A comparison of the Avisoft (v.5.2) and MATLAB Mouse Song Analyzer (v.1.3) vocalization analysis systems in C57BL/6, *Fmr1*-FVB.129, NS-*Pten*-FVB, and 129 mice

Matthew Binder[1,†], **Suzanne O. Nolan**[2,†], **Joaquin N. Lugo**[1,3,*]

[1]Department of Psychology and Neuroscience, Vanderbilt University Medical Center, Nashville, TN, 37232, USA

[2]Department of Pharmacology, Vanderbilt University Medical Center, Nashville, TN, 37232, USA

[3]Institute of Biomedical Studies, Baylor University, Waco, TX 76798, USA

Abstract

Background: Communicative behaviors play a vital role in mammals and are highly relevant to human neurodevelopmental conditions. Mice produce communicative vocalizations that occur in the ultrasonic range, which are commonly analyzed with the Avisoft recording system. Fully automated programs such as the Mouse Song Analyzer in MATLAB, have been developed to analyze USVs in a shorter time period, however, no study has compared the accuracy of MATLAB to Avisoft.

New Method: In order to determine MATLAB's accuracy, we used data from four different mouse strains and assessed whether the total number of USVs detected was similar between systems.

Results: We found that there was a high correlation between systems for the number of USVs emitted from C57BL/6 and NS-*Pten* mice however, Avisoft detected significantly more USVs than MATLAB for both strains. For *Fmr1*-FVB.129 and 129 mice, large correlations were observed between systems and no significant difference was present in the USVs detected. A partial correlation was run to control for the covariates: sex, age, strain, and treatment, and found that only strain substantially influences the relationship between the USVs detected in Avisoft and those detected in MATLAB.

[*]Corresponding author Joaquin N. Lugo, PhD, Baylor University, Department of Psychology and Neuroscience, One Bear Place # 97334, Waco, TX 76798, Phone: 254-710-2389; FAX: 254-710-3033, joaquin_lugo@baylor.edu.
[†] =denotes co-first authorship

**Comparison with Existing Method: T**hese findings demonstrate that there is a high degree of agreement between Avisoft and the Mouse Song Analyzer however, Avisoft does detect significantly more USVs depending on the strain assessed.

Conclusions: Therefore, there are relative advantages and disadvantages with both systems that vocalization researchers should be aware of when interpreting USV results, and when using either system.

Graphical abstract
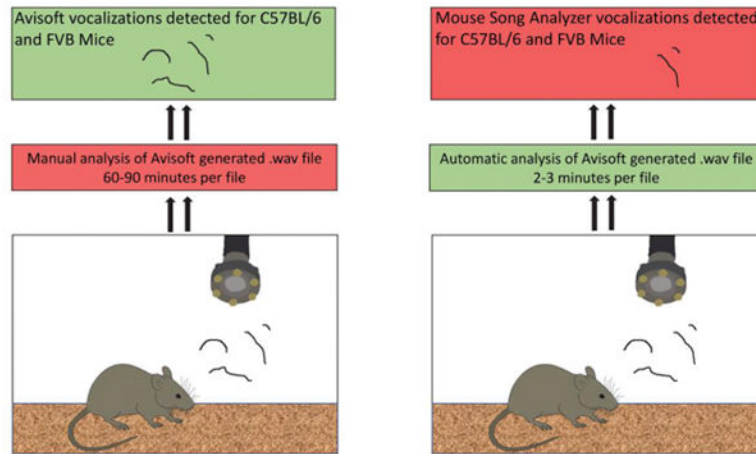
Keywords

Autism spectrum disorder; USV; Epilepsy; Neurodevelopment; Reproducibility

1. Introduction

Communication encompasses a wide a range of behaviors. One aspect of communicative behavior is the production of vocalizations. Vocalizations are emitted in mammals, ranging from mice to humans, and have been shown to convey information about an animals' reproductive status, affective state, and social status, as well as ecologically relevant stimuli such as the presence of food or predators (Barfield and Thomas, 1986; Fendt et al., 2018; Knutson et al., 2002; Moles and D'Amato, 2000; Nyby et al., 1976). In addition to serving basic biological needs, communicative behaviors have also been shown to be a pertinent indicator for the overall health of the animal. Specifically, alterations in communication characterize many neurodevelopmental conditions such as: Autism spectrum disorder (ASD), tuberous sclerosis complex (TSC) and epilepsy in both clinical settings (Esposito et al., 2017; Esposito and Venuti, 2009; McDonald et al., 2017; Viscidi et al., 2014) as well as in preclinical models (Esposito et al., 2017; Esposito and Venuti, 2009; Reynolds et al., 2017; Scattoni et al., 2008; Tsai et al., 2012). Therefore, due to the ubiquitous nature of communication, its importance in mammals, and its relevance to human neurodevelopmental conditions, the study of vocal communication has a diversity of applications and is also highly relevant to human health.

Murine models have been widely used in order to better understand the applications and function of vocal communication. In mice, vocal communication refers to a behavior known as ultrasonic vocalizations (USVs). USVs are aerodynamic whistles that occur between 20–120 kHz and can be readily elicited from both neonatal and adult rodents (Hoffmann et al., 2012; Pasch et al., 2017; Riede, 2011; Riede et al., 2017). Murine vocalizations are commonly recorded with the Avisoft (v.5.2) recording system, which consists of a microphone that detects vocalizations anywhere from 1–125 kHz, and an accompanying software that allows for the analysis of both quantitative and qualitative parameters. While the Avisoft system is comprehensive, it requires a trained experimenter to manually go through each file and remove background noise and ensure that each vocalization is accurately detected. This is a significant time expenditure, as a file of 150–300 USVs takes approximately 60–90 minutes to assess. Therefore, a project with 4 groups and 10 animals per each group would require 2,400 – 3,600 hours to complete. To address this concern, fully automated analysis software has been developed that relies on advanced algorithms to detect vocalizations. One such software program uses the MATLAB computing environment and a freely available MATLAB syntax written by the Jarvis lab called the Mouse Song Analyzer (v.1.3) (http://jarvislab.net/research/mouse-vocal-communication/). Since the Mouse Song Analyzer is automated, it is significantly more time efficient than Avisoft, as a 300 USV file takes approximately 2 minutes to assess. Thus, a project with 4 groups of 10 animals per each group would only require approximately 80 hours to complete, opposed to Avisoft's 2,400. Moreover, the Mouse Song Analyzer is a commonly used tool, which has been utilized in various publications (Arriaga et al., 2012; Chabout et al., 2017; Chabout et al., 2016; Gao et al., 2019; Holy and Guo, 2005). However, it is unknown if the Mouse Song Analyzer's USV count algorithm detects a similar number of USVs as the manually operated Avisoft system.

