Copy number variation in small nucleolar RNAs regulates personality behavior

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

Max-Planck Institute for Evolutionary Biology, Plön, Germany

2: 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

Animals show behavioral traits that can collectively be called personality. We focus here on the role of the Prader-Willi Syndrom gene region in regulating personality behavior. It includes two clusters of tandem repeats coding for small nucleolar RNAs, SNORD115 and SNORD116. SNORD115 is known to regulate splicing of the serotonin receptor *Ht2cr* and SNORD116 is predicted to interact with the transcript of the chromatin regulator *Ankrd11*. We show that both snoRNA clusters display major copy number variation within and between populations, as well as in an inbred mouse strain and that this affects the expression of their specific target genes. Using a set of behavioral scores related to personality in populations of two species of wild mice, guinea pigs and humans, we find a strong correlation between the snoRNA copy number and these scores. Our results suggest that the SNORD clusters are major regulators of personality and correlated traits.

Impact statement

Behavioral variation among individuals is regulated by a tandemly repetitive genetic region that can generate new length variants in every generation.

Introduction

The study of consistent individual differences in behaviour, termed "animal personality", has flourished over the last decades, because it has been recognised as a major contributor to differences in survival and fitness among individuals (Reale, 2007; Reale et al., 2010; Wolf and Weissing, 2012). A major goal is to understand the origin of individual variation in behaviour and the mechanisms generating this variation, especially since it is known to occur also in genetically isogenic strains (Bierbach et al., 2017; Freund et al., 2013; Lewejohann et al., 2011; Wolf and Weissing, 2012).

Personality can be defined as behaviors which are consistent over time and across context (Reale, 2007; Wilson, 1998). Several studies have shown that personality traits have a heritable component and hence, that genetics should play an important role (Dochtermann et al., 2015; Sanchez-Roige et al., 2018; van Oers and Mueller, 2010). However, despite considerable efforts, the genetic pathways that influence personality are largely unknown and there is so far no molecular mechanism that could explain the maintenance of such variation within populations.

A locus that has been implicated in behavioral traits in humans is the Prader–Willi syndrome (PWS) region. PWS is a neurodevelopmental 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* and *Snprn*, of which *Ube3a* is expressed from the maternally provided chromosome and *Snprn* from the paternally provided one. This arrangement has specifically evolved in mammals and is generally conserved among them, including humans (Zhang et al., 2014).

In mice we have identified the PWS region as one of two imprinted regions that evolve particularly fast between natural populations and that could be involved in behavioral differences 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 and metabolism (Cavaille, 2017; Jiang et al., 2010; Nakatani et al., 2009; Qi et al., 2016).

The PWS regions expresses also non-coding RNAs, especially two small nucleolar RNA (SNORD) gene families which are organized in large, tandemly repeated clusters known as SNORD115 and SNORD116 (Cavaille, 2017). The expression of both SNORD115 and SNORD116 is brain-specific and restricted to the alleles on the paternal chromosome.

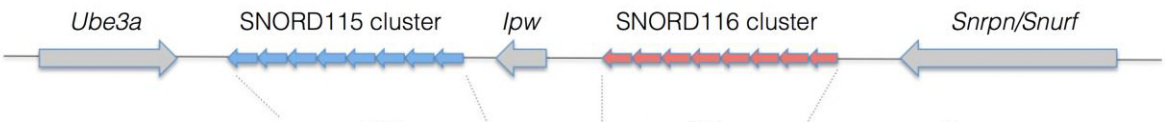
SNORDs are part of a large group of small, metabolically stable snoRNAs which regulate post-transcriptional modification of their target genes (Kiss, 2001). snoRNAs bind via a complementary antisense region to their target RNAs. The SNORD115 antisense element exhibits complementarity to the alternatively spliced exon V of the serotonin receptor *Ht2cr* which is part of the serotonin

regulatory pathway. It has been shown in cell-culture experiments that there is a positive correlation between SNORD115 expression and exon part Vb usage (Kishore and Stamm, 2006). Activation of the receptor by serotonin inhibits dopamine and norepinephrine release in certain areas of the brain (Alex et al., 2005). This regulates mood, anxiety, feeding, and motoneuron functions (see (Stamm et al., 2017) for a recent review). SNORD116 shows a complementary sequence to *Ankrd11* exon X, 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). 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 could influence behavioral traits in mice and other mammals. We used standardized tests that are mostly connected to anxiety profiles, as well as comparable tests for wild guinea pigs and questionnaire-based tests 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 found not only for wildtype strains, but also for the common laboratory inbred strain C57BL/6J. Using transcriptomic analyses, we show that the predicted regulation of *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. These observations confirm a direct causative link. Since our data suggest further that new alleles with different copy numbers are generated at an exceptionally high rate, we conclude that this variation could be the basis for the long sought molecular mechanism for the high variance of personality traits in families and populations.

Results

Given the possible molecular links to the regulation of behavioral responses, we first analyzed copy number variation within the SNORD115 and SNORD116 clusters in house mice (*M. m. domesticus*). Both snoRNAs are part of a larger transcript, from which they are processed (Cavaille, 2017). The repeat unit length for SNORD115 is 1.97kb, for SNORD116 2.54kb. In the mouse reference genome (Waterston et al., 2002), SNORD115 is annotated with 143 copies and SNORD116 with 70 copies, but with an annotation gap. To type copy number differences in this range, we used droplet digital PCR, which we had previously shown to allow accurate copy number determination across a broad range of CNVs in mice (Pezer et al., 2015). We tested individuals derived from two populations of *M. m. domesticus* that were originally caught in the wild in Germany (CB) and France (MC) and 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 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 (Figure 1).



	SNORD115 cluster average (SD)	SNORD115 cluster range	SNORD116 cluster average (SD)	SNORD116 cluster range	R^2
M. m. domesticus CB	202 (54)	95-290	229 (55)	132-325	0.94
M. m. domesticus MC	170 (38)	108-246	188 (53)	86-290	0.87
mouse inbred strain C57BL/6J	140 (47)	55-288	185 (51)	71-302	0.74
Apodemus uralensis (wood mouse)	413 (170)	160-835	n.d.	n.d.	n.a.
Cavia aperea (wild guinea pig)	n.d.	n.d.	10 (2)	6-14	n.a.
humans	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. All are paternally expressed, only *Ube3a* is maternally expressed. Note that the transcript structures are actually more complex and not yet fully resolved (Cavaille, 2017; Lorenc et al., 2014). 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 co-variation regression coefficients (R^2) between the numbers in the two clusters are listed to the right. The full data are provided in Table S1.

The measured copy number variation shows a highly significant regression with the expression of the respective SNORDs in the brain (Figure 2A,B), suggesting a direct relationship between copy number and expression. Note that given that this locus is subject to imprinting, only the paternal copies would be expressed, implying that the correlation should be below 1, even when a perfect relationship exists between copy number and expression.

