Auditory Evoked M100 Response Latency is Delayed in Children with 16p11.2 Deletion but not 16p11.2 Duplication

Julian Jenkins III, *et al.*

July 7, 2015

Abstract

Individuals with the 16p11.2 BP4-BP5 copy number variant (CNV) exhibit a range of behavioral phenotypes that that may include mild impairment in cognition and clinical diagnoses of Autism Spectrum Disorder (ASD). To better understand auditory processing impairments in populations with this chromosomal variation, auditory evoked responses were examined in children with the 16p11.2 deletion, 16p11.2 duplication and age-matched controls. Stimuli consisted of sinusoidal binaural tones presented passively while children underwent recording with magnetoencephalography (MEG). The primary indicator of auditory processing impairment was the latency of the 100ms "M100" auditory response detected by MEG, with the 16p11.2 deletion population exhibiting profoundly delayed M100 latencies relative to controls. This delay remained even after controlling for potential confounds such as age and cognitive ability. No significant difference in M100 latency was observed between 16p11.2 duplication carriers and controls. Additionally, children meeting diagnostic criteria for ASD (16p11.2 deletion carriers) exhibited non-significant latency delays when compared to the corresponding CNV carriers not meeting criteria for ASD. Present results indicate that 16p11.2 deletion is associated with auditory processing delays analogous to (but substantially more pronounced than) those previously reported in "idiopathic" ASD.

Introduction

Individuals with the deletion and duplication of the BP4-BP5 16p11.2 locus (chr16: 29.5-30.1 Mb) have varied behavioral phenotypes, including Autism Spectrum Disorder (ASD) (Hanson et al. 2010; Qureshi et al. 2014), language impairment (Hanson et al. 2010; Shinawi et al. 2010; Zufferey et al. 2012) and developmental delays. Individuals with the 16p11.2 deletion and duplication, particularly duplication carriers, can also have a phenotype "within normal range" although they still tend to demonstrate significant impairments compared with family members who do not carry the deletion or duplication. Mouse models and human data demonstrate that the consequences of deletion of the 16p11.2 CNV may be more severe than the duplication (Horev et al. 2011; Stefansson et al. 2014). The mechanistic linkage between this copy number variant (CNV) and behavioral/clinical/diagnostic findings, however, remains elusive.

Studies have shown delayed auditory evoked neuromagnetic field components (M100: Roberts et al. 2010; M50: Roberts et al. 2013) in conditions such as autism spectrum disorder (ASD). Delayed evoked responses persist after variance associated with cognitive function (IQ) and language ability is considered. It has been speculated that delays in early auditory processing may reveal atypical development of auditory sensory cortex and/or thalamocortical connections (Roberts et al. 2009, 2013) and may functionally underlie subsequent higher order neuronal dysfunction, leading to observed behavioral sequelae.

The purpose of this study is to assess the left and right superior temporal gyrus (STG) auditory evoked neuromagnetic field M100 component, detected by magneto-encephalography (MEG), in children with the 16p11.2 deletion and duplication in comparison to age-matched controls. Specifically, we test the hypothesis that STG M100 latency will be delayed in the genetically-defined cohorts, in particular the deletion carriers.

A hypothesis is that genes or other conserved elements in the 16p11.2 interval are necessary for generation of an age-appropriate M100 response (for example, by coding for synapse formation), and that, as with other CNVs, deletion carriers will be more severely impacted than duplication carriers. An alternative hypothesis predicts that the M100 response, as an indicator of atypical functional activity, more closely ties to phenotype and might thus be atypical in both deletion and duplication carriers with the same neurocognitive phenotype. Secondarily, given that the 16p11.2 deletion and duplication carriers share common breakpoints, it might be expected that observed measurement variability in the M100 will be reduced compared to measures in idiopathic (no known genetic, or other, etiology) ASD populations (Bijlsma et al. 2009). On the other hand, the variable clinical phenotypes associated with the 16p11.2 deletion and duplication (e.g. heterogeneous expression of ASD diagnosis, language impairment, and other phenotypes, ranging from mildly to severely impaired) might be expected to be associated with increased measurement variance of neuronal indices such as the M100.

