Colorado Twin Registry: An Update

Sally-Ann Rhea, Andy A. Gross, Brett C. Haberstick, and Robin P. Corley
Institute for Behavioral Genetics, University of Colorado, Boulder, CO, USA

The Colorado Twin Registry (CTR) is a population-based registry housed at the Institute for Behavioral Genetics at the University of Colorado. Recruitment began in 1982 and includes twins born from 1968 to the present. Four samples are currently drawn from the CTR: The Community Twin Sample, the Longitudinal Twin Sample, the Early Reading Development Sample, and the Colorado Learning Sample. Criteria for enrollment, recruitment strategies, demographic information, and zygosity assignment are explained for each sample. In addition, five studies in which CTR twins are now participating are highlighted. These include studies of cognition, learning ability, and vulnerability to substance abuse and antisocial behavior. The development of the CTR is an ongoing and evolving process, and it has proven to be a valuable registry, relatively representative of the population from which it was drawn.

Keywords: twins, longitudinal, cognition, executive function, learning, substance use

The Colorado Twin Registry (CTR) is housed at the Institute for Behavioral Genetics (IBG) at the University of Colorado at Boulder. In 2006, we reported in this journal details about the CTR and its component projects (Rhea et al., 2006). Since that publication, no significant changes have occurred with respect to recruitment methods. Thus, this report will focus on the continued data collection efforts of several projects, as well as one new data collection effort that uses registry families.

Over the years, the CTR has benefited from its partnerships with the Colorado Department of Health (CDH), Division of Vital Statistics and many public school districts. Although the CDH has provided us with access to birth records for twins born within the state from 1910 forward, we have only contacted twins recruited from 1968 onwards, and only made systematic use of birth records from 1982 to the present. From 1993 through 2003, many public school districts within Colorado provided supplemental recruitment into the registry. Thus, the registry, while primarily comprised of twins born in Colorado, includes some twins who migrated into Colorado during their school years. We continue to follow several samples longitudinally and in many cases twins who were enrolled while living in Colorado who now live outside the state and even outside the country.

For years prior to 1984, large mailings were sent to twins born across several years. From 1984 through 1990, mailings were sent on a monthly basis as we were enrolling families into studies that began in the twins’ infancy (as early as 5 months old) or toddlerhood (14 months). From 1990 to the present, an annual mailing is sent in the fall to all families who gave birth to surviving twins in the preceding year (e.g., in the fall of 2012 we will mail to all 2011 twin families).

In our 2006 report we described several samples and projects drawing on subjects from the CTR. In many cases the samples were or are used across multiple studies. Since 2006, a number of projects have been completed and these include the Toddler Questionnaire Study (Schmitz et al., 1995), Twin Infant Project (Benson et al., 1993), Executive Cognitive Function — Community Twin Study (ECF-CTS), which parallels the ECF Longitudinal Twin Study (ECF-LTS) described below, and the Colorado Adolescent/Adult Twin Study (Siewert et al., 2004). Figure 1 depicts the ongoing studies of the CTR and highlights one project that is newly drawing on registry data. The remaining sections of this report will focus on these five projects.

Registry Enrollment and Characteristics

Initially, the CDH mailed inquiry letters on behalf of IBG to parents of twins, excluding parents for whom CDH matched twin death records, and provided birth record data including names and contact information to IBG staff for all respondents who returned a postage-paid reply card.


ADDRESS FOR CORRESPONDENCE: Robin Corley, Institute for Behavioral Genetics, Campus Box 447, University of Colorado, Boulder, CO 80309, USA. E-mail: Robin.Corley@Colorado.edu
FIGURE 1
Current samples and studies.

FIGURE 2
Average twin birth weight in Colorado by ethnicity.
Note: Data source: Health Statistics Section, Colorado Department of Public Health and Environment, August 2012.

authorizing the CDH to release that data. From 1999 onward, CDH modified this procedure and relied on a ‘negative consent’ process, in that no response was treated as a positive response, and released twin birth information unless the parents returned a card specifically prohibiting the CDH from doing so. The CDH does not release information if the inquiry letter is not deliverable.

