The Nature of Nurture: Using a Virtual-Parent Design to Test Parenting Effects on Children's Educational Attainment in Genotyped Families

Timothy C. Bates,[1] Brion S. Maher,[2] Sarah E. Medland,[3] Kerrie McAloney,[3] Margaret J. Wright,[4,5] Narelle K. Hansell,[4] Kenneth S. Kendler,[6] Nicholas G. Martin,[3] and Nathan A. Gillespie[6]

[1]*Department of Psychology, University of Edinburgh, Edinburgh, UK*
[2]*Bloomberg School of Public Health, Johns Hopkins, Baltimore, MD, USA*
[3]*Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia*
[4]*Queensland Brain Institute, University of Queensland, Brisbane, Queensland, Australia*
[5]*Centre for Advanced Imaging, University of Queensland, Brisbane, Queensland, Australia*
[6]*Virginia Institute of Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, USA*

Research on environmental and genetic pathways to complex traits such as educational attainment (EA) is confounded by uncertainty over whether correlations reflect effects of transmitted parental genes, causal family environments, or some, possibly interactive, mixture of both. Thus, an aggregate of thousands of alleles associated with EA (a polygenic risk score; PRS) may tap parental behaviors and home environments promoting EA in the offspring. New methods for unpicking and determining these causal pathways are required. Here, we utilize the fact that parents pass, at random, 50% of their genome to a given offspring to create independent scores for the transmitted alleles (conventional EA PRS) and a parental score based on alleles *not transmitted* to the offspring (EA VP_PRS). The formal effect of non-transmitted alleles on offspring attainment was tested in 2,333 genotyped twins for whom high-quality measures of EA, assessed at age 17 years, were available, and whose parents were also genotyped. Four key findings were observed. First, the EA PRS and EA VP_PRS were empirically independent, validating the virtual-parent design. Second, in this family-based design, children's own EA PRS significantly predicted their EA ($\beta = 0.15$), ruling out stratification confounds as a cause of the association of attainment with the EA PRS. Third, parental EA PRS predicted the SES environment parents provided to offspring ($\beta = 0.20$), and parental SES and offspring EA were significantly associated ($\beta = 0.33$). This would suggest that the EA PRS is at least as strongly linked to social competence as it is to EA, leading to higher attained SES in parents and, therefore, a higher experienced SES for children. In a full structural equation model taking account of family genetic relatedness across multiple siblings the non-transmitted allele effects were estimated at similar values; but, in this more complex model, confidence intervals included zero. A test using the forthcoming EA3 PRS may clarify this outcome. The virtual-parent method may be applied to clarify causality in other phenotypes where observational evidence suggests parenting may moderate expression of other outcomes, for instance in psychiatry.

■ **Keywords:** educational attainment, parental environment, non-transmitted genotype, virtual-parent design

Home environments and socioeconomic status (SES) correlate moderately with offspring educational attainment (EA) ($r \sim 0.3$; Duncan et al., 2011). However, as parents both transmit genes and create environments for offspring, these correlations are confounded. Measured environments, such as exposure to books or language in the home that are correlated with attainment (Martin et al., 2009) may, therefore, not be causal (cultural transmission), and instead reflect passive gene-environment correlation (Puglisi et al.,

2017). This confounding can take on complex, evocative forms such that environments too are heritable, either due

to passive parental creation of the environment, or active evocation of the environment by offspring (Plomin, 1994). Here, we utilize recent advances in genetics in the form of polygenic risk scores (PRS) for EA, combined with the well-known process of meiosis in which each parent passes a random 50% of their genome to an offspring (with 50% un-transmitted). This allowed use of the pseudo-control method (Cordell et al., 2004) to create polygenic scores for parents consisting only of EA-associated alleles they did not transmit to their offspring (Howey, 2014), a parental control EA2 risk score: PC_EA PRS. This PC_EA PRS was used in a sample 2,333 genotyped twins and their parents to test for evidence of parenting effects by testing the effect of non-transmitted parental alleles on a high-quality measure of offspring EA, un-confounded by offspring genetics.

Below, we briefly describe the causal ambiguity of observational data, the development of PRS, and the pseudo-control method utilized here to create independent genetic scores for parents and offspring. Following this, we test hypotheses regarding how parenting competence, SES niche-construction, and children's genetic potentials impact on EA.

