American Journal on Intellectual and Developmental Disabilities

Genetic subtypes of Phelan-McDermid syndrome exhibit similar rates of change despite differences in level of impairment in developmental constructs

--Manuscript Draft--

Manuscript Number:	AJIDD-D-24-00006R1
Article Type:	Invitation Only - Developmental Synaptopathies
Keywords:	Phelan-McDermid syndrome; genotype-phenotype associations; natural history; shank3; 22q13
Corresponding Author:	Alexander Kolevzon, MD Icahn School of Medicine at Mount Sinai New York, New York UNITED STATES
First Author:	Tess Levy
Order of Authors:	Tess Levy
	Cristan Farmer, PhD
	Siddharth Srivastava, MD
	Kristina Johnson, PhD
	Jadyn Trayvick, BS
	Camille Brune, PhD
	Alexandra Massa, BS
	Hailey Silver, BS
	Paige M. Siper, PhD
	Jessica Zweifach, PhD
	Danielle Halpern, PsyD
	Jennifer H. Foss-Feig, PhD
	Jonathan A. Bernstein, MD
	Elizabeth Berry-Kravis, MD
	Craig M. Powell, MD, PhD
	Mustafa Sahin, MD, PhD
	Latha Valluripalli Soorya, PhD
	Audrey Thurm, PhD
	Joseph D. Buxbaum, PhD
	Alexander Kolevzon, MD
Manuscript Region of Origin:	UNITED STATES
Abstract:	The clinical spectrum of Phelan-McDermid syndrome (PMS) is varied, with wide-ranging degrees of intellectual disability, developmental delays, behavioral abnormalities, and medical features. Different types of genetic variation lead to PMS, and differing genotypes (e.g., size of deletion or type of sequence variant) account for some of this variability, with strong associations between genotype and phenotype observed with degree of intellectual disability, and presence of specific medical features such as renal abnormalities. To date, no studies have assessed how genotype is associated with the natural history of developmental or behavioral features in PMS over time. Here, we report on longitudinal data in developmental and behavioral domains from 154 individuals with PMS, comparing those with Class 1 (minimal) deletions, Class 2 deletions, and sequence variants, assessing both within-subject (individual change over time) and between-subject (across age) differences.

	Consistent with previous results, average scores per group differed in most adaptive and developmental domains, with individuals with Class 1 deletions performing best, followed by individuals with Class 2 deletions and sequence variants, who often performed similarly. However, in most domains of adaptive behavior, intellectual functioning, and behavioral features, genetic groups did not differ in their rate of ...[see text file]

Abstract
The clinical spectrum of Phelan-McDermid syndrome (PMS) is varied, with wide-ranging degrees of intellectual disability, developmental delays, behavioral abnormalities, and medical features. Different types of genetic variation lead to PMS, and differing genotypes (e.g., size of deletion or type of sequence variant) account for some of this variability, with strong associations between genotype and phenotype observed with degree of intellectual disability, and presence of specific medical features such as renal abnormalities. To date, no studies have assessed how genotype is associated with the natural history of developmental or behavioral features in PMS over time. Here, we report on longitudinal data in developmental and behavioral domains from 154 individuals with PMS, comparing those with Class 1 (minimal) deletions, Class 2 deletions, and sequence variants, assessing both within-subject (individual change over time) and between-subject (across age) differences. Consistent with previous results, average scores per group differed in most adaptive and developmental domains, with individuals with Class 1 deletions performing best, followed by individuals with Class 2 deletions and sequence variants, who often performed similarly. However, in most domains of adaptive behavior, intellectual functioning, and behavioral features, genetic groups did not differ in their rate of change over time or in differences in scores across ages. Exceptions, notably in expressive language, existed. These results suggest that, while genotype may be related to overall degree of impairment, individuals with PMS, regardless of genotype, tend to have a similar rate of change over time and age in developmental and behavioral domains. A significant caveat is that sequencing is a relatively recent diagnostic approach, which will bias the results.

Keywords
Phelan-McDermid syndrome, genotype-phenotype associations, natural history, SHANK3, 22q13

Background

Genetic conditions associated with neurodevelopmental disabilities often have a wide range of phenotypic expression. Even within identified syndromes, features can range from language that includes no spoken words to fluent speech, intractable epilepsy to no seizures, and profound intellectual disability to mild or no cognitive impairment (Garber, Visootsak, & Warren, 2008; Levy, Lerman, et al., 2022; Tang et al., 2021; Trelles et al., 2021; Varela, Kok, Otto, & Koiffmann, 2004). Phelan-McDermid syndrome (PMS) is one such syndrome (Burdeus-Olavarrieta et al., 2021; Chen et al., 2022; De Rubeis et al., 2018; Landlust et al., 2022; Nevado et al., 2022; Soorya et al., 2013). The common features of PMS include developmental delays leading to intellectual disability, epilepsy, autism spectrum disorder (ASD) and other behavioral challenges, medical concerns such as renal or cardiac abnormalities, and dysmorphic features. Intellectual disability is essentially ubiquitous, though its severity ranges from profound to mild. Many individuals never develop speech, though this also varies widely (Cusmano-Ozog, Manning, & Hoyme, 2007; Levy, Foss-Feig, et al., 2022; Lindquist et al., 2005; Nevado et al., 2022; Sarasua, Dwivedi, et al., 2014). Epilepsy has been identified in approximately 25% of individuals (Levy, Foss-Feig, et al., 2022; Nevado et al., 2022). Some individuals develop severe neuropsychiatric manifestations and loss of skills in adolescence or later (Kolevzon, Delaby, Berry-Kravis, Buxbaum, & Betancur, 2019). Explanations for the diverse clinical spectrum in PMS have been a source of vigorous research over the last decade, with genotype-phenotype associations being one of the main areas of interest.

The genetics of PMS are as diverse as the phenotype. Although haploinsufficiency of *SHANK3* is the underlying cause of PMS (Wilson et al., 2003), individuals' unique genetic alterations can range from mutations involving a single base pair to deletions that involve millions of base pairs. The study of genotypes has evolved over the years with advances in the resolution of genetic testing technology, with exome sequencing in particular advancing sequence variant identification. Early research focused on approximate deletion sizes or presence

of ring chromosomes, as fluorescence in situ hybridization and karyotypes were used to diagnose microdeletion syndromes (Jeffries et al., 2005; Wilson et al., 2003). Once chromosomal microarrays became the standard of care, the size of deletion was more precisely analyzed (Sarasua et al., 2011; Soorya et al., 2013). With the increasing availability of gene sequencing, sequence variants within *SHANK3* have also been incorporated into genotype-phenotype analyses (De Rubeis et al., 2018; Levy, Foss-Feig, et al., 2022; Serrada-Tejeda, Martínez-Piédrola, Máximo-Bocanegra, Sánchez-Herrera-Baeza, & Pérez-de-Heredia-Torres, 2022; Soorya et al., 2013; Xu et al., 2020). While this evolution in genetic testing methodology adds complexity in comparing literature over time, general themes have emerged from genotype-phenotype associations studies within PMS. Overall, those with larger deletions that extend past *SHANK3* and its neighboring genes tend to have more severe intellectual disability, as well as renal abnormalities (Levy, Foss-Feig, et al., 2022; Nevado et al., 2022; Tabet et al., 2017). Recently, our group used cross sectional data to describe the association between smaller deletions and sequence variants with regression of skills and psychiatric decompensation in adolescent years (Levy, Foss-Feig, et al., 2022). However, understanding this genotype-phenotype relationship may be confounded, since those with smaller deletions or sequence variances had also been found to have more skills generally, while measurement of adolescent decompensation in individuals with Class 2 deletions may be fraught since measuring this phenomenon is harder in those with more profound disability. Patterning of additional features such as ASD, recurrent infections, hypotonia, and cardiac abnormalities have been explored, but clear, replicated, genotype-phenotype relationships have not been found.

Understanding the natural history of PMS, and how genotype plays a role in the phenotypic heterogeneity over time is an essential next step in predicting more accurate prognoses, which is important for both clinical care and clinical trial readiness. To date, the natural history of PMS has not been sufficiently reported in the literature (Schön et al., 2023). In parallel to this manuscript, an analysis of adaptive functioning skills in PMS overall has been

described ([redacted] *in review)*. Over the course of two consecutive years of study visits, individuals showed continued significant impairments relative to the general population. However, mild gains in ability were observed in some domains of adaptive functioning, as measured by the Vineland Adaptive Behavior Scales, 2nd Edition. Whether there are differences in developmental, cognitive, and behavioral trajectories between PMS individuals with Class 1 deletions (including *SHANK3, ARSA, ACR, RABL2B*), Class 2 deletions (including any other genes), and *SHANK3* sequence variants remains unknown. Here, we assessed the relationship between genotype and longitudinal natural history of PMS for the first time by tracking change in key variables over one to five years and describing growth curves as a function of genotype groupings.

Methods

Participants were individuals with PMS who were enrolled in a multi-site prospective longitudinal natural history study within the Rare Disease Clinical Research Network (RDCRN) entitled the Developmental Synaptopathies Consortium. The first phase of this study enrolled participants from 2015 to 2020; the second phase began in 2021 and is ongoing. Because there were too few older adults enrolled in the study at the time of analysis, individuals over the age of 30 years old for Class 1 and Class 2 deletion groups, and age 21 years old for the sequence variant group (*see below for group definition)* were excluded from analysis. Overall, 154 participants were included in the analysis. Sites included Icahn School of Medicine at Mount Sinai (n=59), Boston Children's Hospital (n=28), Rush University Medical Center (n=33), Stanford Medical Center (n=9), National Institute of Mental Health (n=19), and University of Texas Southwestern Medical Center (n=6). For the first study phase (n=99), participants were evaluated annually for 3 years, with additional caregiver forms completed at 6- and 18-month timepoints. A detailed manual of procedures exists to reduce differences in methodological

differences between sites. The second phase includes new participants for three annual visits (n=47) and returning participants from Phase 1 (n=43) for two additional annual visits. Data were also taken from an adult pilot study competed in 2019, which included one timepoint for participants over 18 years of age (n=10). The Adult Pilot Study included the same measures as the other studies. The total group (n=154) is referred to as the combined sample.

Data may be missing from longitudinal studies for a variety of reasons (e.g., failure or inability to comply with study procedures, missed study visit, early drop out, inability to complete in-person evaluations during the COVID-19 pandemic). Over half of participants completed three or more annual timepoints: 48 (31%) participants completed one, 14 (9%) completed two, 51 (33%) completed three, 26 (17%) completed four, and 15 (10%) completed five annual timepoints. In addition, 93 (60%) participants also completed 6-month forms and 79 (51%) completed 18-month forms. The second phase of the study was currently ongoing at the time of analysis and additional visits are expected to be completed. The analysis plan was selected to maximize retention of all possible data (*see below).*

Genetic Grouping

Clinical genetic test reports were collected for each participant; to be included in the natural history study, deletions must have included *SHANK3* or sequence variants in *SHANK3* must have been classified as likely pathogenic or pathogenic. Deletions were all mapped to hg19 and sequence variants to NM_033517.1 (Figure 1). Participants were separated into the following groups: Class 1 deletions are genic deletions that include only *SHANK3* with or without the deletion of *ARSA*, *ACR*, and *RABL2B*. These latter three genes are not expected to contribute to the phenotype of PMS because they are not constrained for protein truncating variants (pLI = 0 in gnomAD database), although *ARSA* is associated with a known autosomal recessive disorder. Class 2 deletions are larger deletions that do not qualify as Class 1 deletions (i.e., including the deletion of any other genes in addition to *SHANK3* and the three mentioned above).

