

Milestone Development in Genetic Conditions from SFARI Registries

Jordan Wickstrom, PhD,¹ Cristan Farmer, PhD,¹ LeeAnne Green Snyder, PhD,² Andrew R. Mitz,

PhD,¹ Stephan J. Sanders, BMBS PhD,³ and Somer Bishop, PhD,^{3†} Audrey Thurm, PhD^{1†}

¹Neurodevelopmental and Behavioral Phenotyping Service, National Institutes of Health,

Bethesda, MD

²SFARI, Simons Foundation, New York, NY

³Department of Psychiatry and Behavioral Sciences, UCSF Weill Institute for Neurosciences,

University of California, San Francisco, San Francisco, CA 94158

Conflicts of Interest: Somer Bishop receives royalties from WPS for sales of the ADOS-2 manual and protocols.

Abstract

Background: Several recent initiatives have sought to better understand characteristic behaviors of rare genetic conditions associated with autism spectrum disorder (ASD). The onset of developmentally expected skills, such as walking and talking, serve as readily quantifiable aspects of the behavioral phenotype. The goals of this study were to describe attainment of major motor and language milestones in genetic conditions implicated in the etiology of ASD and other neurodevelopmental disorders and to compare those phenotypes to idiopathic ASD.

Methods: Participants ages 3 years and older were drawn from two Simons Foundation Autism Research Initiative-funded registries with consistent phenotyping protocols. Inclusion criteria consisted of a confirmed genetic diagnosis for one of 16 specific genetic conditions (Simons Searchlight), and absence of pathogenic genetic findings in idiopathic ASD controls (SPARK). Parent-reported age of acquisition of three motor and two language milestones was described and quantified as on-time or delayed relative to normative expectations.

Results: Delay was more common among participants with rare genetic conditions than idiopathic ASD for all milestones. Compared to the idiopathic ASD group, the median odds of delay among the genetic groups were 8.3 times (IQR 5.8-16.3) higher for sitting, 12.4 times (IQR 5.3-19.5) higher for crawling, 26.8 times (IQR 7.7-41.1) higher for walking, 2.7 times (IQR 1.7-5.5) higher for single words, and 5.7 times (IQR 2.8-18.3) higher for combined words.

Conclusion: Delays in major developmental milestones, particularly in motor skills, may be among the earliest clues that developmental processes may be differentially affected in specific genetically defined conditions versus a behaviorally defined disorder such as idiopathic ASD.

Keywords: copy number variant; behavior; intellectual disability; developmental phenotype

DEVELOPMENTAL MILESTONES 3

Abbreviations: Autism Spectrum Disorder (ASD), Copy Number Variants (CNVs), Intellectual Disability (ID), Neurodevelopmental Disorders (NDDs), Simons Foundation Powering Autism Research (SPARK)

DEVELOPMENTAL MILESTONES 4

Several recent initiatives have emerged to better describe behaviors characteristic of genetic conditions implicated in the etiology of autism spectrum disorder (ASD) and related neurodevelopmental disorders (NDDs). Large-scale collaborations (e.g., Simons Simplex Collection; Fischbach & Lord, 2010; Simons VIP Consortium, 2012) have made available larger cohorts of individuals with specific genetic diagnoses, allowing for a genetics-first approach to examining phenotypes. Defining the phenotypic landscape of these genetic conditions is important for improving identification and clinical management. Additionally, given their established association with ASD, characterizing these conditions may promote understanding of the etiological mechanisms underlying ASD-related phenotypes (Sanders et al., 2019).

Information regarding attainment of developmental milestones, such as walking and talking, is one way to quantify very early aspects of the phenotype. Extensive data on such developmental milestones are available for the general population (Sheldrick et al., 2019; Stør vold, Aarethun, & Bratberg, 2013; Taanila, Murray, Jokelainen, Isohanni, & Rantakallio, 2005; World Health Organization, 2006). Developmental milestone attainment, particularly in motor and language skills, has also been a focal point in ASD research. Among children with ASD, early motor milestones are generally achieved within normal age limits (Bishop, Thurm, Farmer, & Lord, 2016; Havdahl et al., 2020; Matson, Mahan, Kozlowski, & Shoemaker, 2010), while the onset of language milestones is more variable (Eigsti, de Marchena, Schuh, & Kelley, 2011; Mayo, Chlebowski, Fein, & Eigsti, 2013; Mitchell et al., 2006). Behavioral phenotype data regarding early development (“developmental phenotypes”) from recently identified rare genetic conditions associated with ASD are beginning to emerge (Arnett et al., 2020; Bernier et al., 2017; Chawner et al., 2019; Hinton et al., 2013; Winders, Wolter-Warmerdam, & Hickey, 2019). Evidence from a phenotype-first perspective indicates that among children with ASD, those with

identifiable rare genetic conditions walk later than those without known genetic conditions (Bishop et al., 2017).

The goal of the current study was to extend the limited literature on developmental phenotypes in multiple recently-identified genetic conditions implicated in ASD (1q21.1 deletion, 1q21.1 duplication, 16p11.2 deletion, 16p11.2 duplication, *ADNP*, *ASXL3*, *CSNK2A1*, *DYRK1A*, *GRIN2B*, *MED13L*, *PACSI*, *PPP2R5D*, *SCN2A*, *SLC6A1*, *STXBPI*, *SYNGAP1*), using contemporaneously and systematically collected data regarding motor and language milestone acquisition (Simons VIP Consortium, 2012). We compare these developmental phenotypes to those of a sample of children with ASD with no identified genetic condition (SPARK Consortium, 2018). Based on limited reports on milestone attainment in these genetic conditions in the literature, we expected to observe extensive delays relative to idiopathic ASD.

Methods

Cohorts

Simons Searchlight

Simons Searchlight, an initiative launched by the Simons Foundation Autism Research Initiative (SFARI) (previously known as Simons Variation in Individuals Project) (Simons VIP Consortium, 2012), is comprised of groups of people with rare genetic variants implicated in ASD. While nearly all participants have developmental disabilities of varying severity, not all participants have ASD. Individuals with a genetic diagnosis on the Simons Searchlight Gene List (<https://www.simonssearchlight.org/research/what-we-study/>) of 148 gene changes and 20 copy number variants (CNVs) enrolled in the registry from 2014 to September 2020. Participants were

recruited worldwide by the Simons Searchlight community website (<https://www.simonssearchlight.org>) as well as with social media outreach, clinical referrals, and other means of web recruitment (e.g., <http://www.clinicaltrials.gov>) following diagnosis of one of the genetic conditions in Simons Searchlight. The protocol consists of developmental and behavioral surveys in an online study portal, a telephone interview to obtain medical history and administer the Vineland Adaptive Behavior Scales (Sparrow, Cicchetti, & Balla, 2005; Sparrow, Cicchetti, & Saulnier, 2016), and the collection of biospecimens. Files from Version 7 were downloaded on September 29, 2020 from SFARI.org, and included the 1q21.1, 16p11.2, and single gene mutation datasets.

