Introduction

Cleft lip with or without cleft palate (CL/P) is one of the most common congenital malformations in humans involving various genetic and environmental risk factors. The prevalence of CL/P varies according to geographical location, ethnicity, race, gender, and socioeconomic status, affecting approximately 1 in 800 live births worldwide. Genetic studies aim to understand the mechanisms contributory to a phenotype by measuring the association between genetic variants and also between genetic variants and phenotype population. Genome-wide association studies are standard tools used to discover genetic loci related to a trait of interest. Genetic association studies are generally divided into two main design types: population-based studies and family-based studies. The epidemiological population-based studies comprise unrelated individuals that directly compare the frequency of genetic variants between (usually independent) cases and controls. The alternative to population-based studies (case-control designs) includes various family-based study designs that comprise related individuals. An example of such a study is a case-parent trio design study, which is commonly employed in genetics to identify the variants underlying complex human disease where transmission of alleles from parents to offspring is studied. This article describes the fundamentals of case-parent trio study, trio design and its significances, statistical methods, and limitations of the trio studies.
Trio Design

The model of Mendelian inheritance offers a straightforward rationalization of the genetic design of a trait. It prescribes that one gene locus produces the trait in either recessive or dominant pattern in families. However, many traits do not follow such a straightforward model of genetic architecture. Based on Mendelian inheritance principles, every parental allele has a 50% probability of being transmitted. The basic trio design is shown in Fig. 1.

Under the null hypothesis of no association between the disease and the marker, every parent transmits one of their two alleles to each offspring at random with a probability of 50/50 and independently of the other parent and any other offspring. For the example in Fig. 1, the mother can only transmit the A allele. However, the father can transmit either A or B with a probability of 50/50. This holds whenever there is no selection of the offspring related to the marker in question. Thus, when the parent’s genotypes are known, one can easily calculate the distribution of the offspring genotypes under the null hypothesis.

Case–Parent Trio Design

Case–parent trio studies are commonly employed in genetics to identify the variants underlying complex human disease. It involves the diseased/affected child and their parents (father and mother). In these studies, the affected children are selected from a population and then the affected children and their parents are genotyped. Therefore, it does not require control sample data necessarily because the non-transmitted parental alleles or genotypes are used as “controls” for the transmitted alleles or genotypes.

Case–parent trio designs study the effects of a gene marker and gene–environment (GxE) interaction. They are helpful to study the transmission of genetic variants between parents and offspring and how genetic variants differ between the affected individual(s) and the unaffected individuals within a family. Trio designs are popular and an alternate to population-based case–control studies where unrelated individuals are being used for the detection of variants underlying common complex disease risk. Case–parent trio study designs also guard against population stratification and therefore the resulting type I error inflation, which is commonly observed in population-based studies. This technique is most robust to population admixture, and the targeted sequencing of trios is powerful to identify genetic variants that alter the function of the gene.

The family-based association studies were initially proposed by Falk and Rubinstein to avoid false-positive association test results. They collected the parental genotypes of every case used the nontransmitted parental alleles as a control sample because the cases and controls are matched in genetic ancestry and are therefore robust to population stratification. Later these tests were developed by other authors Ott (1989), and Terwilliger and Ott, (1992). Furthermore, Spielman et al. proposed “transmission disequilibrium test” (TDT) to identify preferential transmission of alleles from parent to the affected child.

In most of the genetic association studies, investigators compare an allele or genotype frequencies in unrelated case and control subjects or examine preferential allele transmissions from parents to affected offspring. It is important to contemplate parent of origin effects while studying congenital birth defects such as CL/P because maternal genotype controls the in utero environment of the developing fetus and separates maternal genotypic effects from imprinting effects. The functional activity of some genes or chromosomal regions depends on whether they are transmitted maternally or paternally, and this epigenetic phenomenon is termed as genomic imprinting. Another commonly used method in case–parent trio study is to use the non-transmitted genotypes of parents to affected offspring as control (also known as pseudo controls or complements).