As stated above, from 1993 to 2003 many Colorado school districts provided access to twin families either by providing IBG with lists of putative twins to whom enrollment inquiries were subsequently mailed or by mailing inquiries on our behalf. Most of the subjects who responded were either already in our registry or were in the CDH non-response category, but school mailings enabled us to identify an additional 2,600 pairs who could not be matched to birth records on the basis of information provided.

Subsequent to identification and agreement to contact, a Twin Registry form was mailed to the parents and twins. The form consisting primarily of a demographic survey designed to facilitate further contact with the twins as studies became available. During periods of recruitment for specific studies, follow-up mailings and phone calls were made to the families who met criteria for those studies, resulting in higher levels of enrollment at those times. During the years from 2005 to 2010, not covered in our previous report, we have continued to add 900 twin pairs per year, with 40% of the twin families returning our initial demographic survey.

As shown in Figure 2, between 1984 and 2004, twins in families that reported their ethnicity as Hispanic have a slightly lower average birth weight than those who report being non-Hispanic white. For families reporting their ethnicity as black, the averages are even lower. As shown in
Figure 3, for parental marital status, the contrast is even greater, with both Hispanic and black mothers less likely to be married. Together, these findings may partially explain the greater proportion of white non-Hispanic families relative to all births participating in some studies of the registry, as studies that began in infancy required minimal birth weights, infant health, and the participation of both parents.

Zygosity was determined for the CDH twins from parent-ratings on a zygosity questionnaire based on Nichols and Bilbro (1966). For twins enrolled through the school system, zygosity was tentatively assigned based on two questions: how frequently are the twins mistaken for each other by people who know them, and are they ‘as alike as two peas in a pod’? For twins subsequently seen in-person, DNA has been collected and zygosity confirmed by twin concordance among a minimum of 11 highly informative, short tandem repeat polymorphisms (Smolen, 2005).

Longitudinal Twin Study
The Longitudinal Twin Study is a prospective study designed to investigate the genetic and environmental influences on individual differences in psychological development (Emde & Hewitt, 2001). Information on cognition, temperament, social competence, family relations, and behavioral problems was collected almost yearly from the time the participants were 14 months old through age 16. Subsequent data collection from the sample has taken place through other studies at the Institute: They represent almost one-third of the subjects in the Center for Antisocial Drug Dependence (CADD; described below) and are the complete sample for the ECF-LTS (described below) study.

The sample of the Longitudinal Twin Sample (LTS) was drawn both from the Twin Infant Project sample (244 twin pairs) and from independent registry recruitment (180 twin pairs). Criteria were more stringent for the infant enrollees, but at a minimum, families were required to live within a 3-hour driving range (~300 km), and the twins’ birth weight was at least 1,000 grams (though 96% weighed 1,700 grams or more. A separate group of 59 twin pairs from the infant project born prior to the LTS 14-month testing series were tested on Colorado Adoption Project protocols, which overlapped in many domains with the LTS. From age 9 and older, the protocols were identical, which resulted in a merging of these twins into the LTS sample. Thus, the number of families who have participated in LTS is 483 (male–male twin pairs: 240; female–female twin pairs: 243). Attrition has been low in this sample; at age 16, the culminating age of this portion of the study, 408 (about 85%) remained active.

Although data collection for this portion of the study concluded at age 16, we have continued to code the many video-recorded sessions that were collected at ages 14 months through age 7 and conduct analyses covarying this information with data collected concurrently or from subsequent assessments. For example, archived video data was recently coded to determine levels of self-restraint in twins at ages 14, 20, 24, and 36 months and compared to performance on executive function measures in late adolescence (Friedman et al., 2011). Less-restrained children scored
significantly lower on a common executive function variable, and significantly higher on a shifting-specific factor, suggesting a biological relationship between individual differences in self-restraint and executive functions.

ECF-LTS
The ECF-LTS was designed to examine the genetic and environmental contributions to executive cognitive functioning and its correlates (Friedman et al., 2006, 2008). All members of the LTS sample were invited to participate. At wave 1, when the twins were about 17 years old, 426 monozygotic (MZ) and 371 dizygotic (DZ) individuals did so.