Background

Correlations between parenting and outcomes such as offspring EA are ubiquitous. However, as parents transmit genes and create environments for offspring (Neale & Cardon, 1992), these correlations are confounded. Even factors often favored on intuitive grounds as causes of cognitive development — such as complexity and number of words spoken in the home environment — may not be causal. For instance, Puglisi et al. (2017) conducted a longitudinal study of children at risk for reading disorder and found that while variation in the home literacy environment was a significant predictor of children's reading and language skills, controlling for maternal language ability removed this effect. While mothers with better language and phonological skills had children with higher reading and language skills, and did predict such factors as storybook exposure, factors such as storybook exposure were not significant predictors of offspring reading ability. By contrast, the study found that direct literacy training (such as phonics) did independently predict acquired reading skill, but was unrelated to maternal language ability. This example exemplifies both the possibility of intervention and the risks of confounding leading to incorrect research conclusions. It also indicates that effects on child outcomes may reach beyond genetic transmission from parents to include factors such as literacy training. These are not direct consequences of parental ability, but are instead factors under at least partial parental control and involve constructing niches for children that include such effective interventions. In this way, they reflect effective parenting competence. This competence may both play an important causal role and itself be heritable.

In such cases, not only are genetics transmitted to the child, but their relevant environment, too, is heritable.

The Molecular Era of Research and Polygenic Risk Scores

By contrast, the support for genetic transmission effects is much less controversial. Since the 1970s, a wealth of behavior genetic studies have supported a large role of genetics as a cause of differences in EA (Krapohl et al., 2014; Martin & Martin, 1975). This same evidence base also suggests that attainment is a result not only of a mixture of environmental and genetic influences (Plomin & Loehlin, 1989), but that environments can moderate genetic effects (Tucker-Drob & Bates, 2016). Building on these results, large consortia have been assembled to test for average risk factors in the genome, with initial successes in psychiatric disorder (International Schizophrenia Consortium et al., 2009) and now in studies of EA (Cesarini & Visscher, 2017; Okbay et al., 2016; Rietveld et al., 2013).

Based on the hypothesis that most traits are highly polygenic, that is, that much of the behavioral variation relates to markers of very small effect, each of which in aggregate contribute to an overall polygenic risk, it was recognized that detection of this risk required sample size orders of magnitude greater than previously collected — from around half a million to several million individuals (Zhang et al., 2017). At scale, these molecular genetic methods have identified over 100 gene variants affecting educational outcomes (Okbay et al., 2016) and intelligence (Sniekers et al., 2017). In such samples, markers from an initial large GWAS are ranked according to their support for association, and this set of markers is tested for prediction in an independent replication sample. For any given individual, the PRS is simply the weighted sum of their genetic state at each of many hundreds of alleles associated with the trait (but often weighted by effect size). Such studies have validated twin study results and allowed use of independent samples to establishing shared polygenic risk by testing for genetic correlations with other conditions (Kendler, 2016) and non-behavioral measures, such as brain volumes (Franke et al., 2016), as well as allowing tests of G × E interactions using molecular markers in studies of unrelated individuals (Belsky et al., 2016).

Applied to education, PRS scores account for around 10% of variance in years of education, based purely on genetic associations (Okbay et al., 2016). It is this EA PRS that we utilized in the present study. The EA PRS has been replicated in samples not used in the original discovery, and these suggest EA2 is related to parental SES (i.e., higher EA2 scores in children are associated with higher parental SES; Selzam et al., 2017), as well as (using an earlier, weaker PRS) with high attained SES in offspring (Belsky et al., 2016). Interestingly, Selzam et al. (2017) reported that the EA PRS was associated more strongly with the SES of children's

parents ($R^2 = 7\%$) than it was with the children's cognitive ability ($R^2 = 3.5\%$).

This association of higher parental SES with higher EA scores in their children is compatible with at least two interpretations. The association may reflect effects of parents' genes, which both influence their attainment of higher SES and which, when passed on to their offspring, influence their children's higher attainment (rather than this resulting from the parentally created SES). Alternatively, the environmental niche created by the parents — housing, wealth, and prestige among many other factors (Galobardes et al., 2006) — may foster offspring EA via environmental routes, either independent of the genes transmitted to children (passive G × E covariation), or in interaction with these genes whereby active G × E covariation is created; for instance, when children geared to learn evoke additional support from parents, or where the presence of learning opportunities are capitalized upon more strongly by offspring with higher EA PR scores (Bronfenbrenner & Ceci, 1994; Rowe, 2001; Scarr & McCartney, 1983).

