Deficiency of the paternally inherited gene *Magel2* alters the development of separation-induced vocalization and maternal behavior in mice

Authors:

Gabriela M. Bosque Ortiz [1, 2], Gustavo M. Santana[1, 2], Marcelo O. Dietrich [1, 2, 3, 4 *]

[1] Laboratory of Physiology of Behavior, Department of Comparative Medicine, Yale School of Medicine, New Haven, CT, USA.

[2] Interdepartmental Neuroscience Program, Biological and Biomedical Sciences Program, Graduate School in Arts and Sciences, Yale University, New Haven, CT.

[3] Program in Integrative Cell Signaling and Neurobiology of Metabolism, Department of Comparative Medicine, Yale School of Medicine, New Haven, CT, USA.

[4] Department of Neuroscience, Yale School of Medicine, New Haven, CT, USA.

* Correspondence to: marcelo.dietrich@yale.edu

Abstract

The behavior of offspring results from the combined expression of maternal and paternal genes. Genomic imprinting silences some genes in a parent-of-origin specific manner, a process that, among all animals, occurs only in mammals. How genomic imprinting affects the behavior of mammalian offspring, however, remains poorly understood. Here, we studied how the loss of the paternally inherited gene *Magel2* in mouse pups affects the emission of separation-induced ultrasonic vocalizations (USV). Using quantitative analysis of more than one hundred thousand USVs, we characterized the rate of vocalizations as well as their spectral features from postnatal days 6 to 12 (P6-P12), a critical phase of mouse development during which pups fully depend on the mother for survival. Our analyses show that *Magel2* deficient offspring emit separation-induced vocalizations at lower rates and with altered spectral features. We also show that dams display altered behavior towards their own *Magel2* deficient offspring. In a test to compare the retrieval of two pups, dams retrieve wildtype control pups first and faster than *Magel2* deficient offspring. These results suggest that the loss of *Magel2* impairs the expression of separation-induced vocalization in pups as well as maternal behavior, both of which support the pups' growth and development.

Keywords

Genomic imprinting; offspring-parent conflict; Prader Willi Syndrome; Autism spectrum disorders; behavior development.

Introduction

For the normal development, mammalian offspring need copies from both maternal and paternal genomes. Some genes, however, are expressed in a parent-of-origin specific manner. In other words, some genes are always expressed when inherited from the mother and some genes are always expressed when inherited from the father[1-3]. The process that regulates the expression of genes in a parent-of-origin specific manner is called genomic imprinting[4].

Genomic imprinting depends on epigenetic modifications of the genome. These modifications do not alter the sequence of the DNA but the chemical structure of the DNA, thereby leading to altered gene expression[1,4]. For example, an imprinted gene can be silenced in the maternal genome, and only the paternal allele will be expressed in the offspring (or vice-versa). Thus far, genomic imprinting has only been found in flowering plants and mammals[2,5-7].

The adaptive value of genomic imprinting remains a matter of intense theorization. A prevailing theory on the effects of genomic imprinting, known as the kinship theory[5-7], posits that maternal genes balance the energy investment of the mother between offspring survival and her own while paternal genes favor offspring survival alone. For example, the expression of paternal genes in the offspring would favor growth, while the expression of maternal genes would stunt growth[2,8-11]. The influence of imprinted genes is not restricted to growth, however. Imprinted genes are also primarily involved in brain development and in social behaviors[2,8-10].

Consider, for example, a series of imprinted genes in human chromosome 15 that are paternally inherited[5–9]. The deletion of these paternally inherited genes in chromosome 15 leads to neurodevelopmental disorders, such as Prader-Willi syndrome (PWS). PWS presents with hypotonia and poor feeding early in life, followed by hyperphagia, alteration in social behavior, and cognitive deficits. It should be noted that because PWS involves several genes in human

chromosome 15, it masks the relative contribution of single imprinted genes on the phenotype of the offspring. Among the PWS-related genes, MAGEL2 is a candidate gene for some of the clinical features of PWS. Humans with loss-of-function mutations in MAGEL2 present clinical aspects of PWS and autism spectrum disorders[10–12], suggesting that this single paternally inherited gene supports at least some of the developmental alterations found in PWS. In agreement with the clinical features of MAGEL2 deficiency in humans, *Magel2* deficient mice show impairments in growth and adult social behaviors[10,13,14]. Despite these previous studies, a more systematic investigation of the effects of imprinted genes on behavior is necessary to understand the adaptive value of these genetic modifications in mammals.

We reasoned that investigating the effects of imprinted genes on offspring behavior during the early postnatal period—when the exchange of resources between the mother and the offspring is most important—can help shed light on the adaptive value of genomic imprinting. Towards this end, we studied how the loss of paternally inherited *Magel2* affects the vocal behavior of mouse pups when separated from their dams[15–23], as separation-induced vocalizations signal the needs of the pups to the dams[19,24–29]. In contrast to human babies, mouse pups vocalize in the ultrasonic frequency range (45 - 120 kHz)[20,21,23,30], which humans cannot hear. In order to survey the vocal behavior of mice, we recorded the emission of ultrasonic vocalizations (USVs) when pups were separated from the home nest. We performed these studies at postnatal days 6, 8, 10, and 12, since it is during this phase of mouse development that the peak expression of separation-induced vocalizations typically occurs[15,31,32]. We then used VocalMat[33] to perform quantitative analysis of mouse vocal behavior. Moreover, we employed a maternal retrieval assay to test the effects of the loss of paternally inherited *Magel2* on maternal behavior towards the offspring.[12-17] Our analysis shows that in mouse pups the deficiency of *Magel2* impairs the expression of separation-induced vocalizations. This deficiency also reduces maternal retrieval behavior towards *Magel2* deficient pups compared to non-deficient siblings.

Results

Early waning of vocal behavior in *Magel2*$^{m+/p-}$ deficient pups

To investigate the effects of paternally inherited *Magel2* on the vocal behavior of infant mice, we crossed heterozygote males for *Magel2* deficiency with wildtype females. From this cross, we generated *Magel2* deficient offspring (*Magel2*$^{m+/p-}$) that carry the null allele from the father (p−) and the imprinted allele from the mother (m+). This cross also generates wildtype littermates (*Magel2*$^{m+/p+}$), used as experimental controls. As previously reported[15], *Magel2*$^{m+/p-}$ pups display lower body weight compared to controls (genotype: $F_{1, 181} = 19.71$, $P < 10^{-4}$; age: $F_{3, 181} = 69.64$, $P < 10^{-10}$; genotype x age: $F_{3, 181} = 0.02$, $P = 0.99$; two-way ANOVA; **Figure S1**). Post-hoc analysis shows a significant difference in body weight between genotypes at P12 (control: 6.08 ± 0.14 g; *Magel2*$^{m+/p-}$: 5.51 ± 0.20 g; $P = 0.04$; Holm-Šídák test; **Figure S1**).

We recorded the emission of USVs during 20 minutes of separation from the home nest at different postnatal ages (P6, P8, P10, and P12; **Figure 1A**). First, we analyzed the total number of USVs emitted during the period of separation using two-way ANOVA. We found a significant effect of genotype, age, and interaction between genotype and age (genotype: $F_{1, 181} = 20.61$, $P < 10^{-4}$; age: $F_{3, 181} = 11.80$, $P < 10^{-6}$; genotype x age: $F_{3, 181} = 3.90$, $P = 0.01$; **Figure 1B**). Post-hoc analysis (Holm-Šídák test) shows that the total number of USVs is similar among groups at P6 (control: 991 ± 139 USVs; *Magel2*$^{m+/p-}$: 787 ± 130 USVs; $P = 0.75$), P10 (control: 503 ± 45 USVs; *Magel2*$^{m+/p-}$: 466 ± 70 USVs; $P = 0.98$), and P12 (control: 594 ± 71 USVs; *Magel2*$^{m+/p-}$: 345 ± 46 USVs; $P = 0.06$) (**Figure 1B**). Compared to controls, however, *Magel2*$^{m+/p-}$ pups show a ≈53% reduction in the emission of USVs at P8 (control: 1045 ± 100 USVs; *Magel2*$^{m+/p-}$: 495 ± 65 USVs; $P < 10^{-4}$). We also analyzed the data in 5-minute intervals and found similar effects of *Magel2* deficiency on the rate of vocalizations (**Figure S2**). Moreover, separating our analysis in females and males show similar effects of genotype and age (**Figure 1C-D** and **Table S1**).

