																
4	**Mouse ID**	**Family**	**individual**	SNORD115 (replicate1)	SNORD115 (replicate2)	SNORD115 (replicate3)	**SNORD115 (average)**	**SD**	**copy number classes**		SNORD116(replicate1)	SNORD116 (replicate2)	SNORD116 (replicate3)	**SNORD116 (average)**	**SD**	**copy number classes**		
5	50032086	I	Father	181	195	188	188	7	D		203	216	209	209	7	C		used in behavioral study
6	50034314		Mother	212	222	218	218	4	B		235	241	236	236	5	B		used in behavioral study
7	50042701		pup1	253	259	256	256	3	A		284	292	288	288	4	A		used in behavioral study
8	50042702		pup2	151	161	155	156	5	E		182	174	177	178	4	E		
9	50042709		pup3	187	184	192	188	4	D		193	199	195	196	3	D		
10	50042710		pup4	115	109	112	112	3	H		129	123	126	126	3	H		used in behavioral study
11	50042705		pup5	129	129	131	125	4	G		134	144	138	139	5	G		
12	50042706		pup6	215	211	209	212	3	B		208	211	216	212	4	C		
13	50042707		pup7	158	152	155	155	3	E		177	165	170	170	6	E		used in behavioral study
14	50042708		pup8	201	204	196	200	4	C		224	214	218	219	5	C		used in behavioral study
15	50042703		pup9	224	212	217	218	6	B		230	245	235	236	6	B		used in behavioral study
16	50042704		pup10	135	143	140	139	4	F		153	145	148	158	3	F		used in behavioral study
17									8							8	number of classes in offspring	
18	50026157	II	Father	173	184	178	177	7	C		174	181	168	175	6	B		
19	50026329		Mother	166	158	161	162	4	D		168	164	173	169	5	B		
20	50037155		pup1	189	195	183	189	6	B		162	166	158	162	4	B		
21	50037156		pup2	160	166	162	163	3	D		162	152	158	157	5	B		
22	50037153		pup3	203	197	194	198	5	A		208	194	200	201	7	A		used in behavioral study
23	50037154		pup4	181	169	175	175	6	C		186	174	179	180	3	B		used in behavioral study
24									4							2	number of classes in offspring	
25	50025503	III	Mother	196	184	189	190	6	A		218	212	215	212	6	A		used in behavioral study
26	50028610		Father	120	123	119	121	2	C		137	131	130	134	3	C		used in behavioral study
27	50032425		pup1	152	158	155	155	3	B		151	159	155	155	4	B		
28	50032426		pup2	170	166	160	165	5	B		145	155	151	150	5	B		
29	50032427		pup3	165	160	155	160	5	B		168	162	165	165	3	B		used in behavioral study
30	50032428		pup4	150	160	156	155	5	B		151	159	155	155	4	B		used in behavioral study
31	50032429		pup5	167	177	173	172	5	B		171	167	173	170	3	B		
32									1							1	number of classes in offspring	
33	50025218	IV	Father	192	182	186	187	5	A		208	203	196	202	6	A		used in behavioral study
34	50025912		Mother	143	139	137	140	3	D		141	145	139	142	3	D		used in behavioral study
35	50032078		pup1	177	187	182	182	5	A		201	208	192	200	8	A		
36	50032079		pup2	166	171	163	167	4	B		181	173	176	177	4	B		
37	50032080		pup3	171	179	176	175	4	A		155	165	161	160	5	C		
38	50032082		pup4	158	152	156	155	3	C		147	152	144	148	4	D		
39									3							4	number of classes in offspring	
40	50025515	V	Father	225	218	232	224	8	B		157	167	163	162	5	A		
41	50025695		Mother	136	146	142	141	5	E		161	171	166	166	5	A		used in behavioral study
42	50033121		pup1	262	248	254	255	7	A		143	153	148	148	5	B		
43	50033122		pup2	257	245	250	251	6	A		177	164	170	171	6	A		
44	50033123		pup3	173	185	178	179	6	D		129	137	132	133	4	C		
45	50033124		pup4	221	215	207	214	7	C		170	184	177	177	7	A		
46									3							3	number of classes in offspring	
47	50031232	VI	Father	150	146	144	147	3	E		169	181	174	175	6	B		used in behavioral study
48	50032091		Mother	163	166	158	162	4	D		151	161	156	156	5	C		
49	50041502		pup1	188	195	183	189	6	B		166	161	158	162	4	C		
50	50041501		pup2	160	166	162	163	3	D		154	163	156	157	6	C		used in behavioral study
51	50041503		pup3	209	212	204	208	4	A		210	204	208	207	3	A		
52	50041504		pup4	169	175	181	175	6	C		173	185	179	179	6	B		
53									4							3	number of classes in offspring	
54	50032436	VII	Father	105	110	102	106	4	D		124	127	121	124	3	F		
55	50032116		Mother	191	183	188	187	4	B		204	214	210	209	5	C		
56	50041518		pup1	196	204	199	200	4	A		222	228	225	225	3	B		
57	50041523		pup2	173	167	170	170	3	C		192	188	191	190	2	D		used in behavioral study
58	50041524		pup3	178	176	182	179	3	B		190	198	193	194	4	D		
59	50041521		pup4	203	209	200	206	3	A		246	254	251	250	4	A		
60	50041522		pup5	188	183	191	187	4	B		205	209	208	207	2	C		
61	50041519		pup6	171	174	179	175	4	C		176	182	179	180	3	E		used in behavioral study
62	50041520		pup7	184	191	196	190	2	B		211	218	206	215	3	C		used in behavioral study
63									3							5	number of classes in offspring	
64	50031232	VIII	Father	154	165	159	159	5	C		172	182	178	177	5	B		
65	50032091		Mother	121	129	125	125	4	E		140	146	143	143	3	D		
66	50041498		pup1	181	169	175	175	6	B		182	187	195	188	5	A		used in behavioral study
67	50041499		pup2	183	195	188	189	6	A		190	200	196	195	5	A		
68	50041500		pup3	143	138	135	139	4	D		153	145	150	149	4	D		
69	50041501		pup4	180	178	184	181	3	A		198	194	195	196	2	A		
70	50041502		pup5	146	152	150	149	3	C		166	160	162	163	3	C		
71	50041503		pup6	158	152	154	155	3	C		166	170	169	168	2	B		used in behavioral study
72	50041504		pup7	138	146	141	142	4	D		154	162	159	158	4	C		
73									4							4	number of classes in offspring	
74	50031013	IX	Father	145	135	140	140	5	E		160	150	154	155	5	D		
75	50032115		Mother	196	202	188	195	7	B		182	190	198	190	8	C		
76	50041511		pup1	212	204	207	208	4	A		210	204	207	243	3	A		used in behavioral study
77	50041512		pup2	155	153	149	150	5	D		161	167	165	163	4	D		used in behavioral study
78	50041513		pup3	163	158	155	159	4	D		187	183	185	185	2	C		
79	50041514		pup4	191	185	189	188	3	B		203	206	200	203	3	B		
80	50041515		pup5	177	183	179	180	3	C		202	199	194	198	4	B		
81																4	number of classes in offspring	