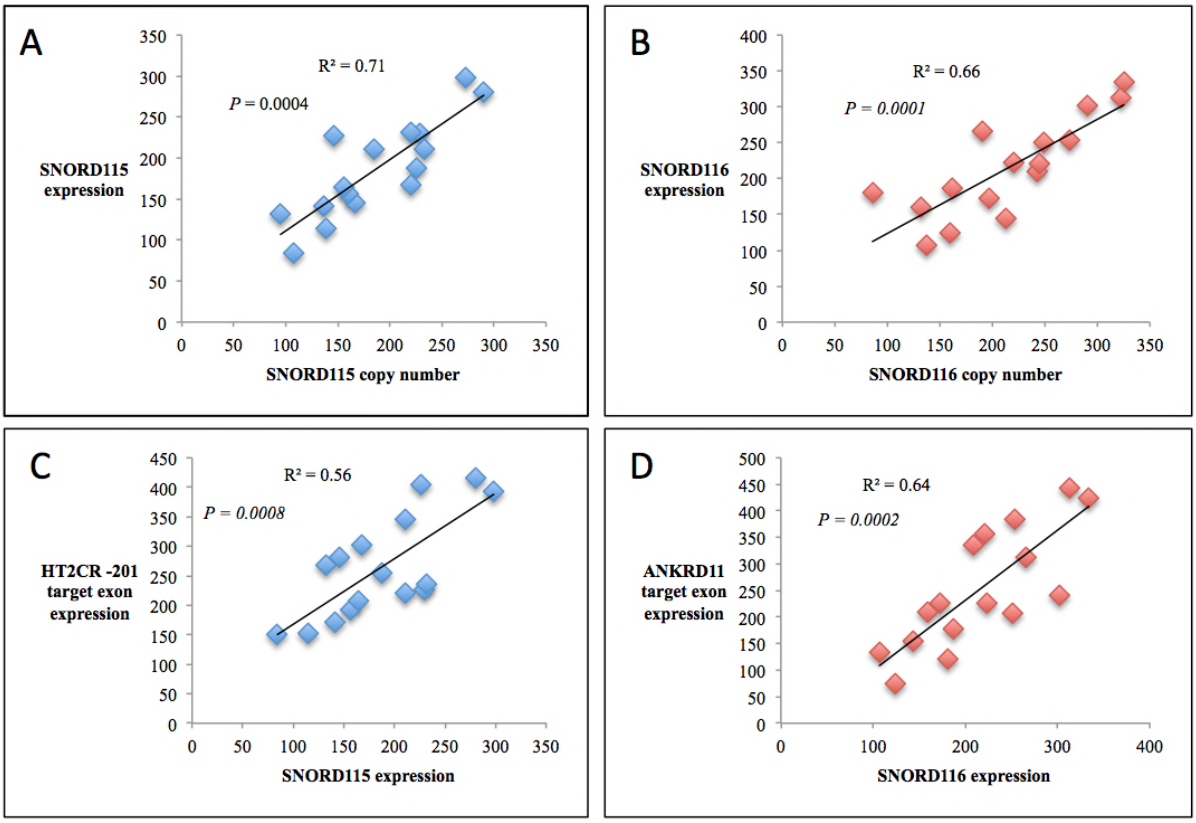


Figure 2: Regression between SNORD115 (blue) and SNORD116 (red) copy numbers and expression. A) and B) Regression with the respective RNAs. C) and D) Regression between snoRNA expression and the splice-site regulated target RNAs. All values were determined by ddPCR in brain RNA preparations from eight outbred animals each of the CB and MC populations. These animals constitute a subset of the 45 animals used for overall copy number variation and were chosen to reflect the spread of the variation. The data were combined, since tests for each population on their own yield similar correlations. The full data are provided in Table S2.

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) and the predicted target of SNORD116 is exon X of *Ankrd11* (Bazeley et al., 2008). When assaying for the expression of these exons, we find indeed a high correlation with the respective SNORD expression levels and exon expression levels (Figure 2C,D). As a control, we assayed also

two transcript variants for each locus that should not be affected by alternative splicing. For these, we find no significant regression with copy numbers of the respective SNORDs (Table S2).

Behavioral tests

To assess whether the SNORD gene copy number variation correlates also with the behavior of the mice, we have employed an aggregate score consisting of three separate standard tests focusing on anxiety. Anxiety is a consistent personality trait in mice (Freund et al., 2013; Lewejohann et al., 2011). Details on the tests and the derivation of the overall score are provided in the Methods section. In short, the development of the score consisted of separate tests (Open Field Test, Dark/Light Box and Elevated Plus Maze). Single test scores were tested for repeatability over the course of the experiment using intra-class correlation coefficients. Those found to be repeatable were clustered and these clusters were subjected to a principle component analysis. The first principle component was then used as a combined single score for anxiety-related behavior.

We find a highly significant regression between these scores and SNORD copy numbers, with coefficients between 0.42-0.52 (Figure 3A-D). The results were analyzed separately for the different populations, since overall copy numbers and overall behavioral scores differ between them. Still, both populations show the same pattern. In each population, animals with higher copy numbers have a higher relative anxiety score.

