**Genotype-phenotype correlations in Cornelia de Lange syndrome: behavioural characteristics and changes with age**

Moss, J., Penhallow, J., Wellham, A., Fitzpatrick, D., Hammond, P and Oliver, C.

**Introduction**

Cornelia de Lange syndrome (CdLS) is a rare multisystem genetic disorder that affects approximately one child in every 40,000-100,000 (O’Brien & Yule, 1995). The syndrome is associated with unusual facial features, limb malformations (Selicorni et al., 2007) and a wide range of health conditions (Hall et al., 2008). Associated intellectual disability (ID) is typically within the severe to profound range, although a proportion of individuals may have moderate or mild ID (Sloneem et al., 2009). Behavioural characteristics include social avoidance, repetitive and self-injurious behaviours and hyperactivity (Berney, Ireland, & Burn, 1999; Hyman Oliver, & Hall, 2002; Moss, Oliver, Arron, Burbidge, & Berg, 2009; Oliver, Arron, Sloneem, & Hall, 2008). Autism spectrum disorder (ASD) characteristics are common, and may be extensive enough to warrant diagnosis of an ASD in 51% -67% of individuals (Basile, Villa, Selicorni, & Molteni, 2007; Berney et al., 1999; Bhuiyan et al., 2006; Moss et al., 2008; Oliver, Berg, Moss, Arron, & Burbidge, 2011; Oliver et al., 2008; Strivastava et al., 2014; Nakanishi et al., 2014). More recently, signs of premature ageing and changes in behaviour, mood and cognition with age have been described (Kline et al., 2009; Oliver et al., 2010; Nelson et al., 2013; Reid, 2010). There is substantial heterogeneity in all aspects of CdLS but very little is known about what predicts phenotypic heterogeneity. Understanding this is crucial to the early identification of individuals with CdLS at higher risk of developing cognitive, behavioural and emotional difficulties and to guide appropriate, targeted intervention and management.

The most common known genetic cause of CdLS is mutation of the NIPBL gene, which accounts for up to 60% of cases (Krantz et al., 2004, Tonkin et al., 2004). However, a number of other less common causal mutations have now been identified. Mutations in SMC1a and SMC3 have been found to account for CdLS in a further 5% of affected individuals (Deardorff et al., 2007; Musio et al., 2006), and more recently, mutations in the HDAC8 and RAD21 genes have been identified in a small number of cases (Deardorff et al., 2012a; Deardorff et al., 2012b). All of these genes are thought to encode proteins related to cohesin complex function.

Studies that have reported genotype-phenotype correlations in CdLS have primarily described variability in clinical and diagnostic characteristics within and between mutation variants. The general consensus from is that individuals with NIPBL mutations are likely to present with more severe clinical features and to have more impaired cognitive function than those with other causal mutations and those for whom mutations have not been identified, although this is not always the case (Gillis et al., 2004; Mannini et al., 2013; Nakanishi et al., 2012). Those with SMC mutations are generally described as presenting with a ‘milder’ CdLS phenotype, moderate cognitive impairment and fewer structural abnormalities than those with NIPBL mutations (Deardorff et al., 2007; Gil-Rodríguez et al., 2015; Píe et al., 2010). Individuals with RAD21 mutations demonstrate a somewhat subtle clinical presentation with a very mild cognitive impairment (Deardorff et al., 2012), while those with HDAC8 mutations are considered to be more similar to those with NIPBL mutations but with fewer limb abnormalities and other possible clinical features that may distinguish them from other individuals with CdLS (Kline et al., 2014; Mannini et al., 2014).

Studies evaluating genotype-phenotype correlations with regard to behavioural characteristics are more limited. Gil-Rodríguez and colleagues (2015) described fewer behavioural problems in those with SMC3 mutations, although no standardised assessments of behaviour were employed and there was no comparison between individuals with different CdLS mutation variants. Nakanishi et al. (2012) describe a trend for higher scores on the Autism Diagnostic Interview-Revised (ADI-R; Lord et al., 1994) in individuals with a NIPBL mutation compared to those without an identified mutation, although this difference was not statistically significant.

In the current study we aimed to evaluate genotype-phenotype associations in relation to a broad range of behavioural features known to be characteristics of CdLS including: challenging behaviour, ASD characteristics, mood and hyperactivity. Specifically, we compared individuals with a confirmed NIPBL mutation to those for whom the NIPBL mutation was not identified and to individuals who were negative for all known CdLS mutations. A secondary aim was to explore the effect of mutation status on potential changes with age that have been reported in the literature.