Sequence variants are pathogenic sequence variants within the *SHANK3* gene. Participants included 41 individuals with Class 1 deletions, 73 with Class 2 deletions, and 40 with sequence variants (Table 1). Individuals were classified based upon their 22q13/SHANK3 alterations, individuals with other genetic findings (e.g., duplications due to unbalanced translocations) were not grouped separately.

Study Procedure

Annual visits included direct testing, caregiver interviews, and caregiver surveys. Adaptive functioning was measured with the Vineland Adaptive Behavior Scales – 2nd Edition (VABS-2) (Sparrow, Cicchetti, & Balla, 2005), and performance was operationalized using subdomain growth scale values (GSVs). GSVs were developed for the purpose of tracking within-person change in capability. GSVs are person-ability scores derived through Rasch analysis (Daniel, 2022). GSVs are a transformation of the raw score; like the raw score, GSVs are unitless, but unlike the raw scores they are measured at the interval level (i.e., a one-unit change has the same meaning across the full scale). Thus, the GSV is the most appropriate expression of absolute performance for quantifying change over time (C. Farmer, Thurm, Troy, & Kaat, 2023). The magnitude of change in GSV can be evaluated relative to its standard error of measurement (SEM), for additional statistical context, or the other norm-referenced scores can be used to assist in its interpretation. GSV conversion tables were not included with the original VABS-2 but are available from the publisher (Daniel, 2022).

Best estimate full-scale, verbal, and nonverbal intellectual quotients (IQ) or developmental quotients (DQ) were calculated from an age- or developmentally appropriate test, either the Stanford Binet Intelligence Scales (Roid & Pomplun, 2012), the Differential Abilities Scales (Elliot, 2007), or Mullen Scales of Early Learning (Mullen, 1995). A DQ (mean of age equivalents divided by chronological age) was calculated from the Mullen Scales since it was used out of age range, which is included as the IQ estimate for this variable. IQ scores are

standard scores with a population mean of 100 and standard deviation of 15. The meaning of a decrease in standard score is ill-defined, as it can result from slower-than-expected acquisition of skills, stability, or true loss of skills (C. A. Farmer et al., 2020).

Challenging behavior was measured by the Aberrant Behavior Checklist (Aman & Singh, 1986), and repetitive behavior and restricted interests were measured by the Repetitive Behavior Scales – Revised (Bodfish, Symons, Parker, & Lewis, 2000), which produce subscale-level raw summed scores.

Vocabulary testing was attempted for participants using Expressive Vocabulary Test -2nd edition (EVT-2) and the Peabody Picture Vocabulary Test – 4th Edition (PPVT-4) (Dunn & Dunn, 2007; Williams, 1997). Due to the significant impairment of many individuals with PMS, a basal score on these assessments was not often reached.

Table 1. Baseline Demographics and Visit Completion [insert here]

Analysis

The goal of this study was to quantify and compare the developmental (age) trajectories between genetic subgroups of PMS for several outcomes. Because participants enrolled at different chronological ages, the effects of age in this study are composed of both between-subject (i.e., using age at enrollment for individual participants) and within-subject (i.e., using the duration of study each individual's participation) information. Proper separation of these sources of variability allows us to determine the extent to which younger study participants can be expected to develop into the phenotype of older participants (within-subject effects), or whether differences between younger and older participants (between-subject effects) are explained by something other than development (known as a cohort effect). Further, it provides a more accurate estimate of variability in these effects than a naïve analysis with only chronological age. In this study, we explored whether each of these effects of age varied by genetic group.

We used general linear mixed effects models with fixed effects for the between-subject effect of age (expressed as a time-invariant average age of participation per person) and the within-subject effect of time (expressed as chronological age minus participant's average age). For example, if an individual had visits at age 6, 7, and 8 years old, their average age would be 7 and their time values would be -1, 0, and +1 (chronological age minus average age). For interpretability, average age was centered at the group mean for each measure, so that main effects pertain to a child of average age for the sample. To explore differences among genetic groups, we included a main effect of genetic group (Class 1 deletions, Class 2 deletions, and sequence variants) and interactions with age and time (group*age, group*time). For GSV scores, given the nonlinear nature of developmental skill acquisition (i.e., skill acquisition in typically developing populations is rapid in early childhood and slower or plateaued as individuals approach adulthood), higher-order terms were included to reflect a potential reduction of age differences at older ages (age^2, group *age^2) and a potential slowing of within-subject change for older participants (age*time, group*age*time). These higher order terms were only used in measures using GSV scores. Because the data were nested within person as well as within site, we included random subject-level intercepts and slopes for time, as well as a random intercept for site. Maximum likelihood estimation allows for all participants with at least one outcome assessment to be included in the analysis.

The tenability of the required assumptions (independence, normality, and constancy of residuals) was evaluated via visual inspection of level 1 and level 2 residuals. In many cases, residuals departed from normality at the tails, indicating more extreme values than expected for a truly normal distribution. This type of non-normality is resistant to transformation because it is likely caused by some unknown and unmodeled causal variables. This is a common limitation in the modeling of data from heterogeneous populations, and the model results must be interpreted with caution.

Parameters of interest were the main effect of genetic group (indicating persistent mean differences between groups), the interaction between genetic group and within-subject time (indicating differences between groups in annualized change), the interaction between genetic group and between-subject age (indicating that the difference between older and younger individuals depends on genetic group), and, when included, interaction between genetic group and between-subject age^2 and three-way interaction between group, between-subject age, and within-subject time. The interpretation for the within-subject time variable is annualized change, or the model-estimated change in score for each year of participation. The interpretation for the between-subject age variable is the model-estimated difference in score for each year of age difference between two participants. While GSVs are the more powerful metric for statistical detection of group differences (C. Farmer et al., 2023), they are unitless scores and the magnitude of differences does not have inherent clinical meaning. In order to aid clinical interpretation of statistical differences, we translate GSV parameter estimates into age equivalent values using the scoring manual.

Consistent with current recommendations (Wasserstein, Schirm, & Lazar, 2019), we present raw (uncorrected) p-values in conjunction with 95% confidence intervals for all parameter estimates and use a nominal p-value threshold of 0.05 only to focus the discussion of results. All uncorrected p-values are available in the main text. All point estimates and 95% confidence intervals are available upon request (*see additional tables).*

Results

Assessment Completion

The number of times participants completed each assessment are in Table 1. Additional specifics on how many times each assessment was completed in each group is available upon request (Additional Table 1). For vocabulary analyses (PPVT-4, EVT-2), details are provided in the corresponding results section.

Adaptive Functioning

Intercepts

Model-derived intercept statistics are shown in the first column of Table 2, values and 95% confidence intervals are available upon request (Additional Table 2). Because of the centering procedures, the model intercepts are the model-estimated value for an individual of average age (11 years) in the middle of their study participation. The genetic groups differed in intercept for almost all subdomains, indicating mean differences in degree of severity/impairment. Within the Communication domain, intercepts were highest for the Class 1 deletion group, followed by the sequence variant group, and then the Class 2 deletion group (Table 2, Additional Table 2, Figure 2). Estimated age equivalents (AEs) for Receptive Language were two years and two months (2y2m), 1y2m, and 1y10m for the Class 1, Class 2, and sequence variant groups, respectively; in Expressive Language the AEs were 2y6m, 1y4m, 1y8m, and in Written, 5y2m, 3y10m, and 4y6m. Within Daily Living Skills, the Personal subdomain intercepts were different between Class 1 and Class 2 deletion groups as well as between Class 1 and sequence variant groups. The intercepts transformed to AEs were 3y6m for the Class 1 group, and 2y4m for both the Class 2 and sequence variant group, indicating over a year difference. In Community, the Class 1 and 2 groups showed a wide difference in age equivalents (4y10, 1y10), and the sequence variant group fell in-between the other groups with an AE of 3y5m. There were no statistical differences between group intercepts in the Domestic subdomain, though AEs were seemingly different at 2y7m, 1y3, and 1y5m for the three groups. Similarly, in the Socialization domain, differences in intercepts were seen between Class 1 and Class 2 deletion groups in all subdomains. The sequence variant group was closer to the Class 2 deletion group in all subdomains. Intercepts transformed to 2y7m, 1y3m, and 1y5m in the Interpersonal Skills subdomain for Class 1, 2, and sequence variant groups, respectively. In Play and Leisure, the AEs were 2y7m, 1y1m, and 1y4m, and in Coping Skills the AEs were 3y5m, 2y1m, and 2y3m for the three groups.

Between Subjects Age (difference across ages)

The between-subjects age variable denotes the difference in score for each year older than the average age of a participant (Table 2, Additional Table 2, Figure 3). No differences were seen between genetic groupings in this age variable. The quadratic term of the between-subjects variable allows the effect of a one-year age difference to change nonlinearly. No differences were seen across genetic groups in the quadratic time variable. Taken together, the lack of differences in between-subject age terms across groups indicates that the difference between younger and older participants is similar across genetic groups.

Within Subject Time (individual change during study participation)

The within-subjects time variable expresses the change of an individual's score over a one-year period of their study participation. Generally, genetic groups exhibited similar within-subject slopes (Figure 3). An exception is the Expressive Language subdomain. For individuals at the average age (i.e., between-subjects age =0), the Class 1 group showed a gain of 2 [-3.71-6.72] GSV points per year, the Class 2 group was shown to remain relatively stable, 0.37 [-2.87 - 3.6], and the sequence variant group was shown to lose around 5 [-10.3-0.88] GSV points per year (Table 2, Figure 3). The Class 1 and Class 2 groups both differed from the sequence variant group but did not differ from each other. When transformed into age equivalents, the Class 1 group gained the equivalent of 1 month of developmental progress over the course of one year (2y6m to 2y7m), the Class 2 group stayed stable at 1y4m, and the sequence variant group lost the equivalent of 2 months of developmental progression over the year (1y8 to 1y6m).

The three-way interaction term addresses whether the difference in within-subjects slope between participants with a one-year difference in age depends on genetic group. There were no differences between the genetic groups apart from Expressive Language. The within-subjects slope stayed relatively stable for both the Class 1 and 2 groups across participant ages, -0.38 [-

1.14-0.37], -0.16 [-0.74-0.41], but the decreasing within-person trajectory slowed down (became less negative) for older participants within the sequence variant group (1.87, 0.74-2.99) (Figure 3). Taken together with the within-subjects time variable, these findings mean that the sequence variant group's negative rate of change (decline in skills) became more positive over age, gradually evening out, as the Class 1 and 2 deletion groups' positive rate of change (acquisition of skills) gradually decreased with age, eventually evening out. Additional figures 1a-c are available upon request and show between- and within- subject terms separately.