115 (Because we did not find differences in the rate of vocalization, we have pooled males and
116 females in all further analysis). In sum, our results thus far show age-specific reductions in the
117 emission of USVs in *Magel2*$^{m+/p-}$ mice, suggesting a non-sex specific effect for paternally
118 inherited *Magel2* on the vocal behavior of the offspring.
119
120 The emission of USVs occurs when the breathing musculature contracts, expelling air from the
121 lungs and propelling it through the larynx[12,18]. Since previous reports found that *Magel2*$^{m+/p-}$
122 mice display hypotonia[19], we considered the hypothesis that the low rate of USV emission in
123 *Magel2* deficient pups is due to a lower capacity to expel air from the lungs. To rule out this
124 hypothesis, we measured the intensity (or volume, in decibels) of the USVs. Since the intensity
125 of the USVs relates to the pressure by which the air is expelled through the larynx[18], a lower
126 intensity is expected in cases of hypotonia. This analysis shows that the intensity of the emitted
127 USVs between *Magel2*$^{m+/p-}$ mice and controls is similar in all ages tested (genotype: $F_{1, 181}$ =
128 2.67, P = 0.10; age: $F_{3, 181}$ = 2.33, P = 0.08; genotype x age: $F_{3, 181}$ = 0.78, P = 0.51; two-way
129 ANOVA; **Figure 1E**). Thus, hypotonia does not seem to be a factor of primary significance for
130 the lower rate of USV emission in *Magel2*$^{m+/p-}$ mice[16].
131
132 ***Magel2*$^{m+/p-}$ mice emit vocalizations with distinct spectral features**
133 In addition to the rate of separation-induced vocalizations, the spectral features of the USVs
134 also correlate with altered maternal care[20]. To test the extent to which *Magel2* deficiency affects
135 the spectral features of USVs across ages, we used two-way ANOVA to analyze the frequency
136 characteristics (pitch) and duration (**Figure 2A**) of USVs[21]. We found significant effects of
137 genotype and age for maximal frequency (genotype: $F_{1, 181}$ = 20.51, $P < 10^{-4}$; age: $F_{3, 181}$ = 3.23,
138 P = 0.02; genotype x age: $F_{3, 181}$ = 6.18, $P < 10^{-3}$; **Figure 2B**), bandwidth (genotype: $F_{1, 181}$ =
139 9.91, $P < 10^{-2}$; age: $F_{3, 181}$ = 6.97, $P < 10^{-3}$; genotype x age: $F_{3, 181}$ = 5.92, $P < 10^{-3}$; **Figure 2E**),
140 and duration (genotype: $F_{1, 181}$ = 4.08, P = 0.04; age: $F_{3, 181}$ = 6.35, $P < 10^{-3}$; genotype x age: $F_{3,}$

141 $_{181}$ = 2.05, P = 0.11; **Figure 2F**). In addition, we found a significant effect of genotype for mean
142 frequency (genotype: $F_{1, 181}$ = 4.86, P = 0.02; age: $F_{3, 181}$ = 0.24, P = 0.87; genotype x age: $F_{3, 181}$
143 = 0.41, P = 0.74; **Figure 2D**), but not for minimal frequency (genotype: $F_{1, 181}$ = 0.47, P = 0.47;
144 age: $F_{3, 181}$ = 1.96, P = 0.12; genotype x age: $F_{3, 181}$ = 0.48, P = 0.69; **Figure 2C**). We then used
145 post-hoc analysis (Holm-Šídák test) and found that at P8—but not at P6, P10, or P12—
146 *Magel2*$^{m+/p-}$ mice show significant differences compared to control mice in maximal frequency,
147 bandwidth, and duration (see panels plotting the 95% confidence intervals in **Figure 2**).
148
149 In addition to the main frequency component, USVs can contain harmonics (**Figure 2G**). We
150 calculated the percentage of USVs with harmonics and found a significantly lower number in
151 *Magel2*$^{m+/p-}$ mice compared to controls at P8 (control: 9.0 ± 1.7 %; *Magel2*$^{m+/p-}$: 3.5 ± 0.8 %; U =
152 145, $P_{2\text{-tailed}}$ = 0.002, Mann-Whitney test; **Figure 2H**) but not at P6 (control: 4.7 ± 0.9 %;
153 *Magel2*$^{m+/p-}$: 3.8 ± 0.6 %; U = 142.5, $P_{2\text{-tailed}}$ = 0.93, Mann-Whitney test; **Figure 2H**), P10
154 (control: 5.8 ± 0.9 %; *Magel2*$^{m+/p-}$: 6.1 ± 1.2 %; U = 240, $P_{2\text{-tailed}}$ = 0.98, Mann-Whitney test;
155 **Figure 2H**), or P12 (control: 3.6 ± 0.7 %; *Magel2*$^{m+/p-}$: 5.0 ± 1.5 %; U = 408.5, $P_{2\text{-tailed}}$ = 0.98,
156 Mann-Whitney test; **Figure 2H**). In sum, these results suggest that the loss of paternally
157 inherited *Magel2* in mice causes discrete changes in the features of separation-induced
158 vocalizations that are most evident at postnatal day eight.
159
160 **Discrete changes in the use of syllable types by *Magel2*$^{m+/p-}$ mice**
161 Mouse pups emit USVs of distinct classes—i.e., syllable types. Thus, the emission of different
162 syllable types could explain the discrete changes in the spectro-temporal features of USVs in
163 *Magel2*$^{m+/p-}$ [21,22]. We used a validated software to automatically categorize each USV into one of
164 eleven syllable types based on the morphology of the main component of the vocalization
165 (**Figure 3A-B**) [21]. The output of the method was the probability for each USV to be of a certain
166 syllable type. The highest probability (P_1) defined the syllable type for a given USV (**Figure 3A**).

167 The percent use of each syllable type per recording was compared between genotypes at each
168 age using two-way ANOVA. Using this approach, we did not find any significant differences in
169 the distribution of syllable types emitted by control and *Magel2*$^{m+/p-}$ mice at P6 (genotype: $F_{1, 374}$
170 = 2.5 x 10^{-16}, $P > 0.99$; class: $F_{10, 374} = 93.25$, $P < 10^{-4}$; genotype x class: $F_{10, 374} = 1.00$, $P = 0.44$)
171 and P12 (genotype: $F_{1, 616} = 7.7 \times 10^{-16}$, $P > 0.99$; class: $F_{10, 616} = 278$, $P < 10^{-4}$; genotype x class:
172 $F_{10, 616} = 0.81$, $P = 0.62$; **Figure 3C, 3F**). At P8, however, *Magel2*$^{m+/p-}$ pups emit vocalizations of
173 different syllable types (genotype: $F_{1, 539} = 2.2 \times 10^{-15}$, $P > 0.99$; class: $F_{10, 539} = 167$, $P < 10^{-6}$;
174 genotype x class: $F_{10, 539} = 9.51$, $P < 10^{-6}$; **Figure 2D)**, which include: 55% less USVs of the type
175 *chevron* (control: 9.0 ± 1.6 %; *Magel2*$^{m+/p-}$: 4.0 ± 0.01 %; $P = 0.02$, Holm-Šídák test); 44% less
176 USVs of the type *step-up* (control: 14.3 ± 1.4 %; *Magel2*$^{m+/p-}$: 8.1 ± 1.3 %; $P < 10^{-3}$, Holm-Šídák
177 test); 36% more USVs of the type *flat* (control: 9.6 ± 0.9 %; *Magel2*$^{m+/p-}$: 15.2 ± 1.2 %; P =
178 0.007, Holm-Šídák test); and 33% more USVs of the type *short* compared to controls (control:
179 23.6 ± 1.8 %; *Magel2*$^{m+/p-}$: 35.5 ± 2.0 %; $P < 10^{-4}$, Holm-Šídák test; **Figure 3D)**. At P10, we also
180 identified discrete differences in the emission of USVs of different syllable types between groups
181 (genotype: $F_{1, 462} = 8.1 \times 10^{-17}$, $P > 0.99$; class: $F_{10, 462} = 149$, $P < 10^{-4}$; genotype x class: $F_{10, 462}$ =
182 2.11, $P = 0.02$; **Figure 2E**): *Magel2*$^{m+/p-}$ pups emit 19% less *down frequency modulation*
183 (control: 31.0 ± 2.1 %; *Magel2*$^{m+/p-}$: 25.0 ± 1.5 %; $P = 0.02$; Holm-Šídák test) and emit 53%
184 more *up frequency modulation* (control: 4.8 ± 0.8 %; *Magel2*$^{m+/p-}$: 10.3 ± 2.0 %; $P = 0.04$; Holm-
185 Šídák test; **Figure 3E**). (**Table S1** provides a detailed analysis of the spectro-temporal features
186 of each syllable type across all ages tested in controls and *Magel2*$^{m+/p-}$ pups).
187
188 In summary, we found that *Magel2*$^{m+/p-}$ mice at P8 use simpler vocalizations that fall under the
189 'flat' and 'short' classifications instead of multicomponent USVs. These findings are in line with
190 our previous results **(Figures 1-2)** demonstrating that the largest differences in vocal behavior
191 occur in eight-day-old *Magel2*$^{m+/p-}$ pups.
192

193 **Altered vocal repertoire of *Magel2*$^{m+/p-}$ mice**

194 As stated above, the vocal analysis pipeline outputs the probability for each USV to be classified
195 as each of the eleven syllable types (P_1, P_2, P_3, … $P_{11;}$ **Figure 3A-B**). This distribution of
196 probabilities allows the qualitative and quantitative comparison of the vocal classification among
197 groups[21]. By considering the distribution of probabilities to classify each USV, it is possible to
198 estimate how similar the vocal repertoire of one group of mice is to another group. To compare
199 the vocal repertoire of mice across all ages studied, we used diffusion maps—a dimensionality
200 reduction technique that decreases the number of dimensions of the probability distribution from
201 eleven classes to three dimensions in a Euclidean space (**Figure 4A**)[21]. Using pairwise
202 comparisons (**Figure 4B**), we estimated the similarity between the vocal repertoire of mice of
203 different ages and genotypes. Using this method, we found that control pups at P6 and P8
204 (Cohen's coefficient: $\kappa = 0.99$) and control pups at P10 and P12 (Cohen's coefficient: $\kappa = 0.95$)
205 display vocal repertoires that are similar to each other (**Figure 4C-D**). These two age groups
206 (P6-P8 and P10-P12), however, present lower pairwise similarities when compared to each
207 other with κ ranging from 0.67 to 0.77 (**Figure 4E**). These results suggest that the vocal
208 repertoire of control pups undergoes significant changes between P8 and P10.