SPARK

The SPARK registry is the largest existing genetic study of ASD (SPARK Consortium, 2018). Individuals from the U.S. with a reported professional ASD diagnosis and their family members are eligible for enrollment in SPARK, regardless of age or genetic status. Families are recruited by 31 sites across the U.S. using an extensive web recruitment strategy. Similar to Simons Searchlight, medical and developmental history, current behavioral data, and biospecimens were collected remotely through an online portal beginning in late 2015. Version 4 was downloaded on April 30, 2020 from SFARI.org.

Sample Selection

The Supplemental Information contains a detailed description of how data were combined, as well as a flow diagram of sample selection (Figure S1). From both cohorts (i.e., Simons Searchlight and SPARK), participants for the current study were probands at least 3 years of age with verified genetic results, and at least one valid response for any of the five

developmental milestones of interest. The minimum age was selected because normative acquisition of the milestones considered in these analyses occurs before the age of 3 years (Centers for Disease Control, June 10, 2020; Sheldrick et al., 2019; World Health Organization, 2006). This allowed for the interpretation of milestones which were reportedly “not yet achieved” as “delayed.”

Cohort-specific inclusion criteria were as follows. Simons Searchlight groups containing at least 15 cases with available walking milestone data were included in the genetic conditions sample, for a total of 479 participants across 16 conditions: four CNVs (1q21.1 deletion, 1q21.1 duplication, 16p11.2 deletion, 16p11.2 duplication) and twelve single genes (*ADNP*, *ASXL3*, *CSNK2A1*, *DYRK1A*, *GRIN2B*, *MED13L*, *PACS1*, *PPP2R5D*, *SCN2A*, *SLC6A1*, *STXBP1*, *SYNGAP1*). Simons Searchlight participants both with and without reported ASD diagnoses were included. Proband from the SPARK sample were included in the idiopathic ASD group (n=3,506) if they underwent genetic analyses and were found to have no pathogenic single gene or CNV events strongly associated with ASD based on SNP genotyping array, whole exome sequencing, and review by the SPARK medical genetics committee (Feliciano et al., 2019). Adults who independently enrolled in Simons Searchlight and SPARK were not included in this study, because reports of milestone attainment are only provided by caregivers of dependent children or adults.

Measures

Upon enrolling in either registry, participants were invited to complete a standard set of questionnaires. All questionnaires for participants in the current study were completed by parents or caregivers. For a list of exact variables and datasets used for each registry, please refer to

Figure S2. Whereas specific measures used for medical history and other data collection varied by registry (see subsections below), the developmental milestone and demographic data were collected via a Background History form in both registries.

The primary outcomes included ages of milestone attainment for sitting, crawling, walking, using single words, and using combined words. For each milestone, respondents chose one of the following: attainment age in months (selected from 1-84 months), achieved after age 7 years, or not yet achieved. However, the response options for the milestone questions changed slightly during the data collection period; the option of “not yet achieved” was introduced in 2018. Prior to this, parents of children who did not acquire a skill may have left the age of milestone attainment blank. To the extent that this is true, the observed data are biased toward children who had acquired a milestone.

For this reason, we explored options for partially imputing missing data on milestone acquisition using scores from the Vineland-II Adaptive Behavior Scales Survey Interview Form (Sparrow et al., 2005). The Vineland-II is a semi-structured interview that can be used across the lifespan. It measures adaptive functioning in four domains (with corresponding subdomains): Communication (Expressive, Receptive, Written), Daily Living (Personal, Domestic, Community), Socialization (Interpersonal Relationships, Play and Leisure Time, Coping Skills), and, in children under age 7, Motor Skills (Gross, Fine). Where available, scores from the Vineland-II were used to estimate whether a participant with missing milestone acquisition data had acquired walking and single words (see Supporting Information for a detailed explanation). Item-level data were not available, so estimated minimum raw scores compatible with the ability to walk (Gross Motor = 27) and use single words (Expressive Language = 19) were used to determine whether or not a participant had attained the milestone. Attainment scores were

DEVELOPMENTAL MILESTONES 9

imputed for 4% of the sample for walking and 14% of the sample for single words. Raw scores less than the threshold were used to code the milestone as “not yet achieved.” If the raw scores exceeded the threshold, the milestone was coded as “achieved, age unknown.”

Finally, “delay” in the acquisition of milestones was operationalized using normative data. Specifically, motor milestones were considered to be delayed if they occurred after 8 months for sitting, 12 months for crawling, and 16 months for walking, based on the 97th percentiles (8.4, 12, and 16, respectively) reported for normative data (World Health Organization, 2006). Language milestones were defined as delayed after 12 months for single words and 24 months for combined words based on the Act Early CDC Recommendations (Centers for Disease Control, June 10, 2020). Milestones which were not attained were considered delayed for the purpose of analysis. For descriptive purposes, the degree of delay was categorized as attainment occurring within 6 months or beyond 6 months (including those who never attained) of the expected age of acquisition.

Simons Searchlight

Parent-reported diagnoses of ASD, Intellectual Disability (ID), Seizure Disorder/Epilepsy, and gestational age were drawn from the medical history interview. Parents or caregivers also participated in the Vineland-II with licensed genetic counselors over the phone at the time of the medical history interview. Raw scores on the Expressive Communication and Gross Motor subdomains at first administration of the Vineland-II were used to guide the imputation of some missing milestone data, as described above.

SPARK

Parent-reported diagnosis of ID, Seizure Disorder/Epilepsy, gestational age, and additional demographics (i.e., ancestry) were drawn from registration, a background history form, and a medical screening survey.

Statistical Analysis

The goals of this study were primarily descriptive. However, it was necessary to quantify the degree of difference in milestone acquisition between the genetic conditions and the idiopathic ASD sample. For the sake of clinical interpretability, we operationalized the outcome as on-time or delayed acquisition of a given milestone as described above. Given the uniform delay amongst some of the genetic conditions, we used the penalized likelihood-based method Firth logistic regression, using the `logistf` package for R version 4.0.2 (Heinze, Ploner, & Jiricka, 2020). Absolute (rate) and relative (odds ratios; OR) probabilities of delay, based on genetic condition, are provided alongside 95% confidence intervals. To further describe the patterns of onset, age of acquisition for those who achieved a milestone prior to age 7 years is summarized for each group using the median and interquartile range. This was possible only for those who acquired the milestone prior to 7 years due to the response options on the form. Proportions of the sample are also presented for those who acquired the skill after age 7 years or at an unknown age, did not acquire the skill, or had missing data regarding acquisition of that skill.