**Procedure**

Participants were identified from a pre-existing database of 252 individuals with CdLS who had taken part in a questionnaire survey as part of a larger research project evaluating behavioural characteristics in neurodevelopmental disorders. These participants had originally been recruited via the CdLS Foundation (UK and Ireland) or via a pre-existing participant database held at the University of Birmingham, UK. A total of 126 individuals provided their consent for the researchers at the University of Birmingham to contact other relevant professionals in order to confirm diagnostic status, including ascertaining the results of mutation analyses, where these had been carried out.

The results from mutation analyses were sought from two clinics in the UK; the MRC Human Genetics Unit, University of Edinburgh and the Northern Regional Genetics Service, Newcastle. These are the only two clinics in the UK where genetic testing for CdLS is routinely conducted. A total of 24 participants had previously been tested at one or other of these clinics (Edinburgh n= 12, Newcastle n=10; DNA sequencing failed in two further participants) and agreed that data could be shared. Of the remaining participants (n=102), 83 were contacted by the research team by phone and by letter and were invited to participate in a genetic screening study at the Human Genetics Unit, Edinburgh (nineteen participants did not have up to date contact details and could not be reached for this purpose). Mutation analyses were performed for a further twelve participants through this screening study and these data shared. This resulted in a total sample of 34 individuals for whom both questionnaire data regarding behavioural characteristics and data from mutation analyses were available. All participants had a confirmed clinical diagnosis of CdLS from a clinical geneticist. The recruitment strategy is summarised in Figure 1.

**Participants**

Participant characteristics are outlined in Table 1. NIPBL mutations were confirmed in seventeen individuals (50.00 % of total sample), one individual had an HDAC8 mutation (2.94%) and two had a SMC1a mutation (5.88%). Six of the participants for whom a NIPBL mutation was not detected had not received further screening for other CdLS mutations because these were not routinely carried out within that particular service. Seven participants were negative for all known CdLS mutations. Participants under the age of five years were excluded from the following analyses because a number of the behavioural measures were not suitable for use in this age group.

**Measures**

Demographic information including date of birth, gender, mobility, verbal ability (i.e. able to communicate more than 30 signs/words) and diagnostic status (by whom and when) was collected using a brief background questionnaire.

**The Wessex Scale(Kushlick et al., 1973)** provides a proxy measure of adaptive behaviour skills. The measure evaluates the physical and social abilities of individuals on subscales including self help skills, continence, mobility, speech and literacy. The measure has good inter-rater reliability with children and adults, at both the item and subscale level (Kushlick et al. 1973; Palmer and Jenkins 1982).

**The Activity Questionnaire (TAQ; Burbidge & Oliver, 2008; Burbidge et al., 2010)** evaluates hyperactivity and impulsivity in individuals with intellectual disability and is suitable for use with both non-verbal and verbal individuals. The questionnaire consists of 18 items across three subscales: impulsivity, over-activity and impulsive speech. Robust internal consistency and reliability has been established (Burbidge et al., 2010)

**The Repetitive Behaviour Questionnaire (RBQ; Moss and Oliver, 2008; Moss et al., 2009)** identifies specific types of repetitive behavior in both children and adults with intellectual disabilities. The questionnaire is made up of nineteen operationally defined and observable behaviors across five subscales: restricted preferences, repetitive speech, insistence on sameness, stereotyped behavior, and compulsive behavior. A five point Likert rating scale is used to record responses which range from ‘never’ to ‘more than once a day’. Other studies have shown the questionnaire to have good reliability and validity (Moss et al., 2009).

**The Challenging Behaviour Questionnaire (CBQ; Hyman, Oliver and Hall, 2002)** is a brief measure designed to assess the presence or absence of challenging behaviours over the past month including physical and verbal aggression, self-injury and destruction of property. Good inter-rater reliability has been established (Hyman, Oliver and Hall, 2002). The CBQ is derived from the Challenging Behaviour Interview which is also reported to have good reliability and validity (Oliver et al., 2003).

**The Mood, Interest and Pleasure Questionnaire Short Version (MIPQ-S; Ross & Oliver, 2003; Ross, Arron & Oliver, 2008)** evaluates two constructs associated with depression in adults and children with intellectual disabilities. Informants are required to rate 12 items based on retrospective observations over a two week period. The questionnaire shows good internal consistency and reliability (Ross and Oliver, 2003).

**The Social Communication Questionnaire (SCQ; Rutter, Bailey & Lord, 2003) i**s a screening tool designed to measure communication and social skills in participants suspected of having ASD. The questionnaire comprises three subscales: communication, social interaction and repetitive and stereotyped behaviors. Higher scores are indicative of more significant ASD characteristics. A cut off score of 15 is thought to be indicative of ASD, while a score of 22 indicates the presence of Autism. The SCQ has been shown to have good concurrent validity with the Autism Diagnostic Observation Schedule and Autism Diagnostic Interview (Berument, Rutter, Lord, Pickles, & Bailey, 1999; Howlin & Karpf, 2004).