209

210 Next, we analyzed the same transitions in the vocal repertoire of *Magel2*$^{m+/p-}$ pups. The
211 comparison between the vocal repertoire of *Magel2*$^{m+/p-}$ pups at P6 and P8 show lower pairwise
212 similarity ($\kappa = 0.80$) compared to control pups (**Figure 4F**). In contrast to littermate controls,
213 *Magel2*$^{m+/p-}$ pups at P8 show a higher pairwise similarity with P10 ($\kappa = 0.84$) and P12 ($\kappa = 0.99$)
214 pups (**Figure 4F-H)**. Finally, we directly compared the vocal repertoire of *Magel2*$^{m+/p-}$ and
215 control pups. *Magel2*$^{m+/p-}$ pups, at P6, show high pairwise similarity when compared to controls
216 at P6 ($\kappa = 1.00$) and P8 ($\kappa = 1.00$), but not at P10 ($\kappa = 0.80$) and P12 ($\kappa = 0.72$). At P8,
217 *Magel2*$^{m+/p-}$ pups show relatively low pairwise similarities with control pups at P6 ($\kappa = 0.76$) and
218 P8 ($\kappa = 0.78$) but show high similarities with control pups at P10 ($\kappa = 1.00$) and P12 ($\kappa = 0.95$).

219 At P10, *Magel2*$^{m+/p-}$ pups show relatively lower pairwise similarities with control pups at P6 (κ =
220 0.67) and P8 (κ = 0.70) than at P10 (κ = 0.82) and P12 (κ = 0.81). This pattern is more evident
221 in P12 Magel2$^{m+/p-}$ pups, which show lower pairwise similarities with control pups at P6 (κ =
222 0.74) and P8 (κ = 0.75) than at P10 (κ = 0.99) and P12 (κ = 0.97).
223
224 Altogether, these analyses suggest a different dynamic for the ontogeny of the vocal repertoire
225 of *Magel2*$^{m+/p-}$ compared to control pups—with *Magel2*$^{m+/p-}$ pups at a younger age (i.e., P8)
226 resembling control pups at an older age (i.e., P10-P12). Thus, the period of development
227 between P8 and P10 seems to mark an important period for the effect of the maternally
228 imprinted gene, *Magel2*, on the vocal behavior of the offspring.
229
230 **Dams prioritize their control offspring compared to *Magel2*$^{m+/p-}$ in a retrieval test**
231 The emission of vocalization by infants draws caregiver's attention and care[12,13,15,17,22,23] while
232 genetically mute mouse pups are neglected by their dams[12]. Together, these observations led
233 us to speculate that the deficiency of *Magel2* in pups, which alters vocal behavior, could lead to
234 altered maternal behavior. In order to test this idea, we used a behavior assay to quantify
235 maternal behavior towards their own control and *Magel2*$^{m+/p-}$ offspring. In this assay, dams are
236 first placed in the middle of a three-chamber apparatus that contains their home nest in the
237 middle chamber (**Figure 5A**). After a period of acclimation, in the next stage, one pup of each
238 genotype is placed at the opposite ends of the apparatus. The time (latency) to retrieve each
239 pup back to the nest is then recorded (**Figure 5A**).
240
241 We initially evaluated whether dams show a preference to retrieve pups from one of the
242 genotypes. This analysis shows that control pups were more likely to be retrieved first (11/15
243 trials; 73.3%) compared to *Magel2*$^{m+/p-}$ pups (4/15 trials; 26.7%; **Figure 5B and Table S2**).
244 Accordingly, on average, the latency to retrieve control pups was less than half of the latency to

245 retrieve *Magel2*$^{m+/p-}$ pups (control: 51.9 ± 12.9 s; *Magel2*$^{m+/p-}$: 104.6 ± 19.1 s; W = 82, $P_{2\text{-tail}}$ =
246 0.02, Wilcoxon test; **Figure 5C**). Because eleven out of fifteen *Magel2*$^{m+/p-}$ pups were retrieved
247 second, we also ran experiments placing wildtype control pups in both ends of the apparatus to
248 compare the latency to retrieve the second pup in each of the experiments. Using this
249 comparison, dams took twice as long to retrieve *Magel2*$^{m+/p-}$ pups compared to control pups
250 (control: 64.57 ± 12.3 s; *Magel2*$^{m+/p-}$: 126.3 ± 22.23 s; W = 22, $P_{2\text{-tail}}$ = 0.08, Mann-Whitney test;
251 **Figure 5D**)**.** In summary, the deficiency of Magel2 in offspring alters not only offspring behavior
252 but also the behavior of the mother towards their own offspring.

Discussion

In this study, we recorded and analyzed vocalizations from *Magel2* deficient pups and their wildtype littermates at postnatal days 6, 8, 10 and 12. Using custom-built software to automatically quantify vocalizations[21], we counted the number of vocalizations and measured the spectro-temporal features of each vocalization, including its intensity, duration, bandwidth, mean frequency, maximum frequency, minimum frequency, and use of harmonic components. We also assigned a syllable type for each vocalization based on its morphological features in the time-frequency plane. We used further quantitative methods to analyze the vocal repertoire of mice across groups and ages. These methods shed light on discrete changes in the development of separation-induced vocalizations in *Magel2* deficient mice.

With regard to the vocal behavior of wildtype mice, our results demonstrate that the emission of separation-induced USVs gradually decreases from P6-P8 to P10-P12. This result agrees with previous studies, which show an inverse-U shape profile for separation-induced vocalizations in mice during the first two weeks of life [15,29,31,32,38]. Based on the large number of vocalizations we analyzed, we also found that wildtype mice use simpler vocalizations at older ages (P10-P12) compared to younger ages (P6-P8). These findings in wildtype mice provide the basis for comparisons with *Magel2* deficient pups.

Magel2 deficient pups show different dynamics for separation-induced vocalizations. At P6, these pups vocalize comparably to wildtype littermates, but at P8, their vocal number and features resemble wildtype littermates that are older (P10-P12). An explanation for these results is that *Magel2* deficient pups are less responsive to certain social cues. In socially isolated pups, therefore, the deprivation of these cues would not induce the behavior to the same degree as in wildtype pups. An alternative explanation for these results is that for *Magel2* deficient

279 pups, vocal behavior does not have the same fitness value compared to wildtype pups. In the
280 latter case, the behavior begins to change at younger ages due to the lack of reinforcement
281 (maternal care). While it is difficult to prove these interpretations experimentally, the delayed
282 latency for dams to retrieve *Magel2* deficient pups compared to wildtype littermates supports the
283 idea that the deficiency of this imprinted gene impairs the fitness of the offspring. Moreover, the
284 fact that *Magel2* deficient pups have lower body weight during early development further
285 suggests a decrease in the fitness of these animals. Whether the change in vocal behavior and
286 the decrease in body weight and maternal behavior are causally related warrants further
287 investigation.
288
289 Interestingly, the deficiency of *Magel2* is not the only example of an imprinted gene affecting the
290 emission of USVs by mouse pups. Consider, for example, previous studies of other imprinted
291 genes. Deletion of the paternally inherited imprinted gene *Peg3* lowers vocal rate in mouse pups
292 [39]. Conversely, deletion of the maternally inherited genes *Gabrb3* and *Ube3a* increases vocal
293 rate[9]. Moreover, duplication of the paternal imprint loci on chromosome 15 increases the
294 separation-induced vocalization of mouse pups[24]. Based on these findings, it is tempting to
295 speculate that imprinted genes inherited from the father increase vocal rate while imprinted genes
296 inherited from the mother decrease vocal rate. To test this generalization more formally, future
297 studies will need to test the effect of *all* imprinted genes on the vocal behavior of the offspring.
298
299 P8 *Magel2*$^{m+/p-}$ vocalize at a low rate, mainly emitting short and flat vocals similar to those of
300 older pups. Moreover, when we estimate the similarity between the vocal repertoire of mutant
301 and wildtype pups across ages tested, we find that P8 *Magel2*$^{m+/p-}$ USVs are most similar to
302 those produced by wildtype P10 and P12 pups. Interestingly, reduction in USV rate is
303 associated to ASD, which is highly comorbid with MAGEL2$^{m+/p-}$ patients[25]. Further studies are
304 necessary to describe whether deviation in USV emission for *Magel2* deficient mice represents

305 some kind of curve shift in the behavior. For example, it could be that the critical phase in which
306 pups use separation-induced vocalization to gain maternal attention is shortened in offspring
307 deficient for *Magel2.* Thus, our findings could suggest that alteration in the emission of USVs is
308 part of a social behavior deficit in *Magel2*$^{m+/p-}$ pups that continues into adulthood[26].
309
310 In broader terms, we posit that our findings support the theory that genomic imprinting evolved
311 to balance the cost of the phenotype for the offspring and for the mother, as well as to balance
312 the best interests of mothers and fathers in altering offspring's phenotype[3,29]. In the case of
313 paternally inherited genes, the expression of these genes increases vocal behavior, thus,
314 increasing maternal care and favoring the use of maternal resources, which is in the best
315 interests of the father. Conversely, the loss of these paternally inherited genes decreases vocal
316 behavior and, consequently, the demand for maternal care, thus, conserving maternal
317 resources, which is in the best interests of the mother[2,5-7,27]. This theoretical framework is
318 supported by our findings in a test that measures the latency to retrieve isolated pups back to
319 the home nest, which demonstrate impaired maternal care towards Magel2 deficient pups
320 compared to wildtype littermates[12,13,15,17,28]. As alluded above, future studies should test this
321 theoretical view more directly by systematically investigating the effect of imprinted genes on the
322 behavior of offspring and on the behavior of mothers towards their offspring. These efforts will
323 help elucidate behavior phenotypes that occur in neurodevelopmental disorders as well as
324 expand our understanding of the evolutionary adaptation of genomic imprinting.