Table 2. Vineland-2 Test Statistics [insert here]

Cognitive Testing

Nonverbal, Verbal, and Full-Scale IQ intercepts, or the model-estimated average for a participant of average age (10 years for IQ), were the highest in the Class 1 deletion group, followed by the sequence variant group, and then the Class 2 deletion group (Table 3, Figure 4, Additional Table 3). The Class 2 and sequence variant group did not differ from each other but did differ compared to the Class 1 group. There was no appreciable effect of between-subject age; there was essentially no difference in IQ score across ages for all groups. Finally, within-subject srate of change within person was a decline of 1-2 IQ points per year of participation.

Table 3. Intellectual and Developmental Quotient Test Statistics [insert here]

Behavior

Few between-group differences were observed in the intercept, between-subject age, or within-subject time effects on ABC scores (Table 4, Additional Table 4). There were differences in the intercepts between the Class 2 deletion and sequence variant groups in Hyperactivity, where the sequence variant group had more severe (higher) scores. Additionally, average scores were

different in Inappropriate Speech between the Class 2 deletion group and the other two groups, where the Class 2 group had less severe (lower) scores than both groups; however, this finding is likely due to lower speech overall rather than specifically less inappropriate speech (Levy, Foss-Feig, et al., 2022). The three genetic groups had similar differences across ages and change over time in all behavior domains.

Table 4. Aberrant Behavior Checklist domain test statistics [insert here]

Repetitive behavior as measured by the RBS-R was also similar across genetic groups (Table 5, Additional Table 6). Average scores only differed in the Ritualistic domain, where the Class 1 deletion group had the most symptoms, followed by the sequence variant and then Class 2 deletion group. The effect of between-subject age differed across groups, such that older participants had lower (less severe) Self Injury scores in the Class 1 group while older participants had higher (more severe) scores in the Class 2 and sequence variant groups. Within-subject analyses revealed similar trajectories over time in repetitive behaviors.

Table 5. Repetitive Behavior Scale – Revised, Domain Test Statistics [insert here]

Vocabulary

The PPVT-4 and EVT-2 were attempted for each participant. Most participants could not complete these assessments due to functional capacity, leading to informative missingness. To illustrate this, Table 6 shows the number of participants who could complete enough of the assessment to receive the base score compared to the number of individuals for whom it was attempted. Table 6 represents participants at all timepoints (e.g., if a participant attempted the assessment three times and achieved a score once, they are counted as being able to achieve). Due to this non-random missingness, analysis or comparison of trajectories is not possible. The

PPVT-4 and EVT-2 are likely not appropriate measures for this population to measure change over time.

Table 6. Completion of PPVT-4 and EVT-2 [insert here]

Discussion

Genotype has been found to explain some of the variability in the clinical phenotype of PMS from studies that report one point in time. How these clinical differences progress over time, however, remains unknown. We aimed to understand the clinical trajectory of individuals with PMS over time, based on three established genetic subgroups that serve as a proxy for deletion size or sequence variant status, and to determine if these trajectories differed depending on the genotype. Overall, we found that the developmental change over time and across ages was largely similar between all three genetic groups, in most domains. However, mean differences between groups in cognitive, adaptive, and language domains were present. While all individuals Showed a similar developmental progression during the years of their participation, individuals with Class 1 deletions had higher baseline scores. This overall finding is not surprising, as our first analyses of an overlapping cohort revealed that those with Class 1 deletions were overall less delayed than the other groups (Levy, Foss-Feig, et al., 2022). However, in this previous cross-sectional analysis, adaptive functioning as measured by Vineland standard scores, did not show significant differences between groups. Moreover, Class 1 deletions and sequence variants were included in the same group for analyses, which may have affected the group-level findings. In the current analysis, GSV scores were used and mean differences, where individuals with Class 1 deletions had higher scores, were identified, which is aligns with both current and prior findings in cognitive and developmental differences. Lastly, behavioral findings were consistent

across analyses where, overall, groups did not differ either in mean scores or trajectories over time.

A finding important for consideration is the variability of trajectory across types of scores. Generally, in this sample, a slight decline in standard scores (e.g., in IQ) and an incline in raw scores were seen across variables. This indicates that individuals are gaining skills over time, but at a rate slower than expected as compared to the general population. The positive slope of raw scores and negative slope of standard scores therefore indicates a slow improvement, not a loss of skills over time. Similar findings are explored in depth in [redacted], *in review*. This pattern has been reported in other longitudinal studies of genetic disorders. For example, in a longitudinal analysis of Fragile X Syndrome, boys with Fragile X without autism showed a decrease in IQ standard score over time but a stable or slight increase in IQ raw score over time (Hernandez et al., 2009). Authors concluded that this pattern was not to due regression, but rather a slower than expected development compared to typically developing peers (Hernandez et al., 2009). This same trend was reported in a longitudinal study of Angelman syndrome (Peters, Horowitz, Barbieri-Welge, Taylor, & Hundley, 2012). In another study of Fragile X syndrome, authors found that the acquisition of new adaptive skills slows as males with the syndrome age (Klaiman et al., 2014). A minority of results in these analyses may indicate a loss of skills, including the Expressive GSV scores in the sequence variant group, which has a downward slope. This coincides with the clinical observations suggesting a proportion of individuals with higher levels of language tend to lose language skills as they age, and the observation of wide variability in the sequence variant group where results may be driven by those changing the most (e.g., regressive episodes). Very few longitudinal studies of rare genetic disorders exist with large enough sample sizes to conduct in depth statistical analyses; this should be considered in future study design.

Individuals with Class 1 deletions appear to gain more skills than those with Class 2 deletions, likely in part because less genetic material is missing. While *SHANK3* is the main

contributor to neuropsychiatric, cognitive, and behavioral features seen in PMS, one could expect deleting more genetic material could produce a more severe neurodevelopmental phenotype (Girirajan et al., 2011). This finding is consistent in both our analyses and independent analyses in the literature (Samogy-Costa et al., 2019; Sarasua, Boccuto, et al., 2014; Sarasua, Dwivedi, et al., 2014; Sarasua et al., 2011; Tabet et al., 2017). In the current analysis, it is evident that individuals with Class 1 deletions have less severe cognitive and adaptive impairments, however the trajectory is similar to the other groups. This indicates that a divergence in developmental trajectory, if present, could have occurred at a younger age than the sample we currently analyzed. Including individuals as young as possible to measure trajectory in infancy and early childhood may help uncover when and how individuals with Class 1 deletions achieve a higher average baseline level. Though this is complicated by the timing of genetic testing, it will likely become more feasible as individuals are offered testing earlier. Indeed, the second phase of this current study includes enrollment of participants as young as 18 months.

A noteworthy observation within this analysis was the similarities between individuals with Class 2 deletions and those with sequence variants, as compared to those with Class 1 deletions. In theory, individuals with Class 1 deletions and sequence variants should be more phenotypically similar, as they are more genetically similar. In contrast, Class 2 deletions can have upwards of 100 additional deleted genes. Though it is possible that there may be true differences between the Class 1 deletion and sequence variant genotypes, many alternative hypotheses exist to explain the differences seen in severity of symptoms.

The first explanation relates to the type of testing required to diagnose each type; microarrays are used to detect deletions and have been the standard of care for ASD and developmental delays for over a decade (Moeschler et al., 2014) and are commonly ordered by many types of clinicians. Sequence variants are diagnosed by gene sequencing (i.e., exome sequencing), which historically has been a test reserved for more severe or complex

neurodevelopmental phenotypes and is typically only ordered by genetics specialists. Therefore, one explanation may be that only the more severely affected individuals with sequence variants are diagnosed, while the less severely affected ones are not referred to genetics or offered exome sequencing and therefore are unaware of their diagnosis, and not involved in research. In addition to being reserved for more complex cases, many insurance companies have specific clinical requirements for coverage for exome sequencing, including unexplained developmental regression. Therefore, individuals are more likely to be offered exome sequencing and diagnosed after a developmental regression – biasing the sample to those who have already lost skills. In fact, in our combined sample, 76% (22/29) of individuals in the sequence variant group who had a regression lost skills before their diagnosis of PMS – likely the trigger for genetic testing. This compares to 47% (7/15) and 32% (6/19) of the Class 1 and Class 2 deletion groups with a regression respectively, who reported that the loss occurred prior to their diagnosis of PMS. Therefore, the sequence variant group may be further biased by having an overabundance of individuals who have lost skills.

Examining an independent sample of individuals with PMS ascertained for loss of skill and psychiatric decompensation shows more individuals than expected with a sequence variant, compared to deletions of either class; 21/48 (44%) had a sequence variant which compares to 40/154, 26%, in the broadly ascertained PMS sample described here *(Extension for Community Healthcare Outcomes, 2023; A. Kolevzon, 2023)*. It is therefore hard to conclude, or even hypothesize, at this time if individuals with sequence variants are more likely to face developmental regression or if sample bias is responsible for most of the differences seen to date. Over time as exome sequencing becomes more widely used due to decreasing costs and increasing awareness, we expect to see more individuals diagnosed with sequence variants and more mildly affected individuals. The former is already being observed; in the first phase of this study, 19 (19%) individuals with sequence variants participated in the study; in the first 3 years

of the second phase we already have 25 new individuals have already enrolled with a sequence variant (49% new enrollment).

Several limitations are important to consider, including floor effects in multiple variables resulting in some unable to be analyzed. Selecting appropriate measures within populations with a wide range of clinical severity is challenging, as most measures are not designed to assess individuals who are profoundly affected. Additional measures have been added to the second phase of the natural history study, specifically in communication, to better assess participants with limited spoken language. As already mentioned, selection bias limiting participants to those who received genetic testing likely resulted in a sample skewed to the more severely affected end of the spectrum of PMS. Finally, these results cannot be generalized to ages outside of those included in analysis; more research is needed on trajectories extending into early adulthood.

Overall, these results illustrate that individuals with PMS, regardless of genetic subtype, tend to consistently gain skills across development, though at a slower rate than the general population. Individuals with Class 1 deletions gained skills at the same rate to the other genetic groups, though they tended to have higher baseline performance in cognitive, adaptive, and language domains within this combined sample. Challenging behavior was largely the same between genetic groups, with the Class 1 deletions having more behavioral symptoms reported in some repetitive behavior domains. Individuals with sequence variants may be at a greater risk of losing skills, particularly in expressive language, as compared to the other genetic subtypes, but this preliminary finding requires further investigation. No differences were found in developmental trajectories that would require different outcome measure analyses in future clinical trials between genetic groups. Future studies should include young participants in order to capture earlier development than available in the current analysis and should follow children over long enough periods of time to detect when changes, including regressions, may show divergences in trajectories both between and within genetic subgroups.

Tables.