The lack of comparisons between Avisoft and the Mouse Song Analyzer is significant, as there is little validation of vocalization systems, which can lead to inconsistent results. For instance, a previous study in our lab compared the quantity of USVs detected in Avisoft to those detected in an older recording program known as Ultravox (2.0) and reported significant differences in USV production and weaker than expected overall correlations between the systems under congruent detection parameters (Binder et al., 2018). Therefore, in a separate study, we have demonstrated that two programs under identical detection parameters can detect significantly different quantities of vocalizations. This is notable, as researchers would draw opposing conclusions from the same data set depending on which system was used in analysis. In light of this, and in order to continue to assess the reproducibility of findings between vocalization programs, the current study directly compared the Avisoft analysis system to the Mouse Song Analyzer's analysis system to assess whether the total vocalizations detected in one program are similar to the total vocalizations detected in the other. Any significant differences in the total vocalizations detected could indicate the need for a reappraisal of previous findings.

2. Materials and methods

2.1 Animals and housing

Data from several projects using C57BL/6, FVB.129, FVB, and 129 mice were used in the present study to evaluate the relative accuracy of the Mouse Song Analyzer's automated USV count algorithm to the manually operated Avisoft program (Binder et al., 2020; Binder and Lugo, 2017; Hodges et al., 2017; Nolan et al., 2019). These specific strains were chosen since they comprise 52% of the most popular mouse strains purchased from Jackson labs, increasing our study's generalizability (Jackson Lab). For specific methods for each project, see individual references. From our previous paper examining high seizure load and USVs on postnatal day (PD) 12 in the C57BL/6J strain, the following numbers were used: male control = 15, female control = 18, male seizure = 17, female seizure = 20 (Nolan et al., 2019). From our previous paper examining the production of neonatal vocalizations in the neuronal subset-specific phosphatase and tensin homolog (NS-*Pten*) knockout mouse on a FVB based background, two timepoints were used (PD 8 and PD 11) and the sample sizes were as follows: male wildtype (WT) = 14, female WT = 15, male knockout (KO) = 12, female KO = 13 (Binder and Lugo, 2017). From our recent study examining vocalizations on PD 8 in *Fmr1*-FVB.129 mice, we had the following numbers: male WT = 17, female WT = 12, male KO = 12, female KO = 12 (Nolan et al, *under review*). From an unpublished dataset examining vocalizations on PD 12 in the 129 background strain after multiple flurothyl seizures across PD 7–11, the respective sample sizes were: male control = 3, female control = 6, male seizure = 6, female seizure = 5. In addition to neonatal USVs, adult USVs were also included, however, due to the nature of the paradigm, only adult male mice were assessed. Specifically, for our previous study examining vocalizations in *Fmr1*-FVB.129 adult mice, we had 19 WT males and 18 KO males (Hodges et al., 2017). From a recent study using this same mating-based paradigm in NS-*Pten* adult mice, we had 14 WT males and 12 KO males (Binder et al., 2020). Altogether, this gave us a final *n* of 314 data files. Across all experiments, animals had *ad libitum* access to food and water. All procedures performed were in accordance with Baylor University's Institutional Animal Care and Use Committee, as well as the National Institutes of Health Guidelines for the Care and Use of Laboratory Animals. All testing was conducted during the light cycle, specifically between 8 am and 5 pm.

2.2 Isolation-induced ultrasonic vocalizations

For projects where subjects were assessed prior to PD 14, vocalizations were recorded using the isolation-induced ultrasonic vocalization paradigm that has been previously described (Scattoni et al., 2008; Shair, 2007). On the day of testing, all of the test pups were brought down from the colony room in their home cages and allowed to habituate to the testing room for 30 minutes prior to the testing period. At the time of testing, pups were separated from their parents and placed into a clean housing cage, warmed to an ambient nesting temperature of approximately 35° C. The pups were then individually removed from the holding cage and placed into a clean cage contained within a 40 cm × 40 cm × 30 cm acrylic, sound-attenuated chamber, which ensured that vocalizations produced by any animal other than the test animal would not be detected by the recording microphone. Vocalizations were recorded using a condenser microphone (CM16/CMPA, Avisoft Bioacoustics,

Germany) that was connected to a recording interface (UltraSoundGate 116Hb, Avisoft Bioacoustics). The microphone is broad spectrum and records USVs from 1–125 kHz. Vocalizations were recorded for a 2-minute duration, after which the pup was removed from the test chamber. Once all mice were tested, they were returned to their parents in the home cage.

2.3 Adult female-urine induced ultrasonic vocalizations

The methods for adult USV collection have been previously described (Wöhr et al., 2011). Briefly, one week prior to testing, male mice were paired with females for a 5-minute duration. This was done to standardize the social experience in males, a necessary prerequisite for eliciting adult vocalizations (Scattoni et al., 2008). Twenty-four hours prior to testing, estrus was induced in females by placing male bedding into the female cage that had been previously introduced to the test males. On the day of testing, fresh urine was collected from females that were in estrus and 20 uL of the urine was then pipetted into the center of a testing chamber. The chamber was sound attenuating to ensure an accurate detection of the test mouse's USVs. Males were placed in the chamber and all vocalizations were recorded for a 5-minute duration using a condenser microphone (CM16/CMPA, Avisoft Bioacoustics, Germany) coupled with a recording interface (UltraSoundGate 116Hb, Avisoft Bioacoustics). After testing, the males were returned to their home cages.