Missing Data

Non-response occurred for 179 participants (4% sample) on one milestone (Simons Searchlight n=46, SPARK n=133), 102 participants (3% sample) on two milestones (Simons Searchlight n=52, SPARK n=50), and 74 participants (2% sample) on three or four milestones (Simons Searchlight n=35, SPARK n=39). We also imposed one logical constraint on the data

that created missing data: where the age of using combined words was younger than the age of using single words, we treated both as missing (Simons Searchlight n=5; SPARK n=15). To the extent possible, missing data were imputed as acquired/not acquired, as described above. Vineland-II data were available only for Simons Searchlight, and were not available for all participants who were missing data. Data were imputed for n=19 of 28 Simons Searchlight participants missing walking and n=69 of 80 missing single word. No other imputation of missing data was performed.

Results

Sample demographics are provided in Table 1. The idiopathic ASD group had a higher proportion of males (82%) compared with the genetic conditions sample (53%). Half the samples maintained an annual household income of \$81,000 or above (54% and 52%, respectively). For the genetic and idiopathic samples, the parent-reported prevalence of ASD was 37% and 100%, ID was 25% and 13%, Seizure Disorder/Epilepsy was 35% and 4%, and premature birth was 12% and 10%, respectively.

Descriptive statistics (quartiles and proportions) for milestone acquisition are provided in Table 2. Means and standard deviations for each milestone, if attained, are provided in Table S1. Overall, probands with genetic conditions were more likely to exhibit both lack of attainment and delayed attainment of milestones than the idiopathic ASD comparison group. Delays were most pronounced for motor milestones. Compared to the idiopathic ASD group, the median odds of delay for the genetic condition groups were higher by 8.3 times (IQR 5.8-16.3) for sitting, 12.4 times (IQR 5.3-19.5) for crawling, and 26.8 times (IQR 7.7-41.1) for walking. For the

language milestones, the median odds of delay for the genetic condition groups were higher than the idiopathic ASD group by 2.7 times (IQR 1.7-5.5) for single words and 5.7 times (IQR 2.8-18.3) for combined words. The rate and extent of delay in milestone attainment is illustrated in Figure 1.

As shown in Table 2 and Figure 1, at least 50% of the sample was delayed in sitting for 10 genetic conditions, crawling for eight conditions, and walking for 12 conditions. All groups, including idiopathic ASD, had a rate of delay $\geq 50\%$ for language milestones. Eight genetic conditions (*ADNP*, *ASXL3*, *GRIN2B*, *MED13L*, *PACSI*, *PPP2R5D*, *STXBPI*, and *SYNGAP1*) had delays for *all* milestones in at least half the sample.

Discussion

This study examined developmental phenotypes across multiple rare genetic conditions, several of which were recently recognized, and compared them to a group identified through genetic analyses as idiopathic ASD. This study extends previous work in the Simons Simplex Collection (Bishop et al., 2017; Buja et al., 2018) by examining multiple early developmental milestones (sitting, crawling, walking, single-word talking, combined-word talking) from systematically-collected data on selected rare genetic conditions associated with ASD, and considering the genetic conditions individually rather than in aggregate. By comparing developmental phenotypes across specific genetic conditions to that of an idiopathic ASD group, we were able to add to a growing literature about diverse patterns of very early motor and language skill acquisition among probands with genetic conditions that are associated with higher rates of ASD and other NDDs. As anticipated, the early developmental milestone profiles

of probands with genetic conditions were marked by more extensive delays (including lack of attainment) than the idiopathic ASD group, especially for the motor milestones.

While phenotypic data are just now becoming available from some of the genetic conditions reported here (Arnett et al., 2020; Berg, Palac, Wilkening, Zelko, & Schust Meyer, 2020; Hanly, Shah, Au, & Murias, 2020), the availability of systematically collected developmental and other behavioral phenotypic data from both the Simons Searchlight and SPARK registries allowed for cross-registry comparisons. Our focus on developmental milestones augments the available literature on cross-sectional cognitive and psychiatric profiles by documenting the earliest manifestations of phenotypes which may cascade into lifelong neurodevelopmental conditions. Examination of key motor and language milestones indicates that most children with the 16 genetic conditions reported here exhibit delays in both of these domains. In particular, varying patterns of delays in motor milestones were found across groups, with greater delays on average found in the single gene conditions compared to the CNVs included in this study. Of note, compared to single gene conditions, microarray testing for CNVs has been available longer, costs less, is more widely available, and is more established as a primary genetic test, all of which may have contributed over time to ascertainment of milder phenotypes in the population. Language delays on the other hand, tended to be consistently delayed across all genetic conditions, although there was much more variability in the complete lack of attainment of language milestones in the single gene conditions (3-56% single words; 12-58% combined words) compared to the CNVs or idiopathic ASD group (0-3% for single words, 0-10% for combined words). The genetic groups with the largest proportion of language delays also had the largest proportion of motor delays, which is consistent with previous literature

suggesting that these systems are strongly related (Ghassabian et al., 2016; Libertus & Hauf, 2017).

The difference between idiopathic ASD and the genetic conditions was most obvious for the motor milestones. While this study is not the first to note early motor impairments associated with rare genetic mutations (Buja et al., 2018), or generally intact motor milestones of children diagnosed with idiopathic ASD (Bishop et al., 2016; Havdahl et al., 2020; Matson et al., 2010; Ozonoff et al., 2008), it extends the findings of others by documenting this pattern across multiple genetic condition groups. These results reveal the great extent and variability of delays, and support the idea that severe motor delays may be indicative of genetic etiologies (Bishop et al., 2017; Satterstrom et al., 2020). Our description of delays using established norms and demarcation of delay exceeding 6 months (including those who never attained) illustrates the magnitude of delay in the selected genetic conditions, and suggests that exacerbated motor delays are a likely predictor of these types of genetic abnormalities. By contrast, delays in language milestones were ubiquitous across all the genetic conditions, as well as the idiopathic ASD group, confirming previous research in ASD that delays in early language are not specific to probands with an identifiable genetic condition.