Acknowledgements

We thank the Dietrich Lab members for critical feedback on the project and the manuscript. Special thanks to: Delva Leão for her contribution to the data analysis; Jeremy Bober for helping with animal maintenance and care; as well as Antonio Fonseca who developed VocalMat software used for the analysis of vocal recordings. We thank David Gillich, Susan Andranovich, Valeria Krizsan, and Vickie Clark for administrative support. M.O.D. was supported by the National Institute of Mental Health of the National Institutes of Health under Award Number (R01MH125008) and by a grant of the Foundation for Prader-Willi Research. G.M.B.O. was supported by the HHMI Gilliam Fellowship. G.M.S. was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES). We thank David Bruin for copyediting the manuscript. The authors declare no conflict of interest.

Author contributions

M.O.D and G.M.B.O. conceived the hypothesis, designed the study and wrote the manuscript. G.M.B.O. performed experiments. G.M.B.O. and G.M.S. analyzed and plotted the data. All authors read and edited the manuscript.

Material and Methods

Experimental models and subject details

All preweaning mice used in the experiments were 6 to 12 days old from both sexes (see table below for total number of recording used under each age specifying sex and genotype). Litters were provided from 9 separate breeding pairs. Separate litters were used for isolation and maternal retrieval tests. Details of litters tested found in **Table S1 and S2**. Dams used were 2 to 6 months old. To generate experimental pups, we used the following cross: *Magel2*$^{m+/p-}$ (Jax #009062) dams bred with C57BL/6J (Jax #000664) males. Offspring from this cross were either *Magel2*$^{m+/p-}$ or wildtype (*Magel2*$^{m+/p+}$). All mice were kept in temperature- and humidity-controlled rooms, in a 12/12 hr. light/dark cycle, with lights on from 7:00 AM–7:00 PM. Studies took place during the light cycle. Food (Teklad 2018S, Envigo) and water were provided ad libitum. All procedures were approved by IACUC (Yale University).

Behavior test

Pups from the same litter were placed individually in a soundproof chamber containing fresh bedding material [19]. An UltraSoundGate Condenser Microphone CM 16 (Avisoft Bioacoustics, Berlin, Germany) was placed 10 cm above the recording chamber and connected to the UltraSoundGate 416 USGH device to record ultrasonic vocalizations. The recording sessions lasted 20 minutes. Four to eight chambers were recorded simultaneously. After testing, mice were placed back in their home cage with the dam. Pups were tested at postnatal days 6, 8, 10, and 12. Because pups that were naïve for the test—only tested at one specific age—show similar results as pups tested at multiple ages, we pooled all mice together for our analysis. Moreover, literature in the field have reported that repeated isolation testing does not affect infant vocal behavior[29,30]. Details of individual behavior measurements are found in **Table S1**.

367 **Vocalization analysis**

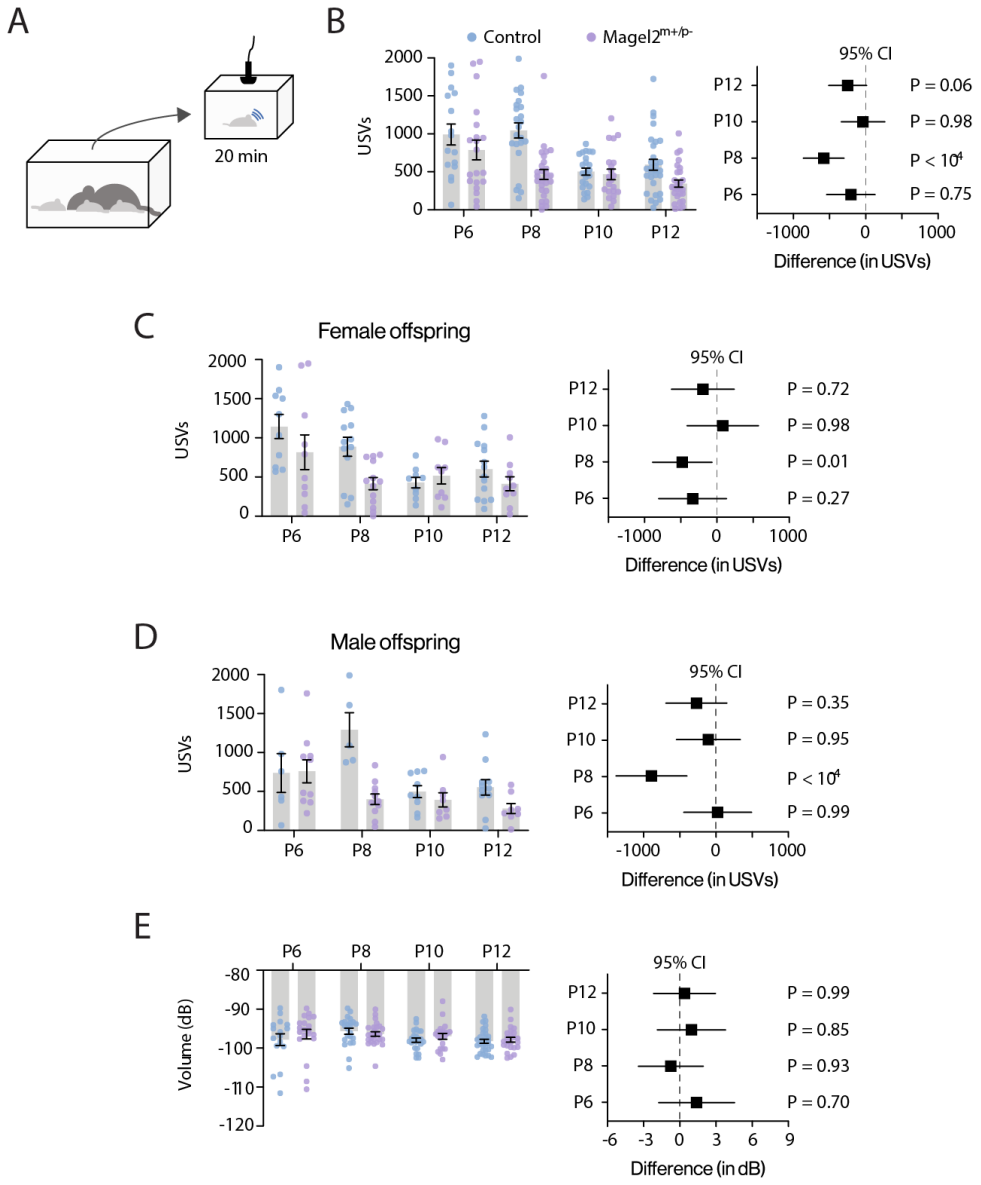
368 Ultrasonic vocalizations were automatically extracted from audio recordings using a custom-built
369 tool [14]. In brief, audio recordings were converted from the time-domain to the frequency-domain
370 using a 1024-point Fast Fourier Transform (FFT) through a 512-width hamming window with
371 50% overlap. Spectrograms were computed from the FFT and processed as images. Each pixel
372 in the spectrogram corresponded to the intensity of each time-frequency component. Next, we
373 applied a series of image-processing techniques (e.g., contrast enhancement, binarization,
374 median filter, and morphological operations) to obtain segmentation of candidate vocalizations.
375 A single spectrogram was generated for each candidate vocalization detected. Candidate
376 vocalizations were classified as noise or real vocalization using a local median noise filter. The
377 remaining vocalization candidates are further labeled under one of eleven call type
378 classifications [20] using a Convolutional Neural Network (CNN), or as noise. The CNN was
379 trained using a curated vocalization dataset, containing over 20,000 noise samples and 40,000
380 vocalization samples. Finally, the tool produces one spectrogram centralized on each
381 vocalization for visual inspection, and a table (*xlsx* format) containing spectro-temporal features
382 for each USV, such as time, duration, bandwidth, frequency, and intensity (minimum, mean, and
383 maximum) values.