S8 Table

	A	B	C	D	E	F	G	H	I	J	K
1	**RNASeq data for target genes**										
2	normalized read coverage for the respective exons										
3											
4		**Htr2c**					**Ankrd11**				
5	mouse ID		SNORD115 copy number	exon Vb chrX:147,169,676-147,169,770	exon VI chrX:147,193,753-147,197,264			SNORD116 copy number	exon IX 5`-part chr8:122,895,985-122,896,156	exon VIII chr8:122,896,560-122,896,707	exon VII chr8:122,898,574-122,898,721
6	50046813	S1	133	36	1612		S1	167	151	163	177
7	50046800	S2	141	14	1412		S2	216	115	112	172
8	50046761	S3	78	31	1624		S3	171	125	133	189
9	50046789	S4	60	15	1823		S4	114	120	119	186
10	50046817	S5	110	29	1805		S5	179	117	122	157
11	50046810	S6	181	66	1704		S6	210	171	107	166
12	50046783	S7	75	23	1849		S7	188	127	132	176
13	50046808	S8	156	31	1849		S8	250	171	122	139
14	50046782	S9	180	44	1642		S9	244	212	153	172
15	50046793	S10	153	29	1626		S10	250	183	142	142
16											
17		**R^2**		**0.43**	**-0.12**		**R^2**		**0.5**	**0.009**	**-0.53**
18		***P-value***		**0.039**	**0.32**		***P-value***		**0.02**	**0.79**	**0.01**
19											
20											
21		**Other predicted SNORD115 target genes**									
22	mouse ID		SNORD115 copy number	TAF1	RALGPS1	PBRM1	CRHR1	DPM2			
23	50046813	S1	133	99	3272.73	24.13	58.05	161			
24	50046800	S2	141	126	3087.3	19.68	82.41	142			
25	50046761	S3	78	123	3258.54	33.54	68.37	151			
26	50046789	S4	60	168	2880.02	18.02	46.64	184			
27	50046817	S5	110	177	2789.58	13.68	55.86	161			
28	50046810	S6	181	107	3082.5	47.95	53.43	122			
29	50046783	S7	75	120	2897.18	11.44	76.72	161			
30	50046808	S8	156	164	3064.71	22.2	35.52	134			
31	50046782	S9	180	168	3376	27	64	171			
32	50046793	S10	153	123	2698.75	24.13	62.23	115			
33											
34		**R^2**		**0.01**	**0.09**	**0.26**	**0.02**	**0.32**			
35		***P-value***		**0.86**	**0.41**	**0.13**	**0.68**	**0.09**			
36											