A second wave of data collection 5 years later, focusing on the relationship of ECF and self-regulation, was initiated in 2008; 663 individual twins have participated as of 31 July, 2012, and we expect to complete the study with 780 pairs altogether. Interestingly, so far, at this later age — when the twins are in their 20s — we have successfully re-enrolled more subjects than the number who have chosen not to continue their participation, resulting in a potentially higher number than the previous wave. Preliminary results from this follow-up study indicate that individual differences in executive functions are quite stable across a 6-year time span. Multivariate twin models suggest that stability is almost entirely due to high genetic correlations across time; there is no new genetic variance at wave 2. Change is due to small non-shared environmental influences at wave 2 and is most significantly correlated with changes in depression symptoms (Friedman et al., in press).

We have proposed conducting a third wave of data collection, again 5 years after the current sessions. The novel approach for this wave will be conducting the session while the subjects are undergoing functional magnetic resonance imaging. We have conducted pilot studies using ECF-CTS subjects and are optimistic that the data collection is feasible and will allow us to link cognitive tests of executive functioning across developmental periods to specific brain region activation and structures.

CADD
The various components of the Center include a sample selected for substance and behavior problems and community samples used both for comparisons to the selected sample and for their usefulness in analyses of genetic and environmental influences on these behavior problems. The majority of the community subjects in the CADD are drawn from two CTR-based twin samples, the LTS, and 1,002 adolescent-enrolled pairs from the CTS. We are currently conducting the third wave of data collection, with each wave occurring at about a 5-year interval. LTS subjects were first interviewed at age 12. The selection criteria for the additional twins were less limiting than for the LTS, requiring only that the twins be between 12 and 18 years of age at initial assessment and having no known abnormalities that would prevent participation, including an inability to provide meaningful consent. In contrast to other twin studies at IBG, opposite-sex twin pairs and a sibling in the same age range were included in the CADD. Consistent with most twin studies, MZ pairs and females are overrepresented, with 46% MZ, 26.5% same-sex DZ, and 27.5% opposite-sex DZ (male–male pairs: 334; female–female pairs: 393; male–female pairs: 275).

In 2008, the Center began developing a measure to assess HIV-related risk behaviors. The resulting instrument is the Modified Risk Behavior Questionnaire (MRBQ) and is composed of 30 questions assessing sexual behaviors, situations, and partner characteristics. The MRBQ is a flexible instrument that facilitates the development of multiple HIV-risk behavior or ‘risky sex’ scales. As a part of the Center’s efforts, we constructed an 11-item risk measure that assesses sexual activity, use of alcohol or drugs before or during sexual activity, and sexual partner characteristics (e.g., sex with an injection drug use or partner with different types of sexually transmitted diseases). The resulting distribution of this scale in our clinical probands and siblings, CTS, and LTS samples is shown in Figure 4. As shown, a greater percentage of clinical probands exhibited high-risk sexual behaviors as compared with their siblings or community twins whereas community samples more often endorsed few sexual risk behaviors. This figure does not correct for age differences, which may account for the similar rates of high-risk behaviors between the clinical probands and LTS twins, who at the time of assessment, were in their early 20s.

In addition to their participation in the primary study, twin subjects are contributing to a new CADD investigation of brain processes involved in risky decision-making and the response to the rewards/loss consequences of such decisions. Imaging data during this assessment are collected from a small subset of twins selected from the CTS for being high or average on measures of behavioral disinhibition (Young et al., 2000), while the entire LTS completes the task in a normal laboratory setting to facilitate the analyses of heritable influence.