GxE interactions have significant scientific and public health implications. In the absence of GxE interactions, using the affected individual and their parents has been proved to be the most robust method to assess disease associations with candidate genes, also as an assessment of both linkage and allelic association with genetic markers. The aim of collecting the genetic data from family members aids in (1) enriching the genetic analysis, (2) increasing the power of study, (3) obtaining correct mutations, (4) addressing the priority regarding population structure bias, and methods of mixing genetic information from unrelated cases and controls.

This design has the most significant advantage of being robust to population structure bias as the estimation of gene transmission would be within the families, conditioning on parental genotypes, and it is an unbiased method as all the subjects are...
from the same family.\textsuperscript{53} Additionally, family data are always advantageous because certain models will test for maternally mediated and parent-of-origin effects that are significantly relevant when studying congenital birth defects in humans.\textsuperscript{54}

**Case–Parent Trio Design Significances**

Several genetic studies focused on numerous complex genetic and epigenetic effects associated with a nonsyndromic CL/P, suggesting that the mother transfers half of her genome to the offspring and provides the environment for the fetus. Variation within the mother's genome could affect the intrauterine environment to the development of the fetus. Parent-of-origin effects, where the effect of inherited DNA depends on whether it is transmitted from the mother or the father, may be difficult to study with a population-based case–control design. More complex genetic effects such as maternal genotype effects, maternal–fetal interactions, and parent-of-origin effects can be best studied using the case–parent trio design, where affected cases and their parents are genotyped.\textsuperscript{55,56}

**Advantages of Case–Parent Trio Design Studies**

The advantages of case–parent trio design studies are as follows:\textsuperscript{15,17,20,21,54,57}

- Robustness in sample collection and false conclusions because of population stratification and ethnic heterogeneousness.
- Testing directly for maternal versus paternal effects.
- It permits the effects of fetal genotype versus parental origin in a very robust manner.
- Correct mutations are often obtained.
- It is unbiased as they share the same genetic ancestry.
- Simplex families (only the affected member) will be used effectively to check for linkage in the presence of disequilibrium.
- Family information will be helpful to check maternal genotype and parent-of-origin effects.
- It minimizes problems with confounding that plague traditional case–control designs as a result of the observed case is always compared with "controls" obtained from the same family.
- It is useful to check for GxE interaction under the conditional logistic or logistic regression framework.

**The Statistical Methods for Case–Parent Trio Studies**

Case–parent trio design studies employ a distinct statistical procedure to examine the linkage and association of a gene with the trait. When studying families, the transmission of alleles from the parent to the offspring is analyzed. In this sense, parents are used as genetic controls for their children.\textsuperscript{38,58–62} There are three main methods in which trios are analyzed, namely: the Transmission Disequilibrium Test (TDT), Conditional on parental genotype (CPG) approach, and combined likelihood method.\textsuperscript{63–67}

- **Transmission Disequilibrium Test (TDT):** it compares the offspring to its antiself, and the unobserved instance of an offspring with each nontransmitted allele.
- **Conditional on parental genotype (CPG):** it compares the proband to unobserved pseudosiblings with all possible transmission patterns. Again, these pseudosiblings are not determined; however, the results of conditioning on the parental genotype are used to make a matched analysis.
- **Combined likelihood approach:** this method collectively models the likelihood of parental and case genotypes.