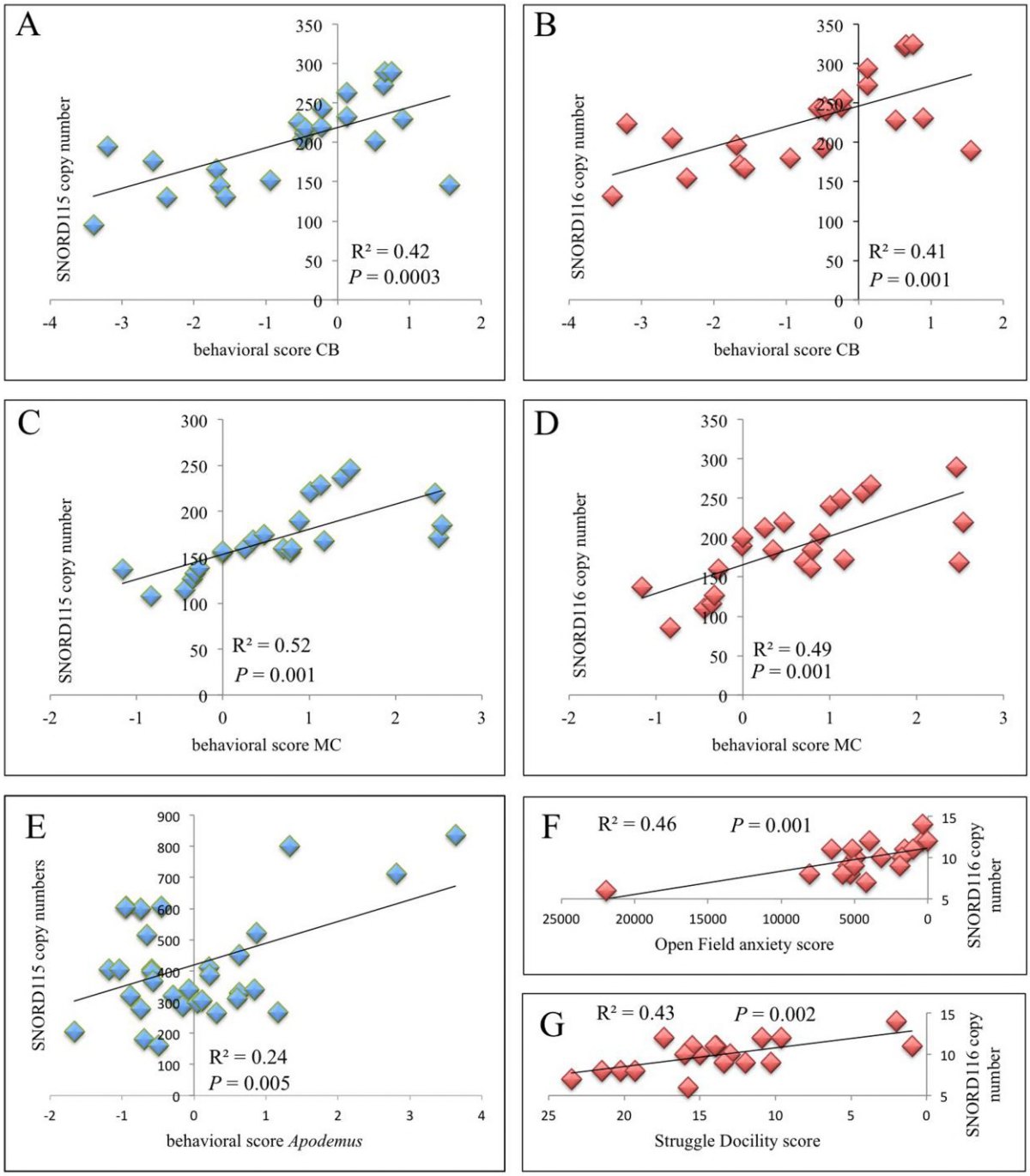


Figure 3: Regression between snoRNA gene copy numbers and behavioral scores. A), C), E) for SNORD115, B), D), F) and G) for SNORD116. A) and B) for the *M. m. domesticus* CB population (N=22), C) and D) for the *M. m. domesticus* MC population (N=23), E) for the wood mouse *Apodemus uralensis* (N=32), F) and G) wild guinea pig *Cavia aperea (N=19)*. Note that the behavioral tests for cavies differ from the tests for the other species. Higher scores represent higher anxiety, i.e. the X-axis is reversed for better comparability. Full data in Table S1.

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, this can not be directly tested, since the haplotypes can not be distinguished by conventional techniques. However, we approached this question by comparing the inheritance of copy numbers in nine families of the MC population. We find that the offspring shows a large variation, exceeding the spread of numbers measured for the parents in 7 out of the 9 families (Figure 4). 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 families with more than four offspring (families 1, 7, 8 and 9) show additional allele classes for each offspring, suggesting that new length alleles have been generated.

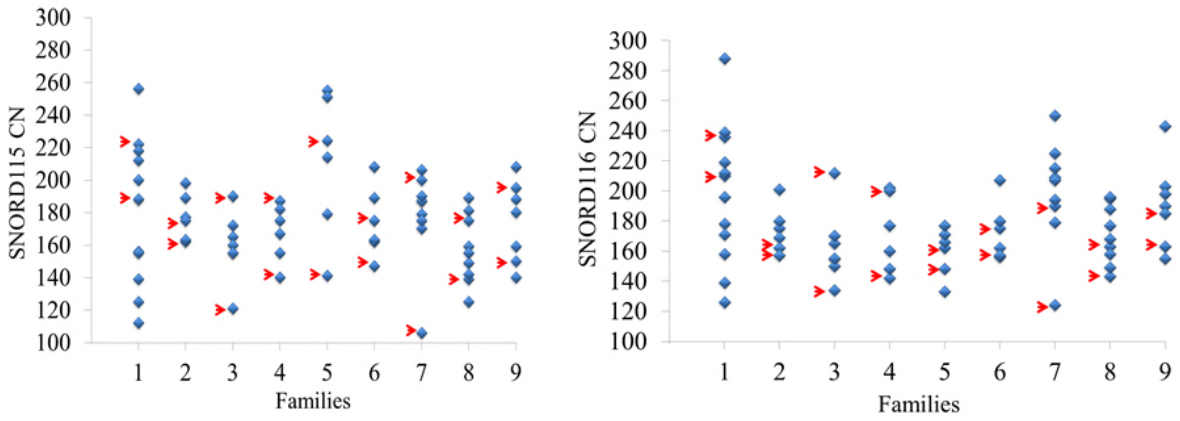


Figure 4: Inheritance patterns of SNORD gene copy numbers (CN) in families. Nine families of the MC population were tested and the copy numbers were determined for each individual (blue diamonds). The parental copy numbers are marked with an arrow. Note that in 7 out of the 9 families the spread of copy numbers exceeds that of the parents. The correlations with the behavioral tests for these animals are provided in Table S3.

Hence, the data show that at least some new haplotypes are generated in every generation. The experiment shows also that new large variation is generated in each generation, in line with the general observation that personality 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 all the offspring to the behavioral tests described above. We find indeed a large variation, as well as the expected significant regression with copy numbers (SNORD115: R^2=0.5, P << 0.001; SNORD116: R^2 = 0.41, P = 0.003) (Table S3).

Further, we asked whether the variability could be so high that somatic variation would become evident. We have therefore scored copy numbers of DNA isolated from different organs of a single

animal, but we did not find significant differences in this case (Table S3). While this does not rule out the possibility of very high somatic variation such that every cell might be different, there is at least no obvious pattern that distinguishes organs, including the brain.

Inbred strain variation

It has been shown previously that personality measures can differ even between individuals in an inbred strain (Freund et al., 2013; Jakovcevski et al., 2008; Lewejohann et al., 2011). Hence, we asked whether inbred mice would also show copy number variability 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 genders (30 female and 30 male). We find that there is indeed a large variation of copy numbers in these inbred mice, comparable to the spread seen in the outbred cohorts (Table 1). 20 mice of each sex were then subjected to the behavioral tests. We found no significant difference between males and females in these tests, but we could confirm the strong correlation with copy numbers (Table S1).

Transcriptome analysis

To investigate transcriptome changes in response to copy number variation of the SNORD genes, we choose ten C57BL/6J individuals representing the spread of copy numbers from low to high and sequenced RNA from their brains. First, we sought to confirm that the target genes would show the expected correlations and the corresponding non-target transcripts of these genes would not show this.

The first analysis focused on *Ht2cr* as target of SNORD115. We find a high positive correlation between SNORD115 copy number and total read count from the target exon in *Ht2cr*-201. As expected, there was no significant correlation between SNORD115 copy number and *Ht2cr*-204 and *Ht2cr*-206 as non-target transcripts. Similar results are observed for the SNORD116 target *Ankrd11*. The target exon of *Ankrd11* shows significant correlation to SNORD116 copy number, while there was no significant correlation between SNORD116 copy number and *Ankrd11*-204 and *Ankrd11*-207 as non-target transcripts (Table S4). Hence, the RNAseq results confirm the previous observations from the ddPCR experiments (see Table S2).

ANKRD11 is a chromatin regulator which influences the expression of downstream genes by binding chromatin modifying enzymes like histone deacetylases (Neilsen et al., 2008; Zhang et al., 2004). An *Ankrd11* knockout study has suggested that several hundred genes are regulated by *Ankrd11* (Gallagher et al., 2015). We used a list of 635 of these target genes for a correlation analysis with the *Ankrd11*-202 transcript abundance. We find that more than 100 of these 635 genes showed a correlation (positive or negative) to *Ankrd11*-202 expression level. Around 70 percent of them are

down-regulated and 30 percent are up-regulated by different *Ankrd11*-202 levels, with a subunit of the GABA A receptor (*Gabrg1*) showing the strongest correlation (Table S4).

Gabrg1 is known to play a crucial role in anxiety regulation in both humans and mice (Nuss, 2015; Tasan et al., 2011). To gain further insights into the function of the *Ankrd11* downstream genes, gene ontology (GO) and KEGG pathway enrichment analysis were performed using the DAVID online tools. This revealed five different functional categories: (I): metabolic pathway (II): proliferation and differentiation (III): anxiety (IV): intellectual ability formation and (V): osteogenesis (Figure 5).