384

385 **Maternal Retrieval Test**

386 The maternal preference test was performed in a three-chamber apparatus (65 x 42 x 23 cm)
387 and comprised of three stages: Stage 1 – acclimation: the dam was allowed to explore the
388 apparatus without the presence of pups for five minutes. In the middle of the apparatus, home
389 cage nesting under an infrared igloo. Stage 2 – exploration: two P8 mice were placed on each
390 side of the apparatus and the dam was allowed to explore the pups for five minutes. Stage 3 -
391 preference: dam was recorded while she retrieved pups back to her home cage nesting. Groups
392 were randomly alternated between both sides to avoid preference for one side of the chamber.

393 Latency to retrieve each pup was timed from videos taken of the retrieval tests. Summary of
394 results found in **Table S2**.
395
396 **Quantification and statistical analysis**
397 Prism 8.0 or above was used to analyze data and plot figures. Shapiro-Wilk normality test was
398 used to assess normal distribution of the data. To analyze differences in the use of harmonics,
399 we used the non-parametric Mann-Whitney Test with Bonferroni correction to find statistically
400 different effects. Comparison of diffusion maps were analyzed through pairwise comparison
401 using Cohen's kappa coefficient. Maternal retrieval latency for retrieval of each pup from the
402 start of test used Wilcoxon test. Analysis of second retrieval latency from first retrieval was
403 calculated using Welch's t test. The rest of the data was analyzed using two-way ANOVA or
404 mixed-effects analysis. Sidak's multiple comparisons test was used to find post hoc differences
405 among groups and to calculate the 95% confidence intervals to report effect size. In the text,
406 values are provided as mean ± SEM. $P < 0.05$ was considered statistically significant and, when
407 necessary and as described above, was corrected using Bonferroni's method. Statistical data
408 are provided in text and in the figures.



409

410 **Figure 1. *Magel2* deficiency affects the emission of ultrasonic vocalizations in mice.** (**A**)
411 Schematic of the protocol used to record separation-induced USVs in mice (from P6 to P12); pups
412 are separated from the home nest in a new chamber equipped with an ultrasonic microphone and
413 recorded for 20 minutes. (**B**) Total number of USVs emitted by control (blue) and *Magel2* deficient
414 (purple) littermates at P6, P8, P10, and P12; right panel denotes the 95% confidence intervals as
415 a measure of effect size. (**C**) Similar to (B), but only considering female pups. (**D**) Similar to (B),
416 but only considering male pups. (**E**) Average intensity of the USVs measured in decibels; right

panel denotes the 95% confidence intervals as a measure of effect size. Bars represent mean value with error bars representing SEM and round symbols representing individual values. When plotting the effect sizes, squared symbols and black lines represent 95% confidence intervals calculated as the different between *Magel2* deficient and control pups. *P* values are provided in the figures as calculated using Sidak's multiple comparison test. The sample sizes for control and *Magel2* deficient pups are: P6, n = 16 and 20; P8, n = 23 and 28; P10, n = 23 and 20; and P12, n = 30 and 28, respectively.

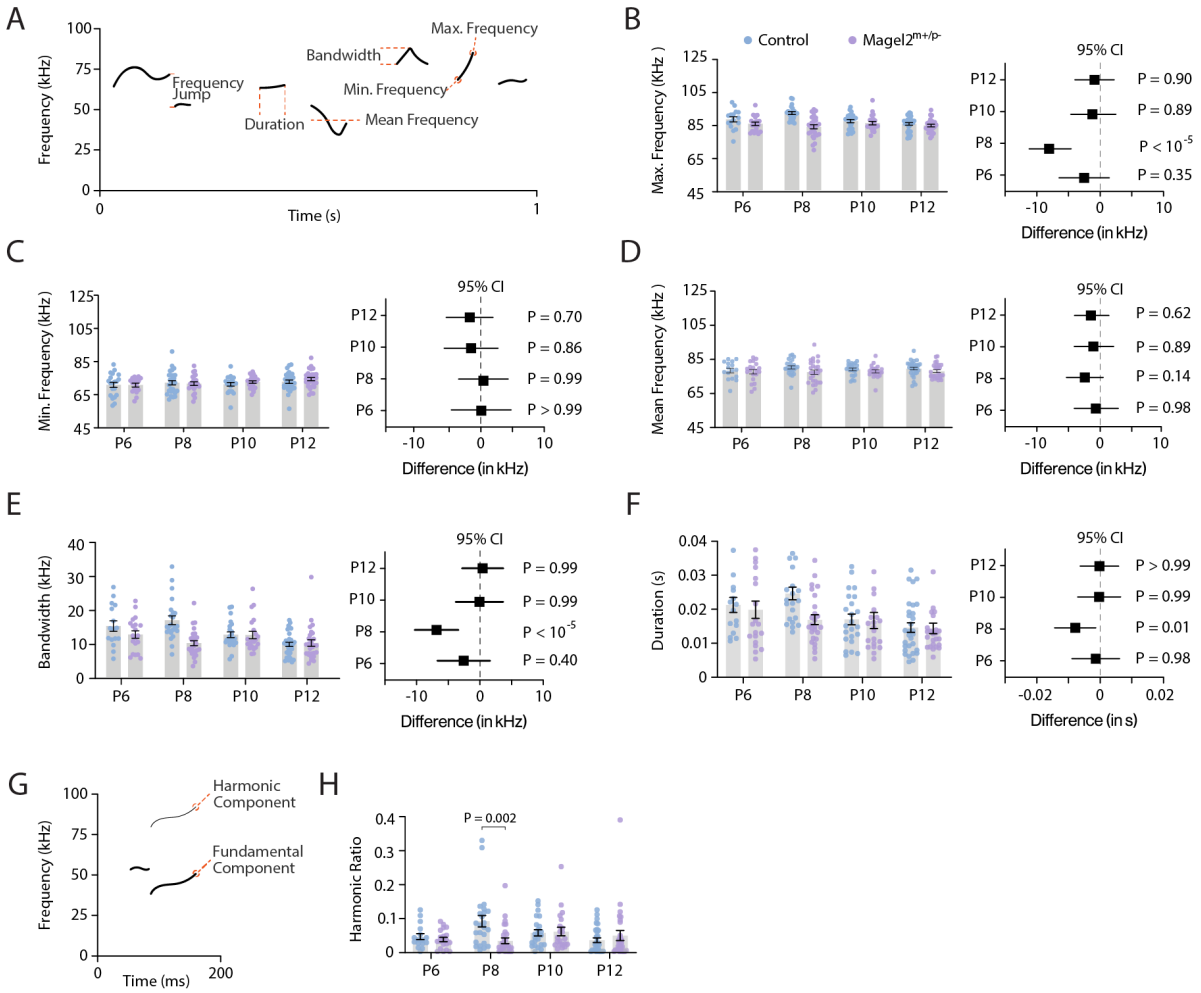
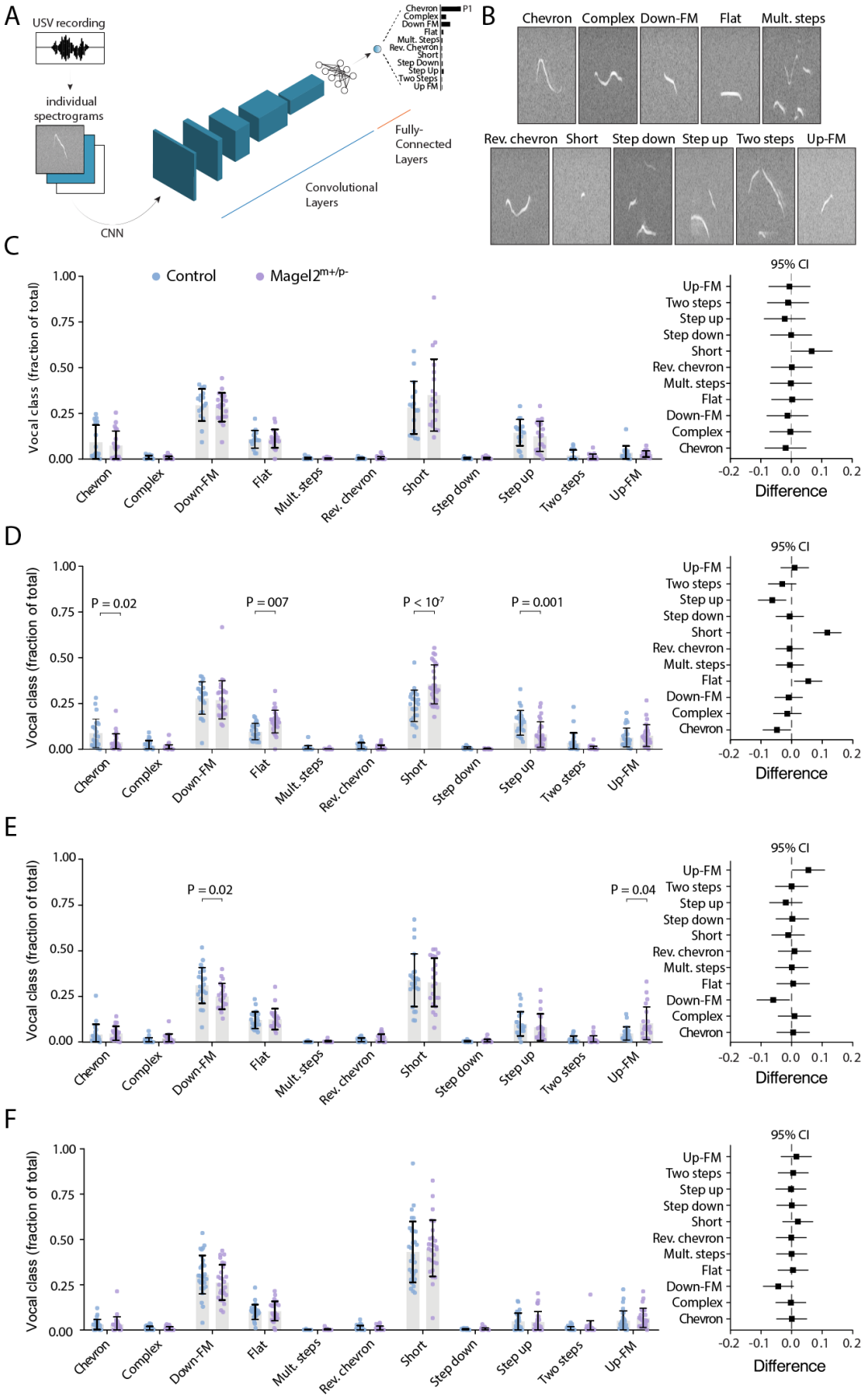