Disentangling Causal Effects of Offspring Genetics From Parent Genetics

The need to disentangle different forms of transmission and methods to enable such disentanglement has been central to the work of behavior geneticists seeking to understand family and cultural effects (Eaves et al., 2014). This seminal work proposed statistical controls to disentangle transmission via parent behavior and parent genetics. These designs require the ability to randomize parent genetics with respect to offspring, the classic example being an adoption design. In the conventional adoption design; however, factors such as large sample sizes, random entrance to the study and, most importantly, random allocation of children to adopting homes (Loehlin, 2000) are seldom possible (Plomin & Loehlin, 1989). Instead, our study uses polygenic scores associated with education attainment, along with the basic biological mechanism of meiosis in which parental gametes (egg and sperm) each contain only a haploid copy of the parental genome to their offspring, with these halves being combined to form the offspring zygote to create a 'virtual parent' design.

Independent Assortment and Independence of Transmitted and Non-Transmitted PRS

Mendel's second law, the 'law of independent assortment', reflects the fact that gamete formation results in a halving of the chromosomal complement, with alleles having a 50% chance of being transmitted to the gamete. Researchers in genetic epidemiology capitalized on this independent assortment to create the concept of a 'pseudo-control', consisting of non-transmitted parental alleles. Pseudo controls have been widely used in the study of gene-environment interactions and parent-of-origin effects among factors (Cordell et al., 2004), and software such as PseudoCons (Howey, 2014) exists to readily generate these pseudo-control genomes. Importantly, whereas offspring and parent genotypes share 50% of their DNA, the offspring genotype and pseudo-con or non-transmitted genotype are uncorrelated. Moreover, this pseudo-control genotype, consisting of the non-transmitted parental alleles, can itself be scored using PRS, leading to the creation of uncorrelated offspring- and non-transmitted parental risk scores.

In the present method, two PRS are formed: one for the offspring, consisting of the EA-associated alleles they have inherited (EA PRS) and a second parental-control risk score (PC_EA PRS) consisting of the EA-associated alleles in the parental genomes that were not passed on to their offspring. Here, we use these two independent scores to address the following question: To what extent do non-transmitted genetic variants in parents create environments influencing the attainment of their offspring?

To implement this design, we used a high-quality measure of attainment and the well-validated PRS for EA (Okbay et al., 2016), assessed in 2,335 offspring and their 2,145 parents where both offspring and both parents had been genotyped, and SES recorded for parents, along with a measure of highest educational outcome and a comprehensive seven-hour assessment of EA (terminal high-school core-skills scores) undertaken by the offspring. This allowed computing both the conventional PRS for EA, parental pseudo-control PC_EA PRS, with measured SES and educational outcomes. These data allowed us to test the following four hypotheses:

1. That children's EA PRS is significantly associated with their attainment.
2. That parents' EA PRS is associated with the SES they attain.
3. That parent SES is associated with offspring EA.
4. That offspring EA is associated with non-transmitted parental EA PRS, reflecting cultural transmission via creation of EA-linked environments.

Methods
Subjects
The sample consisted of all 2,335 children and their genotyped parents available within the Brisbane Adolescent Twin Study (Wright & Martin, 2004) for whom EA data and genotype data were available, along with genotype and SES data for their parents: 1,333 mothers (mean age at testing of twins, 17.15 years, $SD = 0.39$) and 1,002 fathers (mean age at test 17.2 years, $SD = 0.41$). Exclusion criteria for entry to the cohort were significant head injury, neurological, or psychiatric illness, substance dependence, or chronic use of medications with central nervous system effects. Subjects were genotyped on the 610K Illumina genome-wide SNP

platform (Medland et al., 2009). The study was approved by the Human Research Ethics Committee at QIMR Berghofer Medical Research Institute.

Phenotypic Measures

The Queensland Core Skills Test (QCST; Queensland Studies Authority, 2003) is a seven-hour comprehensive assessment of EA taken over two consecutive days and sat by most (~85%) Queensland school-leavers in their final year of schooling, typically aged 17. It is used in high-stakes outcomes, such as selection for entry into tertiary education. As such, it was designed specifically to assess general scholastically acquired higher-order skills and provides a global index of academic achievement across a very diverse range of skills, including mathematical problem solving, comprehension and explanation, interpretation (e.g., cartoons, photographs, and flow charts), grasp of scientific methodology, reading graphs, spelling and basic calculations, understanding spatial and mechanical relationships, and producing written prose, as well as creative ability and presentational skills (Queensland Studies Authority, 2003). The QCST exam differs from year to year, and to allow analysis, scores were standardized using the means and standard deviation of the entire Queensland sample within each year. In previous work using this sample, and taking advantage of the twin design, we have shown QCST scores to have a heritability of 0.76 (from bivariate twin modeling) and with genetic correlations with measures of cognitive ability ranging from 0.64 to 0.91 depending on the specific cognitive test (Wainwright et al., 2005).