Single gene conditions, such as those included in this study, may be induced in animal and cell culture model systems, providing a mechanism to investigate pathogenesis of ASD-related phenotypes. Questions regarding whether these models can actually provide insight into *all* cases of idiopathic ASD, which models might be best-suited to this objective, and whether ASD neurobiology can be distinguished from broader developmental disabilities, have yet to be addressed. Answering these questions will require a more detailed understanding of the relationships between developmental delays and ASD; however, thus far both gene discovery

and genotype-phenotype analyses support a model in which single gene mutations seem to vary in their contributions to developmental delay versus ASD (Satterstrom et al., 2020). Our results are consistent with this model. Single gene disorders with high rates of significant delays in early milestones (e.g. *PPP2R5D*, *ASXL3*) may be observed more frequently in developmental delay/ID cohorts than in ASD cohorts, while those with relatively lower (but still clinically significant) rates of delayed milestones (e.g. *SLC6A1*, *SCN2A*) may be observed at similar rates in developmental delay/ID cohorts *as well as* in ASD cohorts (Satterstrom et al., 2020). Of note, none of the genes observed at a higher rate in ASD than developmental delay (e.g., *PTEN*, *NRXN1*) from Satterstrom et al. (2020) were included in this study. With larger sample sizes and consistent ascertainment, developmental phenotype results across gene groups are likely to provide critical benchmarks to orient the results of model system analyses (Sanders et al., 2019).

Limitations

A main limitation of this study is that the sample was drawn from registries based on self-referral, and small sample sizes precluded the inclusion of some groups (only 16 of the 29 genetic conditions with data currently available in Simons Searchlight are represented here). Further, participants in the Simons Searchlight registry were enrolled based on genetic testing, which could have been precipitated by delays in early milestones, as developmental delay is one of the main indications for genetic testing. These results may therefore reflect an ascertainment bias in the genetic condition group. These registries are not population-based, so it is possible that individuals with more severe phenotypes are not represented. In this way, the extent to which ascertainment strategies may affect these results likely varies by the different conditions and diagnoses. At present, we do not consider a multitude of factors that may be affecting

variability within genetic subgroups, including inheritance and variant effect (e.g., missense vs. protein-truncating variant). In addition, the idiopathic ASD group included here may go on to receive a genetic diagnosis as new genetic discoveries are made. Still, a major strength of this study is that the idiopathic group was defined not by the lack of any “known genetic syndrome,” as is common in ASD research, but rather confirmed to not have significant genetic findings based on the current state of the field.

Given the high socioeconomic status and the distribution and missing data on race of the participants in the registries, these results may not be generalizable to the larger population of probands with these conditions. In addition, the age of milestone acquisition was not collected past the age of 7 years, and missing data were common, so the summary statistics for acquisition may be biased.

Conclusions

For children with genetic conditions associated with ASD, findings here provide evidence that delays in motor and language milestones are among the early indicators of a diverging developmental trajectory. Motor delays may distinguish children with specific identifiable genetic etiologies from those with idiopathic ASD and should prompt early referral for genetic testing. Further, given its different profile, currently defined “idiopathic” ASD may reflect differing etiologies and mechanisms from rare genetic variants. Broader and deeper data are needed to investigate the possibility of distinguishing among genetic conditions on the basis of phenotypic profiles.

Acknowledgements

This research was supported (in part) by the Intramural Research Program of the NIMH (1ZICMH002961), an NIMH grant (U01 MH111662, SJS), and by an Autism Science Foundation (SJS, SB) grant. We are grateful to all of the families at the participating Simons Searchlight sites as well as the Simons Searchlight Consortium. We appreciate the opportunity to obtain access to Simons Searchlight phenotypic data on the SFARI Base. Approved researchers can obtain the Simons Searchlight population dataset described in this study <https://www.sfari.org/resource/simons-searchlight/> by applying at <https://base.sfari.org>. In addition, we are thankful for the chance to gain access to the SPARK phenotypic data on the SFARI Base. We are grateful to all of the families in SPARK, the SPARK clinical sites and SPARK staff. Approved researchers can obtain the SPARK population dataset described in this study <https://www.sfari.org/resource/spark/> by applying at <https://base.sfari.org>.

- **What's Known**
 - There is an increasing number of genes and genetic conditions being identified that impart risk for autism spectrum disorder and other neurodevelopmental disorders.
 - Charting early developmental milestones in genetic conditions, based on norms acquired from typically developing children, provides valuable information regarding developmental phenotypes that characterize these conditions.
- **What's New**
 - Compared to children with idiopathic autism spectrum disorder and typically developing children, children with the 16 rare genetic conditions reported here generally display more pronounced delays in early motor and expressive language milestones.
- **What's Relevant**
 - Description of early developmental phenotypes in genetic conditions associated with autism spectrum disorder may be helpful for referral and identification, as well as for the search for mechanisms for potential interventions.

†Somer Bishop and Audrey Thurm should be considered joint senior author.

Correspondence

Audrey Thurm, NBPP/NIMH, 10 Center Drive MSC 1255, Building 10 Room 1C250, Bethesda, MD 20892-1255, Fax: 301-402-8497, Phone: 301-496-6768, Email: athurm@mail.nih.gov.

References

- Arnett, A. B., Beighley, J. S., Kurtz-Nelson, E. C., Hoekzema, K., Wang, T., Bernier, R. A., & Eichler, E. E. (2020). Developmental predictors of cognitive and adaptive outcomes in genetic subtypes of Autism Spectrum Disorder. *Autism Res*, 13(10), 1659-1669. doi:10.1002/aur.2385
- Berg, A. T., Palac, H., Wilkening, G., Zelko, F., & Schust Meyer, L. (2020). SCN2A-developmental and epileptic encephalopathies: Challenges to trial-readiness for non-seizure outcomes. *Epilepsia*. doi:10.1111/epi.16750
- Bernier, R., Hudac, C. M., Chen, Q., Zeng, C., Wallace, A. S., Gerdts, J., . . . Chung, W. K. (2017). Developmental trajectories for young children with 16p11.2 copy number variation. *Am J Med Genet B Neuropsychiatr Genet*, 174(4), 367-380. doi:10.1002/ajmg.b.32525
- Bishop, S. L., Farmer, C., Bal, V., Robinson, E. B., Willsey, A. J., Werling, D. M., . . . Thurm, A. (2017). Identification of developmental and behavioral markers associated with genetic abnormalities in Autism Spectrum Disorder. *Am J Psychiatry*, 174(6), 576-585. doi:10.1176/appi.ajp.2017.16101115
- Bishop, S. L., Thurm, A., Farmer, C., & Lord, C. (2016). Autism spectrum disorder, intellectual disability, and delayed walking. *Pediatrics*, 137(3), e20152959. doi:10.1542/peds.2015-2959
- Buja, A., Volfovsky, N., Krieger, A. M., Lord, C., Lash, A. E., Wigler, M., & Iossifov, I. (2018). Damaging de novo mutations diminish motor skills in children on the autism spectrum. *Proceedings of the National Academy of Sciences*, 115(8), E1859. doi:10.1073/pnas.1715427115
- Centers for Disease Control. (June 10, 2020). CDC's Developmental Milestones. Retrieved from <https://www.cdc.gov/ncbddd/actearly/milestones/index.html>
- Chawner, S., Doherty, J., Moss, H., Bearden, C., Chung, W., Curran, S., . . . Van Den Bree, M. (2019). 7 - A genetic first approach to dissecting the heterogeneity of autism: Phenotypic comparison of autism risk copy number variants. *European Neuropsychopharmacology*, 29, S783-S784. doi:<https://doi.org/10.1016/j.euroneuro.2017.08.008>
- Eigsti, I.-M., de Marchena, A. B., Schuh, J. M., & Kelley, E. (2011). Language acquisition in autism spectrum disorders: A developmental review. *Research in*