Figure 2. *Magel2* deficient pups emit ultrasonic vocalizations of distinct spectro-temporal features. (**A**) Illustration of a spectrogram with the spectro-temporal features measured for each USV. (**B**) Maximum frequency of the USVs emitted by control and *Magel2* deficient littermates at P6, P8, P10, and P12; right panel denotes the 95% confidence intervals as a measure of effect size. (**C**) Similar to (B) but plotting the minimum frequency of the USVs. (**D**) Similar to (B) but plotting the mean frequency of the USVs. (**E**) Similar to (B) but plotting the bandwidth of the USVs. (**F**) Similar to (B) but plotting the duration of the USVs. (**G**) Illustration of the spectrogram of a single USV with a harmonic component. (**H**) Ratio of harmonic across all USVs emitted by control and *Magel2* deficient littermates. Bars represent mean value with error bars representing SEM and round symbols representing individual values. When plotting the effect sizes, squared

436 symbols and black lines represent 95% confidence intervals calculated as the different between
437 *Magel2* deficient and control pups. In B-F, *P* values are provided in the figures as calculated
438 using Sidak's multiple comparison test post hoc analysis from two-way ANOVA test. In H, *P*
439 values are provided as calculated using Mann-Whitney test. The sample sizes for control and
440 *Magel2* deficient pups are: P6, n = 16 and 20; P8, n = 23 and 28; P10, n = 23 and 20; and P12,
441 n = 30 and 28, respectively.
442


A
USV recording
individual spectrograms
CNN
Convolutional Layers
Fully-Connected Layers
Chevron
Complex
Down FM
Flat
Mult. Steps
Rev. Chevron
Short
Step Down
Step Up
Two Steps
Up FM
P1
B
Chevron
Complex
Down-FM
Flat
Mult. steps
Rev. chevron
Short
Step down
Step up
Two steps
Up-FM
C
Control
Magel2m+/p-
Vocal class (fraction of total)
1.00
0.75
0.50
0.25
0.00
95% CI
Difference
-0.2
-0.1
0.0
0.1
0.2
D
P = 0.02
P = 007
P < 10-7
P = 0.001
E
P = 0.02
P = 0.04
F

443

444 **Figure 3. *Magel2* deficient pups emit ultrasonic vocalizations with discrete changes in the**
445 **distribution of syllable types**. (**A**) Illustration of the convolutional neural network used to
446 classify each USV into one of eleven syllable types based on their morphology in spectrograms.
447 (**B**) Spectrograms representing each of the eleven syllable types. (**C**) Distribution of syllable
448 types in P6 pups—control in blue and *Magel2* deficient in purple. Data are showed as fraction of
449 the total number of USVs; right panel denotes the 95% confidence intervals as a measure of
450 effect size. (**D**) Similar to (C), but for P8 pups. (**E**) Similar to (C), but for P10 pups. (**F**) Similar to
451 (C), but for P12 pups. Bars represent mean value with error bars representing SEM and round
452 symbols representing individual values. When plotting the effect sizes, squared symbols and
453 black lines represent 95% confidence intervals calculated as the different between *Magel2*
454 deficient and control pups. *P* values are provided in the figures as calculated using Sidak's
455 multiple comparison test as a post hoc analysis after two-way ANOVA. The sample sizes for
456 control and *Magel2* deficient pups are: P6, n = 16 and 20; P8, n = 23 and 28; P10, n = 23 and
457 20; and P12, n = 30 and 28, respectively.

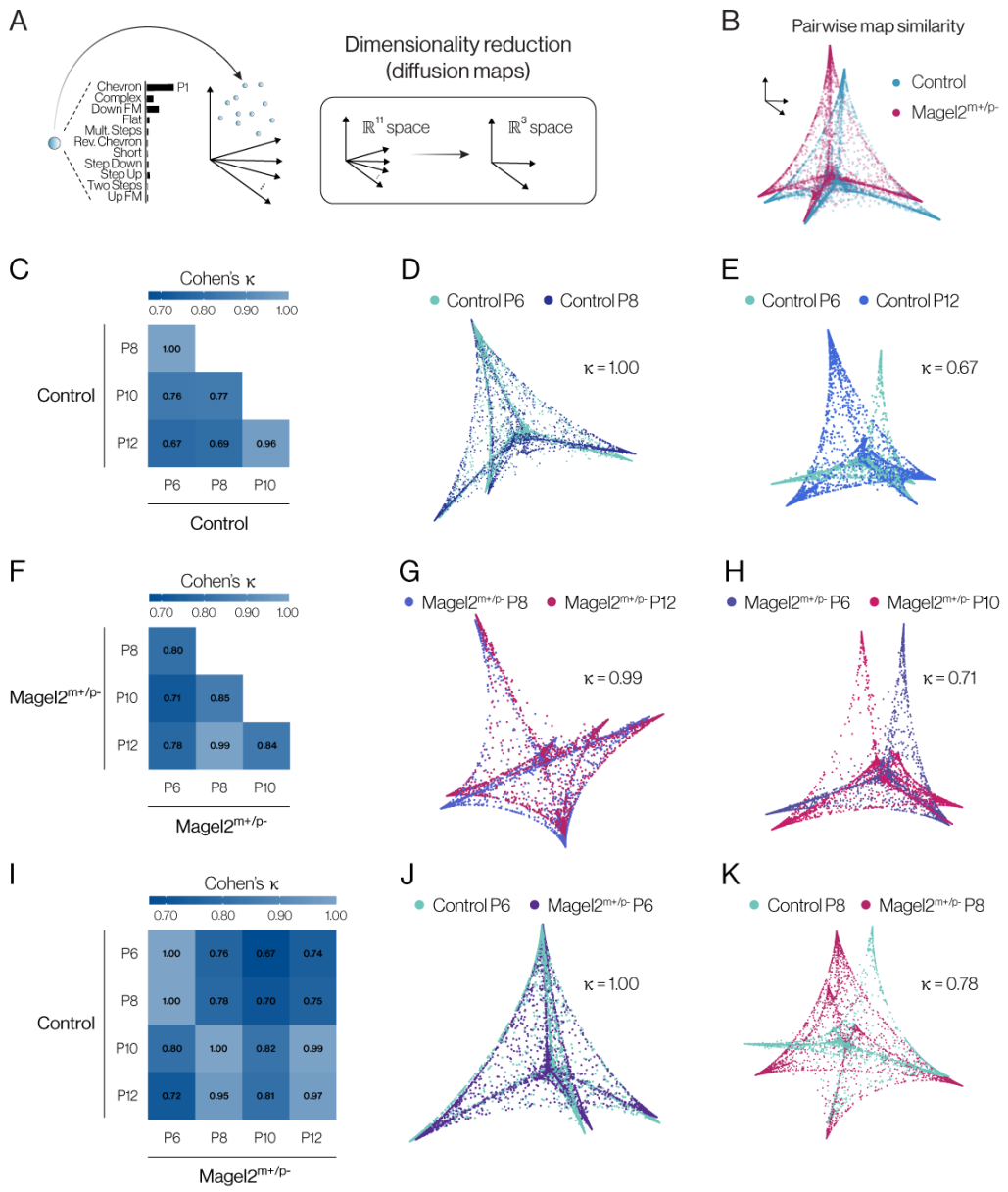


Figure 4. Analysis of the vocal repertoire of pups across ages. (**A**) Illustration of the output of the convolutional neural network, with a distribution of eleven probabilities for vocal classification (one probability for each of the eleven syllable types, with the highest probability defining the syllable type). Using diffusion maps, a dimensionality reduction technique, these eleven dimensions are reduced to three dimensions in the Euclidian space. (**B**) Illustration of a pairwise comparison of the vocal repertoire of pups using diffusion maps and 3D alignment of