SES was assessed using the Australian Socioeconomic Index 2006 (AUSEI06) occupational status scale (McMillan et al., 2009). Where an individual's occupation was outside the current definitions of the labor force (e.g., housewife, unemployed), their partner's occupation was used to determine their SES. In the case where occupation information could not be ascertained, then an occupational score, based on years of education completed, was imputed.

Forming Polygenic Risk Scores for Educational Attainment in Offspring and Parents

For parents, the PseudoCons application (Howey, 2014) was used to create non-transmitted parental genomes. This program creates a control parent genome consisting only of those genetic variants that are not present in the offspring, a pseudo-control (Cordell et al., 2004).

PRS for EA2 were calculated in the conventional manner using the LDpred application (Vilhjalmsson et al., 2015) using SNP effect sizes from 100% of SNPs in the second EA GWAS (Okbay et al., 2016), recomputed leaving out the BATS sample (which formed part of the original EA2 project), and using European population LD information from the 1,000 Genomes reference set. Polygenic scores were standardized to have mean of zero and a standard de-

viation of 1. EA2 PRS were computed for the offspring and parental pseudo-control genomes in the same way.

Results

As an initial validity check, the independence of the non-transmitted and transmitted EA polygenic scores was verified using a linear model. The association did not differ significantly from zero, $\beta = 0.01$ (95% CI [-0.04, 0.05]), $t = 0.31$, $p = .755$ (see Figure 1). EA is, of course, highly heritable creating a large non-independence in attainment nested within families, $\chi^2(1) = 242.69$, $p < .0001$, ICC = 0.57, so family ID and zygosity were entered as random variables in all models. Thus, models that fail to account for these covariances among the subjects will not be able to distinguish genuine parental-gene effects from confounding due to assortative mating and other factors. Two modeling approaches were used to control for this nesting and patterns of gene and environment sharing. In our initial analyses, we used a multilevel framework, using the linear mixed-effects (lme) function from the nlme package (Pinheiro et al., 2017). In addition, we tested the core hypothesis in a structural equation model, where we were able to model the full nesting within families via a multigroup approach, and also include genetic associations among twins and siblings. For instance, the non-transmitted EA2 PRS of one DZ twin, while uncorrelated with their own transmitted EA2 PRS, has expected correlations of 0.5 with the transmitted EA2 PRS of their co-twin and other siblings. Multilevel model results with nesting in family and zygosity are presented first.

1. Are children's EA polygenic scores associated with their EA? We first tested the prediction that in this family-control design, children with higher EA PRS would score higher on the QCST. This was confirmed in a multilevel model with QCST scores as the dependent variable and offspring EA PRS score as the predictor, controlling for age at which the QCST was taken, sex, and parental SES. Offspring EA PRS was a highly significant predictor of offspring QCST, $\beta = 0.15$ (95% CI [0.09, 0.21]), $\chi^2(1) = 17.8$, $p = < .001$: AIC = 1,736.797 (see Figure 2, panel A). This confirmed that EA PRS was associated with EA and that this association was not due to stratification or other artifacts of gene frequency, which are controlled in this within-family design.

2. Are parents' EA PRS associated with the SES they attain and provide to offspring? We next tested whether parents' own EA PR scores were positively associated with their own SES. In a linear model, mean parental EA PRS was a highly significant predictor of parental SES, $F(1, 1,245) = 52.55$, $p = 7.32*10^{-13}$, with a standardized $\beta = 0.20$ (95% CI [0.15, 0.26]). This association indicates a significant effect of parents' EA PRS on their own SES.