**Data analysis**

To evaluate genotype-phenotype correlations, between group comparisons were conducted contrasting the clinical characteristics (self help skills, mobility, hearing, vision and speech) and scores on the behavioural assessments of those participants who had a confirmed mutation in NIPBL (NIPBL positive) and those for whom a mutation in NIPBL had not been identified (NIPBL negative; includes participants who were positive for SMC1a and HDAC8 mutations). A more refined, exploratory comparison contrasting individuals who were NIPBL positive with those who were negative for all known CdLS mutations (no mutation detected; NMD) was also conducted. Chi squared tests were conducted for categorical data and independent samples t- tests (or nonparametric equivalent when data were not normally distributed) or analysis of covariance were conducted for continuous variables (or variables which could be treated as continuous), with chronological age as a covariate for all comparisons and self help skills as a covariate for NIPBL vs. NMD comparisons. Given previous reports within the literature of changes with age in CdLS, Pearson correlations were performed between chronological age and behavioural variables within each mutation status group (NIPBL positive, NIPBL negative, NMD).

**Results**

*Clinical characteristics*

Participant characteristics and results from between group analyses are described in Table 2. The NIPBL positive group was significantly younger and had significantly lower self-help skills than the NMD group. Differences in chronological age and self-help skills between the NIPBL positive and NIPBL negative groups also approached significance (*p* =.05). When chronological age was controlled for statistically, the difference in self-help skills between the NIPBL positive and NMD groups approached significance (F1,22=4.30 ; *p* =.05) and there was no longer a significant difference between the NIPBL positive and NIPBL negative groups (F1,29 = 3.14; *p*=.09). There were no significant group differences with regard to gender ratio, hearing, vision, speech or mobility between the NIPBL positive and NIPBL negative groups or between the NIPBL positive and NMD groups.

*Behavioural characteristics*

Table 3 describes the scores on each of the behavioural questionnaires completed and the results of the between group analyses. The NIPBL positive group scored significantly lower on the mood subscale of the MIPQ-S than the NIPBL negative group. This difference remained significant when chronological age was controlled for (F1,28 =9.36; *p* =.01). Stereotyped behaviour was reported as being significantly higher in the NIPBL positive group compared to the NIPBL negative group, only when chronological age was controlled for statistically (F1,28 = 5.26; *p* =.03). There were no significant differences between the NIPBL positive and NMD groups, with the exception of the mood subscale of the MIPQ-S which was significantly lower in the NIPBL positive group, only when chronological age was controlled for (F1,21 = 6.80; *p*=.02) and the communication score on the Social Communication Questionnaire, which was significantly higher in the NMD group, only when self help skills were controlled for (F1,20 = 5.36; *p*=.03).

*Changes with age*

Table 4 shows the results from the Pearson correlation analyses between chronological age and scores on behavioural measures in each of the three mutation status groups. A significant negative correlation between age and scores on the MIPQ-S (indicating lower mood, interest and pleasure with older age) was identified in the NIPBL positive group only. A significant positive correlation between age and stereotyped behaviour (indicating increased frequency of stereotyped behaviour with age) was identified in the NIPBL negative group only.

**Discussion**

In the current study we aimed to describe genotype-phenotype correlations in CdLS, with a specific focus on behavioural characteristics, using standardised behavioural measures. The current study utilised pre-existing databases in two specialist genetics centres (the only two centres within the UK that screen for CdLS) and an extensive behavioural database at the University of Birmingham. In total, 34 participants for whom both behavioural and genetic data were available and able to be shared were identified.

Individuals with a confirmed mutation in the NIPBL gene had significantly lower self help scores than those who were NIPBL mutation negative and those who were negative for all known CdLS mutations. This is consistent with previous reports of greater severity of cognitive impairment in individuals with NIPBL (Gillis et al., 2004; Mannini et al., 2013; Nakanishi et al., 2012). However, these differences in self-help skills were largely accounted for by the fact that the NIPBL positive group was also significantly younger than the NIPBL negative and NMD groups. No other differences in clinical characteristics including vision, hearing, mobility and speech were identified.

The NIPBL positive group was reported to show significantly lower levels of mood when compared to the NIPBL negative group and the NMD group. Furthermore, those with a confirmed NIPBL mutation showed significantly more frequent stereotyped behaviour relative to those who were NIPBL negative. This is the first study to assess and identify specific differences in these areas of behaviour according to mutation status in CdLS. The findings are consistent with the suggestion that individuals with NIPBL mutations are, broadly speaking, more severely affected by the syndrome than those who are NIPBL negative (Gillis et al., 2004).