2.4 Avisoft analysis

All ultrasonic vocalizations collected were analyzed using Avisoft SASLab Pro software (Avisoft SASLabPro, RRID:SCR_014438). The parameters consisted of: a fast Fourier transformation (FFT) length of 1024, time window overlap of 75%, a 100% hamming window, a time resolution of 1 ms, and a sampling frequency of 250 kHz, per established protocol (Scattoni et al., 2008). Files were cleaned by a trained experimenter that was blinded to the treatment condition and identity of the animal. Each file was manually assessed, and each vocalization was identified using the Avisoft software. Any extraneous noise was manually removed from the spectrogram.

2.5 MATLAB Mouse Song Analyzer analysis

The MATLAB Mouse Song Analyzer analysis procedure has been previously described (Arriaga et al., 2012; Chabout et al., 2017). Briefly, all .wav files were converted to sonograms using freely available MATLAB syntax developed by the Jarvis lab (http://jarvislab.net/research/mouse-vocal-communication/). Once the files were converted, the sonograms were processed using the graphical user interface with a density inter-syllable interval (ISI) of 0.25. The minimum frequency was set to 15,000 Hz, the max frequency to 125,000 Hz, the sampling frequency to 250 kHz, and the threshold to 0.3. The parameters used were the same as those used in previous studies and are comparable to Avisoft's USV count parameters (Arriaga et al., 2012; Chabout et al., 2017). All other parameters assessed can be found in Chabout, 2017.

2.6 Statistical analysis

All data was analyzed using IBM SPSS Statistics 21.0 (IBM, USA) or GraphPad Prism 7 software (La Jolla, CA). To determine the accuracy of the Mouse Song Analyzer's automated processing to the manually operated Avisoft program, correlations for each strain assessed were performed. Additionally, to examine if Avisoft and the Mouse Song Analyzer detected a significantly different quantity of calls from one another, independent *t*-tests or nonparametric Mann-Whitney *U* tests (when homogeneity of variance was violated) per each strain were run. A partial correlation analysis was then conducted to control for the following covariates: sex (male, female), strain (C57BL/6, *Fmr1*-FVB.129, NS-*Pten*-FVB, 129), age (PD 8,11,12,49, 60), and treatment (genotype or seizure status). A correlation that did not control for covariates (a zero-order correlation) was also calculated to determine the influence the covariates had on the relationship in USV detection between Avisoft and the Mouse Song Analyzer. A value of $p < .05$ was considered significant for each statistical test and the figures depict the mean ± standard error of the mean (SEM). The effect sizes for the correlations are as follows: small effect $r = 0.1$, medium effect $r = 0.3$, large effect $r = 0.5$.

3. Results

3.1 Association and number of USVs detected between Avisoft and the Mouse Song Analyzer

For the C57/BL6 strain, there was a large correlation between the total number of USVs detected by the Mouse Song Analyzer and the total USVs detected by Avisoft ($r(69) = .71$, $p < .001$) (Figure 1A). A Mann-Whitney *U* test was then used to assess if one program detected a significantly greater quantity of USVs than the other and found that Avisoft detected significantly more vocalizations than the Mouse Song Analyzer ($U = 640$, $p < .001$) (Figure 1B). For the *Fmr1*-FVB.129 model, there was a large correlation in the USVs detected between the two programs ($r(88) = .95$, $p < .001$) (Figure 1C). There was no significant difference in the quantity of vocalizations detected between Avisoft and the Mouse Song Analyzer ($t(176) = .99$, $p = .32$) (Figure 1D). In the NS-*Pten*-FVB model, there was a large correlation between the Mouse Song Analyzer and Avisoft for the total USVs that were detected ($r(132) = .61$, $p < .001$) (Figure 1E). However, upon analysis with a Mann-Whitney *U* test, Avisoft was found to detect significantly more vocalizations than the Mouse Song Analyzer ($U = 7539$, $p < .001$) (Figure 1F). In the 129 strain, there was again a large correlation present for the quantity of USVs detected in the Mouse Song Analyzer and those detected in Avisoft ($r(19) = .99$, $p < .001$) (Figure 1G). No significant difference in USV detection was found between the two systems ($t(38) = .30$, $p = .77$) (Figure 1H).

3.2 Correlation controlling for variables of assessment

A partial correlation was performed to determine the strength of the relationship between the number of USVs detected by the two systems whilst controlling for sex (male, female) (Figure 2A), strain (C57BL/6, *Fmr1*-FVB.129, NS-*Pten*-FVB,129) (Figure 2B) age of the animal (PD8,11,12,49,60) (Figure 2C) and treatment (genotype or seizure status) (Figure 2D). There was a large positive correlation between the number of USVs detected across the two systems, which was statistically significant ($r(307) = 0.69$, $p < 0.001$). The zero-order correlation also showed that there was a large statistically significant correlation between the

number of USVs detected across the two systems (r(310) = 0.66, p < 0.001), indicating that these covariates had little influence in controlling for the relationship between the two systems (Figure 2E).

Due to our earlier comparisons, which found a statistically significant difference in the number of USVs detected between systems for the C57BL/6 and NS-*Pten*-FVB strains, another partial correlation was run with the strain variable only including the C57BL/6 and NS-*Pten*-FVB mice. This was done to assess whether the extremely strong correlations between programs for 129 mice were biasing the earlier partial correlation and masking a weaker relationship between systems for the C57BL/6 and NS-*Pten*-FVB mice. Upon running this analysis, we observed a large positive correlation between systems while controlling for sex (Figure 3A), strain (C57BL/6, NS-*Pten*-FVB) (Figure 3B), age (Figure 3C), and treatment (Figure 3D) (r(201) = 0.66, p < 0.001). However, the zero-order correlation was notably weaker (r(201) = 0.54, p < 0.001) with strain accounting for .36 of the variance (Figure 3E). Therefore, while the sex, age, and treatment covariates have little influence on USV detection between systems, the strain of the mouse does substantially influence the relationship between USVs detected in Avisoft and those detected in the Mouse Song Analyzer.