Early Reading Development Study (ERDS)
The ERDS was a collaborative effort between researchers in Australia, Sweden, and the United States, utilizing a longitudinal twin study to investigate the genetic and environmental influences on pre-reading and early reading-related skills; it was also known as the International Longitudinal Twin Study (ILTS). For the Colorado component, from 2000 through 2005, parents of 4-year-old same-sex twins with no known abnormalities who lived within a 2-hour driving radius of Boulder were recruited from the CTR. As shown in Table 1, 489 families were enrolled for the preschool visit. Although no effort was made to balance gender and zygosity status was unknown at the time of recruitment, the sample is roughly equal for both.
Subsequently, three waves of data were collected in the summers after the first 3 years of formal schooling. For the international study a fifth and final wave was collected after completion of fourth grade. However, in the United States, a sixth wave of data collection, which has been expanded to include measures of executive function, Attention Deficit Hyperactivity Disorder (ADHD), and math skills, has been initiated. Results from previous waves have shown substantial heritability for reading, inattention, and hyperactivity-impulsivity across the first three school years (Ebejer et al., 2010) and this new study will provide an opportunity to extend this research to a broader developmental range. Two years of data collection have been completed and seems to be on track to maintain the remarkable retention rate (96% or better at the earlier waves).

**Colorado Learning Disabilities Research Center (CLDRC)**

In our previous report we stated that there was one twin study at IBG, which had not drawn its sample from the CTR, although some of its participants were also participants in CTR studies. Now, however, the CLDRC, an ongoing project focused on the genetic and environmental etiologies of reading deficits (DeFries et al., 1997) is drawing some of its subjects from the registry. Subjects had previously been recruited from schools specifically for having reading problems or being appropriate controls. In the last few years, privacy concerns and resource issues have resulted in fewer schools being willing to undertake the recruitment responsibility. In 2009, the CLDRC began recruiting directly from the registry, first identifying families with appropriate age twins (8- to 18-year-olds) who also live within a 2-hour driving range of Boulder, then interviewing a parent and obtaining school records to assess whether the twins meet study criteria.

Twins who have had difficulty in reading or who have shown symptoms of ADHD, as well as twins who have been learning without undue difficulty and who have not shown symptoms of ADHD, are recruited. Additionally, they must have no vision or hearing impairments, and no history of seizure disorders. Following initial contact and explanation
of the study, interested parents fill out consents and ADHD questionnaires, mail in copies of reading scores, and request their twins’ teachers to fill out ADHD questionnaires. Parents next participate in phone interviews regarding reading, ADHD, and health history. Twin pairs who meet criteria based on all of the above are invited to participate. Reading scores, parent and teacher ADHD questionnaires, and parent phone interviews determine selection and grouping of participants.

Since this supplemental recruitment began, more than 50 pairs, about 30% of the CLDRC sample ascertained from 2009 to the present, have been enrolled from the CTR rather than from schools. It is anticipated that this proportion may increase as school recruitment becomes increasingly difficult.

Recent analyses on the etiology of reading disability and ADHD show that both disorders share a common cognitive deficit in processing speed that is primarily due to genetic influences (Willcutt et al., 2010).

### Molecular Genetics

In most of the registry studies, DNA saliva samples have been collected and are stored in molecular genetics facilities for use in any of the studies. Table 2 shows some of the polymorphisms currently available, followed by several examples illustrating these uses.

<table>
<thead>
<tr>
<th>Polymorphism</th>
<th>Gene</th>
<th>Variable number tandem repeats (VNTR)</th>
<th>Single nucleotide polymorphisms (SNP)</th>
<th>Whole genome SNP array (WGA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHTTLPR</td>
<td>Serotonin transporter</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>DAT1</td>
<td>Dopamine transporter</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>DRD4</td>
<td>Dopamine receptor</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>MAOA-uVNTR</td>
<td>Monoamine oxidase A</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>rs4650</td>
<td>COMT</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>rs1799971</td>
<td>Mu opioid receptor</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>rs25531</td>
<td>Serotonin transporter</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>rs12945042</td>
<td>Serotonin transporter</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>rs1042173</td>
<td>Serotonin transporter</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>rs8023462</td>
<td>CHRNA5/A3/B4 cluster</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>rs16969968</td>
<td>CHRNA5/A3/B4 cluster</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>rs4950</td>
<td>CHRNA5/A3/B4 cluster</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>rs1800497</td>
<td>Dopamine receptor D2</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>rs1049353</td>
<td>Cannabinoid receptor 1</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>rs237889</td>
<td>Oxytocin receptor</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>1 Million Affymetrix</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Note: * Colorado Twin Registry; **Longitudinal Twin Sample; *Colorado Twin Sample.