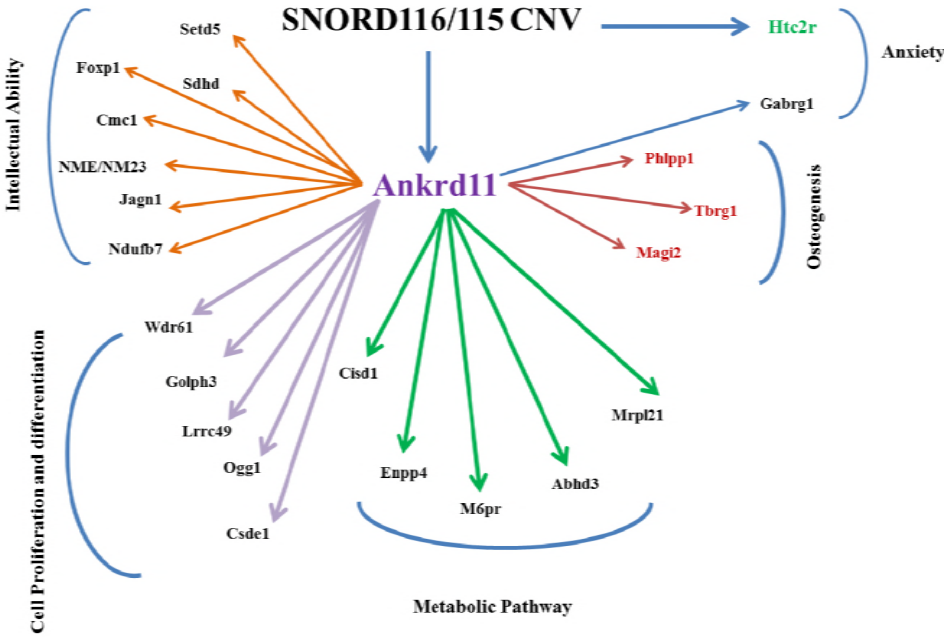


Figure 5: Overview and depiction of genes and processes that are influenced by SNORD115/116 copy number variations (CNV) at the transcriptional level. SNORD115 CNV targets directly the expression of the serotonin receptor *Ht2cr* and would this is known to be involved in anxiety responses. SNORD116 CNV, on the other hand, has as direct target the chromatin regulator *Ankrd11* and this affects a whole range of target genes. Among them is the GABA receptor *Gabrg1*, which is also known to be involved in anxiety responses. Three other groups of genes regulated by *Ankrd11* have effects on behavioral phenotypes as well, one group has general effects on cell differentiation and development and one group on osteogenesis. The figure is drawn based on the results provided in Table S4.

In humans, splice site variant mutations skipping exons IX, X and XI of *Ankrd11* cause the KGB syndrome (Low et al., 2017; Sirmaci et al., 2011).This is a rare genetic disorder characterized by intellectual disability, autism spectrum disorder, and craniofacial abnormalities (Ka and Kim, 2018). Hence, this provides further support that the splice-site regulation of *Ankrd11* through SNORD116 has a direct functional relevance for behavioral traits.

SNORD116 copy number and craniofacial features

Given that one of the categories of genes affected by *Ankrd11* expression level relates to osteogenesis (see above) and 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 have an effect on craniofacial features. We used a 3D landmarking approach to quantify differences in skull shapes based on CT scans, following the procedures described in (Pallares et al., 2014; Pallares et al., 2016). A principle component analysis was then used to assess the results. We find that PC1, which explains 35% of the variance, correlates indeed highly significantly with SNORD116 copy number (Figure 6). This suggests that not only behavioral, but also craniofacial traits are influenced by copy number variation at SNORD116.

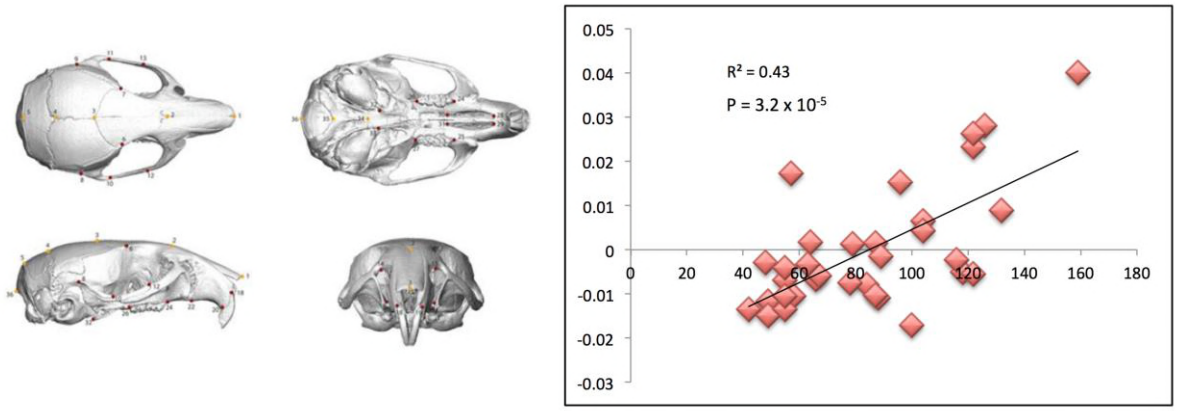


Figure 6: Skull shape changes in response to SNORD116 copy number variation. Left: depiction of landmarks used for morphometric analysis, right: plot of SNORD116 copy numbers (x-axis) versus PC1 values of the principal component analysis. Landmark descriptions an plot data are provided in Table S5.

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 correlation between copy number variation and personality traits also in other species.

We have first analyzed a second mouse species, the wood mouse *Apodemus uralensis*, which is separated from *Mus m. domesticus* 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. From the currently available genomic data, we could only retrieve sufficient information for the SNORD115 cluster, but this has even higher copy numbers than found in *Mus musculus* (Figure

1). We find indeed a significant regression with SNORD115 copy numbers and behavioral scores in the wood mice as well (Figure 3E).

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 behavioral tests that can be considered to reflect anxiety behavior (Guenther and Trillmich, 2015). 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 anxiety scores from the two tests (Figure 3F, G).

Humans

To test a possible correlation between SNORD copy numbers and personality traits in humans, we tested a subset of a cohort of 541 healthy individuals that had taken part in a study based on the Tridimensional Personality Questionnaire (TPQ). From these individuals, we chose the top 48 each for low and high anxiety scores 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 1) (Table S1). Particularly significant are SNORD115 and SNORD116_2. The latter is the variant that is predicted to bind to *Ankrd11* exon X, while the possible target genes for the other two SNORD116 variants are not yet clear. These latter ones show generally only little copy number variation (Table 1). 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 1: 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 (SD)	28 (8.0)	6 (1.1)	16 (4.4)	8 (1.4)	57 (12.6)
high anxiety and activity group (SD)	21 (4.5)	6 (0.8)	12 (1.7)	7 (1.1)	47 (6.8)
p-value[#]	2.8×10^{-5}	0.26	7.6×10^{-5}	0.072	1.7×10^{-5}

paired t-test, two-tailed

Discussion

Our data establish a link between behavioral scores and copy number variation at the SNORD115 and SNORD116 loci within the PWS region across four species of mammals. This includes also an isogenic mouse strain, which is not expected to carry many other genetic variants apart of the copy number variants (Zurita et al., 2011). Further, our data show that new copy number alleles are generated every generation, which provides an explanation for the high variances observed between parents and offspring, as well as the low general heritability. Importantly, this observation implies also that an artificial manipulation of copy numbers by engineering mouse strains would not be possible, since strains with defined number could not be established. On the other hand, it is known that whole deletions of parts of the PWS region including the snoRNAs lead to complex phenotypes, including behavior, both in mice and humans. Further, the interaction of the snoRNAs with their target genes via sequence complementarity provides a mechanistic explanation for the observed regulatory effects. This, together with the consistency of the correlation across four species establishes a causative link between snoRNA copy number and behavioral traits. Hence, our results establish a new paradigm in the regulation of behavioral variance, namely the involvement of highly mutable regulatory snoRNA clusters. We note that current association mapping protocols would not be able to establish this, since they rely on a stable association with neighboring SNPs, which is not given in the light of the fast generation of new copy number alleles.