S9 Table

copy number comparison to predicted Ankrd11 target genes

correlation with Ankrd11-Exon IX 5´-fragment expression

the P-values refer to each correlation

Ensemble gene ID	Gene ID	R^2 ANKRD11	*P-value*	R^2 SNORD116	*P-value*
negative correlation					
ENSMUSG00000001260	Gabrg1	0.94	0.0001	0.78	0.001
ENSMUSG00000021486	Prelid1	0.86	0.0001	0.77	0.001
ENSMUSG00000017404	Rpl19	0.78	0.001	0.74	0.001
ENSMUSG0000005125		0.76	0.001	0.7	0.01
ENSMUSG00000044340	Phlpp1	0.71	0.01	0.71	0.01
ENSMUSG00000061559	Wdr61	0.68	0.01	0.66	0.01
ENSMUSG00000011114	Tbrg1	0.67	0.01	0.64	0.01
ENSMUSG00000025410	Dctn2	0.67	0.01	0.66	0.01
ENSMUSG00000023961	Enpp4	0.67	0.01	0.66	0.01
ENSMUSG00000024829	Mrpl21	0.66	0.01	0.65	0.01
ENSMUSG00000017858	Ift52	0.66	0.01	0.65	0.01
ENSMUSG00000020694	Tlk2	0.65	0.01	0.63	0.01
ENSMUSG00000020088	Sar1a	0.65	0.01	0.61	0.01
ENSMUSG00000048503	Tmem136	0.64	0.01	0.61	0.01
ENSMUSG00000047766	Lrrc49	0.63	0.01	0.62	0.01
ENSMUSG00000002635	Pdcd2l	0.62	0.01	0.62	0.01
ENSMUSG00000017897	Eya2	0.62	0.01	0.58	0.01
ENSMUSG00000002635	Pdcd2l	0.62	0.01	0.57	0.01
ENSMUSG00000026872	Zeb2	0.62	0.01	0.62	0.01
ENSMUSG00000007458	M6pr	0.62	0.01	0.59	0.01
ENSMUSG00000029363	Rfc5	0.61	0.01	0.66	0.01
ENSMUSG00000061758	Akr1b10	0.60	0.01	0.59	0.01
ENSMUSG00000022450	Ndufa6	0.59	0.01	0.58	0.01
ENSMUSG00000021256	Vash1	0.59	0.01	0.57	0.01
ENSMUSG00000060923	Acyp2	0.59	0.01	0.58	0.01
ENSMUSG00000022792	Yars2	0.58	0.01	0.58	0.01
ENSMUSG00000043004	Gng2	0.56	0.02	0.58	0.01
ENSMUSG00000001445	Mrpl10	0.56	0.02	0.55	0.01
ENSMUSG00000064115	Cadm2	0.55	0.02	0.5	0.02
ENSMUSG00000064115	Cadm2	0.55	0.02	0.52	0.02
ENSMUSG00000032531	Amotl2	0.55	0.02	0.54	0.01
ENSMUSG00000021589	Rhobtb3	0.55	0.02	0.59	0.01
ENSMUSG00000048065	Cyb5r2	0.55	0.02	0.53	0.01
ENSMUSG00000050174	Nudt6	0.54	0.02	0.51	0.01
ENSMUSG00000047084	Ngm	0.54	0.02	0.52	0.01
ENSMUSG00000021629	Slc30a5	0.53	0.02	0.47	0.03
ENSMUSG00000065990	Aurkaip1	0.53	0.02	0.49	0.03
ENSMUSG00000065990	Aurkaip1	0.53	0.02	0.44	0.03
ENSMUSG00000042198	Chchd7	0.52	0.03	0.45	0.03
ENSMUSG00000068739	Sars	0.52	0.03	0.44	0.03
ENSMUSG00000033938	Ndufb7	0.51	0.03	0.53	0.02
ENSMUSG00000049511	Htr1b	0.51	0.03	0.48	0.03
ENSMUSG00000025630	Hprt	0.50	0.03	0.47	0.03
ENSMUSG00000028284	Map3k7	0.49	0.03	0.48	0.03
ENSMUSG00000040034	Nup43	0.49	0.03	0.46	0.03
ENSMUSG00000032184	Lysmd2	0.49	0.03	0.4	0.05
ENSMUSG00000020070	Rufy2	0.49	0.03	0.4	0.05
ENSMUSG00000020070	Rufy2	0.49	0.03	0.41	0.05
ENSMUSG00000022200	Golph3	0.49	0.03	0.42	0.05