- Autism Spectrum Disorders*, 5(2), 681-691.
doi:<https://doi.org/10.1016/j.rasd.2010.09.001>
- Feliciano, P., Zhou, X., Astrovskaya, I., Turner, T. N., Wang, T., Brueggeman, L., . . . Chung, W. K. (2019). Exome sequencing of 457 autism families recruited online provides evidence for autism risk genes. *NPJ Genom Med*, 4, 19.
doi:10.1038/s41525-019-0093-8
- Fischbach, G. D., & Lord, C. (2010). The Simons Simplex Collection: A resource for identification of autism genetic risk factors. *Neuron*, 68(2), 192-195.
doi:10.1016/j.neuron.2010.10.006
- Ghassabian, A., Sundaram, R., Bell, E., Bello, S. C., Kus, C., & Yeung, E. (2016). Gross motor milestones and subsequent development. *Pediatrics*, 138(1).
doi:10.1542/peds.2015-4372
- Hanly, C., Shah, H., Au, P. Y. B., & Murias, K. (2020). Description of neurodevelopmental phenotypes associated with 10 genetic neurodevelopmental disorders: A scoping review. *Clin Genet*. doi:10.1111/cge.13882
- Havdahl, A., Farmer, C., Schjølberg, S., Øyen, A. S., Surén, P., Reichborn-Kjennerud, T., . . . Bishop, S. (2020). Age of walking and intellectual ability in autism spectrum disorder and other neurodevelopmental disorders: a population-based study. *J Child Psychol Psychiatry*. doi:10.1111/jcpp.13369
- Heinze, G., Ploner, M., & Jiricka, L. (2020). logistf: Firth's Bias-Reduced Logistic Regression. R package version. R Package version 1.24. Retrieved from <https://CRAN.R-project.org/package=logistf>
- Hinton, R., Budimirovic, D. B., Marschik, P. B., Talisa, V. B., Einspieler, C., Gipson, T., & Johnston, M. V. (2013). Parental reports on early language and motor milestones in fragile X syndrome with and without autism spectrum disorders. *Dev Neurorehabil*, 16(1), 58-66. doi:10.3109/17518423.2012.704414
- Libertus, K., & Hauf, P. (2017). Editorial: Motor skills and their foundational role for perceptual, social, and cognitive development. *Frontiers in Psychology*, 8(301). doi:10.3389/fpsyg.2017.00301
- Matson, J. L., Mahan, S., Kozlowski, A. M., & Shoemaker, M. (2010). Developmental milestones in toddlers with autistic disorder, pervasive developmental disorder--not otherwise specified and atypical development. *Dev Neurorehabil*, 13(4), 239-247. doi:10.3109/17518423.2010.481299
- Mayo, J., Chlebowski, C., Fein, D. A., & Eigsti, I.-M. (2013). Age of first words predicts cognitive ability and adaptive skills in children with ASD. *Journal of Autism and Developmental Disorders*, 43(2), 253-264. doi:10.1007/s10803-012-1558-0
- Mitchell, S., Brian, J., Zwaigenbaum, L., Roberts, W., Szatmari, P., Smith, I., & Bryson, S. (2006). Early language and communication development of infants later diagnosed with autism spectrum disorder. *J Dev Behav Pediatr*, 27(2 Suppl), S69-78. doi:10.1097/00004703-200604002-00004
- Ozonoff, S., Young, G. S., Goldring, S., Greiss-Hess, L., Herrera, A. M., Steele, J., . . . Rogers, S. J. (2008). Gross motor development, movement abnormalities, and early identification of autism. *J Autism Dev Disord*, 38(4), 644-656.
doi:10.1007/s10803-007-0430-0
- Sanders, S. J., Sahin, M., Hostyk, J., Thurm, A., Jacquemont, S., Avillach, P., . . . Bearden, C. E. (2019). A framework for the investigation of rare genetic disorders

- in neuropsychiatry. *Nat Med*, 25(10), 1477-1487. doi:10.1038/s41591-019-0581-5
- Satterstrom, F. K., Kosmicki, J. A., Wang, J., Breen, M. S., De Rubeis, S., An, J. Y., . . . Buxbaum, J. D. (2020). Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. *Cell*, 180(3), 568-584.e523. doi:10.1016/j.cell.2019.12.036
- Sheldrick, R. C., Schlichting, L. E., Berger, B., Clyne, A., Ni, P., Perrin, E. C., & Vivier, P. M. (2019). Establishing new norms for developmental milestones. *Pediatrics*, 144(6), e20190374. doi:10.1542/peds.2019-0374
- Simons VIP Consortium. (2012). Simons Variation in Individuals Project (Simons VIP): A genetics-first approach to studying autism spectrum and related neurodevelopmental disorders. *Neuron*, 73(6), 1063-1067. doi:10.1016/j.neuron.2012.02.014
- SPARK Consortium. (2018). SPARK: A US cohort of 50,000 families to accelerate autism research. *Neuron*, 97(3), 488-493. doi:10.1016/j.neuron.2018.01.015
- Sparrow, S. S., Cicchetti, D., & Balla, D. A. (2005). *Vineland adaptive behavior scales: Second edition (Vineland-II), survey interview form*. Livonia, MN: Pearson Assessments.
- Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2016). *Vineland adaptive behavior scales: Third edition (Vineland-3)*. Bloomington, MN: NCS Pearson, Inc.
- Størvold, G. V., Aarethun, K., & Bratberg, G. H. (2013). Age for onset of walking and prewalking strategies. *Early Hum Dev*, 89(9), 655-659. doi:10.1016/j.earlhumdev.2013.04.010
- Taanila, A., Murray, G. K., Jokelainen, J., Isohanni, M., & Rantakallio, P. (2005). Infant developmental milestones: A 31-year follow-up. *Dev Med Child Neurol*, 47(9), 581-586.
- Winders, P., Wolter-Warmerdam, K., & Hickey, F. (2019). A schedule of gross motor development for children with Down syndrome. *J Intellect Disabil Res*, 63(4), 346-356. doi:10.1111/jir.12580
- World Health Organization. (2006). WHO motor development study: Windows of achievement for six gross motor development milestones. *Acta Paediatr Suppl*, 450, 86-95. doi:10.1111/j.1651-2227.2006.tb02379.x