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 variations of the cluster. Both the family studies, 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. 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. This suggests at minimum the generation of one new variant every few generations. Our family studies suggest that this may even occur every generation, given that we find more than four length variants among the offspring. Imprinted regions are known to include hotspots of meiotic recombination (Paigen and Petkov, 2010) and this has also been shown for the PWS region in humans (Robinson and Lalande, 1995). However, such an extreme rate of change has not been described so far.

Another very unusual 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 mice and humans, for which we could obtain the respective data. 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 pathways

Our transcript analysis data in response to SNORD115 copy number variation confirm the known interaction and regulation of the alternatively spliced exon Vb of the serotonin receptor *Ht2cr* through this snoRNA (Kishore and Stamm, 2006). We suggest that the corresponding predicted interaction of SNORD116 with *Ankrd11* may also be a direct interaction, based on our data. This can be inferred from the specificity of the effect on the predicted interacting exon X, in comparison to the transcripts of the locus that do not contain this exon and which do not show a dependence on copy number variation.

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 increased anxiety (Ding et al., 2008). A further 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 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 is under the control of the SNORD116 gene cluster copy number variation. For humans it has been suggested that there is indeed a link between personality and facial characteristics (Kramer and Ward, 2010; Squier and Mew, 1981), 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, it is unclear whether they would even recognize different craniofacial shapes among their conspecifics. However, there could be 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*.

383 *Implications for the genetics of personality traits*

384 Our findings can resolve a long standing controversy about the genetics and plasticity of personality
385 traits. Our data suggest that these traits are indeed genetically specified, but through a hypervariable
386 locus that would escape conventional genetic mapping approaches. Hence, it is not necessary to
387 invoke developmental or secondary epigenetic effects in explaining variance in personality traits,
388 given that there is a mechanistic pathway of modulation of target RNA expression through different
389 concentrations of snoRNAs.

390 The fact that we find significant regressions between personality scores and SNORD copy numbers
391 also in other rodents and even in humans, implies that the system is conserved throughout mammals.
392 In humans, we observed this even within a psychiatrically healthy control group, where we find that
393 extremes of personality traits for anxiety and activity are associated with significant differences in
394 copy numbers. The effects are somewhat weaker than for the rodents, but humans are evidently also
395 more subjected to environmental influences than the animals that were bred under controlled
396 conditions. Most intriguingly, however, the effect in humans is in the opposite direction, i.e. more
397 anxious individuals harbor lower copy numbers. This would suggest that the copy number variation
398 acts only as a general regulator, but the actual behavioral consequences are modified by downstream
399 pathways. The existence of such modifiers is also suggested in the comparison between the two
400 mouse populations in our study. MC mice have on average lower copy numbers than CB mice, but
401 they are overall more anxious when compared at a population level (Linnenbrink et al. unpublished
402 observations). Also the fact that a complete deletion of the SNORD116 cluster in inbred mice leads to
403 very anxious animals (Ding et al., 2008) suggests that modifiers must play a role as well. We suggest
404 that it should be possible to identify such modifiers in association studies when one controls for the
405 SNORD115 and SNORD116 copy numbers as a covariate. Because of the variation induced by these
406 snoRNAs, they may have been hidden in the background so far. In fact, it may generally become
407 necessary for behavioral studies to include the SNORD115 and SNORD116 copy numbers as
408 covariates, to reveal new patterns independent of this variation.

409

410 *Evolutionary implications*

411 One can raise the question of how and why such a hypervariable system could have evolved and how
412 it is maintained. This question has in fact been posed since a long time and there have been a number
413 of attempts to propose evolutionary models that could explain the large variability in behavioral traits.
414 These include drift effects in a neutral context (Tooby and Cosmides, 1990), mutation-selection
415 balance (Zhang and Hill, 2005) or balancing selection processes (Dingemanse and Wolf, 2010; Penke
416 and Jokela, 2016). But none of these models has taken the possibility into account that a hypervariable
417 locus could control this variation. Evidently, we need to caution that that the specific mechanism that

418 is revealed by our study is only applicable to mammals, while the question of variability of
419 personality traits is an issue across all animal taxa. Still the system found in mammals could provide a
420 guidance towards finding comparable systems also in other taxa.

421

422 **Acknowledgements**

423 We thank Chen Xie for his valuable advice with transcriptome analysis, Stefan Dennenmoser,
424 Cemalettin Bekpen and Guy Reeves for their advice on SNORDs copy number analysis, Sven Künzel
425 for sequencing, Ellen McConnell, Nicole Thomsen and Cornelia Burghardt for their great lab-related
426 support, Elke Blohm-Sievers for her help in CT scanning, the entire mouse team at MPI for
427 evolutionary biology, in particular Christine Pfeifle, Susanne Holz and Camilo Medina for taking care
428 of the mice in this study and Heike Harre and Fikri Tuğberk Kara for their help with behavioral tests
429 and brain dissection. Animals were kept according to FELASA (Federation of European Laboratory
430 Animal Science Association) guidelines, with the permit from the Veterinäramt Kreis Plön: 1401-
431 144/PLÖ-004697and the Veterinäramt Bielefeld for cavies, respectively. The experiments involving
432 the behavioral analyses was approved and registered under V244-71173/2015 through the local
433 authorities (Ethik-Kommission beim Umwelt- und Landwirtschaftsministerium Schleswig-Holstein).
434 The animal welfare officer was informed about the sacrifice of the animals for the molecular analyses.
435 The study on humans was approved by the ethics committee of the Ludwig-Maximilians-University
436 (LMU) in Munich and written informed consent was obtained from all subjects.

437

Methods

Mouse strains

Wildtype mice used in this study were offspring of mice that originated from wild populations sampled in the Massif Central region of France (MC) and the Cologne/Bonn region of Germany (CB) in 2004 and 2005 and then held under outbreeding conditions at the Max-Planck-Institute for Evolutionary Biology in Plön (see supplementary files in (Harr et al., 2016) for details). The C57BL/6J inbred strain was purchased at the age of 3 weeks from Jackson Laboratory.

Mice were usually kept in type III cages (Bioscape, Germany) and were weaned at the age of 3 weeks. 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 prior to experiments was at 20–24°C, 50–65% humidity and on a 12:12 light-dark schedule with lights on at 7 am.

DNA and RNA extraction

All dissections were done by 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 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 Assay kit (Invitrogen) reagent.

RNA extraction was done by using Trizol reagent. 1mL Trizol per 40mg tissue 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 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 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.

The quality of the RNA samples were 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

Total RNA which is enriched in small RNA was extracted by using the mirVana miRNA Isolation Kit. 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 to be used for the ddPCR. 1 µg of total RNA which was enriched in small RNA was used as input.

RNAseq analysis

Poly-A^+ RNA was used for cDNA synthesis and Illumina library preparation by using the Truseq stranded RNA HT kit. The libraries passing quality control were subjected to sequencing on an Illumina NextSeq 500 sequencing system. 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 (Waterston et al., 2002) by using Hisat2 (Kim et al., 2015). Htseq was used for counting reads overlapping with a specific feature (gene) (Anders et al., 2015). Differential expression analysis was performed with the DESeq2 package (Love et al., 2014) in 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. GO and KEGG pathway enrichment analyses were performed using DAVID online tools (Version 6.8, https://david-d.ncifcrf.gov/), with the classification stringency set to "medium" P value of <0.05

Droplet digital PCR

Digital PCR is a method enabling absolute quantification of DNA targets without the need to construct a calibration curve as used commonly 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 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.