ENSMUSG00000034175	Rhbdd3	0.49	0.03	0.4	0.05
ENSMUSG00000062257	Opcml	0.48	0.03	0.45	0.03
ENSMUSG00000031774	Fam192a	0.47	0.03	0.44	0.03
ENSMUSG00000050164	Mchr1	0.47	0.03	0.44	0.03
ENSMUSG00000050088	1600012H06Rik	0.47	0.03	0.43	0.03
ENSMUSG00000040003	Magi2	0.47	0.03	0.47	0.03
ENSMUSG00000042709	Atpaf2	0.46	0.03	0.4	0.05
ENSMUSG00000052738	Suclg1	0.46	0.03	0.4	0.05
ENSMUSG00000002520		0.45	0.03	0.4	0.05
ENSMUSG00000040822	1700123O20Rik	0.45	0.03	0.4	0.05
ENSMUSG00000027305	Ndufaf1	0.45	0.03	0.4	0.05
ENSMUSG00000033713	Foxn3	0.45	0.03	0.4	0.05
ENSMUSG00000067713	Prkag1	0.44	0.03	0.4	0.05
ENSMUSG00000067713	Prkag1	0.44	0.03	0.4	0.05
ENSMUSG00000009356	Lpo	0.44	0.03	0.4	0.05
ENSMUSG00000059796	Eif4a1	0.44	0.03	0.4	0.05
ENSMUSG00000024352	Spata24	0.42	0.04	0.42	0.05
ENSMUSG00000051236	Msrb3	0.41	0.05	0.42	0.05
ENSMUSG00000050812	AI314180	0.41	0.05	0.43	0.04
ENSMUSG00000005360	Slc1a3	0.41	0.05	0.4	0.05
ENSMUSG00000026176	Ctdsp1	0.41	0.05	0.48	0.03
ENSMUSG00000034171	Faah	0.40	0.05	0.41	0.05
ENSMUSG00000015790	Surf1	0.40	0.05	0.4	0.05
positive correlation					
ENSMUSG00000034269	Setd5	0.81	0.001	0.75	0.01
ENSMUSG00000021240	Abcd4	0.72	0.01	0.74	0.01
ENSMUSG00000018167	Stard3	0.67	0.01	0.66	0.01
ENSMUSG00000001105	Ift20	0.66	0.01	0.67	0.01
ENSMUSG00000025340	Rabgef1	0.65	0.01	0.64	0.01
ENSMUSG00000047412	Zbtb44	0.60	0.01	0.61	0.01
ENSMUSG00000037747	Phyhipl	0.59	0.01	0.6	0.01
ENSMUSG0000002581		0.58	0.01	0.62	0.01
ENSMUSG00000062580	Timm17a	0.57	0.01	0.58	0.01
ENSMUSG00000018677	Slc25a39	0.57	0.01	0.59	0.01
ENSMUSG00000000171	Sdhd	0.56	0.01	0.55	0.02
ENSMUSG00000002475	Abhd3	0.56	0.01	0.52	0.02
ENSMUSG00000039601	Rcan2	0.55	0.02	0.57	0.01
ENSMUSG00000041959	S100a10	0.54	0.02	0.51	0.02
ENSMUSG0000006020		0.52	0.02	0.5	0.03
ENSMUSG00000041079	Rwdd2b	0.51	0.02	0.5	0.03
ENSMUSG00000096935	1700113A16Rik	0.51	0.02	0.47	0.03
ENSMUSG00000034384	Barhl2	0.49	0.03	0.47	0.03
ENSMUSG00000038279	Nop2	0.49	0.03	0.57	0.01
ENSMUSG00000096054	Syne1	0.48	0.03	0.54	0.02
ENSMUSG0000003212		0.48	0.03	0.53	0.02
ENSMUSG00000031616	Ednra	0.47	0.03	0.6	0.01
ENSMUSG00000000402	Egfl6	0.47	0.03	0.47	0.03
ENSMUSG00000031189	Aff2	0.47	0.03	0.44	0.03
ENSMUSG00000032033	Barx2	0.44	0.03	0.41	0.05
ENSMUSG00000024327	Slc39a7	0.44	0.03	0.41	0.05
ENSMUSG00000024927	Rela	0.44	0.03	0.43	0.04
ENSMUSG0000004559		0.43	0.04	0.4	0.05
ENSMUSG00000027207	Galk2	0.43	0.04	0.4	0.05
ENSMUSG00000019795	Pcmt1	0.42	0.04	0.4	0.05
ENSMUSG00000030682	Cdipt	0.42	0.04	0.4	0.05
ENSMUSG00000000827	Tpd52l2	0.41	0.05	0.4	0.05