Table 1. Baseline Demographics and Visit Completion

		Class 1 Deletion	Class 2 Deletion	Sequence Variant
Total Participants		**41**	**73**	**40**
Baseline age, years (mean, SD)		13.04 (7.6)	9.34 (6.4)	9.27 (4.7)
Age at PMS diagnosis (mean, SD)		7.34 (5.2)	4.63 (5.1)	7.77 (4.7)
Sex (% female)		16 (39%)	38 (52%)	15 (37%)
Baseline Full Scale IQ (mean, SD)		36.13 (17.8)	23.42 (16.8)	27.91 (21.4)
Verbal (% verbal)		33/41 (80%)	27/73 (37%)	22/40 (55%)
Skill Loss History (% with regression)		15/41 (37%)	19/73 (26%)	29/40 (73%)
Phase 1	Baseline	27	53	19
	Month 6	24 (11%)	51 (4%)	18 (5%)
	Year 1	26 (4%)	53 (0%)	16 (16%)
	Month 18	22 (19%)	44 (17%)	13 (32%)
	Year 2	24 (11%)	51 (4%)	17 (11%)
Phase 2 Returning	Visit 4	11	20	11
	Visit 5	4	5	6
Phase 2 New	Baseline	10	16	21
	Year 1	1	3	3
Adult Pilot	Baseline	6*	4	0

Table 1: Baseline characteristics of participants. Verbal classification made from clinician interview. Skill Loss data from the Autism Diagnostic Interview – Revised and clinician report of regression. For Phase 1, Phase 2, and Adult Pilot rows, the number represents the participants in each group per timepoint, along with the percentage of missing visits. This number represents whether the participant was present for the visit, not specific assessment completion. Assessment completion counts are located in Additional Table 1, which is available upon request. Phase 2 Visit 5 for returning participants and Phase 2 Year 1 for new participants were not completed enrollment at the time of analysis; therefore, the visits are not considered missing. *2/6 continued to DSC2 New Enrollment

Table 2. Vineland-2 Test Statistics

	Intercept	Between-subj age	Between subj-age ^2	Within subj time	Between*Within
Receptive					
Class 1 vs. Class 2	-2.78 (103.25), p=0.006	-0.10 (106.37), p=0.92	1.09 (104.40), p=0.279	-0.71 (75.71), p=0.48	0.02 (70.55), p=0.987
Class 1 vs. Sequence variant	-0.41 (106.38), p=0.68	-0.95 (107.00), p=0.344	-1.42 (107.18), p=0.16	-0.89 (64.67), p=0.379	1.71 (73.89), p=0.091
Class 2 vs. Sequence variant	1.77 (105.90), p=0.08	-0.97 (104.73), p=0.333	-1.90 (106.34), p=0.06	-0.32 (60.66), p=0.747	1.82 (63.39), p=0.073
Expressive					

Class 1 vs. Class 2	-3.51 (102.48), p=0.001	-0.08 (104.56), p=0.938	0.83 (102.95), p=0.409	-0.56 (73.68), p=0.577	0.45 (71.88), p=0.652
Class 1 vs. Sequence variant	-1.88 (104.77), p=0.062	-0.67 (104.93), p=0.505	-0.13 (105.15), p=0.899	-2.23 (63.96), p=0.029	3.26 (75.51), p=0.002
Class 2 vs. Sequence variant	0.86 (104.49), p=0.394	-0.68 (103.33), p=0.501	-0.51 (104.50), p=0.611	-2.04 (62.26), p=0.045	3.15 (64.93), p=0.002
Written					
Class 1 vs. Class 2	-2.61 (103.66), p=0.01	-0.95 (104.62), p=0.342	0.91 (103.22), p=0.367	-0.45 (79.54), p=0.656	1.48 (80.34), p=0.142
Class 1 vs. Sequence variant	-0.93 (104.99), p=0.354	-1.69 (104.33), p=0.094	-0.73 (104.92), p=0.468	-1.96 (68.78), p=0.055	-0.39 (82.78), p=0.701
Class 2 vs. Sequence variant	1.12 (104.89), p=0.267	-0.98 (102.88), p=0.329	-1.14 (104.35), p=0.258	-1.81 (71.78), p=0.074	-1.54 (72.58), p=0.129
Personal					
Class 1 vs. Class 2	-2.58 (107.09), p=0.011	-0.43 (108.30), p=0.67	0.32 (97.63), p=0.752	-1.13 (88.66), p=0.263	1.08 (85.97), p=0.285
Class 1 vs. Sequence variant	-1.98 (108.79), p=0.051	-1.55 (107.02), p=0.125	0.26 (102.40), p=0.793	-0.73 (77.78), p=0.466	1.05 (89.75), p=0.296
Class 2 vs. Sequence variant	0.04 (107.75), p=0.971	-1.32 (103.51), p=0.19	0.11 (100.08), p=0.91	0.29 (75.67), p=0.775	0.31 (79.34), p=0.754
Domestic					
Class 1 vs. Class 2	-1.50 (102.81), p=0.137	0.82 (105.80), p=0.417	0.05 (102.81), p=0.962	0.20 (57.75), p=0.845	2.25 (61.82), p=0.028
Class 1 vs. Sequence variant	-1.46 (106.06), p=0.149	-0.37 (106.22), p=0.715	-0.34 (106.36), p=0.735	0.49 (49.57), p=0.627	1.26 (63.95), p=0.213
Class 2 vs. Sequence variant	-0.30 (105.61), p=0.765	-1.19 (103.67), p=0.237	-0.36 (105.19), p=0.722	0.37 (53.7), p=0.712	-0.34 (54.68), p=0.737
Community					
Class 1 vs. Class 2	-3.40 (101.40), p=0.001	-0.55 (104.60), p=0.584	0.97 (102.76), p=0.333	-0.89 (194.79), p=0.375	1.08 (194.30), p=0.28
Class 1 vs. Sequence variant	-0.98 (104.77), p=0.328	-1.35 (105.36), p=0.181	-0.82 (105.83), p=0.414	-1.69 (194.43), p=0.092	-0.57 (194.64), p=0.568
Class 2 vs. Sequence variant	1.67 (104.40), p=0.098	-0.99 (103.12), p=0.326	-1.26 (104.92), p=0.211	-1.05 (194.50), p=0.297	-1.46 (193.62), p=0.147
Interpersonal					

Class 1 vs. Class 2	-2.88 (102.29), p=0.005	-0.07 (104.62), p=0.945	0.90 (103.18), p=0.372	-1.16 (65.20), p=0.252	-0.96 (67.29), p=0.341
Class 1 vs. Sequence variant	-1.70 (104.79), p=0.092	-0.89 (105.15), p=0.374	-0.60 (105.44), p=0.548	-1.97 (56.30), p=0.054	0.18 (70.02), p=0.855
Class 2 vs. Sequence variant	0.54 (104.50), p=0.59	-0.94 (103.56), p=0.351	-1.01 (104.81), p=0.315	-1.13 (58.43), p=0.265	0.92 (59.71), p=0.36
Play and Leisure					
Class 1 vs. Class 2	-2.43 (102.04), p=0.017	0.27 (104.88), p=0.79	0.76 (103.25), p=0.447	-0.29 (65.29), p=0.77	-1.13 (61.42), p=0.262
Class 1 vs. Sequence variant	-1.31 (105.03), p=0.193	-1.25 (105.57), p=0.213	-0.49 (105.90), p=0.623	0.33 (55.92), p=0.746	-0.70 (65.02), p=0.489
Class 2 vs. Sequence variant	0.58 (104.71), p=0.564	-1.66 (103.69), p=0.099	-0.84 (105.16), p=0.403	0.67 (52.44), p=0.505	0.11 (55.51), p=0.917
Coping					
Class 1 vs. Class 2	-2.62 (100.95), p=0.01	-0.670 (105.71), p=0.505	1.20 (102.39), p=0.232	0.83 (45.10), p=0.413	-0.88 (47.03), p=0.381
Class 1 vs. Sequence variant	-1.65 (105.69), p=0.102	-1.35 (106.75), p=0.182	-0.03 (107.04), p=0.98	-0.64 (38.22), p=0.524	0.08 (48.81), p=0.934
Class 2 vs. Sequence variant	0.37 (104.99), p=0.711	-0.88 (103.04), p=0.384	-0.58 (105.51), p=0.563	-1.61 (40.83), p=0.116	0.76 (40.82), p=0.451
Fine Motor					
Class 1 vs. Class 2	-2.65 (102.37), p=0.009	-0.01 (105.13), p=0.991	0.73 (102.14), p=0.466	0.58 (41.60), p=0.566	-0.49 (44.83), p=0.63
Class 1 vs. Sequence variant	-0.50 (104.89), p=0.621	-0.79 (105.72), p=0.43	-0.76 (105.16), p=0.449	-1.12 (36.10), p=0.27	0.21 (57.61), p=0.836
Class 2 vs. Sequence variant	1.590 (104.22), p=0.116	-0.88 (103.49), p=0.382	-1.09 (104.18), p=0.279	-1.90 (39.06), p=0.064	0.57 (50.09), p=0.571
Gross Motor					
Class 1 vs. Class 2	-1.45 (101.94), p=0.15	0.85 (103.81), p=0.396	0.04 (101.83), p=0.968	2.1 (51.54), p=0.041	-1.49 (49.31), p=0.142
Class 1 vs. Sequence variant	0.09 (103.62), p=0.926	-0.79 (103.98), p=0.433	-0.41 (103.68), p=0.682	0.12 (43.69), p=0.904	-0.93 (53.27), p=0.358
Class 2 vs. Sequence variant	1.24 (103.13), p=0.22	-1.70 (102.50), p=0.092	-0.42 (103.09), p=0.673	-2.02 (42.02), p=0.05	0.14 (44.46), p=0.892

Table 2. The t statistic, degrees of freedom, and p value are listed for each comparison. Note: the between-subjects centered age was 11 years.

Table 3. Intellectual and Developmental Quotient Test Statistics

	Intercept	Between-subj age	Within subj time
Nonverbal			
Class 1 vs. Class 2	-3.20 (106.94), p=0.002	-0.33 (107.43), p=0.74	0.41 (94.66), p=0.68
Class 1 vs. Sequence variant	-2.03 (108.45), p=0.044	-1.25 (109.84), p=0.213	1.04 (75.14), p=0.303
Class 2 vs. Sequence variant	0.81 (108.53), p=0.422	-0.96 (108.77), p=0.337	0.80 (75.43), p=0.425
Verbal			
Class 1 vs. Class 2	-3.69 (107.41), p=0	-0.07 (108.17), p=0.943	-0.21 (88.51), p=0.837
Class 1 vs. Sequence variant	-2.19 (108.83), p=0.031	-0.82 (110.52), p=0.417	-0.45 (66.91), p=0.653
Class 2 vs. Sequence variant	1.12 (108.90), p=0.266	-0.75 (109.16), p=0.457	-0.33 (65.65), p=0.745
Full Scale			
Class 1 vs. Class 2	-3.55 (107.29), p=0.001	-0.19 (108.06), p=0.849	0.33 (98.23), p=0.742
Class 1 vs. Sequence variant	-2.16 (108.62), p=0.033	-1.04 (110.21), p=0.302	0.56 (76.74), p=0.574
Class 2 vs. Sequence variant	1.01 (108.67), p=0.316	-0.87 (108.81), p=0.388	0.35 (75.22), p=0.731

Table 3. The t statistic, degrees of freedom, and p value are listed for each comparison. Note: the between-subject age was centered at 10 years.