A two-center multisite approach was adopted, as part of the broader Simons Variation in Individuals Project (Simons VIP; http://sfari.org/resources/simons-vip). The analysis examined the left and right STG auditory cortex M100 latencies in response to 200, 300, 500 and 1000 Hz sinusoidal tones. These analyses compared 16p11.2 deletions and duplications to age-matched controls and secondarily explored M100 latencies in the subset of 16p11.2 deletion and duplication carriers with ASD. We observe that children with the 16p11.2 deletion exhibit M100 delays relative to age-matched controls, of even greater magnitude than previously reported in idiopathic ASD compared to age-matched controls. Children with the 16p11.2 duplication exhibit M100 latencies not significantly different from age-matched controls. Meeting diagnostic criteria for ASD in 16p11.2 CNV probands does not significantly further prolong latency.

Materials and Methods

Genetic status confirmation and neuropsychological assessment

16p11.2 deletion and duplication pediatric participants included individuals with the same recurrent 600kb deletion (chr16:29,652,999 - 30,199,351; hg19) without other pathogenic CNVs or known genetic diagnoses (Zufferey et al. 2012). Probands with the 16p11.2 deletion and duplication were identified through routine clinical chromosome microarrays and recruited through the Simons VIP Connect website to participate (Simons VIP Consortium 2012). Cascade genetic testing of family members using whole genome high-resolution oligonucleotide arrays (Agilent 244k, G4411B, Agilent Technologies, Palo Alto, CA) determined if the deletion or duplication was de novo or inherited, to determine if carriers had other clinically significant CNVs (which would be exclusionary) and to identify other 16p11.2 CNV carriers within the family (Baldwin et al. 2008). Age-matched controls underwent chromosome microarrays to rule out pathological CNVs at the 16p11.2 locus and throughout the genome. Eight (out of 35) 16p11.2 deletion participants and three (out of 16) 16p11.2 duplication participants met criteria for ASD as established by clinical impression using DSM-IV-TR criteria (American Psychiatric Association 2000).

Following screening, families participated in initial data collection at one of four Simons VIP phenotyping sites (Boston Children's Hospital, Baylor College of Medicine, University of Washington, Children's Hospital of Philadelphia) for a comprehensive and standardized multi-day evaluation. The study was approved by the Institutional Review Board at each participating institution; all participants provided informed consent prior to data collection. All diagnostic interviewing and cognitive testing was videotaped for later review. Standardization of measurements across sites included mandatory formalized, standardized training on all measures through in-person training sessions and webinars for all clinicians, cross-site reliability and maintenance through monthly clinician conference calls and periodic videotape review, and validation and diagnostic confirmation through data review and observation of video recorded sessions by independent consultants.

Experienced, licensed clinicians gave best-estimate, clinical DSM-IV-TR diagnoses using all information obtained during the research evaluation. Information was based on the standardized interview, questionnaire, and observation processes described below as well as results from standardized administration of the Diagnostic Interview Schedule for Children (Shaffer et al. 2000), SCL-90 (Derogatis 1977) and review of available medical records and prior testing. To capture the range of psychiatric presentation, exclusionary criteria for

diagnoses were not considered (e.g., if a child met criteria for ADHD and ASD, both diagnoses were considered). Autism-specific diagnostic measures included the Autism Diagnostic Observation Scale (Lord et al. 2000) and the Autism Diagnostic Interview – Revised (Rutter et al. 2003); both measures were administered by research-reliable clinicians.

Participants were administered cognitive measures by experienced and licensed child psychologists via the standardization procedure described above. The Social Responsiveness Scale was used as a continuous measure of social and behavioral problems with high scores thought to correspond with greater likelihood of ASD (Constantino and Gruber 2005). Cognitive and language measures included: Differential Abilities Scale, Second Edition (Elliott 2007), Wechsler Abbreviated Scales of Intelligence (Wechsler 2003), the Clinical Evaluation of Language Fundamentals, Fourth Edition (Semel et al. 2003), Comprehensive Test of Phonological Processing- Nonword Repetition subtest (Wagner et al. 1999). Standard scores were used, or when standard scores were not available, ratio intelligence quotient (IQ) scores were calculated, for Full Scale IQ (FSIQ), Verbal IQ (VIQ), and Nonverbal IQ (NVIQ). Control exclusion criteria included known neurologic or psychiatric diagnosis in the participant or any sibling, English as a second language, drug use or incidental findings on MRI.