466 the manifolds (see methods for more details). (**D**) Comparison of the pairwise distance matrix
467 between control pups at different ages using Cohen's Kappa coefficient. (**D-E**) Examples of two
468 pairwise comparisons with high and low alignment. (**F-H**) Similar to (**D-E**), but for *Magel2*
469 deficient pups. (**I-K**) Similar to (C-H), but comparing control and *Magel2* deficient pups across
470 ages. The sample sizes for control and *Magel2* deficient pups are: P6, n = 16 and 20; P8, n = 23
471 and 28; P10, n = 23 and 20; and P12, n = 30 and 28, respectively.
472

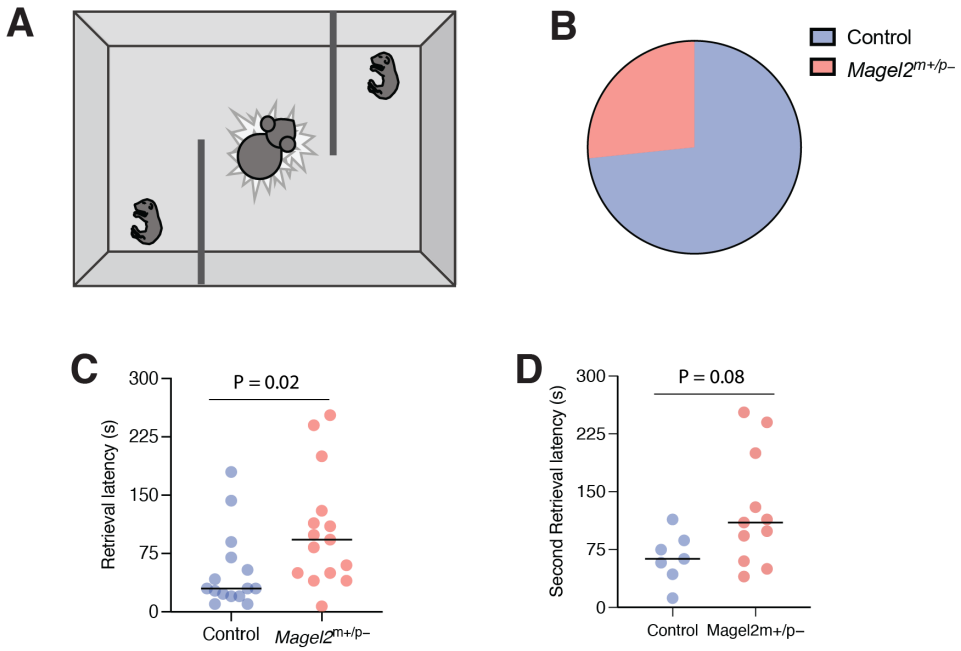


Figure 5. Maternal retrieval behavior is biased towards wildtype versus *Magel2* deficient pups. (**A**) Diagram of maternal retrieval protocol with P8 pups on opposite ends of the apparatus and mother at nest in the middle compartment. (**B**) Pie chart showing percentages of first retrieval of control and *Magel2* deficient pups from all trials. (**C**) Welch's t-test comparing the retrieval latency of control and *Magel2* deficient pups. (**D**) Welch's t-test comparing the latency for the retrieval of the second pup comparing *Magel2* deficient pups and wildtype controls that were picked second in the wildtype versus wildtype trial. Significant *P* values are shown in the graphs. The number *Magel2* deficient pups and controls trials are: P8, n = 15 and n = 7.

485 **References**

486 1. Barlow DP, Bartolomei MS. Genomic imprinting in mammals. *Cold Spring Harb Perspect*
487 *Biol.* 2014;6(2).

488 2. Perez JD, Rubinstein ND, Dulac C. New Perspectives on Genomic Imprinting, an
489 Essential and Multifaceted Mode of Epigenetic Control in the Developing and Adult Brain. *Annu*
490 *Rev Neurosci.* 2016;39:347-384.

491 3. Ferguson-Smith AC. Genomic imprinting: the emergence of an epigenetic paradigm. *Nat*
492 *Rev Genet.* 2011;12(8):565-575.

493 4. Morison IM, Ramsay JP, Spencer HG. A census of mammalian imprinting. *Trends*
494 *Genet.* 2005;21(8):457-465.

495 5. Haig D. Coadaptation and conflict, misconception and muddle, in the evolution of
496 genomic imprinting. *Heredity (Edinb).* 2014;113(2):96-103.

497 6. Patten MM, Ross L, Curley JP, Queller DC, Bonduriansky R, Wolf JB. The evolution of
498 genomic imprinting: theories, predictions and empirical tests. *Heredity (Edinb).* 2014;113(2):119-
499 128.

500 7. Spencer HG, Clark AG. Non-conflict theories for the evolution of genomic imprinting.
501 *Heredity (Edinb).* 2014;113(2):112-118.

502 8. Li L, Keverne EB, Aparicio SA, Ishino F, Barton SC, Surani MA. Regulation of maternal
503 behavior and offspring growth by paternally expressed Peg3. *Science.* 1999;284(5412):330-
504 333.

505 9. Jiang YH, Pan Y, Zhu L, et al. Altered ultrasonic vocalization and impaired learning and
506 memory in Angelman syndrome mouse model with a large maternal deletion from Ube3a to
507 Gabrb3. *PLoS One.* 2010;5(8):e12278.

508 10. Hernandez A, Martinez ME, Fiering S, Galton VA, St Germain D. Type 3 deiodinase is
509 critical for the maturation and function of the thyroid axis. *J Clin Invest.* 2006;116(2):476-484.

510 11. Martinez ME, Charalambous M, Saferali A, et al. Genomic imprinting variations in the
511 mouse type 3 deiodinase gene between tissues and brain regions. *Mol Endocrinol.*
512 2014;28(11):1875-1886.

513 12. Hernandez-Miranda LR, Ruffault PL, Bouvier JC, et al. Genetic identification of a
514 hindbrain nucleus essential for innate vocalization. *Proc Natl Acad Sci U S A.*
515 2017;114(30):8095-8100.

516 13. Zimmerberg B, Rosenthal AJ, Stark AC. Neonatal social isolation alters both maternal
517 and pup behaviors in rats. *Dev Psychobiol.* 2003;42(1):52-63.

518 14. Uematsu A, Kikusui T, Kihara T, et al. Maternal approaches to pup ultrasonic
519 vocalizations produced by a nanocrystalline silicon thermo-acoustic emitter. *Brain Res.*
520 2007;1163:91-99.

521 15. Hahn ME, Lavooy MJ. A review of the methods of studies on infant ultrasound
522 production and maternal retrieval in small rodents. *Behav Genet.* 2005;35(1):31-52.

523 16. Asaba A, Kato M, Koshida N, Kikusui T. Determining Ultrasonic Vocalization
524 Preferences in Mice using a Two-choice Playback Test. *J Vis Exp.* 2015(103).

525 17. Noirot E. Ultrasounds and maternal behavior in small rodents. *Dev Psychobiol.*
526 1972;5(4):371-387.

527 18. Riede T. Subglottal pressure, tracheal airflow, and intrinsic laryngeal muscle activity
528 during rat ultrasound vocalization. *J Neurophysiol.* 2011;106(5):2580-2592.

529 19. Kamaludin AA, Smolarchuk C, Bischof JM, et al. Muscle dysfunction caused by loss of
530 Magel2 in a mouse model of Prader-Willi and Schaaf-Yang syndromes. *Hum Mol Genet.*
531 2016;25(17):3798-3809.

532 20. Zimmer MR, Fonseca AHO, Iyilikci O, Pra RD, Dietrich MO. Functional Ontogeny of
533 Hypothalamic Agrp Neurons in Neonatal Mouse Behaviors. *Cell.* 2019;178(1):44-59 e47.

534 21. Fonseca AH, Santana GM, Bosque Ortiz GM, Bampi S, Dietrich MO. Analysis of
535 ultrasonic vocalizations from mice using computer vision and machine learning. *Elife.* 2021;10.

536 22. Grimsley JM, Monaghan JJ, Wenstrup JJ. Development of social vocalizations in mice.
537 *PloS one.* 2011;6(3):e17460.

538 23. Branchi I, Santucci D, Alleva E. Ultrasonic vocalisation emitted by infant rodents: a tool
539 for assessment of neurobehavioural development. *Behav Brain Res.* 2001;125(1-2):49-56.

540 24. Nakatani J, Tamada K, Hatanaka F, et al. Abnormal behavior in a chromosome-
541 engineered mouse model for human 15q11-13 duplication seen in autism. *Cell.*
542 2009;137(7):1235-1246.

543 25. Takahashi T, Okabe S, Broin PO, et al. Structure and function of neonatal social
544 communication in a genetic mouse model of autism. *Mol Psychiatry.* 2016;21(9):1208-1214.

545 26. Fountain MD, Tao H, Chen CA, Yin J, Schaaf CP. Magel2 knockout mice manifest
546 altered social phenotypes and a deficit in preference for social novelty. *Genes Brain Behav.*
547 2017;16(6):592-600.

548 27. Wilkins JF, Haig D. What good is genomic imprinting: the function of parent-specific
549 gene expression. *Nat Rev Genet.* 2003;4(5):359-368.