Table 4. Aberrant Behavior Checklist domain test statistics [insert here]

	Intercept	Between-subj age	Within subj time
Irritability			
Class 1 vs. Class 2	-1.59 (107.41), p=0.115	0.648 (111.22), p=0.52	-0.676 (70.11), p=0.50
Class 1 vs. Sequence variant	-0.13 (115.77), p=0.90	0.30 (123.68), p=0.762	0.23 (51.57), p=0.817
Class 2 vs. Sequence variant	1.48 (117.10), p=0.141	-0.26 (119.44), p=0.798	0.91 (55.92), p=0.366
Social Withdrawal			
Class 1 vs. Class 2	-0.28 (111.52), p=0.779	0.01 (114.67), p=0.99	-1.52 (51.86), p=0.135
Class 1 vs. Sequence variant	0.83 (117.59), p=0.406	0.24 (123.39), p=0.812	-0.43 (38.58), p=0.667
Class 2 vs. Sequence variant	1.25 (118.48), p=0.214	0.22 (120.60), p=0.825	0.97 (40.62), p=0.34
Stereotypy			
Class 1 vs. Class 2	-0.47 (108.08), p=0.641	0.73 (111.68), p=0.467	-0.19 (46.05), p=0.851
Class 1 vs. Sequence variant	1.29 (115.03), p=0.2	0.44 (121.83), p=0.659	-0.03 (33.08), p=0.98
Class 2 vs. Sequence variant	1.96 (116.11), p=0.052	-0.19 (118.54), p=0.846	0.16 (36.63), p=0.875
Hyperactivity			
Class 1 vs. Class 2	-1.77 (108.73), p=0.079	-0.27 (111.23), p=0.789	-0.57 (47.98), p=0.574
Class 1 vs. Sequence variant	0.89 (114.05), p=0.378	0.01 (119.14), p=0.994	-0.74 (34.94), p=0.463
Class 2 vs. Sequence variant	2.86 (114.88), p=0.005	0.24 (116.70), p=0.812	-0.28 (36.23), p=0.779
Inappropriate Speech			

Class 1 vs. Class 2	-3.79 (103.79), p=0	0.32 (106.92), p=0.752	-0.20 (55.81), p=0.844
Class 1 vs. Sequence variant	-2.05 (110.49), p=0.043	0.94 (116.69), p=0.351	-0.07 (40.53), p=0.947
Class 2 vs. Sequence variant	1.52 (111.50), p=0.121	0.64 (113.52), p=0.521	0.12 (43.19), p=0.908

The t statistic, degrees of freedom, and p value are listed for each comparison. Note: the between-subjects centered age was at 11 years.

Table 5. Repetitive Behavior Scale – Revised, Domain Test Statistics

	Intercept	Between-subj age	Within subj time
Stereotypic			
Class 1 vs. Class 2	0.11 (106.54), p=0.91	1.31 (104.89), p=0.193	0.40 (47.23), p=0.688
Class 1 vs. Sequence variant	0.63 (113.17), p=0.527	1.06 (112.82), p=0.292	0.29 (41.17), p=0.776
Class 2 vs. Sequence variant	0.63 (115.55), p=0.529	-0.15 (115.72), p=0.883	-0.05 (41.89), p=0.961
Self Injury			
Class 1 vs. Class 2	-1.30 (99.57), p=0.198	2.18 (93.42), p=0.032	0.66 (73.44), p=0.515
Class 1 vs. Sequence variant	-1.08 (108.74), p=0.281	2.16 (104.93), p=0.033	-0.20 (63.48), p=0.839
Class 2 vs. Sequence variant	0.06 (111.99), p=0.951	0.22 (110.03), p=0.824	-0.81 (65.39), p=0.421
Compulsive			
Class 1 vs. Class 2	-1.92 (103.56), p=0.057	0.62 (102.02), p=0.534	-1.32 (40.56), p=0.196
Class 1 vs. Sequence variant	-1.38 (109.24), p=0.172	0.86 (108.89), p=0.39	-1.39 (36.10), p=0.173
Class 2 vs. Sequence variant	0.38 (111.29), p=0.702	0.33 (111.47), p=0.744	-0.37 (38.28), p=0.714
Ritualistic			
Class 1 vs. Class 2	-3.55 (108.11), p=0.001	-0.23 (106.39), p=0.815	0.13 (25.85), p=0.898
Class 1 vs. Sequence variant	-2.67 (115.12), p=0.009	-0.30 (114.91), p=0.767	-1.89 (21.82), p=0.072
Class 2 vs. Sequence variant	0.54 (117.75), p=0.588	-0.09 (118.15), p=0.926	-2.17 (23.03), p=0.04
Sameness			
Class 1 vs. Class 2	-1.92 (109.87), p=0.058	0.14 (108.13), p=0.888	-1.02 (38.30), p=0.313
Class 1 vs. Sequence variant	-1.87 (116.37), p=0.064	-0.49 (116.04), p=0.626	-1.01 (30.79), p=0.321
Class 2 vs. Sequence variant	-0.21 (118.99), p=0.834	-0.68 (119.15), p=0.495	-0.19 (33.09), p=0.851
Restricted Interests			
Class 1 vs. Class 2	-1.52 (96.61), p=0.132	-0.64 (95.15), p=0.522	-0.12 (38.22), p=0.905
Class 1 vs. Sequence variant	-0.07 (104.59), p=0.948	-0.49 (104.34), p=0.626	0.28 (33.58), p=0.781
Class 2 vs. Sequence variant	1.50 (107.67), p=0.138	0.10 (109.18), p=0.922	0.42 (33.90), p=0.681

The t statistic, degrees of freedom, and p value are listed for each comparison. Note: the between-subjects centered age was 10 years.

Table 6. Completion of PPVT-4 and EVT-2 [insert here

	PPVT-4	EVT-2
Class 1	22/41 (54%)	21/40 (53%)
Class 2	22/73 (30%)	15/72 (21%)
Sequence variant	19/39 (49%)	14/39 (36%)
Chi square p value	0.027	0.003

Proportion of participants who could complete the PPVT-4 and EVT-2 at any timepoint.

Tables and Figures available upon request

1. Assessment Completion. This table includes the number of participants in each genetic group that completed each clinical assessment.
2. Vineland Domain Intercepts and 95% Confidence Intervals. This table includes the point estimate and 95% confidence interval for each genetic group in each Vineland domain for the intercept, between-subjects age, between-subjects quadratic term, within-subjects time, and the between*within subjects interaction term.
3. Intellectual Quotient Intercepts and 95% Confidence Intervals. This table includes the point estimate and 95% confidence interval for each genetic group for the Full Scale, Verbal, and Nonverbal IQ for the intercept, between-subjects age, and within-subjects time.
4. ABC Domain Intercepts and 95% Confidence Intervals. This table includes the point estimate and 95% confidence interval for each genetic group in each ABC domain for the intercept, between-subjects age, and within-subjects time. RBS-R Domain Intercepts and 95% Confidence Intervals
5. RBS-R Domain Intercepts and 95% Confidence Intervals. This table includes the point estimate and 95% confidence interval for each genetic group in each RBS-R domain for the intercept, between-subjects age, and within-subjects time. RBS-R Domain Intercepts and 95% Confidence Intervals

1. Vineland within and between subjects effects displayed separately, for each subdomain.

Acknowledgements
This study was funded, in part, by the National Institute of Health Developmental Synaptopathies Consortium (U54NS092090, NCT02461420) and the Intramural Research Program of the National Institute of Mental Health (1ZICMH002961). The Developmental Synaptopathies Consortium (U54NS092090) is part of the Rare Diseases Clinical Research Network (RDCRN), an initiative of the Office of Rare Diseases Research (ORDR), National Center for Advancing Translational Sciences (NCATS).

References

Aman, M. G., & Singh, N. N. (1986). Aberrant behavior checklist. *East Aurora, NY: Slosson*.

Bodfish, J. W., Symons, F. J., Parker, D. E., & Lewis, M. H. (2000). Varieties of repetitive behavior in autism: Comparisons to mental retardation. *Journal of autism and developmental disorders, 30*, 237-243.

Burdeus-Olavarrieta, M., San José-Cáceres, A., García-Alcón, A., González-Peñas, J., Hernández-Jusdado, P., & Parellada-Redondo, M. (2021). Characterisation of the clinical phenotype in Phelan-McDermid syndrome. *J Neurodev Disord, 13*(1), 26. doi:10.1186/s11689-021-09370-5

Chen, L., Yao, Z. Y., Wu, X., He, S. R., Liu, Y. M., Wang, X. Y., . . . Feng, Z. C. (2022). Phelan-McDermid Syndrome in Pediatric Patients With Novel Mutations: Genetic and Phenotypic Analyses. *Front Pediatr, 10*, 888001. doi:10.3389/fped.2022.888001

Cusmano-Ozog, K., Manning, M. A., & Hoyme, H. E. (2007). 22q13.3 deletion syndrome: a recognizable malformation syndrome associated with marked speech and language delay. *Am J Med Genet C Semin Med Genet, 145c*(4), 393-398. doi:10.1002/ajmg.c.30155

Daniel, M. V. L. (2022). [Growth scale value (GSC): Theory, development, and applications.].

De Rubeis, S., Siper, P. M., Durkin, A., Weissman, J., Muratet, F., Halpern, D., . . . Kolevzon, A. (2018). Delineation of the genetic and clinical spectrum of Phelan-McDermid syndrome caused by SHANK3 point mutations. *Molecular autism, 9*, 31-31. doi:10.1186/s13229-018-0205-9

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4: Peabody picture vocabulary test*: Pearson Assessments.

Elliot, C. (2007). Differential abilities scale—2nd edition (DAS-II) manual. *Harcourt Assessment*.

Extension for Community Healthcare Outcomes, P. M. S. C. G. (2023). *Phelan McDermid Syndrome Neuropsychiatric Consultation Group*.

Farmer, C., Thurm, A., Troy, J. D., & Kaat, A. J. (2023). Comparing ability and norm-referenced scores as clinical trial outcomes for neurodevelopmental disabilities: a simulation study. *J Neurodev Disord, 15*(1), 4. doi:10.1186/s11689-022-09474-6

Farmer, C. A., Kaat, A. J., Thurm, A., Anselm, I., Akshoomoff, N., Bennett, A., . . . Miller, J. S. (2020). Person Ability Scores as an Alternative to Norm-Referenced Scores as Outcome Measures in Studies of Neurodevelopmental Disorders. *Am J Intellect Dev Disabil, 125*(6), 475-480. doi:10.1352/1944-7558-125.6.475

Garber, K. B., Visootsak, J., & Warren, S. T. (2008). Fragile X syndrome. *Eur J Hum Genet, 16*(6), 666-672. doi:10.1038/ejhg.2008.61

Girirajan, S., Brkanac, Z., Coe, B. P., Baker, C., Vives, L., Vu, T. H., . . . Silengo, M. (2011). Relative burden of large CNVs on a range of neurodevelopmental phenotypes. *PLoS genetics, 7*(11), e1002334.