The most common statistical procedure used in case–parent trio design study is the TDT introduced by Spielman et al. It is a robust test for the linkage and disequilibrium in a variety of complex diseases that are associated with genetic etiology. It is more sensitive than haplotype sharing tests and requires only simplex families.\textsuperscript{68}

The CPG approach developed by Schaid and Sommer models the probability of an affected child's genotype CPG as a function of the genotype relative risks (GRRs) of the child.\textsuperscript{69,70} This CPG likelihood approach permits versatile modelling of the GRRs, which may be quantifiable using standard maximum-likelihood procedures. GRR is also helpful in estimating the magnitude of the gene–disease association in case–parent trio studies.\textsuperscript{71,72}

There are various commercially available software suites for case–parent trio study designs such as Golden Helix (http://www.goldenhelix.com) and freely accessible software system PLINK.\textsuperscript{73,74} PLINK is arguably the most commonly used software and a very powerful environment to carry out genetic association studies. It includes an integrated module to carry out the allelic TDT introduced by Spielman et al. to assess the marginal SNP effects on the phenotype in case–parent trio studies.

Unlike the case–control approach, family-based methods are not subject to inflation of results because of population substructure, as they examine the transmission between parents and offspring and remove any potential impact caused by population allele frequency variations. However, family-based methods have less statistical power than the population-based case–control designs for a similar number of individual studied.\textsuperscript{75}

**Disadvantages of Case–Parent Trio Design Studies**

The disadvantages of case–parent trio design studies are as follows:\textsuperscript{20,48,52,76–79}

- Not all case–parent trios are informative (only heterozygous parents give information), some are intrinsically uninformative about allele transmission.
- They provide a robust assessment of gene effects; however, they do not impart complete information on GxE interaction.
- Not all of the parental genotypes (one or both) are available for the study because of death, separated parents (divorced), or false paternity.
- Some families/parents may refuse to participate in clinical research.
- The cost of genotyping is expensive for the case–parent trios (three persons) than for the case–controls (two
persons) without a corresponding linear increase in statistical power.
• Difficulty in recruiting large samples and time-consuming.
• Discarding those families with missing one or both parental genotypes will cause statistical power loss.
• If there is no deviation from multilocus Hardy–Weinberg equilibrium, the case–parent trio design will have no power to discover even tight linkage.

Conclusion
Case–parent trio studies are commonly used in genetics to identify the variants underlying complex human disease. They are useful to study the transmission of genetic variants between parents and offspring and how genetic variants differ between the affected individual(s) and the unaffected individuals within a family. This design has the most significant advantage of being robust to population structure bias and can test for more complex genetic effects such as maternal genotype effects, maternal–fetal interactions, and parent-of-origin effects that are significantly relevant when studying congenital birth defects.

In the absence of GxE interactions, using cases and their parents has proved to be a powerful test to assess disease associations with candidate genes. Case–parent trio study designs are adequate to estimate the genetic effects, and therefore the aim of collecting genetic information from family members enrich the genetic analysis, increase the power of study, help in obtaining correct mutations, and address the priority regarding population structure bias and methods of combining genetic information from unrelated cases and controls.

Conflict of Interest
None declared.