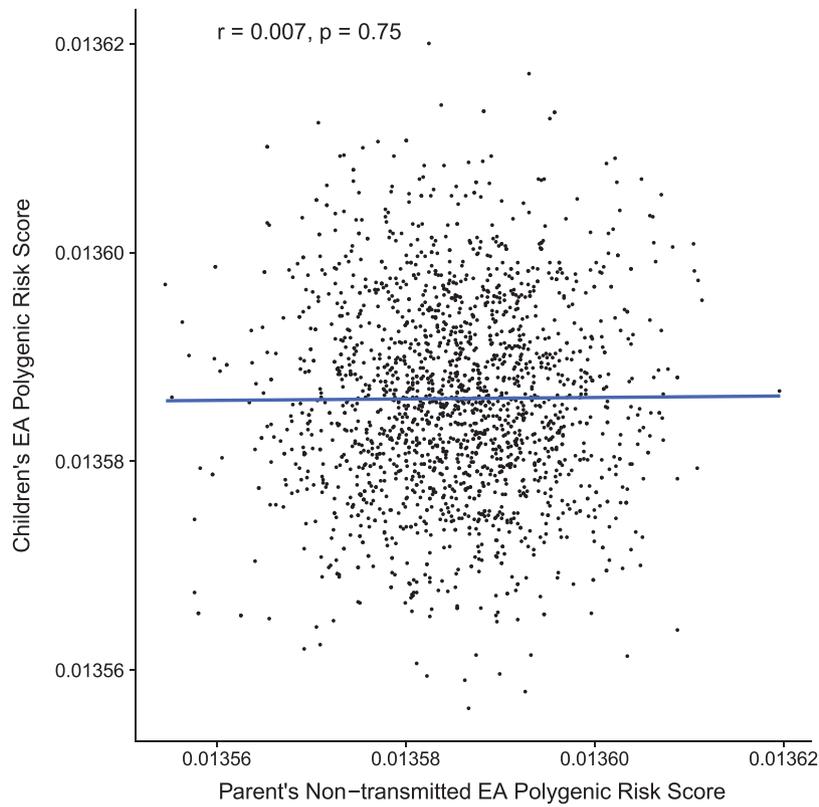**3. Are parental SES and offspring EA associated?** As shown in Figure 3, in the same model, SES was also a highly

FIGURE 1

(Colour online) Validating the independence of the parental non-transmitted EA PRS (*x*-axis) and children's EA PRS (*y*-axis).
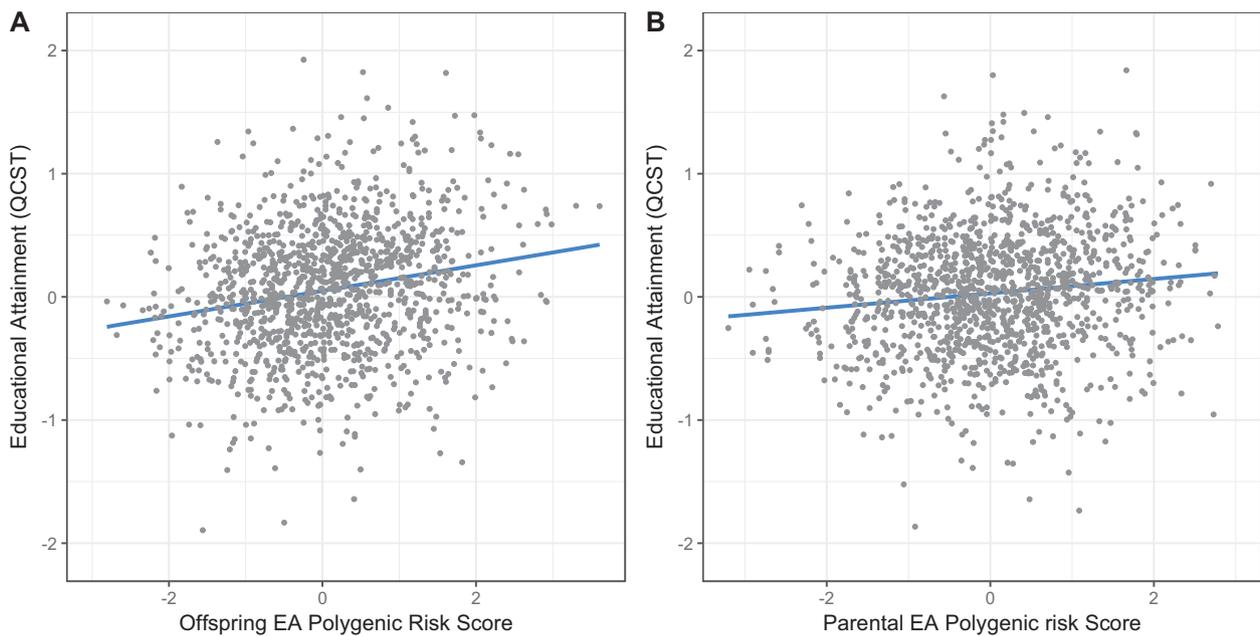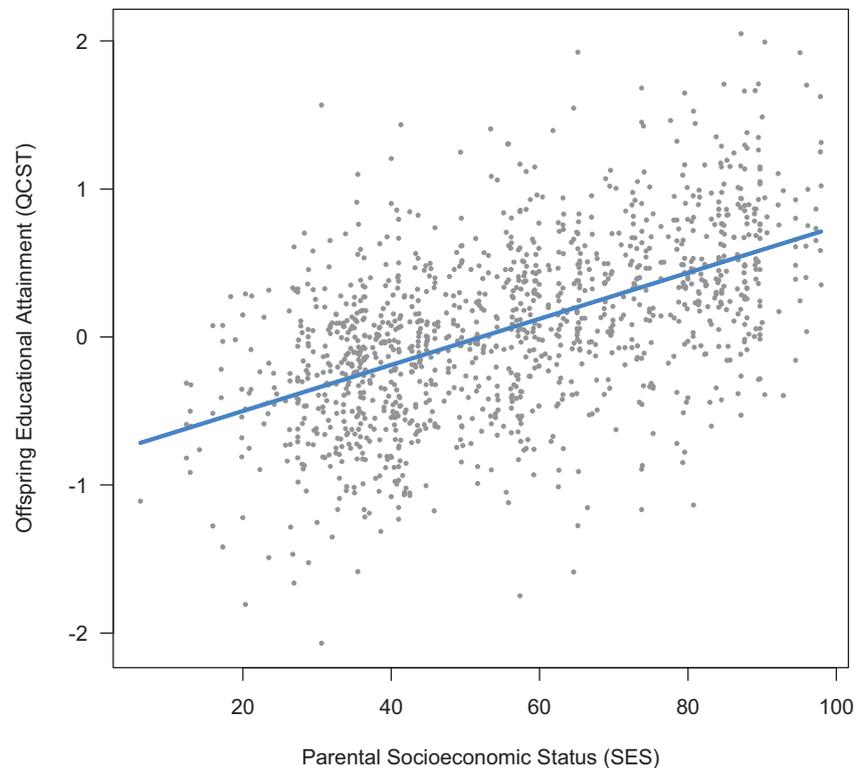