Based on geographical proximity, participants proceeded to MEG evaluation at one of two sites (Children's Hospital of Philadelphia, or University of California, San Francisco).

Participants

One hundred thirty-seven child participants (Children's Hospital of Philadelphia (CHOP): 63, University of California, San Francisco (UCSF): 74) were recruited (46 16p11.2 deletions, 25 duplications, 66 controls). Of this initial pool, 19 were excluded based on eligibility criteria (e.g., psychological/neurological profile, drug use, incidental findings from MRI). Of the residual pool of 118, who underwent MEG, 19 were found inevaluable due primarily to motion artifact (6 controls, 7 deletions, 6 duplications). Thus, of the evaluable 99 participants (CHOP: 47, UCSF: 52) forty-eight were age-matched controls, thirty-five 16p11.2 deletion carriers and sixteen 16p11.2 duplication carriers. Of the age-matched controls there were nineteen females and twenty-nine males, with five left-handed, thirty-five right-handed and eight ambidextrous (Oldfield 1971). Of the 16p11.2 deletion carriers there were fifteen females and twenty males, with nine left-handed, twenty-two right-handed and four ambidextrous. Of the 16p11.2 duplication carriers, there were six females and ten males, with thirteen right-handed and three ambidextrous.

Stimuli, procedure and delivery

200 Hz, 300 Hz, 500 Hz and 1000 Hz sinusoidal tones were passively presented. Tones were generated with LabView, sampled at 44.1 kHz with 16-bit resolution and were of 400 ms duration with 10 ms linear onset and offset ramps. Prior to data acquisition, an auditory threshold test was conducted with 1000 Hz tones of 300 ms duration and 10 ms rise time, binaurally presented (starting at a comfortable hearing level) and decreased until reaching auditory threshold for each ear. The experimental tones were presented 45 dB above threshold, with each tone presented 130 times. After completion of MEG scanning, structural MRIs were acquired using a 1mm isotropic ME-MP-RAGE 3D T1 sequence (Siemens Trio™ 3T, Siemens Medical Solutions, Erlangen, Germany).

Experimental stimuli were presented using an Edirol UA-1x external D-to-A converter using E-Prime stimulus presentation software. Stimuli were delivered binaurally through a TDT SA1 power amp, a pair of PA5 attenuators and a MA5 microphone amplifier to Eartone ER3A transducers and non-magnetic air tubes and eartip inserts. The ISI varied pseudo-randomly in the range 900 to 1100 ms.

MEG recording

Data were acquired using either a 275-channel whole-head magnetometer (CHOP) or a 272-channel whole-head magnetometer (UCSF). Using anti-aliasing filters, recording bandwidth was DC-300 Hz, sampled at 1200 Hz/channel. Prior to scanning and data acquisition, three head-position indicator coils attached to the

participant's scalp at nasion and left and right pre-auricular points provided continuous information on head position and orientation relative to the MEG sensors. Electrodes attached to the left and right clavicles and to the bipolar oblique (upper and lower left sites) recorded electrocardiogram (ECG) and electro-oculogram (EOG).

MEG analysis

Source space analyses were implemented using BESA 5.2 (BESA Research, Gräfelfing, Germany). Prior to evoked response analysis (offline averaging, filtering and baseline correction), the data were downsampled to 500 Hz with a zero-phase high-pass filter (cutoff frequency: 166.7 Hz, -24dB/octave slope). Eyeblink and heart artifact correction was implemented as described in Roberts et al. (2010). Epochs with artifacts other than blinks and heartbeats were rejected on the basis of amplitude and gradient criteria (amplitude > 1200fT/cm, gradients > 800fT/cm/sample). For each condition (i.e., 200, 300, 500, and 1000 Hz tones), epochs from the continuous recording were defined 500 ms pre- and post-trigger onset.