550 28. Farrell WJ, Alberts JR. Maternal responsiveness to infant Norway rat (Rattus
551 norvegicus) ultrasonic vocalizations during the maternal behavior cycle and after steroid and
552 experiential induction regimens. *J Comp Psychol.* 2002;116(3):286-296.

553 29. Barnes TD, Rieger MA, Dougherty JD, Holy TE. Group and Individual Variability in
554 Mouse Pup Isolation Calls Recorded on the Same Day Show Stability. *Front Behav Neurosci.*
555 2017;11:243.

30. Verjat A, Rodel HG, Feron C. Isolation calls in house mouse pups: Individual consistency across time and situations. *Dev Psychobiol.* 2019;61(8):1135-1145.

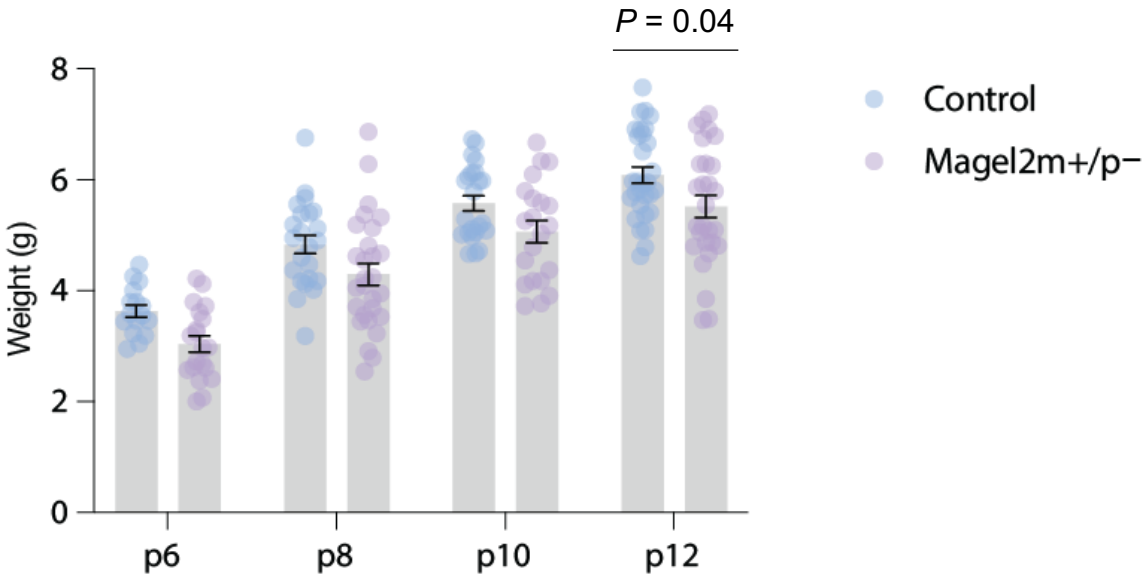


Figure S1. Body weight of pups across tested ages. Individual body weights of control (blue) and *Magel2* deficient (purple) littermates at P6, P8, P10, and P12. Line with error bars represent mean and SEM. Round symbols representing individual values. *P* values are provided in the main text. The sample sizes for control and *Magel2* deficient pups are: P6, n = 16 and 20; P8, n = 23 and 28; P10, n = 23 and 20; and P12, n = 30 and 28, respectively.

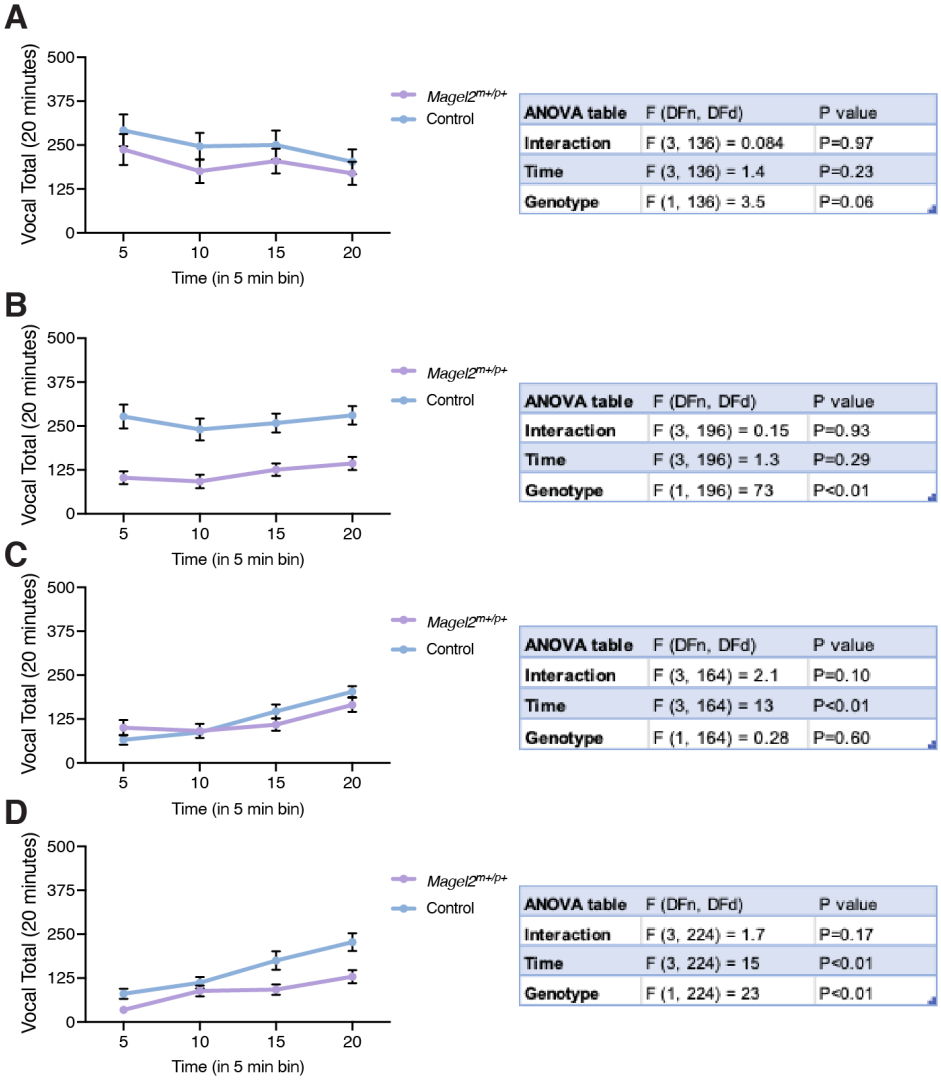


Figure S2. Vocal rate in 5-minute bins across ages. Average USV emission of control (blue) and *Magel2* deficient (purple) littermates at (A) P6, (B) P8, (C) P10, and (D) P12. Connecting lines trace USV means for each genotype across 5-minute bins from the 20-minute isolation recording. Line with error bars represent mean and SEM of weights. Round symbols represent mean values. 2way ANOVA results for each age are found on the tables on the right of each graph. The sample sizes for control and *Magel2* deficient pups are: P6, n = 16 and 20; P8, n = 23 and 28; P10, n = 23 and 20; and P12, n = 30 and 28, respectively.

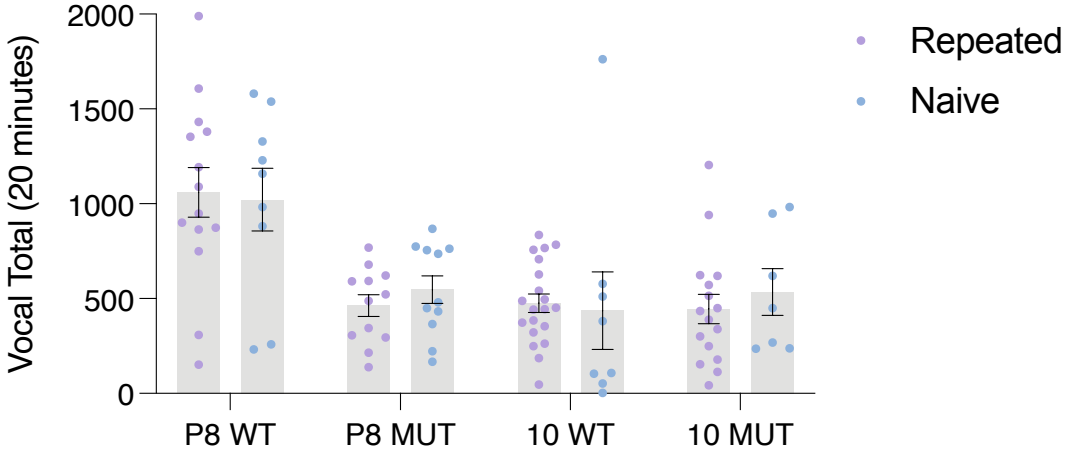


Figure S3. Rate of USVs in naïve versus experienced mouse pups. Average USV emission of pups tested repeatedly (purple) and naïve pups (blue). Bars represent USV means for each genotype and condition. Line with error bars represent mean and SEM. Round symbols represent rate for individual pups. Multiple Mann-Whitney test was used to compare rates. *P* values are the following: P8 WT = 0.36; P8 MUT = 0.46; P10 WT = 0.46; P10 MUT = 0.58.