4. Discussion

The present study compared the Avisoft analysis system to the MATLAB Mouse Song Analyzer analysis system in order to assess the reproducibility of findings between programs. The USV detection parameters (sampling frequency, high and low cut offs etc.) were congruent between systems in order to ensure a fair comparison. The C57BL/6, *Fmr1*-FVB.129, NS-*Pten*-FVB, and 129 strains were assessed since they are commonly used, comprising 52% of the most popular mouse strains purchasable from Jackson labs (Jackson Lab). Upon assessment, large correlations were found between systems for each strain. However, accompanying tests did indicate that Avisoft selectively detected significantly more USVs than the Mouse Song Analyzer. The strain of mouse assessed was shown to significantly influence the relationship between recording systems, while the variables of sex, age, and treatment did not significantly affect USV detection for Avisoft or the Mouse Song Analyzer.

Although large correlations between Avisoft and the Mouse Song Analyzer were detected for the total quantity of USVs emitted in C57BL/6 and NS-*Pten*-FVB mice, the correlations were lower than expected, as they were .71 and .60, respectively. This is less than ideal, since the standard in the literature for two systems under analogous detection parameters is a correlation between .9–1 (McHugh, 2012). Notably, data that is less highly correlated may lead to USV findings that are not consistent across analysis programs. Our study supports this, as we found that Avisoft detected significantly more USVs than the Mouse Song Analyzer for both C57BL/6 and NS-*Pten*-FVB mice. Additionally, upon further assessment with a partial correlation, these two strains were shown to substantially influence the relationship between Avisoft and the Mouse Song Analyzer for USV detection. Altogether, this indicates that results attained in one system for either the C57BL/6 or NS-*Pten*-FVB mouse strains may not be reproducible in the other system.

Interestingly, while lower than expected correlations were observed for C57BL/6 and NS-*Pten*-FVB mice, this was not the case for the *Fmr1*-FVB.129 and 129 mice. Each of these strains exhibited correlations greater than .9 (.95 and .99 respectively). Furthermore, there was no statistically significant difference in the total quantity of USVs detected between programs and neither strain was shown to substantially influence the relationship between Avisoft and the Mouse Song Analyzer. Therefore, for the 129 background strain, results appear to be reproducible across systems.

The selective variability in USV detection between Avisoft and the Mouse Song Analyzer may best be attributed to the cut off parameters of the automated software. Specifically, for C57BL/6 and NS-*Pten*-FVB mice, the MATLAB algorithm that breaks down clusters of USVs may do so in a manner that occasionally counts 2 USVs that occur close together in time as one larger USV. Since Avisoft requires an operator to manually examine each vocalization, such a potential pitfall could be corrected or outright avoided, resulting in a larger total USV count for Avisoft. Specifically, the Avisoft program avoids the problem of USV clustering by rendering a two-dimensional representation of each vocalization in addition to transducing USVs to put them into the human auditory spectrum. Therefore, Avisoft provides both visual and auditory cues to the user in order to ensure a valid and accurate detection of the USVs. However, for the 129 strain, it is possible that these mice simply vocalize in a pattern that better conforms to the automated detection algorithm, perhaps displaying a larger interval of time between vocalizations than either the C57BL/6 or NS-*Pten*-FVB strains, thereby avoiding any discrepancies in USV detection. Furthermore, a previous study by Wiaderkiewicz (2013) found that 129 mice emit fewer clusters of USVs than C57/BL6 mice. While this difference was non-significant, any reduction in the clusters of vocalizations may decrease the complexity of USV detection, and in turn better conform to the Mouse Song Analyzer's algorithm (Wiaderkiewicz et al., 2013). Altogether, several factors may contribute to the strain-specific differences in vocalization detection between the Mouse Song Analyzer and Avisoft. Due to the growing prominence of automatic detection software, the elucidation of potential factors that may bias automatic USV detection represents a compelling future direction for further study.

Previous studies have found that a significant amount of published scientific literature is unable to be reproduced. Specifically, past studies assessed 170 papers published in psychology, behavioral, or medical journals and found that only 36–61% percent of the assessed studies were able to be replicated (Camerer et al., 2018; Ioannidis, 2005; OCS, 2015). Due to the cumulative nature of science, any inability to reproduce prior findings can have a long-lasting impact for basic, translational, and clinical research. Failures in replication have spurred the NIH to initiate a charge to enhance rigor and reproducibility in science (NIH). While our study takes a valuable step towards fulfilling this charge, more studies need to follow in order to ensure consistent results in vocal communication research, as there are numerous USV analysis programs available that are often not externally validated. Thus, future studies could continue to compare Avisoft and the Mouse Song Analyzer across other commonly used mouse strains, as our study indicates that strain is a more important variable in USV detection than the age, sex, or treatment (knockout, or seizure status) of the mice. Future investigations could also compare the reliability of automatic and manual programs when assessing the spectral and temporal characteristics of

USVs, as well as the types of USVs being emitted, as these parameters have also been shown to be relevant to communication and neurodevelopmental disorders (Binder and Lugo, 2017; Esposito et al., 2017; Esposito and Venuti, 2009; Esposito and Venuti, 2010; Scattoni et al., 2008). Additionally, studies comparing Avisoft to other recently developed USV analysis programs such as the DeepSqueak or MUPET, automated and semi-automated programs, would provide pertinent points of comparison as well as further validation, and could help researchers make more informed decisions when selecting data extraction methods (Coffey et al., 2019; Van Segbroeck et al., 2017). These studies could also be extended to include rats as well, since any inconsistency between USV analysis programs would similarly affect rat communication research. Overall, this process of internal validation is of the utmost necessity and will allow for the reproducibility of vocalization results to be independent from the software used in analysis.