additional genes from the whole transcriptome showing correlations with SNORD copy numbers

correlation with SNORD115 copy number

Ensemble gene ID	Gene ID	SNORD115	*P-value*
negative correlation			
ENSMUSG00000008520	Gm12927	0.90	0.0001
ENSMUSG00000005807	Eml1	0.89	0.0001
ENSMUSG00000006098	Galnt13	0.87	0.0001
ENSMUSG00000004337	Hexim2	0.87	0.0001
ENSMUSG00000006211	Scfd2	0.85	0.0001
ENSMUSG00000009020	4930503B20Ri	0.83	0.0009
ENSMUSG00000007375	Gm10570	0.82	0.0009
ENSMUSG00000003178	Katnb1	0.80	0.001
ENSMUSG00000007163	Cebpd	0.78	0.001
ENSMUSG00000004967	Morm4	0.77	0.001
ENSMUSG00000003137	Zrsr2	0.76	0.001
ENSMUSG00000007547	Deaf1	0.75	0.01
ENSMUSG00000004742	Dlk2	0.74	0.01
ENSMUSG00000002048	Xbp1	0.73	0.01
ENSMUSG00000004003	Nup43	0.71	0.01
ENSMUSG00000009740	9230116N13R	0.70	0.01
ENSMUSG00000002541	Dctn2	0.68	0.01
ENSMUSG00000002086	Luc7l3	0.66	0.01
ENSMUSG00000002873	Aldh4a1	0.64	0.01
ENSMUSG00000004713	Cd24a	0.63	0.01
ENSMUSG00000004903	Tmem121	0.62	0.01
ENSMUSG00000003693	4921524J17Ri	0.61	0.01
ENSMUSG00000004148	Sbx3	0.60	0.01
ENSMUSG00000001740	Stac2	0.57	0.02
ENSMUSG00000002127	Fdft1	0.55	0.02
ENSMUSG00000009672	Hmgb1-ps9	0.54	0.02
ENSMUSG00000009764	Gm26862	0.53	0.02
ENSMUSG00000002638	3110009E18Ri	0.52	0.02
ENSMUSG00000002641	Tnni1	0.50	0.02
ENSMUSG00000003184	Comp	0.49	0.03
ENSMUSG00000003377	Npbwr1	0.47	0.03
ENSMUSG00000004699	Spsb4	0.46	0.03
ENSMUSG00000004711	Fam221a	0.45	0.03
ENSMUSG00000004783	Cdca4	0.44	0.04
ENSMUSG00000006165	Spry3	0.43	0.04
ENSMUSG00000007877	Zfp59	0.42	0.05
positive correlation			
ENSMUSG00000006882	Csde1	0.86	0.0001
ENSMUSG00000000019	Nalcn	0.85	0.0001
ENSMUSG00000004168	Amot	0.85	0.0001
ENSMUSG00000003803	Igsf8	0.85	0.0001
ENSMUSG00000003201	Nectin1	0.85	0.0001
ENSMUSG00000004408	S100a1	0.83	0.0009
ENSMUSG00000003258	Bsn	0.83	0.0009
ENSMUSG00000002426	Syt4	0.82	0.0009
ENSMUSG00000000765	Gabrb2	0.82	0.0009
ENSMUSG00000005900	Grin2a	0.81	0.001
ENSMUSG00000000083	Fmr1	0.81	0.001
ENSMUSG00000004174	Utp15	0.81	0.001
ENSMUSG00000003481	Grip1	0.80	0.001
ENSMUSG00000005753	Dtnbp1	0.80	0.001
ENSMUSG00000003243	Dync1li1	0.80	0.001
ENSMUSG00000002018	Mdm2	0.80	0.001
ENSMUSG00000027771	Acad9	0.69	0.01
ENSMUSG00000002495	Esrra	0.65	0.01
ENSMUSG00000002403	Rsph1	0.61	0.01
ENSMUSG00000007456	Gcnt7	0.60	0.01
ENSMUSG00000011443	Gm48488	0.59	0.01
ENSMUSG00000002128	Eif5	0.58	0.01
ENSMUSG00000006661	Zfp932	0.58	0.01
ENSMUSG00000004509	S1pr1	0.57	0.01
ENSMUSG00000000283	Plin4	0.56	0.01
ENSMUSG00000004942	Gm9837	0.51	0.02
ENSMUSG00000002839	Ptprd	0.49	0.02
ENSMUSG00000003400	Scaper	0.46	0.03
ENSMUSG00000003179	Mmp15	0.45	0.03
ENSMUSG00000002618	Ppm1f	0.43	0.04
ENSMUSG00000003577	Krt20	0.42	0.05