Primers were designed with 50-70 bp amplicon length, GC content <50% and with very low potential for primer dimer structure.

Primers used in this study

Gene name	Forward sequence 5′→3′	Reverse sequence 5′→3′
SNORD 66	GTGTCTGGGCCACTGAGAC	TTCCTCAGGTCCTCAATCCCA
SNORD 115	GGGCCTCAGCGTAATCCTAT	ACCCAATGTCATGAAGAAAGGTG
SNORD 116	ACCTCAGTTCCGATGAGAGTG	TTCCCAGTCAAACATTCCTTGG
β-catenin	GGAAAAGAGCCTCAGGGCAT	CTGCCTGACGGCCAGG
Htr2c-201	ATTCGCGGACTAAGGCCATC	GAAGTTCGGGTCATTGAGCAC
HTR2C_204	CAAAATGGCACCCTGACCTG	GCCCTGGGTTCAATATCTGTTAC
HTR2C_206	ACAGAGTTGCTTGTTTGGTGTG	TCATGCCTGACAGTGGCATAG
Ankrd11-202	TATTGCCATCGACGGAGCTG	ACCTCCTGTTAGGCAAAGGC
Ankrd11_207	GGGAAGGAGAGGAGCAGAAAC	GGGAGGATCTGTCAGTTGCTATG

Ankrd11_204	CAGGGTTCTCCTTGAGCAAGAC	TGAAATCGAGGCAGCTGGTG

The QX100™ Droplet Digital™ PCR System (Bio-Rad, Hercules, CA, USA) was used in this study 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, 1ng DNA or cDNA and water.

For tandem copy separation, sample viscosity reduction and improved template accessibility, DNA digestion was done with EcoRI for SNORD115 and BamHI for SNORD116. These enzymes cut only once in the repeated units. 5 units of restriction enzymes were added to each sample. Then 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 (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.

Behavioral Tests

Behavioral tests were performed on the *M. m. domesticus* mice starting at an age of 24 weeks, on C57BL/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, an Open Field and a Dark/Light Box test, all adapted towards the use of wild mice to avoid escape. Test setups were cleaned with 30% ethanol between individual tests.

The Elevated Plus Maze consisted of four arms, 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 (Holmes et al., 2000). 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 grey PVC walls and they were allowed to explore it for 5 min (Reale, 2007; Wilson et al., 1994; Yuen et al., 2015). The speed of the mouse, the distance travelled and time spent within 10 centimeters off the wall vs in the central area were measured.

For the Dark/Light Box, the focal mouse was placed in a test apparatus containing a small dark shelter with two exits. During the first five minutes, the time until the mouse pokes its nose out of the shelter and the first time the tail is visible, was recorded. At five minutes, a set of keys was dropped next to the test apparatus from 136 cm height, and the second part of the experiment began. The time it took for the mouse to first look out and when the entire mouse was visible were measured. If mice did not

559 come out at all, the time was set to be 600 seconds. This test was adapted from tests in inbred mice
560 and common voles (Herde and Eccard, 2013; Young and Johnson, 1991).

561 The behavioral tests were filmed using a TSE camera (TSE system, Germany). To score the videos
562 from each test, all the videos were transferred to Videomot2 system (TSE system, Germany). Mice
563 were detected by the software in 3 points (head/center/tail base tracking) and then the software
564 automatically generates the numerical data of the time that each mouse spent at zones of interest.

565 Since personality is defined as consistent behavioral traits, we needed only those measurements which
566 were consistent over the course of the experiment. Hence, each behavioral test was repeated every 4
567 weeks for three times.

568 Statistical analysis were carried out using R 3.3.3 and R 3.3.2 . As the repeatability of a behavior is a
569 key component for the identification of personality trait, all single measurements assessed in the
570 behavioral tests were subjected to repeatability analysis. Repeatability was calculated using "rptR"
571 package (Nakagawa and Schielzeth, 2010). Normally distributed data were calculated using rpt (anova
572 based) and for count data rpt.poisGLMM was used. ID was used as a random effect in these models.

573 To determine whether individual behavioral measurements are correlated, a Spearman correlation
574 matrix was generated. P-values were corrected using the Holm method. Behaviors were clustered
575 using the protocol from (Herde and Eccard, 2013). An hierarchical cluster function was used from the
576 R package "cluster", specifically "agnes", to determine the relationship between the measurements.
577 All measurements were clustered using Manhattan clustering with complete linkage (Gyuris et al.,
578 2011; Herde and Eccard, 2013; Tremmel and Muller, 2013). Further details on the validation of the
579 statistics are provided in the thesis of Rebecca Krebs (Krebs 2018, University of Kiel).

580 The behavioral tests on the wild cavies (*Cavia aperea*) were conducted as described in (Guenther and
581 Trillmich, 2015). In short, to measure struggle docility, the animal was turned on its back and held in
582 the hand of an observer. For 30 s, the time the animal struggled to actively escape that situation was
583 recorded. To measure open field anxiety, the distance moved (cm) when individuals were exposed to
584 an open field for 20 min, was scored. The first 10 min, a shelter was present in the arena under which
585 animals could hide. For the second 10 min, this shelter was lifted out of the arena.

586

587 *Craniofacial shape analysis*

588 The morphometric analysis was performed as described previously (Pallares et al., 2015). Briefly,
589 mouse heads were scanned using a computer tomograph (micro-CT—vivaCT 40; Scanco,
590 Bruettisellen, Switzerland) at a resolution of 48 cross-sections per millimeter. Using the TINA
591 landmarking tool (Schunke et al., 2012), 36 three-dimensional landmarks were positioned in the skull.
592 The raw 3D landmark coordinates obtained in TINA tool were exported to MorphoJ (Klingenberg,
593 2011) for further morphometric analyses.

594 The symmetric component of the skull was obtained following (Klingenberg et al., 2002). In short, a
595 mirror image of the landmark configuration of each individual was generated, and a full GPA was
596 performed with the original and mirror configurations. Again, the resulting con-figurations were
597 averaged to obtain the symmetric component of shape variation. The new landmark coordinates
598 generated by the GPA are called "Procrustes coordinates".

599 Shape features were computed in MorphoJ principal components (PCs) from the n x 3k covariance
600 matrix of Procrustes coordinates, where n is the number of samples and k is the number of landmarks;

601 3k represents the number of Procrustes coordinates. PC loadings computed in this analysis are defined
602 as morphological score in this study.

603

604 *Human cohort and evaluation of personality traits.*

605 541 healthy controls were randomly selected from the Munich registry of residents and interviewed
606 for the presence of DSM-IV anxiety, affective, somatoform, eating, alcohol dependence, drug abuse
607 or dependence, disorders using a modified version of the Munich Composite International Diagnostic
608 Interview (Wittchen and Pfister, 1997) at the MPIP. Only individuals negative for the above-named
609 disorders were included in the context of a large study on the genetics of major depression using
610 EDTA blood as the source for DNA (see (Heck et al., 2009) for more detail). Five hundred and
611 fortyone individuals also completed the Tridimensional Personality Questionnaire (TPQ), a validated
612 personality measure (Weyers et al., 1995). To match the behavioral assessments in mice, single items
613 of the TPQ reflecting anxiety and activity were used to select individuals on the extremes of this
614 distribution. Individuals can score from 0 to 6 on the activity and anxiety scale, respectively and their
615 sum was used to identify the extreme groups, while matching for age and sex. The high anxiety and
616 activity group (N = 48) had a mean additive score of 7.35 (SD 1.04). The low anxiety and activity
617 group (N = 48) had a mean additive score of 0.73 (SD 0.49). The mean age of the high anxiety group
618 was 42.3 year and 41.6 years for the low anxiety group with 33.3% and 62.5% females, respectively.
619

Literature

Alex, K.D., Yavanian, G.J., McFarlane, H.G., Pluto, C.P., and Pehek, E.A. (2005). Modulation of dopamine release by striatal 5-HT2C receptors. Synapse *55*, 242-251.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq-a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166-169.