4.1 Conclusion

Altogether, large correlations between Avisoft and the Mouse Song Analyzer were observed for each strain assessed. Furthermore, there were near perfect correlations between programs for the 129 mice indicating a high degree of reproducibility when USV counts are low. However, for the C57BL/6 and NS-*Pten*-FVB strains, which emit a greater quantity of USVs, Avisoft was shown to detect significantly more vocalizations than the Mouse Song Analyzer, which indicates that some findings in these strains may not be reproducible with both programs. Therefore, while large correlations between systems were present, the overall data indicates that Avisoft is superior to the Mouse Song Analyzer when assessing the total USV count, particularly when USV counts are high, which may be attributed to different cut off parameters between systems. However, while Avisoft is selectively more sensitive, it also comes at a significant time expenditure that is largely avoided in the Mouse Song Analyzer program. Therefore, both systems have relative advantages and disadvantages that should be taken into consideration by researchers in the vocalization field before either program is utilized or results interpreted.

Acknowledgements

References

Jackson Labs. https://www.jax.org/jax-mice-and-services/find-and-order-jax-mice/most-popularjax-mice-strains accessed 5/21/2020.

Arriaga G, Zhou EP, Jarvis ED. Of Mice, Birds, and Men: The Mouse Ultrasonic Song System Has Some Features Similar to Humans and Song-Learning Birds. PLOS ONE, 2012; 7: e46610.

Barfield RJ, Thomas DA. The Role of Ultrasonic Vocalizations in the Regulation of Reproduction in Ratsa. Annals of the New York Academy of Sciences, 1986; 474: 33–43. [PubMed: 3555231]

Binder MS, Hernandez-Zegada CJ, Potter CT, Nolan SO, Lugo JN. A comparison of the Avisoft (5.2) and Ultravox (2.0) recording systems: Implications for early-life communication and vocalization research. Journal of Neuroscience Methods, 2018; 309: 6–12. [PubMed: 30118724]

Binder MS, Jones DG, Hodges SL, Lugo JN. NS-Pten adult knockout mice display both quantitative and qualitative changes in urine-induced ultrasonic vocalizations. Behavioural Brain Research, 2020; 378: 112189.

Binder MS, Lugo JN. NS-Pten knockout mice show sex- and age-specific differences in ultrasonic vocalizations. Brain and behavior, 2017; 7: e00857-e.

Camerer CF, Dreber A, Holzmeister F, Ho T-H, Huber J, Johannesson M, Kirchler M, Nave G, Nosek BA, Pfeiffer T, Altmejd A, Buttrick N, Chan T, Chen Y, Forsell E, Gampa A, Heikensten E, Hummer L, Imai T, Isaksson S, Manfredi D, Rose J, Wagenmakers E-J, Wu H. Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. Nature Human Behaviour, 2018; 2: 637–44.

Chabout J, Jones-Macopson J, Jarvis ED. Eliciting and analyzing male mouse ultrasonic vocalization (USV) songs. Journal Of Visualized Experiments, 2017.

Chabout J, Sarkar A, Patel SR, Radden T, Dunson DB, Fisher SE, Jarvis ED. A Foxp2 Mutation Implicated in Human Speech Deficits Alters Sequencing of Ultrasonic Vocalizations in Adult Male Mice. Front Behav Neurosci, 2016; 10.

Coffey KR, Marx RG, Neumaier JF. DeepSqueak: a deep learning-based system for detection and analysis of ultrasonic vocalizations. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology, 2019; 44: 859–68. [PubMed: 30610191]

Esposito, Hiroi, Scattoni. Cry, Baby, Cry: Expression of Distress As a Biomarker and Modulator in Autism Spectrum Disorder. International Journal of Neuropsychopharmacology, 2017; 20: 498–503.

Esposito, Venuti. Comparative Analysis of Crying in Children with Autism, Developmental Delays, and Typical Development. Focus on Autism and Other Developmental Disabilities, 2009; 24: 240–7.

Esposito G, Venuti P. Developmental changes in the fundamental frequency (f0) of infants' cries: A study of children with Autism Spectrum Disorder. Early Child Development and Care, 2010; 180: 1093–102.

Fendt M, Brosch M, Wernecke KEA, Willadsen M, Wöhr M. Predator odour but not TMT induces 22-kHz ultrasonic vocalizations in rats that lead to defensive behaviours in conspecifics upon replay. Scientific Reports, 2018; 8: 11041. [PubMed: 30038341]

Gao S-C, Wei Y-C, Wang S-R, Xu X-H. Medial Preoptic Area Modulates Courtship Ultrasonic Vocalization in Adult Male Mice. Neuroscience Bulletin, 2019; 35: 697–708. [PubMed: 30900143]

Hodges SL, Nolan SO, Reynolds CD, Lugo JN. Spectral and temporal properties of calls reveal deficits in ultrasonic vocalizations of adult Fmr1 knockout mice. Behavioural Brain Research, 2017; 332: 50–8. [PubMed: 28552599]

Hoffmann F, Musolf K, Penn DJ. Ultrasonic courtship vocalizations in wild house mice: spectrographic analyses. Journal of ethology, 2012; 30: 173–80.

Holy TE, Guo Z. Ultrasonic Songs of Male Mice. PLOS Biology, 2005; 3: e386.

Ioannidis JPA. Contradicted and Initially Stronger Effects in Highly Cited Clinical Research. Jama, 2005; 294: 218–28. [PubMed: 16014596]

Knutson B, Burgdorf J, Panksepp J. Ultrasonic vocalizations as indices of affective states in rats. Psychological Bulletin, 2002; 128: 961–77. [PubMed: 12405139]

McDonald NM, Varcin KJ, Bhatt R, Wu JY, Sahin M, Nelson CA 3rd, Jeste SS. Early autism symptoms in infants with tuberous sclerosis complex. Autism research : official journal of the International Society for Autism Research, 2017.

McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb), 2012; 22: 276–82. [PubMed: 23092060]

Moles A, D'Amato FR. Ultrasonic vocalization by female mice in the presence of a conspecific carrying food cues. Animal Behaviour, 2000; 60: 689–94. [PubMed: 11082239]

NIH. National Institutes of Health. Rigor and Reproducibility. https://www.nih.gov/researchtraining/rigor-reproducibility accessed 5/23/2018.

Nolan SO, Hodges SL, Condon SM, Muhammed IDA, Tomac LA, Binder MS, Reynolds CD, Lugo JN. High seizure load during sensitive periods of development leads to broad shifts in ultrasonic

vocalization behavior in neonatal male and female C57BL/6J mice. Epilepsy & Behavior, 2019; 95: 26–33. [PubMed: 31022661]

Nyby J, Dizinno GA, Whitney G. Social status and ultrasonic vocalizations of male mice. Behavioral Biology, 1976; 18: 285–9. [PubMed: 999582]

OCS. Open Science Collaboration. Estimating the reproducibility of psychological science. Science, 2015; 349: aac4716.

Pasch B, Tokuda IT, Riede T. Grasshopper mice employ distinct vocal production mechanisms in different social contexts. Proc Biol Sci, 2017; 284: 20171158.

Reynolds, Nolan, Huebschman, Hodges, Lugo. Early-life status epilepticus acutely impacts select quantitative and qualitative features of neonatal vocalization behavior: Spectrographic and temporal characterizations in C57BL/6 mice. Epilepsy & behavior : E&B, 2017; 72: 58–62.

Riede T. Subglottal pressure, tracheal airflow, and intrinsic laryngeal muscle activity during rat ultrasound vocalization. Journal of neurophysiology, 2011; 106: 2580–92. [PubMed: 21832032]

Riede T, Borgard HL, Pasch B. Laryngeal airway reconstruction indicates that rodent ultrasonic vocalizations are produced by an edge-tone mechanism. R Soc Open Sci, 2017; 4: 170976-.

Scattoni ML, Gandhy SU, Ricceri L, Crawley JN. Unusual Repertoire of Vocalizations in the BTBR T +tf/J Mouse Model of Autism. PLoS ONE, 2008; 3: e3067.

Shair HN. Acquisition and expression of a socially mediated separation response. Behavioural Brain Research, 2007; 182: 180–92. [PubMed: 17379325]

Tsai PT, Hull C, Chu Y, Greene-Colozzi E, Sadowski AR, Leech JM, Steinberg J, Crawley JN, Regehr WG, Sahin M. Autistic-like behaviour and cerebellar dysfunction in Purkinje cell Tsc1 mutant mice. Nature, 2012; 488: 647–51. [PubMed: 22763451]

Van Segbroeck M, Knoll AT, Levitt P, Narayanan S. MUPET-Mouse Ultrasonic Profile ExTraction: A Signal Processing Tool for Rapid and Unsupervised Analysis of Ultrasonic Vocalizations. Neuron, 2017; 94: 465–85.e5. [PubMed: 28472651]

Viscidi EW, Johnson AL, Spence SJ, Buka SL, Morrow EM, Triche EW. The association between epilepsy and autism symptoms and maladaptive behaviors in children with autism spectrum disorder. Autism : the international journal of research and practice, 2014; 18: 996–1006. [PubMed: 24165273]

Wiaderkiewicz J, Głowacka M, Grabowska M, Barski J-J. Ultrasonic vocalizations (USV) in the three standard laboratory mouse strains: developmental analysis. Acta Neurobiol Exp, 2013; 73: 557–63.

Wöhr, Roullet FI, Crawley JN. Reduced scent marking and ultrasonic vocalizations in the BTBR T +tf/J mouse model of autism. Genes, Brain and Behavior, 2011; 10: 35–43.

Highlights

- There were significant strain-specific differences in USV detection across systems