correlation with SNORD116 copy number

Ensemble gene ID	Gene ID	R^2 SNORD116	*P-value*
negative correlation			
ENSMUSG0	Ticrr	0.94	0.0001
ENSMUSG0	Eepd1	0.92	0.0001
ENSMUSG0	Nelfcd	0.91	0.0001
ENSMUSG0	Msrb2	0.90	0.0001
ENSMUSG0	Gm29764	0.90	0.0001
ENSMUSG0	Ift74	0.89	0.0001
ENSMUSG0	Ybx3	0.88	0.0001
ENSMUSG0	Gm47235	0.87	0.0001
ENSMUSG0	Filip1	0.87	0.0001
ENSMUSG0	Erp44	0.87	0.0001
ENSMUSG0	Pigc	0.86	0.0001
ENSMUSG0	Nubp2	0.86	0.0001
ENSMUSG0	Plcl2	0.86	0.0001
ENSMUSG0	Gabrg1	0.78	0.001
ENSMUSG0	Prelid1	0.77	0.001
ENSMUSG0	Rpl19	0.74	0.01
ENSMUSG0	Phlpp1	0.71	0.01
ENSMUSG000000051		0.7	0.01
ENSMUSG0	Wdr61	0.66	0.01
ENSMUSG0	Dctn2	0.66	0.01
ENSMUSG0	Enpp4	0.66	0.01
ENSMUSG0	Rfc5	0.66	0.01
ENSMUSG0	Mrpl21	0.65	0.01
ENSMUSG0	Ift52	0.65	0.01
ENSMUSG0	Tbrg1	0.64	0.01
ENSMUSG0	Tlk2	0.63	0.01
ENSMUSG0	Lrrc49	0.62	0.01
ENSMUSG0	Pdcd2l	0.62	0.01
ENSMUSG0	Zeb2	0.62	0.01
ENSMUSG0	Sar1a	0.61	0.01
ENSMUSG0	Tmem136	0.61	0.01
ENSMUSG0	M6pr	0.59	0.01
ENSMUSG0	Akr1b10	0.59	0.01
ENSMUSG0	Rhobtb3	0.59	0.01
ENSMUSG0	Eya2	0.58	0.01
ENSMUSG0	Ndufa6	0.58	0.01
ENSMUSG0	Acyp2	0.58	0.01
ENSMUSG0	Yars2	0.58	0.01
ENSMUSG0	Gng2	0.58	0.01
ENSMUSG0	Pdcd2l	0.57	0.01
ENSMUSG0	Vash1	0.57	0.01
ENSMUSG0	Mrpl10	0.55	0.01
ENSMUSG0	Amotl2	0.54	0.02
ENSMUSG0	Cyb5r2	0.53	0.02
ENSMUSG0	Ndufb7	0.53	0.02
ENSMUSG0	Cadm2	0.52	0.02
ENSMUSG0	Ngm	0.52	0.02
ENSMUSG0	Nudt6	0.51	0.02
ENSMUSG0	Cadm2	0.5	0.02
ENSMUSG0	Aga	0.46	0.03
ENSMUSG0	Prr5	0.44	0.03
ENSMUSG0	Slc26a11	0.43	0.04
ENSMUSG0	Il7	0.41	0.05
positive correlation			
ENSMUSG0	Cisd2	0.92	0.0001
ENSMUSG0	Gm22408	0.92	0.0001
ENSMUSG0	Tarf2ip	0.84	0.0001
ENSMUSG0	Cbln1	0.84	0.0001
ENSMUSG0	Slc39a9	0.81	0.001
ENSMUSG0	Zfp369	0.81	0.001
ENSMUSG0	Setd5	0.75	0.01
ENSMUSG0	Abcd4	0.74	0.01
ENSMUSG0	Ift20	0.67	0.01
ENSMUSG0	Stard3	0.66	0.01
ENSMUSG0	Rabgef1	0.64	0.01
ENSMUSG000000025		0.62	0.01
ENSMUSG0	Zbtb44	0.61	0.01
ENSMUSG0	Phyhipl	0.60	0.01
ENSMUSG0	Ednra	0.60	0.01
ENSMUSG0	Slc25a39	0.59	0.01
ENSMUSG0	Timm17a	0.58	0.01
ENSMUSG0	Rcan2	0.57	0.01
ENSMUSG0	Nop2	0.57	0.01
ENSMUSG0	Sdhd	0.55	0.01
ENSMUSG0	Syne1	0.54	0.02
ENSMUSG000000032		0.53	0.02
ENSMUSG0	Abhd3	0.52	0.02
ENSMUSG0	S100a10	0.51	0.02
ENSMUSG000000060		0.50	0.02
ENSMUSG0	Rwdd2b	0.50	0.02
ENSMUSG0	Ugp2	0.47	0.03
ENSMUSG0	Dock7	0.47	0.03
ENSMUSG0	Filip1l	0.47	0.03
ENSMUSG0	Gm48342	0.47	0.03
ENSMUSG0	Hmgn5	0.46	0.03
ENSMUSG0	Epha7	0.46	0.03
ENSMUSG0	Arhgap6	0.44	0.04
ENSMUSG0	Arid1a	0.44	0.04
ENSMUSG0	**Ankrd11**	**0.43**	0.04
ENSMUSG0	Gm36445	0.41	0.05

S10 Table

	A	B	C	D	E	F	G	H	I
1	**Behavioral score data for *Cavea***								
2	**animal ID**	**sex**	**SNORD116 (replicate 1)**	**SNORD116 (replicate 2)**	**SNORD116 (replicate 3)**	**average**	**SD**	**anxiety score in the Open Field test (distance run in cm)**	**anxiety scoreStruggle Docility (seconds)**
3	72A-C0B7	Male	10	12	11	11	1.0	6560	14
4	74F-E335	Male	9	10	12	10	1.5	3168	13
5	74F-B698	Male	12	13	10	12	1.5	5417	13
6	72A-BA76	Male	13	11	12	12	1.0	322	11
7	72A-8511	Male	7	8	10	8	1.5	8034	20
8	75B-460E	Male	11	10	8	10	1.5	4896	15
9	72A-808E	Male	8	6	10	8	2.0	21942	16
10	72A-ABED	Male	13	12	11	12	1.0	3942	10
11	75B-72C0	Male	9	6	10	8	2.1	5284	19
12	3433	Female	13	10	9	11	2.1	5152	14
13	4246	Female	15	16	12	14	2.1	356	2
14	D080	Female	7	9	8	8	1.0	5783	22
15	4F77	Female	9	10	13	11	2.1	1585	1
16	7CB4	Female	9	13	10	11	2.1	997	16
17	8326	Female	9	11	15	12	3.1	0	17
18	36C2	Female	11	9	8	9	1.5	4987	10
19	8341	Female	10	12	9	10	1.5	1763	16
20	390D	Female	6	8	8	7	1.2	4212	24
21	0FCF	Female	8	8	10	9	1.2	1873	12