No other significant group differences were identified with regard to challenging behaviour hyperactivity and the presence of ASD characteristics, with the exception of the communication domain of the Social Communication Questionnaire which was found to be significantly higher in the NMD group relative to the NIPBL group, when self-help skills were controlled for. The findings regarding ASD characteristics are broadly consistent with previous study findings which also indicate no statistically significant association between ASD behaviours and type of mutation (Bhuiyan et al., 2006), although Nakanishi and colleagues (2012) report a difference approaching significance. Nakanishi et al., (2012) also report a high rate of false negatives when using the Social Communication Questionnaire in CdLS and therefore, further assessment of ASD characteristics using direct measures is required.

Scores on the Mood, Interest and Pleasure questionnaire were significantly, negatively correlated with chronological age (indicating lower scores for older participants) in the NIPBL positive group. Interestingly, this was not identified in those who did not have the NIPBL mutation or within the smaller group of individuals for whom a mutation had not been identified. In the NIPBL negative group, a significant positive correlation with age was identified for stereotyped behaviour, indicating increasing frequency of stereotyped behaviours in this group. Previous studies have described significant changes in mood and repetitive behaviour with age in CdLS (Oliver et al., 2010; Nelson et al., 2013), alongside a number of other behavioural and physical changes (Kline et al., 2009). Given that the distribution of ages varies between the groups, these findings should be interpreted with caution. However, this exploratory analysis suggests that there may be a degree of specificity for these changes with genetic variation. Variability in the nature and degree to which changes with age manifest across different genetic variations of CdLS has not previously been described but has prominent clinical implications. Identifying those most at risk for changes with age in CdLS would enable early detection, intervention and management. Furthermore, the suggestion that the nature of these changes with age may be different in those with different genetic mutations may be important for understanding the aetiology of this change and the relevance of genetic mechanism in this pathway. These findings should be evaluated further in a larger study sample in order to confirm the pattern of variability.

The study findings should be considered in the context of a number of limitations. Interpretation of the findings is somewhat limited by the small sample size. However, analyses identified significant mutation group differences and associations despite the small sample size, suggesting that statistical power was sufficient. Previous studies have demonstrated heterogeneity *within* the group of individuals identified as having NIPBL mutations, with missense mutations resulting in a milder presentation than deletion, nonsense and splicing mutations (Bhuiyan et al., 2006; Gillis et al., 2004; Mannini et al., 2013; Píe et al., 2010). The nature of NIPBL mutations in the current study sample are outlined in Table 1. However, the sample was not sufficiently large enough to enable group comparisons across these subtypes. It is likely that the small sample size results from the data collection strategy employed. A retrospective approach was employed in order to ‘pool’ resources among specialist UK based centres working with individuals with CdLS and their families. This was considered to be the most efficient approach to data collection because it utilised existing data and consequently reduced the burden on families (avoiding repeat DNA analysis and repeating behavioural surveys). Surprisingly, it proved more difficult than expected to combine existing data sets across different research and clinical groups (largely as a result of ethical restrictions imposed on the databases at each research/clinical site), resulting in a relatively small sample size (relative to the size of existing databases and the number of participants who had previously provided consent for information to be shared across different research/clinical groups). Given the rarity of the syndrome and to avoid participant fatigue, researchers and clinicians should collectively consider ways in which this approach might be maximised more effectively in the future. A centralised, national database might be one way in which this could be achieved. A further limitation is the fact that the three mutation status groups differed according to chronological age, with those individuals with a confirmed NIPBL mutation being significantly younger than those for whom there was no NIPBL mutation or other CdLS related mutation identified. Although chronological age was controlled for statistically in order to rule out this confound, the variability in age across the mutation status groups might reflect a cohort effect, whereby the accuracy of diagnosis, based on clinical features, may have improved following the identification of the NIPBL causal gene and following improved availability for genetic testing more broadly. In the current study, participant diagnosis (based on clinical features) had been previously confirmed by one of two specialist genetics clinics with extensive experience of CdLS. Consequently, we feel confident that those included in the study sample are representative of the syndrome.

In summary, the findings from this study suggest that there may be subtle differences in the broad behavioural phenotype and in the developmental trajectory of behaviours, according to genetic mutation status in CdLS. In particular, individuals with NIPBL mutations show lower levels of mood compared to those who are NIPBL negative and those for whom no mutation has been identified. The findings also suggest that those individuals who are NIPBL positive might be at greater risk for experiencing a decline in mood with age. These findings require replication in a larger study sample.