Hernandez, R. N., Feinberg, R. L., Vaurio, R., Passanante, N. M., Thompson, R. E., & Kaufmann, W. E. (2009). Autism spectrum disorder in fragile X syndrome: a longitudinal evaluation. *American Journal of Medical Genetics Part A, 149*(6), 1125-1137.

Jeffries, A. R., Curran, S., Elmslie, F., Sharma, A., Wenger, S., Hummel, M., & Powell, J. (2005). Molecular and phenotypic characterization of ring chromosome 22. *American Journal of Medical Genetics Part A, 137A*(2), 139-147.

Klaiman, C., Quintin, E. M., Jo, B., Lightbody, A. A., Hazlett, H. C., Piven, J., . . . Reiss, A. L. (2014). Longitudinal profiles of adaptive behavior in fragile X syndrome. *Pediatrics, 134*(2), 315-324. doi:10.1542/peds.2013-3990

Kolevzon, Delaby, E., Berry-Kravis, E., Buxbaum, J. D., & Betancur, C. (2019). Neuropsychiatric decompensation in adolescents and adults with Phelan-McDermid syndrome: a systematic review of the literature. *Molecular Autism, 10*, 50. doi:10.1186/s13229-019-0291-3

Kolevzon, A. (2023). [Phelan-McDermid Syndrome Extension for Community Healthcare Outcomes].

Landlust, A. M., Visser, L., Flapper, B. C. T., Ruiter, S. A. J., Zwanenburg, R. J., van Ravenswaaij-Arts, C. M. A., & van Balkom, I. D. C. (2022). Understanding Behavior in Phelan-McDermid Syndrome. *Front Psychiatry, 13*, 836807. doi:10.3389/fpsyt.2022.836807

Levy, T., Foss-Feig, J. H., Betancur, C., Siper, P. M., Trelles-Thorne, M. D. P., Halpern, D., . . . Kolevzon, A. (2022). Strong evidence for genotype-phenotype correlations in Phelan-McDermid syndrome: results from the developmental synaptopathies consortium. *Hum Mol Genet, 31*(4), 625-637. doi:10.1093/hmg/ddab280

Levy, T., Lerman, B., Halpern, D., Frank, Y., Layton, C., Zweifach, J., . . . Kolevzon, A. (2022). CHAMP1 disorder is associated with a complex neurobehavioral phenotype including autism, ADHD, repetitive behaviors and sensory symptoms. *Hum Mol Genet, 31*(15), 2582-2594. doi:10.1093/hmg/ddac018

Lindquist, S. G., Kirchhoff, M., Lundsteen, C., Pedersen, W., Erichsen, G., Kristensen, K., . . . Brøndum-Nielsen, K. (2005). Further delineation of the 22q13 deletion syndrome. *Clinical Dysmorphology, 14*(2), 55-60. doi:10.1097/00019605-200504000-00001

Mullen, E. M. (1995). *Mullen scales of early learning manual*: American Guidance Service.

Nevado, J., García-Miñaúr, S., Palomares-Bralo, M., Vallespín, E., Guillén-Navarro, E., Rosell, J., . . . Lapunzina, P. (2022). Variability in Phelan-McDermid Syndrome in a Cohort of 210 Individuals. *Front Genet, 13*, 652454. doi:10.3389/fgene.2022.652454

Peters, S. U., Horowitz, L., Barbieri-Welge, R., Taylor, J. L., & Hundley, R. J. (2012). Longitudinal follow-up of autism spectrum features and sensory behaviors in Angelman syndrome by deletion class. *Journal of Child Psychology and Psychiatry, 53*(2), 152-159.

Roid, G. H., & Pomplun, M. (2012). *The stanford-binet intelligence scales*: The Guilford Press.

Samogy-Costa, C. I., Varella-Branco, E., Monfardini, F., Ferraz, H., Fock, R. A., Barbosa, R. H. A., . . . Passos-Bueno, M. R. (2019). A Brazilian cohort of individuals with Phelan-McDermid syndrome: genotype-phenotype correlation and identification of an atypical case. *Journal of Neurodevelopmental Disorders, 11*(1), 13-13. doi:10.1186/s11689-019-9273-1

Sarasua, S. M., Boccuto, L., Sharp, J. L., Dwivedi, A., Chen, C. F., Rollins, J. D., . . . DuPont, B. R. (2014). Clinical and genomic evaluation of 201 patients with Phelan-McDermid syndrome. *Human Genetics, 133*(7), 847-859. doi:10.1007/s00439-014-1423-7

Sarasua, S. M., Dwivedi, A., Boccuto, L., Chen, C.-F., Sharp, J. L., Rollins, J. D., . . . DuPont, B. R. (2014). 22q13.2q13.32 genomic regions associated with severity of speech delay, developmental delay, and physical features in Phelan–McDermid syndrome. *Genetics in Medicine, 16*(4), 318-328. doi:10.1038/gim.2013.144

Sarasua, S. M., Dwivedi, A., Boccuto, L., Rollins, J. D., Chen, C. F., Rogers, R. C., . . . Collins, J. S. (2011). Association between deletion size and important phenotypes expands the genomic region of interest in Phelan-McDermid syndrome (22q13 deletion syndrome). *Journal of Medical Genetics, 48*(11), 761-766. doi:10.1136/jmedgenet-2011-100225

Schön, M., Lapunzina, P., Nevado, J., Mattina, T., Gunnarsson, C., Hadzsiev, K., . . . Hennekam, R. C. (2023). Definition and clinical variability of SHANK3-related Phelan-McDermid syndrome. *European Journal of Medical Genetics, 66*(7), 104754. doi:https://doi.org/10.1016/j.ejmg.2023.104754

Serrada-Tejeda, S., Martínez-Piédrola, R. M., Máximo-Bocanegra, N., Sánchez-Herrera-Baeza, P., & Pérez-de-Heredia-Torres, M. (2022). Descriptive Analysis of Adaptive Behavior in Phelan-McDermid Syndrome and Autism Spectrum Disorder. *Front Neurosci, 16*, 893003. doi:10.3389/fnins.2022.893003

Soorya, L., Kolevzon, A., Zweifach, J., Lim, T., Dobry, Y., Schwartz, L., . . . Buxbaum, J. D. (2013). Prospective investigation of autism and genotype-phenotype correlations in 22q13 deletion syndrome and SHANK3 deficiency. *Molecular Autism, 4*(1), 18. doi:10.1186/2040-2392-4-18

Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland adaptive behavior scales Vineland-II: Survey forms manual*: Pearson Minneapolis, MN.

Tabet, A. C., Rolland, T., Ducloy, M., Lévy, J., Buratti, J., Mathieu, A., . . . Bourgeron, T. (2017). A framework to identify contributing genes in patients with Phelan-McDermid syndrome. *NPJ Genomic Medicine, 2*, 32. doi:10.1038/s41525-017-0035-2

Tang, L., Levy, T., Guillory, S., Halpern, D., Zweifach, J., Giserman-Kiss, I., . . . Grice, D. E. (2021). Prospective and detailed behavioral phenotyping in DDX3X syndrome. *Mol Autism, 12*(1), 36. doi:10.1186/s13229-021-00431-z

Trelles, M. P., Levy, T., Lerman, B., Siper, P., Lozano, R., Halpern, D., . . . Buxbaum, J. (2021). Individuals with FOXP1 syndrome present with a complex neurobehavioral profile with high rates of ADHD, anxiety, repetitive behaviors, and sensory symptoms. *Mol Autism, 12*(1), 61. doi:10.1186/s13229-021-00469-z

Varela, M. C., Kok, F., Otto, P. A., & Koiffmann, C. P. (2004). Phenotypic variability in Angelman syndrome: comparison among different deletion classes and between deletion and UPD subjects. *Eur J Hum Genet, 12*(12), 987-992. doi:10.1038/sj.ejhg.5201264

Wasserstein, R. L., Schirm, A. L., & Lazar, N. A. (2019). Moving to a World Beyond "$p < 0.05$". *The American Statistician, 73*(sup1), 1-19. doi:10.1080/00031305.2019.1583913

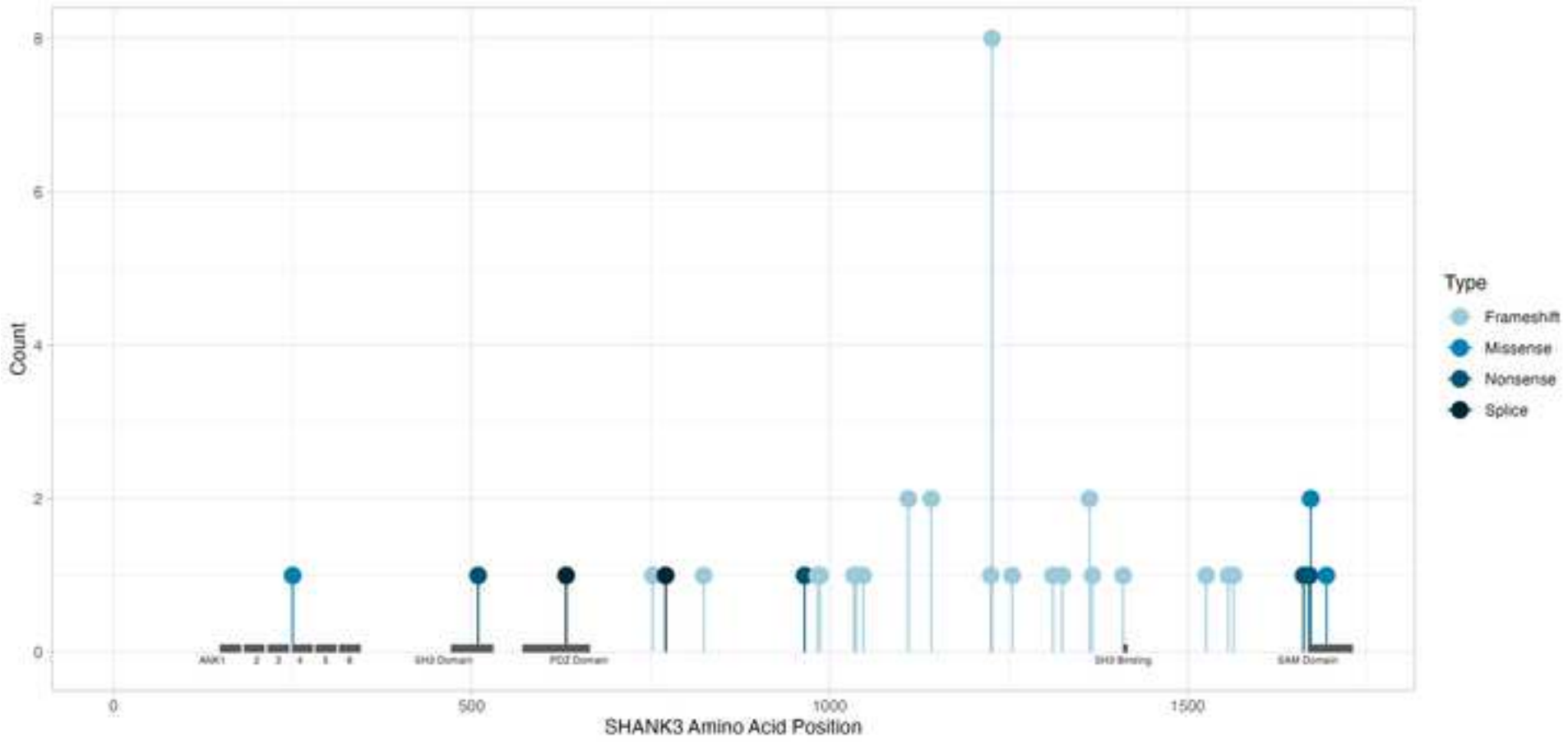
Williams, K. T. (1997). Expressive vocabulary test second edition (EVT™ 2). *J. Am. Acad. Child Adolesc. Psychiatry, 42*, 864-872.