Sample Information	Overall Genetic Sample (Searchlight)	Overall Control Sample (SPARK)
Total Sample	479 (100%)	3506 (100%)
16p11.2 Deletion	100 (21%)	--
16p11.2 Duplication	35 (7%)	--
1q21.1 Deletion	17 (4%)	--
1q21.1 Duplication	26 (5%)	--
<i>ADNP</i>	16 (3%)	--
<i>ASXL3</i>	26 (5%)	--
<i>CSNK2A1</i>	17 (4%)	--
<i>DYRK1A</i>	18 (4%)	--
<i>GRIN2B</i>	25 (5%)	--
<i>MED13L</i>	16 (3%)	--
<i>PACSI</i>	16 (3%)	--
<i>PPP2R5D</i>	31 (7%)	--
<i>SCN2A</i>	47 (10%)	--
<i>SLC6A1</i>	34 (7%)	--
<i>STXBP1</i>	36 (8%)	--
<i>SYNGAP1</i>	19 (4%)	--
Age		
Mean \pm SD	9.1 \pm 5.6	9.5 \pm 5.3
[Min, Max]	[3.0, 39.1]	[3.0, 46.8]
Sex		
Female	223 (47%)	645 (18%)
Male	256 (53%)	2861 (82%)
Annual Household Income		
50,000 or below	112 (23%)	834 (24%)
51,000-80,000	94 (20%)	787 (22%)
81,000-130,000	124 (26%)	1006 (29%)
131,000 or above	134 (28%)	797 (23%)
Unknown (Missing)	15 (3%)	82 (2%)

Inheritance Status		
<i>de novo</i> (1)	291 (61%)	--
Germline-Mosaicism (2)	3 (<1%)	--
Inherited (3)	66 (14%)	--
Unknown (4)	119 (25%)	--
Race		
Asian	6 (1%)	91 (3%)
Black	4 (<1%)	85 (2%)
Native American	0 (0%)	6 (<1%)
Pacific Islander	1 (<1%)	1 (<1%)
White	265 (55%)	2465 (70%)
Other	0 (0%)	31 (<1%)
Mixed	38 (8%)	818 (23%)
Unknown (Missing)	165 (34%)	9 (<1%)
Diagnostic History		
ASD-only	120 (25%)	3066 (87%)
ID-only	63 (13%)	0 (0%)
Comorbid ASD+ID	58 (12%)	440 (13%)
Seizure Disorder/Epilepsy	166 (35%)	145 (4%)
Pre-Term	58 (12%)	359 (10%)

Table 1. Sample demographics are provided for all genetic conditions in Simons Searchlight combined and the idiopathic ASD group from SPARK.

Genetic Status, Sample Size (n), Age in Months (Mean ± SD)	Milestone	Milestone Acquisition								Milestone Delay	
		Yes < Age 7				Yes > Age 7	Yes, Age Unknown (Vineland)	No	Not Reported	Total Delay	Odds of Delay for Genetic Conditions vs. ASD
		Q1	Q2	Q3	%	%	%	%	%	%	Odds Ratios [95% CI]
16p11.2 DELETION n=100 age=9.2 ± 4.5	Sat	6.00	7.00	9.00	99%	0%	0%	0%	1%	32%	3.19 [2.05, 4.86]
	Crawled	8.00	10.00	12.00	95%	0%	0%	0%	5%	18%	3.06 [1.74, 5.10]
	Walked	13.00	15.00	18.00	98%	0%	1%	0%	1%	45%	5.08 [3.36, 7.62]
	Single Words	12.00	18.00	25.00	91%	2%	4%	1%	2%	71%	1.53 [0.99, 2.43]
	Combined Words	24.00	36.00	43.50	82%	4%	0%	2%	12%	70%	1.49 [0.95, 2.39]
16p11.2 DUPLICATION n=35 age=9.1 ± 4.5	Sat	6.00	8.00	10.00	97%	0%	0%	0%	3%	47%	5.90 [2.99, 11.57]
	Crawled	9.00	10.00	12.00	94%	3%	0%	0%	3%	18%	3.13 [1.21, 6.97]
	Walked	13.00	17.00	21.50	94%	0%	3%	0%	3%	58%	8.36 [4.23, 16.9]
	Single Words	12.00	15.50	24.00	86%	3%	0%	3%	9%	72%	1.54 [0.75, 3.45]
	Combined Words	24.00	30.00	42.00	83%	6%	0%	6%	6%	76%	1.89 [0.9, 4.38]
1q21.1 DELETION n=17 age=5.9 ± 2.9	Sat	5.00	6.00	7.00	100%	0%	0%	0%	0%	6%	0.60 [0.07, 2.40]
	Crawled	8.00	9.00	10.50	100%	0%	0%	0%	0%	6%	1.25 [0.14, 5.00]
	Walked	12.00	14.00	17.00	100%	0%	0%	0%	0%	29%	2.73 [0.91, 7.17]
	Single Words	12.50	17.00	20.25	94%	0%	6%	0%	0%	75%	1.73 [0.63, 5.75]
	Combined Words	24.00	26.50	36.00	94%	0%	0%	0%	6%	63%	1.02 [0.39, 2.87]
1q21.1 DUPLICATION n=26 age=7.3 ± 4.1	Sat	6.00	7.00	9.00	96%	0%	0%	0%	4%	36%	3.81 [1.64, 8.37]
	Crawled	9.50	11.00	12.00	96%	0%	0%	0%	4%	20%	3.68 [1.28, 8.91]
	Walked	13.00	14.00	18.00	96%	0%	4%	0%	0%	32%	3.02 [1.26, 6.71]
	Single Words	12.00	16.00	24.00	92%	0%	0%	0%	8%	71%	1.45 [0.64, 3.65]
	Combined Words	24.00	36.00	42.00	85%	4%	0%	0%	12%	61%	0.96 [0.43, 2.26]
ADNP n=16 age=10.5 ± 5.9	Sat	8.00	9.00	11.00	94%	0%	0%	0%	6%	73%	16.91 [6.01, 56.79]
	Crawled	13.00	15.00	18.00	94%	0%	0%	0%	6%	87%	74.03 [22.29, 378.35]
	Walked	22.00	29.00	40.50	100%	0%	0%	0%	0%	100%	205.12 [27.74, 26175.00]
	Single Words	13.50	20.00	27.00	56%	6%	13%	13%	13%	83%	2.62 [0.76, 13.58]
	Combined Words	39.00	48.00	57.00	56%	6%	0%	13%	25%	100%	15.75 [2.07, 2018.14]
ASXL3 n=26 age=12.2 ± 8.2	Sat	8.00	11.00	15.00	96%	0%	0%	0%	4%	72%	16.32 [7.21, 40.90]
	Crawled	12.00	15.00	22.50	85%	4%	0%	4%	8%	63%	22.37 [10.02, 52.61]
	Walked	24.00	33.00	51.00	81%	0%	0%	8%	12%	96%	93.24 [24.06, 838.35]
	Single Words	21.00	33.00	49.50	31%	15%	0%	54%	0%	100%	33.03 [4.63, 4190.11]
	Combined Words	37.50	48.00	67.50	15%	19%	0%	58%	8%	100%	30.86 [4.3, 3918.04]
CSNK2A1 n=17 age=8.2 ± 4.2	Sat	6.00	8.00	11.00	100%	0%	0%	0%	0%	47%	5.92 [2.28, 15.14]
	Crawled	9.25	11.00	14.75	94%	0%	0%	0%	6%	38%	8.49 [2.99, 22.37]
	Walked	17.50	22.00	32.00	100%	0%	0%	0%	0%	88%	38.54 [11.89, 194.86]
	Single Words	12.00	24.00	27.50	82%	6%	6%	6%	0%	75%	1.73 [0.63, 5.75]
	Combined Words	35.25	36.00	48.00	59%	12%	0%	12%	18%	100%	18.26 [2.44, 2334.79]
DYRK1A n=18 age=8.3 ± 5.4	Sat	7.75	10.00	11.25	100%	0%	0%	0%	0%	72%	16.24 [6.3, 48.21]
	Crawled	10.00	12.00	15.25	100%	0%	0%	0%	0%	44%	11.10 [4.32, 27.71]
	Walked	18.00	18.50	20.50	100%	0%	0%	0%	0%	94%	72.52 [18.27, 657.31]