Molecular Genetics

In most of the registry studies, DNA saliva samples have been collected and are stored in molecular genetics facilities for use in any of the studies. Table 2 shows some of the polymorphisms currently available, followed by several examples illustrating these uses.

Sakai et al. (2012) examined the role of variation in the oxytocin receptor (OXTR) gene in a case–control sample of adolescents with and without conduct disorder and their parents. Variation in the OXTR gene has been implicated in affiliative behaviors characterized by emotional exchanges and pro-social behaviors. Family-based tests revealed only a trending association for one single nucleotide polymorphism (SNP, $p = .06$), rs2139184. Case-control analyses of adolescents with and without CD implicated rs2268492 ($p = .05$), while patient-parent versus control-parent analyses implicated rs237889 ($p = .004$). None of these suggested SNPs, however, remained significant after correcting for multiple testing.

Munn-Chernoff et al. (2012) adopted a gene-based approach to testing variation in and around the serotonin transporter (SLC6A4) in a sample of adolescent and young adult female twins and female non-twin siblings. In particular, the authors examined the role of seven polymorphisms in weight and shape concerns and behaviors and binge eating. Results from family-based tests failed to implicate any of this variation in the etiology of these two disordered eating behaviors.

Hartman et al. (2009) examined the genetic association between variation in the cannabinoid receptor 1 gene and cannabis dependence symptoms. Adolescents and young adults were recruited from community and clinical populations within the CADD study. Seven SNPs were characterized and examined, using both case–control and family-based designs. Both case–control and family-based tests provided mixed results implicating rs1049353 with having one or more dependence symptom. This study did not replicate a previous finding (Hopfer et al., 2006) that implicated rs806380 in a smaller sample drawn from CADD. Further work to clarify the relationship between rs1049353 and cannabis dependence is suggested by these results.

### Discussion

The CTR has proven an incredibly valuable collaboration between the CDH, the school districts of Colorado, and the IBG at the University of Colorado, Boulder since its inception 30 years ago. Five active studies that span ages from infancy to young adulthood, and domains as divergent as early reading, emotional development, and substance involvement, have made use of this resource to develop study samples. This is in addition to earlier concluded studies that have drawn subjects from the CTR, and which were described in some detail in the earlier report from our group (Rhea et al., 2006). Continuing enrollment in the CTR remains strong, and families have continued to express interest in being contacted for both new and ongoing studies, as shown by parental willingness to complete a basic demographic survey without compensation. A number of grants that would make use of these willing participants are currently in preparation for submission over the next year. But as shown in the discussion of current studies, new avenues for research and collaboration with other groups outside Colorado continue to make use of the CTR samples. Thus, the ERDS/ILTS project is currently collecting measures that parallel those of Dr. Stephen Petrill of Ohio State University (http://wrrp.psy.ohio-state.edu/). Cognitive data from CTR-derived twins were used in the recent multinational...
study by Haworth et al. (2010) that found increasing genetic influence on general cognitive ability from childhood to young adulthood. Our group has been actively engaged in documenting the extensive wealth of data collected from the CTR-derived studies in such a way that interested potential collaborators can search for overlap with their own studies. We are proud of the many ways the CTR twin samples have been used, and we look forward to their continuing usefulness in the years ahead.

Acknowledgments

The CTR owes its existence to the invaluable cooperation of the CDH, Division of Vital Statistics and the many participating families. It has been funded across the years by a CRCW grant from the University of Colorado; National Institutes of Health grants HD19802, HD010333, HD18426, MH43899, DA05131, DA011015, and HD027802; the John D. and Catherine T. MacArthur Foundation; and Veterans Administration grant 1296.07.1629B. The founders and directors were R. P. Plomin, J. C. DeFries, R. N. Emde, D. W. Fulker, J. K. Hewitt, and R. K. Olson, who were assisted by many fellow investigators and able staff.

References