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

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

Bierbach, D., Laskowski, K.L., and Wolf, M. (2017). Behavioural individuality in clonal fish arises despite near-identical rearing conditions. Nature Communications *8*.

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

Cassidy, S.B., Schwartz, S., Miller, J.L., and Driscoll, D.J. (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*.

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

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

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

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

Gallagher, D., Voronova, A., Zander, M.A., Cancino, G.I., Bramall, A., Krause, M.P., Abad, C., Tekin, M., Neilsen, P.M., Callen, D.F.*, et al.* (2015). Ankrd11 Is a Chromatin Regulator Involved in Autism that Is Essential for Neural Development. Developmental Cell *32*, 31-42.

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

Gyuris, E., Fero, O., Tartally, A., and Barta, Z. (2011). Individual behaviour in firebugs (Pyrrhocoris apterus). Proceedings of the Royal Society B-Biological Sciences *278*, 628-633.

Harr, B., Karakoc, E., Neme, R., Teschke, M., Pfeifle, C., Pezer, Z., Babiker, H., Linnenbrink, M., Montero, I., Scavetta, R.*, et al.* (2016). Genomic resources for wild populations of the house mouse, Mus musculus and its close relative Mus spretus. Scientific Data *3*.

Heck, A., Lieb, R., Ellgas, A., Pfister, H., Lucae, S., Erhardt, A., Himmerich, H., Horstmann, S., Kloiber, S., Ripke, S.*, et al.* (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.

Herde, A., and Eccard, J.A. (2013). Consistency in boldness, activity and exploration at different stages of life. Bmc Ecology *13*.

Holmes, A., Parmigiani, S., Ferrari, P.F., Palanza, P., and Rodgers, R.J. (2000). Behavioral profile of wild mice in the elevated plus-maze test for anxiety. Physiology & Behavior *71*, 509-516.

Jakovcevski, M., Schachner, M., and Morellini, F. (2008). Individual variability in the stress response of C57BL/6J male mice correlates with trait anxiety. Genes Brain and Behavior *7*, 235-243.

671 Jiang, Y.H., Pan, Y.Z., Zhu, L., Landa, L., Yoo, J., Spencer, C., Lorenzo, I., Brilliant, M., Noebels, J., and
672 Beaudet, A.L. (2010). Altered Ultrasonic Vocalization and Impaired Learning and Memory in
673 Angelman Syndrome Mouse Model with a Large Maternal Deletion from Ube3a to Gabrb3.
674 Plos One *5*.
675 Ka, M.H., and Kim, W.Y. (2018). ANKRD11 associated with intellectual disability and autism regulates
676 dendrite differentiation via the BDNF/TrkB signaling pathway. Neurobiology of Disease *111*,
677 138-152.
678 Kim, D., Landmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory
679 requirements. Nature Methods *12*, 357-U121.
680 Kishore, S., and Stamm, S. (2006). The snoRNA HBII-52 regulates alternative splicing of the serotonin
681 receptor 2C. Science *311*, 230-232.
682 Kiss, T. (2001). Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. Embo
683 Journal *20*, 3617-3622.
684 Klingenberg, C.P. (2011). MorphoJ: an integrated software package for geometric morphometrics.
685 Molecular Ecology Resources *11*, 353-357.
686 Klingenberg, C.P., Barluenga, M., and Meyer, A. (2002). Shape analysis of symmetric structures:
687 Quantifying variation among individuals and asymmetry. Evolution *56*, 1909-1920.
688 Kramer, R.S.S., and Ward, R. (2010). Internal facial features are signals of personality and health.
689 Quarterly Journal of Experimental Psychology *63*, 2273-2287.
690 Lewejohann, L., Zipser, B., and Sachser, N. (2011). "Personality" in Laboratory Mice Used for
691 Biomedical Research: A Way of Understanding Variability? Developmental Psychobiology *53*,
692 624-630.
693 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R.,
694 and Genome Project Data, P. (2009). The Sequence Alignment/Map format and SAMtools.
695 Bioinformatics *25*, 2078-2079.
696 Lorenc, A., Linnenbrink, M., Montero, I., Schilhabel, M.B., and Tautz, D. (2014). Genetic
697 Differentiation of Hypothalamus Parentally Biased Transcripts in Populations of the House
698 Mouse Implicate the Prader-Willi Syndrome Imprinted Region as a Possible Source of
699 Behavioral Divergence. Molecular Biology and Evolution *31*, 3240-3249.
700 Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for
701 RNA-seq data with DESeq2. Genome Biology *15*.
702 Low, K.J., Hills, A., Williams, M., Duff-Farrier, C., McKee, S., and Smithson, S.F. (2017). A splice-site
703 variant in ANKRD11 associated with classical KBG syndrome. American Journal of Medical
704 Genetics Part A *173*, 2844-2846.
705 Nakagawa, S., and Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: a
706 practical guide for biologists. Biological Reviews *85*, 935-956.
707 Nakatani, J., Tamada, K., Hatanaka, F., Ise, S., Ohta, H., Inoue, K., Tomonaga, S., Watanabe, Y., Chung,
708 Y.J., Banerjee, R.*, et al.* (2009). Abnormal Behavior in a Chromosome-Engineered Mouse
709 Model for Human 15q11-13 Duplication Seen in Autism. Cell *137*, 1235-1246.
710 Neilsen, P.M., Cheney, K.M., Li, C.W., Chen, J.D., Cawrse, J.E., Schulz, R.B., Powell, J.A., Kumar, R.,
711 and Callen, D.F. (2008). Identification of ANKRD11 as a p53 coactivator. Journal of Cell Science
712 *121*, 3541-3552.
713 Nicholls, R.D., Saitoh, S., and Horsthemke, B. (1998). Imprinting in Prader-Willi and Angelman
714 syndromes. Trends in Genetics *14*, 194-200.
715 Nuss, P. (2015). Anxiety disorders and GABA neurotransmission: a disturbance of modulation.
716 Neuropsychiatric Disease and Treatment *11*, 165-175.
717 Paigen, K., and Petkov, P. (2010). Mammalian recombination hot spots: properties, control and
718 evolution. Nature Reviews Genetics *11*, 221-233.
719 Pallares, L.F., Carbonetto, P., Gopalakrishnan, S., Parker, C.C., Ackert-Bicknell, C.L., Palmer, A.A., and
720 Tautz, D. (2015). Mapping of Craniofacial Traits in Outbred Mice Identifies Major
721 Developmental Genes Involved in Shape Determination. Plos Genetics *11*.