Wilson, H. L., Wong, A. C., Shaw, S. R., Tse, W. Y., Stapleton, G. A., Phelan, M. C., . . . McDermid, H. E. (2003). Molecular characterisation of the 22q13 deletion syndrome supports the role of haploinsufficiency of SHANK3/PROSAP2 in the major neurological symptoms. *Journal of Medical Genetics, 40*(8), 575-584. doi:10.1136/jmg.40.8.575

Xu, N., Lv, H., Yang, T., Du, X., Sun, Y., Xiao, B., . . . Yu, Y. (2020). A 29 Mainland Chinese cohort of patients with Phelan–McDermid syndrome: genotype–phenotype correlations and the role of SHANK3 haploinsufficiency in the important phenotypes. *Orphanet Journal of Rare Diseases, 15*(1), 335. doi:10.1186/s13023-020-01592-5

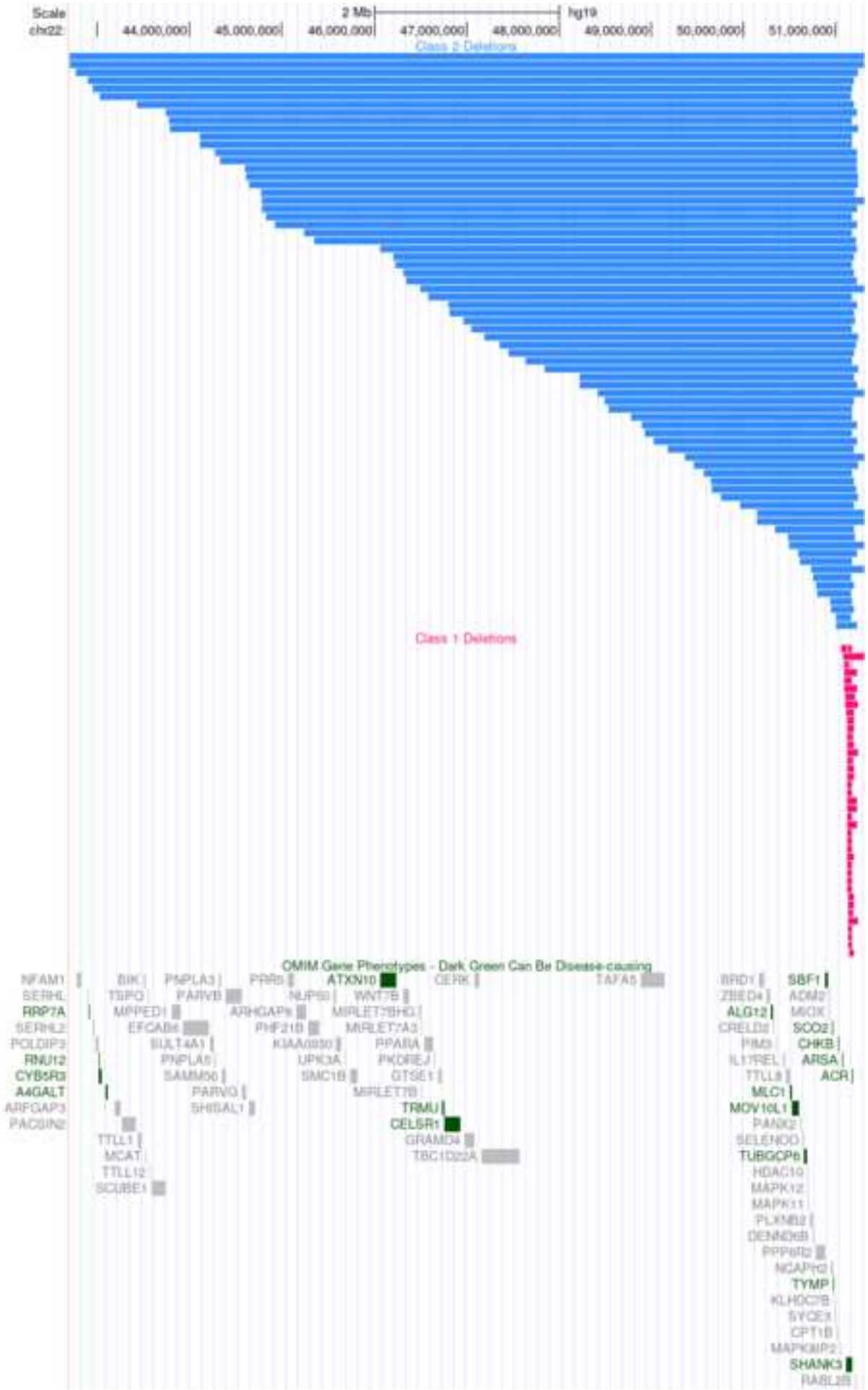
1a

Click here to access/download;Figure;Fig 1a.png



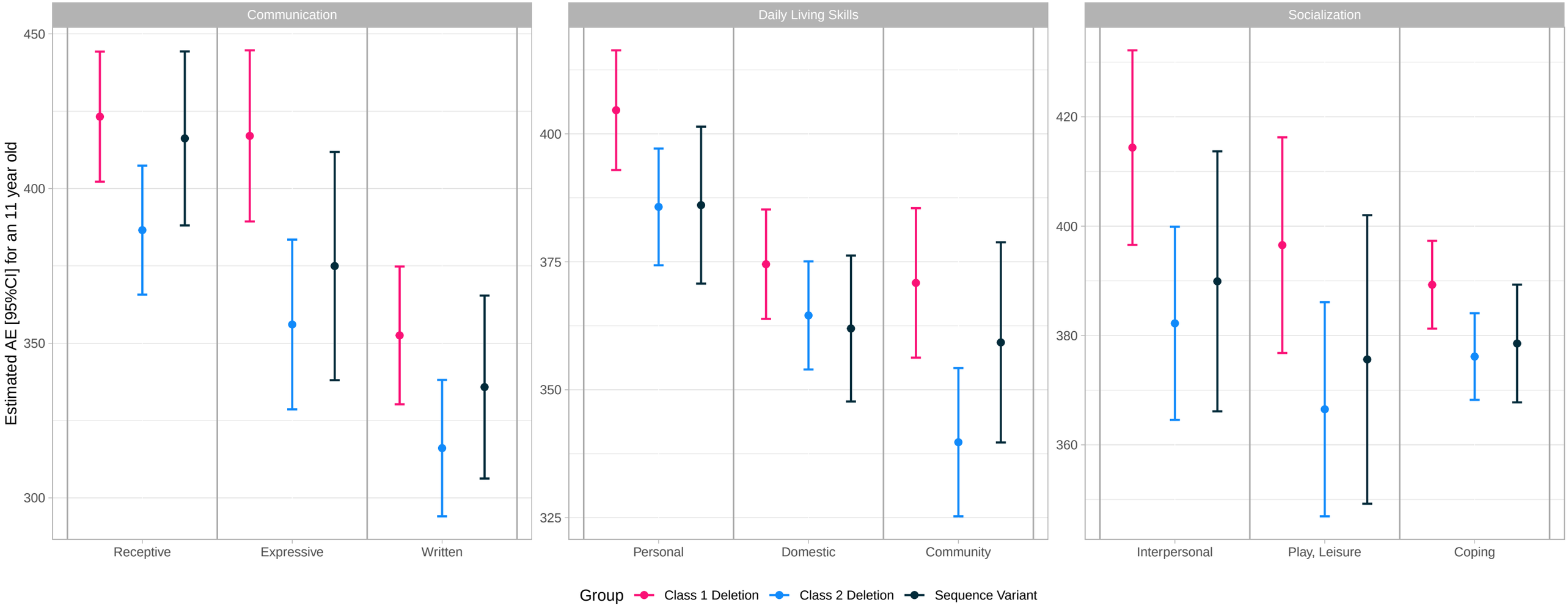
1b

Click here to access/download;Figure;Fig 1b.png

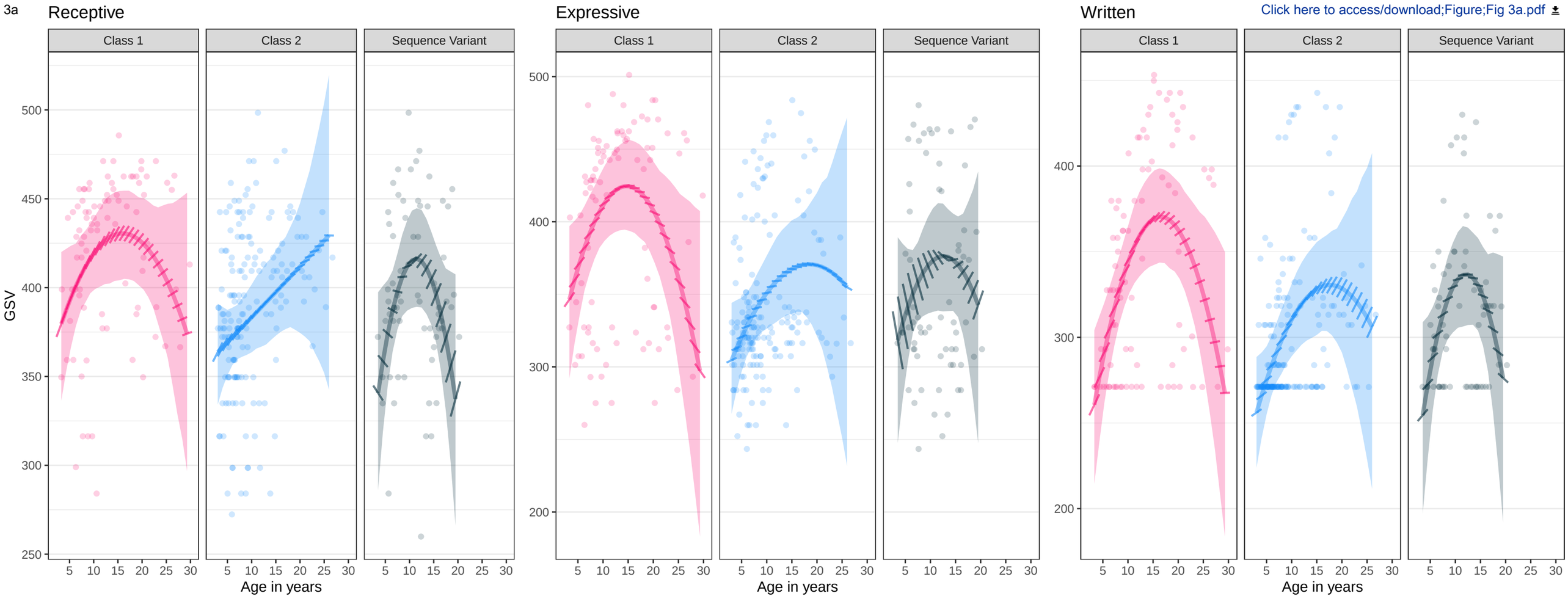


Representation of Vineland Subdomain Intercepts Using Age Equivalents

Click here to access/download;Figure;Fig 2.pdf

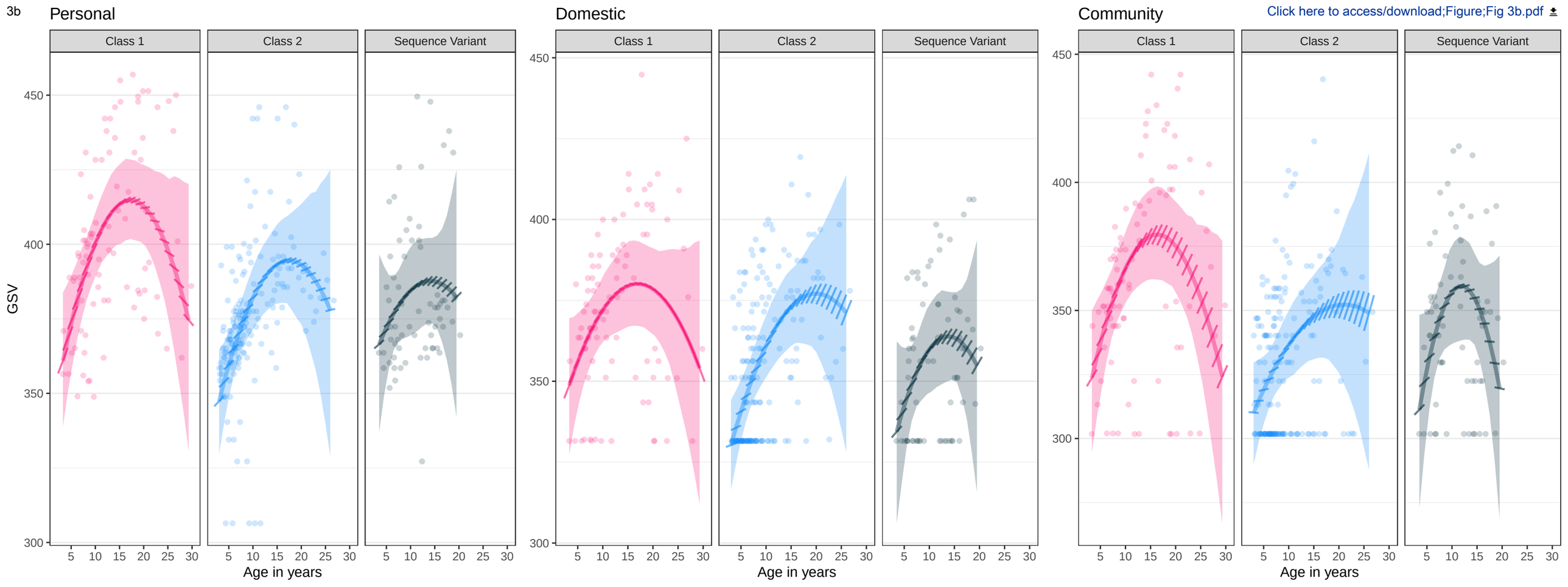


3a



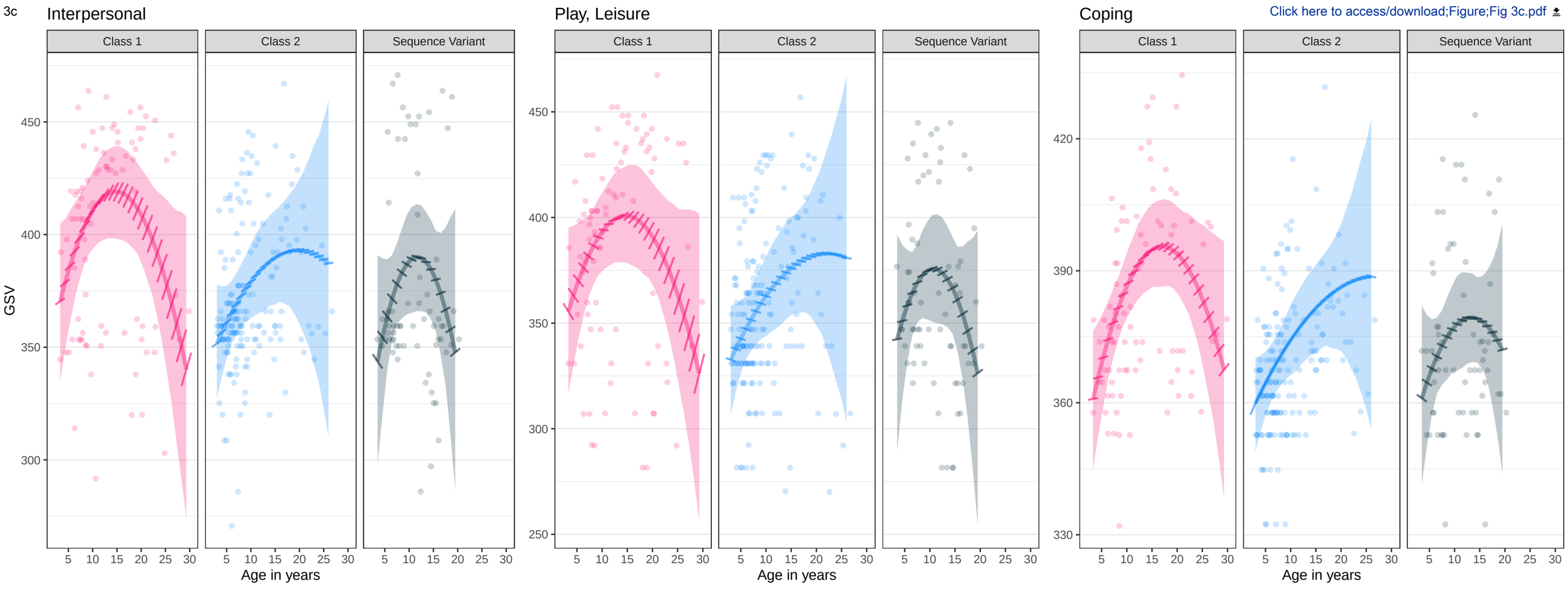
Click here to access/download;Figure;Fig 3a.pdf

3b



Click here to access/download;Figure;Fig 3b.pdf

3c



Click here to access/download;Figure;Fig 3c.pdf

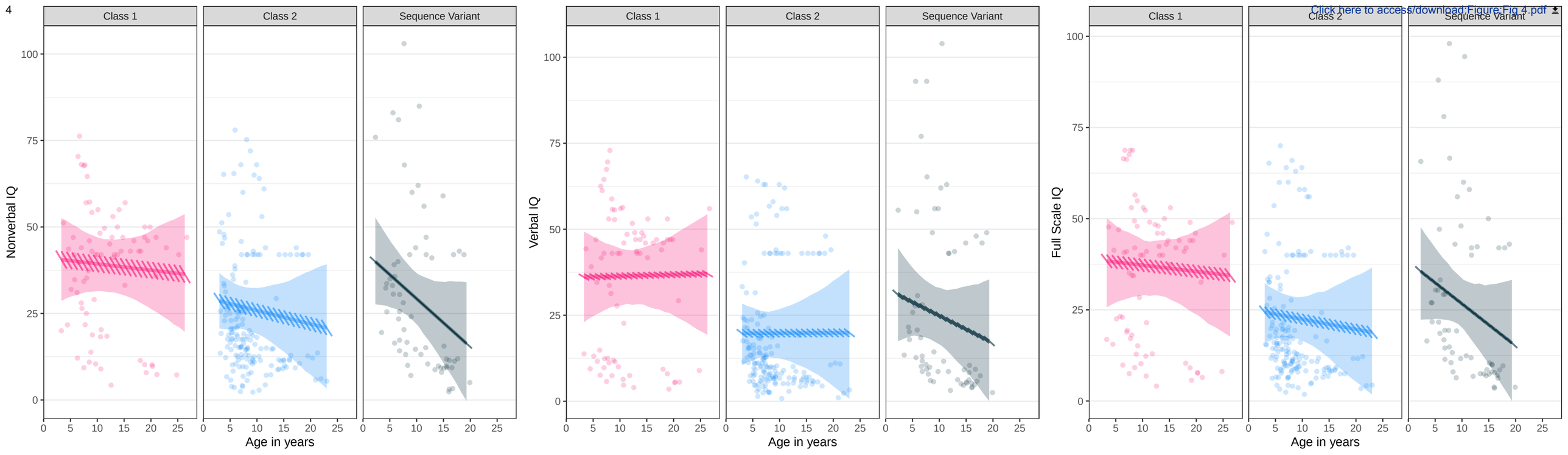


Figure 1. Participant Genetics
Legend: Genetic representation of the cohort. A. Sequence variants are shown in a lollipop plot based on amino acid position and frequency of variant; colors indicate different types of variants. B. Participant deletions are mapped onto hg19 chromosome 22q13. Each horizonal bar represents one participant. Blue represents Class 2 deletions; pink represents Class 1 deletions. OMIM genes are displayed underneath, representing genes in the area.