1180

S11 Table

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	**copy number correlation with test score in humans**														
2															
3	NID	SNORD115	SNORD116 _1	SNORD116_2	SNORD116_3	all SNORD116	age	k_sex	marco	activity_TPQ	anxiety_TPQ	curiosity_TPQ	anxandactivity_TPQ	low	combined
4	MPIPSYKL_004850	32	6	23	9	37	65	1	0	0	0	0	0	1	0
5	MPIPSYKL_004865	31	5	23	8	36	58	2	0	0	0	0	0	1	0
6	MPIPSYKL_005017	31	5	21	8	35	61	2	0	0	0	0	0	1	0
7	MPIPSYKL_005034	22	5	14	9	26	57	2	0	0	0	0	0	1	0
8	MPIPSYKL_005131	28	5	13	7	27	27	1	0	0	0	0	0	1	0
9	MPIPSYKL_005229	21	6	14	9	27	24	2	0	0	0	0	0	1	0
10	MPIPSYKL_005399	17	6	13	7	28	32	2	0	0	0	0	0	1	0
11	MPIPSYKL_005413	49	7	25	9	38	37	2	0	0	0	0	0	1	0
12	MPIPSYKL_005013	29	4	11	6	20	40	2	0	0	0	1	0	1	1
13	MPIPSYKL_005164	27	5	10	5	22	40	2	0	0	0	1	0	1	1
14	MPIPSYKL_005310	19	6	12	7	28	29	1	0	0	0	1	0	1	1
15	MPIPSYKL_005414	32	7	23	10	39	30	2	0	0	0	1	0	1	1
16	MPIPSYKL_005713	29	6	14	9	29	18	2	0	0	0	1	0	1	1
17	MPIPSYKL_004910	31	8	12	9	27	51	2	0	0	1	0	1	1	2
18	MPIPSYKL_004912	39	9	23	7	39	38	1	0	0	1	0	1	1	2
19	MPIPSYKL_004931	33	5	20	7	33	25	2	0	0	1	0	1	1	2
20	MPIPSYKL_004965	34	7	23	8	36	38	2	0	0	1	0	1	1	2
21	MPIPSYKL_004983	34	4	20	6	30	30	1	0	0	1	0	1	1	2
22	MPIPSYKL_005097	26	5	11	6	23	63	1	0	1	0	0	1	1	2
23	MPIPSYKL_005102	34	6	12	7	27	25	2	0	0	1	0	1	1	2
24	MPIPSYKL_005120	41	7	15	9	31	28	2	0	1	0	0	1	1	2
25	MPIPSYKL_005149	34	7	17	9	35	60	2	0	1	0	0	1	1	2
26	MPIPSYKL_005236	43	8	14	11	29	33	1	0	0	1	0	1	1	2
27	MPIPSYKL_005257	21	5	11	7	23	34	2	0	1	0	0	1	1	2
28	MPIPSYKL_005272	22	6	12	7	26	51	2	0	1	0	0	1	1	2
29	MPIPSYKL_005292	22	6	15	8	30	30	2	0	0	1	0	1	1	2
30	MPIPSYKL_005294	26	8	16	9	32	48	1	0	0	1	0	1	1	2
31	MPIPSYKL_005320	21	6	14	8	31	53	1	0	0	1	0	1	1	2
32	MPIPSYKL_005333	42	8	14	11	28	56	1	0	0	1	0	1	1	2
33	MPIPSYKL_005395	22	6	12	6	28	57	2	0	0	1	0	1	1	2
34	MPIPSYKL_005496	23	5	14	10	25	30	2	0	0	0	2	0	1	2
35	MPIPSYKL_005589	17	6	13	6	28	39	2	0	0	1	0	1	1	2
36	MPIPSYKL_004968	36	6	24	9	37	42	2	0	0	1	1	1	1	3
37	MPIPSYKL_004981	28	6	22	7	38	51	1	0	0	1	1	1	1	3
38	MPIPSYKL_005018	36	6	22	10	36	33	2	0	0	1	1	1	1	3
39	MPIPSYKL_005056	26	6	14	8	27	39	1	0	0	1	1	1	1	3
40	MPIPSYKL_005084	23	6	15	7	28	32	1	0	0	1	1	1	1	3
41	MPIPSYKL_005146	30	6	12	7	26	48	1	0	0	1	1	1	1	3
42	MPIPSYKL_005185	23	6	11	8	25	53	2	0	0	1	1	1	1	3
43	MPIPSYKL_005266	24	6	14	8	29	25	2	0	0	1	1	1	1	3
44	MPIPSYKL_005377	29	7	17	9	30	48	2	0	0	1	1	1	1	3
45	MPIPSYKL_005426	20	4	13	6	24	58	2	0	1	0	1	1	1	3
46	MPIPSYKL_005446	15	6	12	7	25	31	2	0	0	1	1	1	1	3
47	MPIPSYKL_005028	18	5	13	7	28	39	1	0	0	1	2	1	1	4
48	MPIPSYKL_005140	40	7	17	10	30	30	1	0	0	1	2	1	1	4
49	MPIPSYKL_005265	14	6	12	6	24	48	2	0	0	1	2	1	1	4
50	MPIPSYKL_005551	17	6	10	6	24	70	1	0	1	1	0	2	1	4