medRxiv preprint doi: <https://doi.org/10.1101/2021.04.07.21254183>; this version posted April 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

	Single Words	22.50	36.00	63.00	72%	0%	11%	17%	0%	94%	6.44 [1.60, 58.60]
	Combined Words	37.50	60.00	72.00	44%	11%	0%	22%	22%	100%	18.26 [2.44, 2334.79]
GRIN2B n=25 age=8.2 ± 3.9	Sat	9.00	11.00	17.00	100%	0%	0%	0%	0%	80%	24.67 [10.24, 70.61]
	Crawled	11.00	14.00	18.00	76%	0%	0%	0%	24%	58%	18.55 [7.61, 47.03]
	Walked	19.00	24.00	42.00	92%	0%	4%	4%	0%	83%	28.32 [11.09, 90.58]
	Single Words	12.00	20.00	30.00	76%	0%	4%	20%	0%	79%	2.21 [0.91, 6.33]
	Combined Words	27.75	42.00	60.00	56%	8%	0%	24%	12%	86%	3.51 [1.26, 13.27]
MEDI3L n=16 age=9.4 ± 8.5	Sat	8.00	10.00	12.00	94%	0%	0%	0%	6%	67%	12.63 [4.63, 38.58]
	Crawled	10.50	13.50	17.75	88%	0%	0%	0%	13%	57%	17.93 [6.41, 52.54]
	Walked	18.00	23.00	26.00	94%	0%	0%	6%	0%	88%	36.05 [11.02, 183.01]
	Single Words	17.00	24.00	42.00	81%	0%	0%	19%	0%	94%	6.44 [1.60, 58.60]
	Combined Words	24.00	42.00		19%	25%	0%	31%	25%	92%	4.83 [1.15, 44.55]
PACSI n=16 age=6.8 ± 4.4	Sat	7.00	9.00	11.00	94%	0%	0%	0%	6%	53%	7.50 [2.77, 20.73]
	Crawled	10.75	12.50	14.75	88%	0%	0%	0%	13%	50%	13.71 [4.84, 38.87]
	Walked	19.50	21.00	24.50	81%	6%	6%	6%	0%	87%	33.56 [10.15, 171.17]
	Single Words	24.00	34.50	45.00	50%	6%	6%	25%	13%	92%	5.19 [1.26, 47.71]
	Combined Words	39.00	42.00	69.00	56%	6%	0%	25%	13%	100%	18.26 [2.44, 2334.79]
PPP2R5D n=31 age=8.9 ± 5.2	Sat	9.00	11.00	18.00	100%	0%	0%	0%	0%	84%	31.89 [13.64, 89.78]
	Crawled	13.75	18.00	26.25	84%	0%	0%	6%	10%	82%	58.58 [24.61, 166.71]
	Walked	23.00	27.50	45.00	84%	0%	3%	6%	6%	90%	48.84 [18.18, 181.75]
	Single Words	18.00	25.00	42.00	61%	3%	6%	29%	0%	97%	11.84 [3.12, 105.75]
	Combined Words	36.00	54.00	72.00	48%	3%	0%	32%	16%	100%	33.38 [4.67, 4234.69]
SCN2A n=47 age=8.8 ± 5.3	Sat	6.00	8.00	12.00	94%	0%	0%	2%	4%	44%	5.32 [2.92, 9.58]
	Crawled	10.00	11.00	17.75	85%	0%	0%	6%	9%	42%	9.95 [5.32, 18.28]
	Walked	14.00	18.00	22.50	81%	0%	2%	17%	0%	65%	11.49 [6.35, 21.54]
	Single Words	12.00	18.00	36.00	57%	9%	0%	32%	2%	83%	2.82 [1.41, 6.37]
	Combined Words	24.00	31.00	58.50	26%	13%	0%	40%	21%	86%	3.72 [1.63, 10.35]
SLC6A1 n=34 age=9.0 ± 5.5	Sat	6.00	9.00	11.00	94%	0%	0%	0%	6%	56%	8.44 [4.23, 17.18]
	Crawled	8.75	10.50	13.00	100%	0%	0%	0%	0%	29%	5.88 [2.70, 11.94]
	Walked	12.75	16.00	19.50	100%	0%	0%	0%	0%	47%	5.54 [2.81, 10.86]
	Single Words	12.00	18.00	24.00	91%	0%	0%	3%	6%	72%	1.54 [0.75, 3.45]
	Combined Words	25.00	36.00	48.00	68%	6%	0%	18%	9%	84%	3.03 [1.30, 8.53]
STXBPI n=36 age=11.1 ± 8.4	Sat	7.00	9.50	13.00	94%	3%	0%	0%	3%	60%	9.81 [5.05, 19.63]
	Crawled	10.75	14.50	18.00	72%	6%	0%	8%	14%	68%	28.07 [13.56, 61.79]
	Walked	16.50	25.00	36.00	72%	6%	0%	11%	11%	81%	25.34 [11.35, 65.70]
	Single Words	9.75	12.00	20.00	28%	14%	0%	56%	3%	83%	2.83 [1.28, 7.27]
	Combined Words	23.00	36.00	54.00	14%	14%	0%	56%	17%	97%	12.39 [3.27, 110.54]
SYNGAP1 n=19 age=9.2 ± 5.1	Sat	7.00	9.00	12.00	89%	5%	0%	0%	5%	56%	8.17 [3.29, 20.92]
	Crawled	9.50	12.00	16.50	89%	5%	0%	0%	5%	50%	13.71 [5.44, 34.54]
	Walked	15.50	21.50	24.00	95%	5%	0%	0%	0%	74%	16.39 [6.44, 48.28]
	Single Words	21.75	36.00	48.00	63%	5%	5%	21%	5%	88%	3.86 [1.20, 19.51]
	Combined Words	42.00	60.00	60.00	37%	21%	0%	26%	16%	94%	6.51 [1.62, 59.22]
Idiopathic ASD	Sat	5.00	6.00	7.00	98%	0%	NA	0%	2%	13%	NA