- Avisoft requires more time but detects more USVs than the Mouse Song Analyzer

- The Mouse Song Analyzer takes less time but detects fewer vocalizations

- There are relative advantages and disadvantages with both programs

- Not all vocalization findings may be reproducible across systems
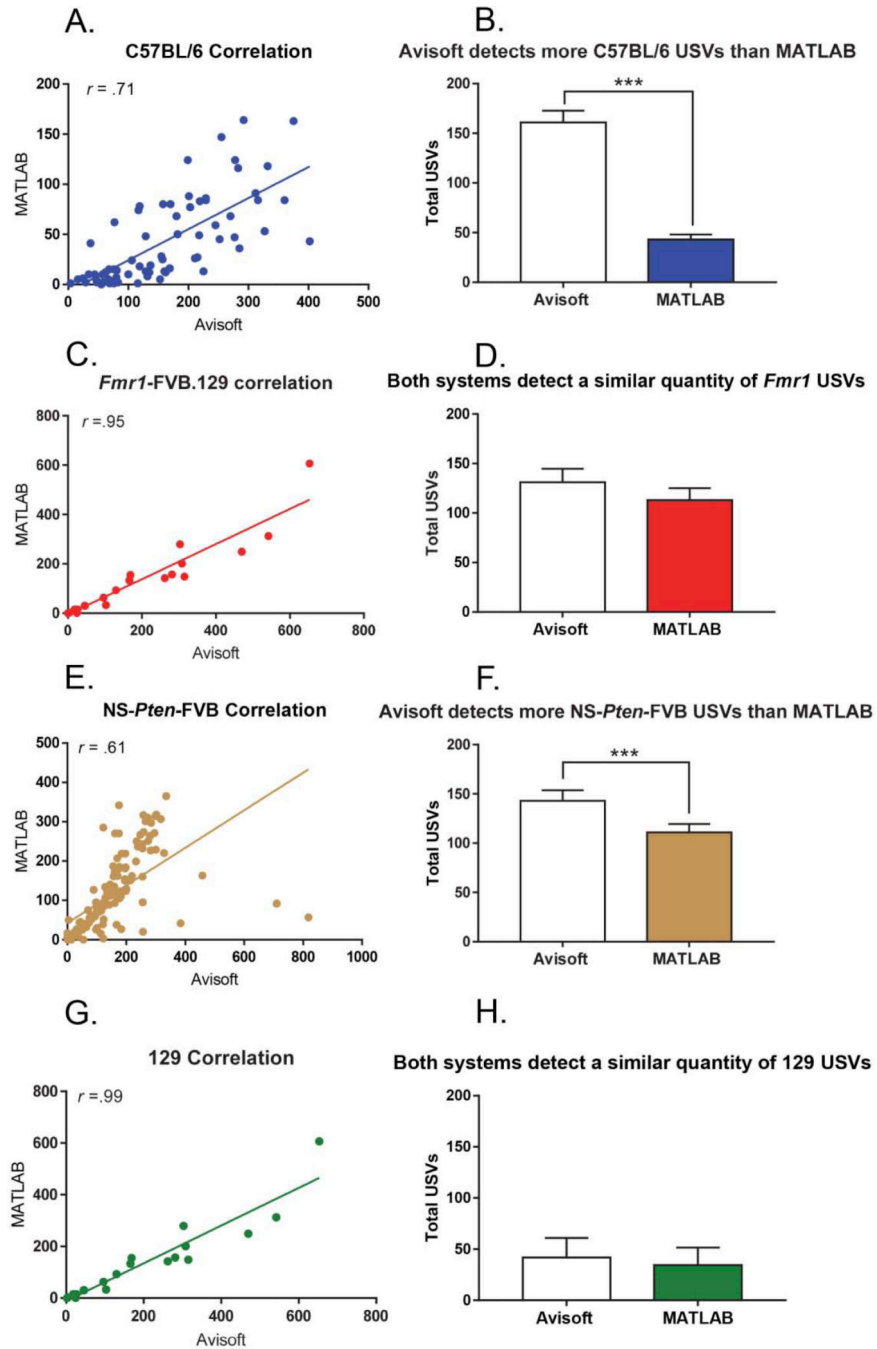
Figure 1. The Avisoft system detects more vocalizations in the C57BL/6J and NS-*Pten*-FVB background strains, but not the *Fmr1*-FVB.129 and 129 background strains, when compared to the Mouse Song Analyzer.

(A). There is a large correlation between systems for C57BL/6 mice. (B). Avisoft detected significantly more USVs than the Mouse Song Analyzer for C57BL/6 mice. (C). There is a large correlation between systems for *Fmr1*-FVB.129 mice. (D). There is no statistically significant difference in the quantity of calls detected between Avisoft and the Mouse Song Analyzer for *Fmr1*-FVB.129 mice. (E). There is a large correlation between systems for the NS-*Pten*-FVB mice. (F) Avisoft detected significantly more USVs than the Mouse Song

Analyzer for NS-*Pten*-FVB mice. (G). There is a large correlation between systems for 129 mice. (H). There is no statistically significant difference in the quantity of calls detected between Avisoft and the Mouse Song Analyzer for 129 mice. The data points represent the mean and the error bars represent the standard error of the mean. $* = p < .05$, $** = p < .01$, $*** = p < .001$.
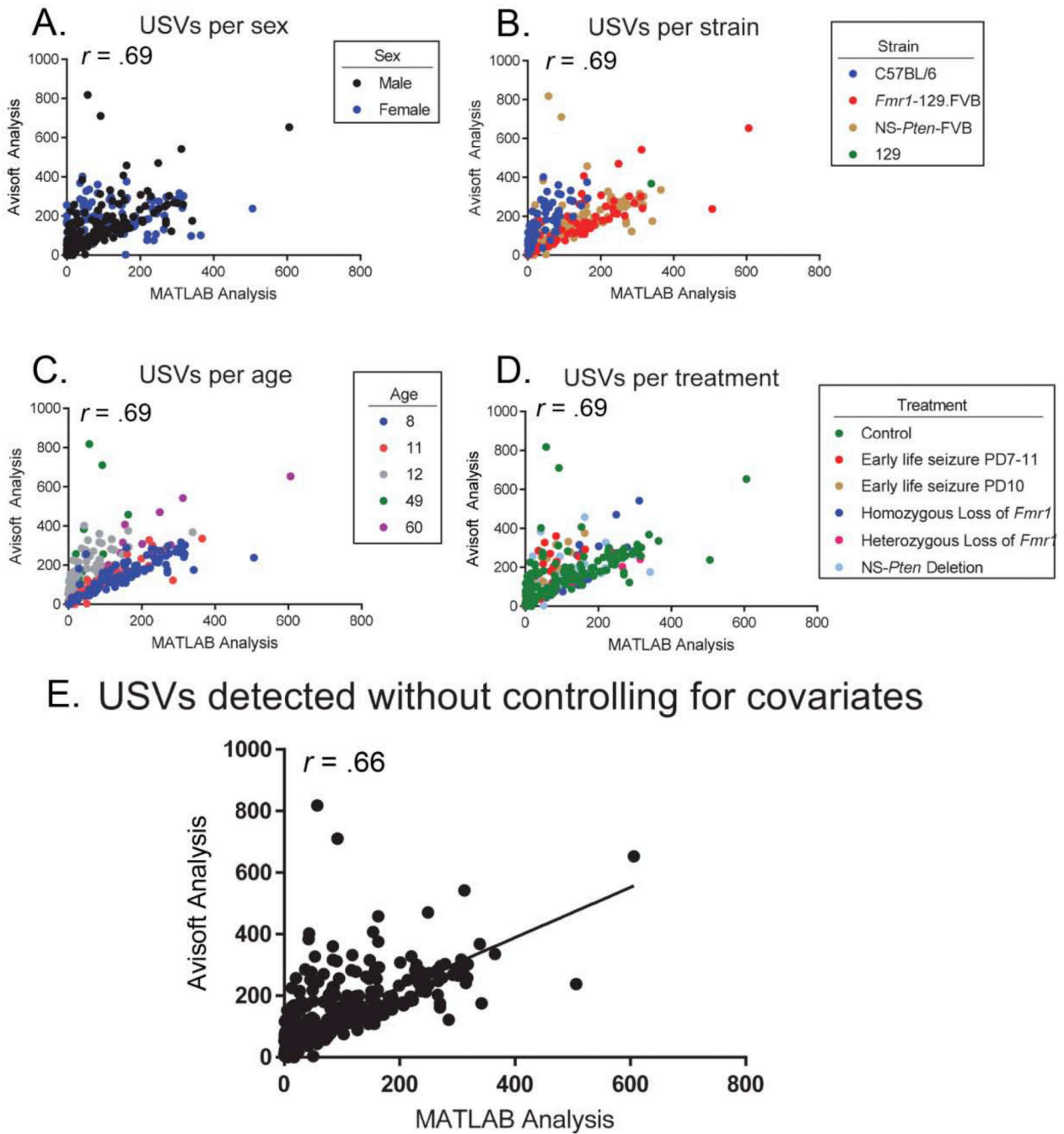
Figure 2. Overall, sex, strain, age, and treatment do not moderate the relationship between vocalizations detected by the MATLAB and Avisoft systems.