722 Pallares, L.F., Harr, B., Turner, L.M., and Tautz, D. (2014). Use of a natural hybrid zone for
723 genomewide association mapping of craniofacial traits in the house mouse. Molecular Ecology
724 *23*, 5756-5770.
725 Pallares, L.F., Turner, L.M., and Tautz, D. (2016). Craniofacial shape transition across the house
726 mouse hybrid zone: implications for the genetic architecture and evolution of between-
727 species differences. Development Genes and Evolution *226*, 173-186.
728 Penke, L., and Jokela, M. (2016). The evolutionary genetics of personality revisited. Current Opinion
729 in Psychology *7*, 104-109.
730 Pezer, Z., Harr, B., Teschke, M., Babiker, H., and Tautz, D. (2015). Divergence patterns of genic copy
731 number variation in natural populations of the house mouse (Mus musculus domesticus)
732 reveal three conserved genes with major population-specific expansions. Genome Research
733 *25*, 1114-1124.
734 Qi, Y., Purtell, L., Fu, M., Lee, N.J., Aepler, J., Zhang, L., Loh, K., Enriquez, R.F., Baldock, P.A.,
735 Zolotukhin, S.*, et al.* (2016). Snord116 is critical in the regulation of food intake and body
736 weight. Scientific Reports *6*.
737 Reale, D. (2007). What do we really know about selection on personality? European Journal of
738 Personality *21*, 618-620.
739 Reale, D., Dingemanse, N.J., Kazem, A.J.N., and Wright, J. (2010). Evolutionary and ecological
740 approaches to the study of personality. Philosophical Transactions of the Royal Society B-
741 Biological Sciences *365*, 3937-3946.
742 Robinson, W.P., and Lalande, M. (1995). SEX-SPECIFIC MEIOTIC RECOMBINATION IN THE PRADER-
743 WILLI/ANGELMAN SYNDROME IMPRINTED REGION. Human Molecular Genetics *4*, 801-806.
744 Sanchez-Roige, S., Gray, J.C., MacKillop, J., Chen, C.H., and Palmer, A.A. (2018). The genetics of
745 human personality. Genes Brain and Behavior *17*.
746 Sato, M. (2017). Early Origin and Evolution of the Angelman Syndrome Ubiquitin Ligase Gene Ube3a.
747 Frontiers in Cellular Neuroscience *11*.
748 Schunke, A.C., Bromiley, P.A., Tautz, D., and Thacker, N.A. (2012). TINA manual landmarking tool:
749 software for the precise digitization of 3D landmarks. Frontiers in Zoology *9*.
750 Sirmaci, A., Spiliopoulos, M., Brancati, F., Powell, E., Duman, D., Abrams, A., Bademci, G., Agolini, E.,
751 Guo, S.R., Konuk, B.*, et al.* (2011). Mutations in ANKRD11 Cause KBG Syndrome, Characterized
752 by Intellectual Disability, Skeletal Malformations, and Macrodontia. American Journal of
753 Human Genetics *89*, 289-294.
754 Squier, R.W., and Mew, J.R.C. (1981). THE RELATIONSHIP BETWEEN FACIAL STRUCTURE AND
755 PERSONALITY-CHARACTERISTICS. British Journal of Social Psychology *20*, 151-160.
756 Stamm, S., Gruber, S.B., Rabchevsky, A.G., and Emeson, R.B. (2017). The activity of the serotonin
757 receptor 2C is regulated by alternative splicing. Human Genetics *136*, 1079-1091.
758 Tasan, R.O., Bukovac, A., Peterschmitt, Y.N., Sartori, S.B., Landgraf, R., Singewald, N., and Sperk, G.
759 (2011). ALTERED GABA TRANSMISSION IN A MOUSE MODEL OF INCREASED TRAIT ANXIETY.
760 Neuroscience *183*, 71-80.
761 Tooby, J., and Cosmides, L. (1990). ON THE UNIVERSALITY OF HUMAN-NATURE AND THE
762 UNIQUENESS OF THE INDIVIDUAL - THE ROLE OF GENETICS AND ADAPTATION. Journal of
763 Personality *58*, 17-67.
764 Tremmel, M., and Muller, C. (2013). Insect personality depends on environmental conditions.
765 Behavioral Ecology *24*, 386-392.
766 van Oers, K., and Mueller, J.C. (2010). Evolutionary genomics of animal personality. Philosophical
767 Transactions of the Royal Society B-Biological Sciences *365*, 3991-4000.
768 Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R.,
769 Ainscough, R., Alexandersson, M., An, P.*, et al.* (2002). Initial sequencing and comparative
770 analysis of the mouse genome. Nature *420*, 520-562.

771 Weyers, P., Krebs, H., and Janke, W. (1995). RELIABILITY AND CONSTRUCT-VALIDITY OF THE GERMAN
772 VERSION OF CLONINGERS TRIDIMENSIONAL PERSONALITY QUESTIONNAIRE. Personality and
773 Individual Differences *19*, 853-861.
774 Wilson, D.S. (1998). Adaptive individual differences within single populations. Philosophical
775 Transactions of the Royal Society of London Series B-Biological Sciences *353*, 199-205.
776 Wilson, D.S., Clark, A.B., Coleman, K., and Dearstyne, T. (1994). SHYNESS AND BOLDNESS IN
777 HUMANS AND OTHER ANIMALS. Trends in Ecology & Evolution *9*, 442-446.
778 Wittchen, H.U., and Pfister, H. (1997). DIA-X-Interview. Instruktionsmanual zur Durchführung
779 vonDIA-X-Interviews. In (Swets & Zeitlinger, Frankfurt).

780 Wolf, M., and Weissing, F.J. (2012). Animal personalities: consequences for ecology and evolution.
781 Trends in Ecology & Evolution *27*, 452-461.
782 Young, R., and Johnson, D.N. (1991). A FULLY AUTOMATED LIGHT DARK APPARATUS USEFUL FOR
783 COMPARING ANXIOLYTIC AGENTS. Pharmacology Biochemistry and Behavior *40*, 739-743.
784 Yuen, C.H., Pillay, N., Heinrichs, M., Schoepf, I., and Schradin, C. (2015). Personality does not
785 constrain social and behavioural flexibility in African striped mice. Behavioral Ecology and
786 Sociobiology *69*, 1237-1249.
787 Zhang, A.H., Yeung, P.L., Li, C.W., Tsai, S.C., Dinh, G.K., Wu, X.Y., Li, H., and Chen, J.D. (2004).
788 Identification of a novel family of ankyrin repeats containing cofactors for p160 nuclear
789 receptor coactivators. Journal of Biological Chemistry *279*, 33799-33805.
790 Zhang, X.S., and Hill, W.G. (2005). Genetic variability under mutation selection balance. Trends in
791 Ecology & Evolution *20*, 468-470.
792 Zhang, Y.J., Yang, J.H., Shi, Q.S., Zheng, L.L., Liu, J., Zhou, H., Zhang, H., and Qu, L.H. (2014). Rapid
793 Birth-and-Death Evolution of Imprinted snoRNAs in the Prader-Willi Syndrome Locus:
794 Implications for Neural Development in Euarchontoglires. Plos One *9*.
795 Zhao, Y., Xia, Q.Y., Yin, Y.P., and Wang, Z.K. (2016). Comparison of Droplet Digital PCR and
796 Quantitative PCR Assays for Quantitative Detection of Xanthomonas citri Subsp citri. Plos One
797 *11*.
798 Zurita, E., Chagoyen, M., Cantero, M., Alonso, R., Gonzalez-Neira, A., Lopez-Jimenez, A., Lopez-
799 Moreno, J.A., Landel, C.P., Benitez, J., Pazos, F*., et al.* (2011). Genetic polymorphisms among
800 C57BL/6 mouse inbred strains. Transgenic Research *20*, 481-489.

801

802