n=3506 age= 9.5 ± 5.3	Crawled	7.00	8.00	10.00	97%	0%	NA	1%	2%	7%	NA
	Walked	11.00	13.00	15.00	100%	0%	NA	0%	0%	14%	NA
	Single Words	11.00	15.00	24.00	93%	0%	NA	3%	3%	62%	NA
	Combined Words	18.00	30.00	42.00	84%	1%	NA	10%	4%	61%	NA

Table 2. Quartiles for milestone acquisition, as well as proportion of the sample, are provided for those who achieved a milestone prior to age 7. Proportions of the sample are also presented for those who acquired the skill after age 7 or at an unknown age (age is unknown because Vineland scores were used to replace missing data where possible for walking and single words), did not acquire the skill, or had no reported data. Proportions of the sample that exhibited delays for each milestone are provided, as well as the odds ratios and confidence intervals for the odds of delay, for each respective genetic condition in comparison to the idiopathic ASD (SPARK) group.

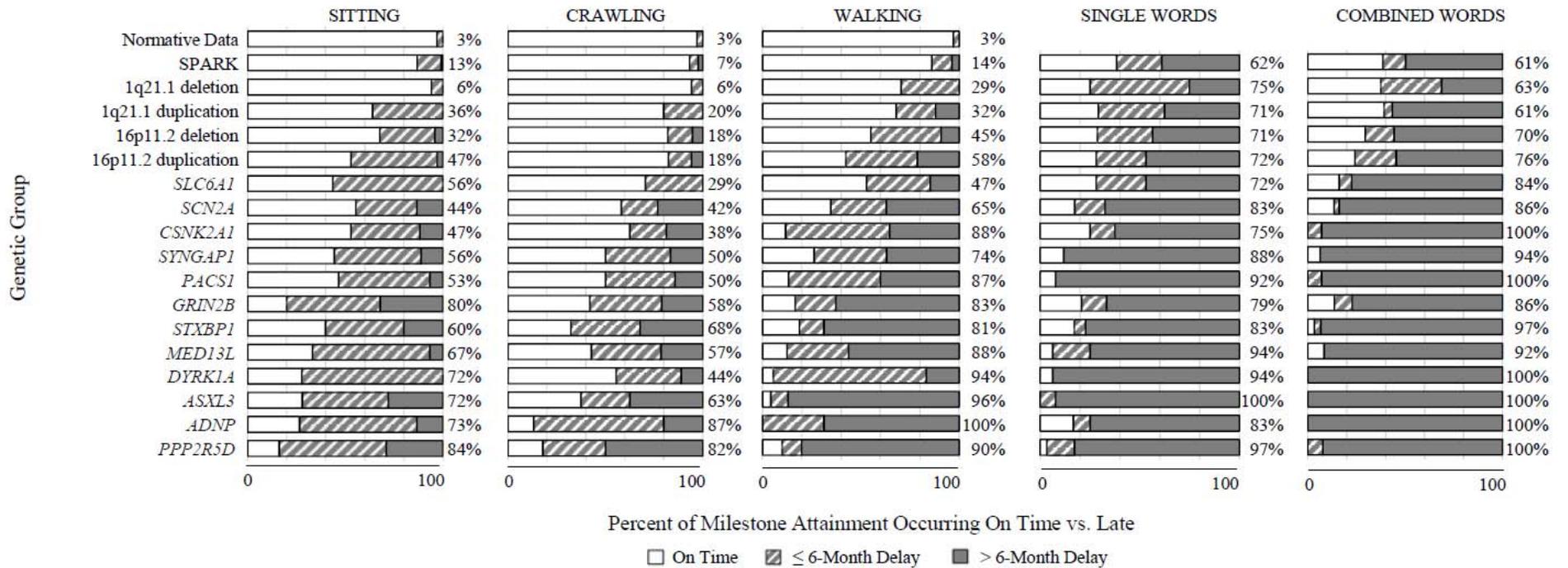


Figure 1. For each genetic group, the percent of milestone attainment is provided for sitting, crawling, walking, and single and combined words. This figure indicates the percentage of on-time versus late attainment, with late attainment divided by delays of up to six months versus beyond six months (including those who never attained the skill). Delays were defined as >8 months for sitting (WHO), >12 months for crawling (WHO), >16 months for walking (WHO), >12 months for single words (CDC), and >24 months for combined words (CDC). Total percentages for late attainment are provided in the right margin of each bar graph.