Figure 2. Representation of Vineland Subdomain Intercepts Using Age Equivalents
Legend: An illustration of the estimated age equivalents corresponding to the GSV intercepts (i.e., estimated average score for a genetic group at the mean age of 11 years or 132 months). Estimated age equivalents are shown in months.

Figure 3. Vineland-2 Subdomain GSV Personal Trajectory and Effect of Age
Legend: The plots show the estimated within-person change as a function of between-subject age. Individual observed scores are represented by scatter points. The bolded line shows the between-subjects effect of age (i.e., model-estimated score given the average age on the X-axis). Individual lines illustrate the model-estimated within-person trajectories (i.e., the estimated rate of change in score for an individual during study participation). A. Communication subdomains, including Receptive, Expressive and Written Communication Skills. B. Daily Living Skills subdomains including Personal, Domestic, and Community Daily Living Skills. C. Socialization subdomains including Interpersonal, Play and Leisure, and Coping Socialization Skills.

Figure 5. Intellectual Quotient Personal Trajectory and Effect of Age
Legend: The plots show the estimated within-person change as a function of between-subject age. Individual observed scores are represented by scatter points. The bolded line shows the between-subjects effect of age (i.e., model-estimated score given the average age on the X-axis). Individual lines illustrate the model-estimated within-person trajectories (i.e., the estimated rate of change in score for an individual during study participation).