51	MPIPSYKL_005290	24	6	14	8	28	42	1	0	0	1	3	1	1	5
52	MPIPSYKL_005344	14	5	14	8	27	30	2	0	2	4	0	6	0	12
53	MPIPSYKL_004896	25	7	12	8	26	54	2	0	4	3	0	7	0	14
54	MPIPSYKL_004917	25	8	13	7	27	41	2	0	2	5	0	7	0	14
55	MPIPSYKL_004967	17	5	12	6	25	37	1	0	2	5	0	7	0	14
56	MPIPSYKL_004976	25	5	14	8	28	22	2	0	4	2	2	6	0	14
57	MPIPSYKL_005033	25	7	14	9	30	39	1	0	5	2	0	7	0	14
58	MPIPSYKL_005043	26	7	13	9	28	24	2	0	1	5	2	6	0	14
59	MPIPSYKL_005222	18	6	12	8	28	25	2	0	1	5	2	6	0	14
60	MPIPSYKL_005246	24	6	13	10	25	37	2	0	5	2	0	7	0	14
61	MPIPSYKL_005361	18	6	13	6	26	48	2	0	3	4	0	7	0	14
62	MPIPSYKL_005468	25	6	14	7	28	59	1	0	2	5	0	7	0	14
63	MPIPSYKL_005472	21	6	13	8	24	31	2	0	2	4	2	6	0	14
64	MPIPSYKL_004900	13	5	9	5	22	49	1	0	5	2	1	7	0	15
65	MPIPSYKL_004938	20	7	14	8	29	28	1	0	2	5	1	7	0	15
66	MPIPSYKL_005047	28	6	11	8	25	61	1	0	2	5	1	7	0	15
67	MPIPSYKL_005129	22	6	14	8	28	40	2	0	3	4	1	7	0	15
68	MPIPSYKL_005138	25	7	14	8	27	62	1	0	4	3	1	7	0	15
69	MPIPSYKL_005385	16	5	12	6	25	30	2	0	4	3	1	7	0	15
70	MPIPSYKL_005486	24	6	13	8	27	44	2	0	2	5	1	7	0	15
71	MPIPSYKL_004866	25	7	14	8	28	37	1	0	5	2	2	7	0	16
72	MPIPSYKL_004942	17	5	11	7	23	59	1	0	2	5	2	7	0	16
73	MPIPSYKL_005006	17	5	7	7	19	50	1	0	5	2	2	7	0	16
74	MPIPSYKL_005024	24	5	13	7	25	48	1	0	3	4	2	7	0	16
75	MPIPSYKL_005107	22	6	13	7	26	53	1	0	3	4	2	7	0	16
76	MPIPSYKL_005109	21	6	13	7	27	56	1	0	3	4	2	7	0	16
77	MPIPSYKL_005123	24	6	13	8	28	40	1	0	4	3	2	7	0	16
78	MPIPSYKL_005196	26	7	13	9	25	28	1	0	6	1	2	7	0	16
79	MPIPSYKL_005201	15	4	10	5	22	44	2	0	4	3	2	7	0	16
80	MPIPSYKL_005204	24	6	11	8	26	48	1	0	6	2	0	8	0	16
81	MPIPSYKL_005248	33	7	16	9	30	49	1	0	3	4	2	7	0	16
82	MPIPSYKL_005353	19	6	12	7	25	47	1	0	4	3	2	7	0	16
83	MPIPSYKL_004862	16	5	12	7	24	41	1	0	5	2	3	7	0	17
84	MPIPSYKL_004909	19	6	10	7	22	52	2	0	2	5	3	7	0	17
85	MPIPSYKL_005162	17	5	10	6	21	50	1	0	5	3	1	8	0	17
86	MPIPSYKL_005174	24	6	10	6	24	47	1	0	3	5	1	8	0	17
87	MPIPSYKL_005210	22	5	12	8	24	48	1	0	6	1	3	7	0	17
88	MPIPSYKL_005525	15	5	12	7	23	41	1	0	4	4	1	8	0	17
89	MPIPSYKL_005557	14	5	12	6	22	45	1	0	3	4	3	7	0	17
90	MPIPSYKL_005568	15	6	11	5	25	52	1	0	3	4	3	7	0	17
91	MPIPSYKL_005029	22	6	13	8	26	26	1	0	4	4	2	8	0	18
92	MPIPSYKL_005176	31	6	11	7	25	44	1	0	6	2	2	8	0	18
93	MPIPSYKL_005421	19	5	12	8	23	37	1	0	3	5	2	8	0	18
94	MPIPSYKL_004954	25	5	10	6	23	35	1	0	5	3	3	8	0	19
95	MPIPSYKL_005319	26	6	14	8	27	47	2	0	6	3	1	9	0	19
96	MPIPSYKL_005089	22	6	14	7	28	23	2	0	3	6	2	9	0	20
97	MPIPSYKL_005714	22	6	15	8	28	18	1	0	5	4	3	9	0	21
98	MPIPSYKL_005386	22	5	12	7	25	45	1	0	6	5	0	11	0	22
99	MPIPSYKL_005317	22	6	15	8	29	59	1	0	6	5	1	11	0	23
100	**average**	25	6	14	8	27									
101	**SD**	7	1	4	1	4									
102	t-test	0.000007	0.28	0.000011	0.059	0.000004									
103															