References
1 Schutte BC, Murray JC. The many faces and factors of orofacial clefts. Hum Mol Genet 1999;8(10):1853–1859
2 Bender PL. Genetics of cleft lip and palate. J Pediatr Nurs 2000;15(04):242–249
3 Spritz RA. The genetics and epigenetics of orofacial clefts. Curr Opin Pediatr 2001;13(06):556–560
4 Murray JC. Gene/environment causes of cleft lip and/or palate. Clin Genet 2002;61(04):248–256
5 Wong FK, Hagg U. An update on the aetiology of orofacial clefts. Hong Kong Med J 2004;10(05):331–336
6 Dixon MJ, Marazita ML, Beatty TH, Murray JC. Cleft lip and palate: understanding genetic and environmental influences. Nat Rev Genet 2011;12(03):167–178
7 Kohli SS, Kohli VS. A comprehensive review of the genetic basis of cleft lip and palate. J Oral Maxillofac Pathol 2012;16(01):64–72
8 Funato N, Nakamura M. Identification of shared and unique gene families associated with oral clefts. Int J Oral Sci 2017;9(02):104–109
9 Grant SF, Wang K, Zhang H, et al. A genome-wide association study identifies a locus for nonsyndromic cleft lip with or without cleft palate on 8q24. J Pediatr 2009;155(06):909–913
10 Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 2009;5(06):e1000529
11 Shi M, Mostowska A, Jugessur A, et al. Identification of microdeletions in candidate genes for cleft lip and/or palate. Birth Defects Res A Clin Mol Teratol 2009;85(01):42–51
12 Beatty TH, Murray JC, Marazita ML, et al. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. Nat Genet 2010;42(06):525–529
13 Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. PLoS Biol 2010;8(01):e1000294
14 Laurie CC, Doheny KF, Mirel DB, et al. Quality control and quality assurance in genotypic data for genomewide association studies. Genet Epidemiol 2010;34:591–602
15 Khoury MJ. Case-parental control method in the search for disease-susceptibility genes. Am J Hum Genet 1994;55(02):414–415
16 Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science 1996;273(5281):1516–1517
17 Wilcox AJ, Weinberg CR, Lie RT. Distinguishing the effects of maternal and offspring genes through studies of “case-parent triads.” Am J Epidemiol 1998;148(09):893–901
18 Weinberg CR. Methods for detection of parent-of-origin effects in genetic studies of case-parents triads. Am J Hum Genet 1999;65(01):229–235
19 Umbach DM, Weinberg CR. The use of case-parent triads to study joint effects of genotype and exposure. Am J Hum Genet 2000;66(01):251–261
20 Beatty TH, Hetmanski JB, Zeiger JS, et al. Testing candidate genes for non-syndromic oral clefts using a case-parent trio design. Genet Epidemiol 2002;22(01):1–11
21 Sull JW, Liang KY, Hetmanski JB, et al. Maternal transmission effects of the PAX genes among cleft case-parent trios from four populations. Eur J Hum Genet 2009;17(06):831–839
22 Epstein MP, Veal CD, Trembath RC, Barker JN, Li C, Satten GA. Genetic association analysis using data from triads and unrelated subjects. Am J Hum Genet 2005;76(04):592–608
23 Beatty TH, Taub MA, Scott AF, et al. Confirming genes influencing risk to cleft lip with/without cleft palate in a case-parent trio study. Hum Genet 2013;132(07):771–781
24 Li YM, Xiang Y. Detecting disease association with rare variants in case-parents studies. J Hum Genet 2017;62(05):549–552
25 Bateson W. Mendel’s Principles of Heredity. Cambridge: Cambridge University Press; 1909
26 Laird NM, Lange C. Family-based methods for linkage and association analysis. Adv Genet 2008;60:219–252
27 Laird NM, Lange C. The role of family-based designs in genomewide association studies. Stat Sci 2009;24:388–397
28 Sull JW, Liang KY, Hetmanski JB, et al. Differential parental transmission of markers in RUNX2 among cleft case-parent trios from four populations. Genet Epidemiol 2008;32(06):505–512
29 Schwender H, Li Q, Neumann C, et al. Detecting disease variants in case-parent trio studies using the bioconductor software package trin. Genet Epidemiol 2014;38(06):516–522
30 Zhu H, Kartiko S, Finnell RH. Importance of gene-environment interactions in the etiology of selected birth defects. Clin Genet 2009;75(05):409–423
31 Cordell HJ. Estimation and testing of gene-environment interactions in family-based association studies. Genomics 2009a93(01):5–9
32 Cordell HJ. Detecting gene-gene interactions that underlie human diseases. Nat Rev Genet 2009b10(06):392–404
33 Chen YH, Lin HW. Simple association analysis combining data from trios/sibships and unrelated controls. Genet Epidemiol 2008;32(06):520–527
34 Laird NM, Horvath S, Xu X. Implementing a unified approach to family-based tests of association. Genet Epidemiol 2000;19(Suppl 1):S36–S42
35 Zhao H. Family-based association studies. Stat Methods Med Res 2000;9(06):563–587