(A-D). Upon controlling for the variables of sex, strain, age, and treatment, a large correlation of $r = .69$ was found. (E). The zero-order correlation of .66 indicates that none of these covariates substantially influence the relationship between USVs detected in Avisoft and those detected in the Mouse Song Analyzer ($p < .05$).
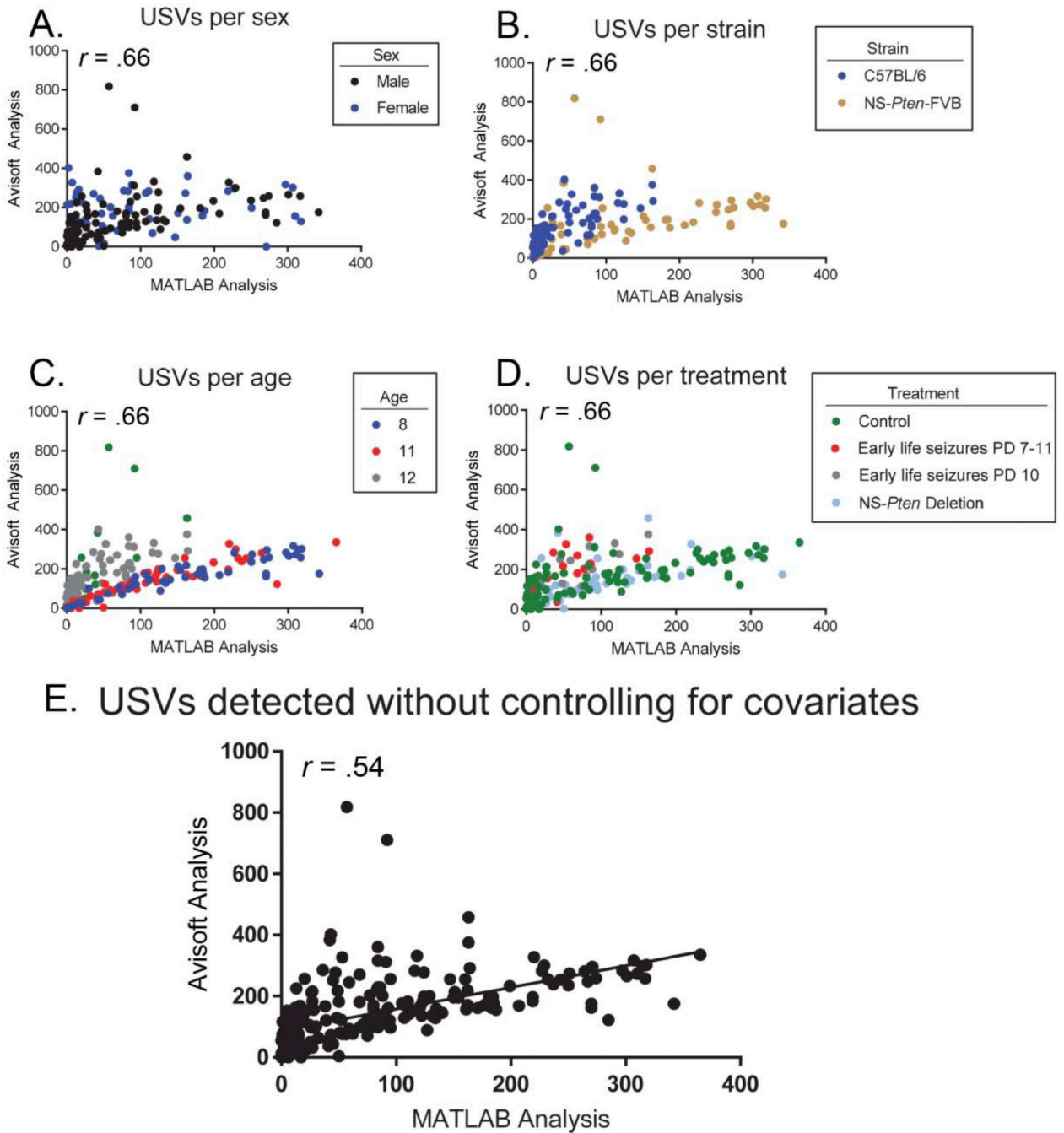
Figure 3. When assessing the C57BL/6 and NS-*Pten*-FVB strains, strain substantially moderates the relationship between Avisoft and MATLAB but sex, age, and treatment do not.

(A-D). Upon controlling for the variables of sex, strain, age, and treatment in the C57BL/6 and NS-*Pten*-FVB strains, a large correlation of $r = .66$ was found. (E). The zero-order correlation of .54 revealed that strain substantially influences the relationship between USVs detected in Avisoft and those detected in the Mouse Song Analyzer, whereas the sex, age, and treatment variables do not substantially influence this relationship ($p < .05$).