The presence of a M100 response in the left and right STG was determined using a standard dipole source model that transformed the averaged and filtered MEG sensor data into brain space (MEG data co-registered to the Montreal Neurologic Institute (MNI) averaged brain) using a dipole model with multiple sources (Scherg and von Cramon 1985; Scherg 1990; Scherg and Berg 1996). Specifically, the source model included left and right STG regional sources positioned at Heschl's gyrus and nine fixed regional sources modeling brain background activity and acting as probe sources for additional activity. Each participant's eye-blink and heartbeat source vectors were included in the individual source models (Lins et al. 1993; Berg and Scherg 1994).

Auditory evoked responses were analyzed after applying three Butterworth filters: (i) a 1 Hz forward-phase high-pass (-6dB/octave slope); (ii) a zero-phase 40Hz low-pass filter (-48dB/octave slope) and (iii) a 60 Hz notch filter (5 Hz width). For each tone, the evoked response was baseline corrected over the pre-trigger interval. M100 latency peak for the left and right STG dipoles was determined based on identification of an M100 in the sensor and source waveforms within an acceptable range (85-185 ms), resemblance to canonical M100 magnetic field topography and dipole goodness of fit (GOF). Adjustments (extensions) to the acceptable latency range were considered based on the age of the participants (younger participants typically have longer latencies; see Roberts et al. 2010; Roberts et al., 2013; Edgar et al. 2013). If a participant did not exhibit an auditory cortex magnetic field topography for a given condition, those observations were excluded from further analysis. Typically, the M100 either followed a detectable M50 or preceded a detectable M200 (or both). A total of nineteen participants (two controls, seventeen deletions; four meeting ASD criteria) in the age range 7.98 – 12.60 years had the M100 search latency window extended.

Statistical analysis

Statistical analyses were performed using R 3.1.2 (R Core Team 2014). A given result was considered significant for (i) $p < 0.05$ (linear mixed models, Wald Type II Chi-square tests, correlation tests, Tukey HSD) and (ii) z > 2 (z-statistic, Tukey HSD). Results were visualized using the `ggplot2` package (Wickham 2009).

Linear mixed models (LMMs) with a dependent variable of M100 latency (ms) were performed with random effect of Subject, fixed effects of Case (16p11.2 deletion vs. 16p11.2 duplication vs. neurotypical control), Hemisphere, Stimulus Condition (200, 300, 500, 1000 Hz) and Acquisition Site (CHOP vs. UCSF) with subject Age as a covariate and random slopes of Stimulus Condition and Hemisphere. LMMs were fit via maximum likelihood using the `lme4` package (Bates et al. 2014). Significance of the fixed effects was assessed using Wald Type II Chi-square tests via the `car` package (Fox and Weisberg 2011) and multiple comparisons, significance of individual factor levels and effect sizes were assessed using Tukey HSD via the `multcomp` package (Hothorn et al. 2008). The primary model evaluated was a main effects model of `Hemisphere + Stimulus Condition + Case + Age + Site + (Stimulus Condition + Hemisphere | Subject)`. Additional models with NVIQ and VIQ as a covariate were also assessed. Correlation tests (Spearman's ρ) between neuropsychological assessment scores (VIQ, NVIQ, SRS total score, CTOPP non-word repetition) and M100

latency were assessed in 16p11.2 deletion and duplication carriers. Comparisons of 16p11.2 deletion and duplication carriers meeting vs. not meeting ASD diagnostic criteria were assessed using a LMM of Hemisphere + Stimulus Condition + ASD status + Age + Site; random effects structures were the same as in previous analyses.

Results

M100 recording success rate

No significant difference in sensation level was observed between groups with mean differences < 5dB ($\chi^2 = 3.810, df = 2, p = 0.148). In age-matched controls evaluable data was obtained in 45 out of 48 participants. In terms of success elicit based rejection : controls : 45/54; deletions : 35/42; duplications : 16/22, \chi^2 = 1.33, p = 0.51. Not all subjects had 8 evaluable res stimulus success rates indicated the success rate for 16p11.2 duplication carriers was, however, less than that of 16p11.2 deletion ca 0.281, SE = 0.222, z = -4.231, p < 1e - 04) or age - matched controls (estimate = 0.314, SE = 0.209, z = -3.748, p < 0.001). Success rates for 16p11.2 deletion carriers vs. age-matched controls did not differ (estimate = 0.539, SE = 0.174, z = 0.902, p = 0.637). Summaries of age and psychological score distributions for participants with evaluable$