FIGURE 2

(Colour online) Plot of measured EA (Queensland Core-Skills Test scores) on the *y*-axis, predicted by, in panel A, offspring polygenic risk scores for educational attainment (EA PRS) and, in panel B, parental control (non-transmitted allele) EA polygenic risk scores (PC_EA PRS).

FIGURE 3

(Colour online) Parental measured SES as a predictor of offspring educational outcomes (Queensland Core Skills Test).

significant predictor of QCST, $\beta = 0.33$ [0.27, 0.4], $t(647) = 10.21$, $p < .001$. Because G × SES interactions have been reported in which SES moderates the impact of genetic differences on attained cognitive ability (Tucker-Drob & Bates, 2016), we next tested whether offspring genetic potential (EA PRS) interacted with SES in predicting QCST scores. This was tested by adding an EA PRS × SES interaction term to the model. This term was not significant and was estimated at zero, $\beta = -0.001$ (95% CI [-0.004, 0.002]), $t(594) = -0.754$, $p = .451$, suggesting that in this sample there was no evidence for SES moderating the impact of EA2 PRS on EA (see Figure 4).

4. Are parents' EA PRS alleles associated with higher offspring EA, even when not transmitted to their children? Finally, we moved to our core hypothesis in the present paper: testing for evidence that EA2 PRS has its effects on EA, at least in part, via a parenting phenotype. Evidence for such a trait would be found if alleles contributing to the EA PRS and present in the parents, but not transmitted to their offspring are associated with higher offspring EA, despite not being transmitted across the generations. This would constitute support for a parental-competence phenotype impacting offspring EA, un-confounded by parent–offspring genetic relatedness.

This hypothesis was tested using a multilevel model with QCST as the dependent variable, with sex, age-at-test, children's own EA PRS and, finally, parental scores for their non-transmitted alleles (PC_EA PRS) as predictors. As shown in Figure 2, panel B, non-transmitted parental EA-alleles had a significant effect, $F(1, 606) = 4.245$, $p = .0397$, $\beta = 0.058$ (95% CI [0.003, 0.114]): approximately 38% of the effect size of children's own EA PRS. To test the idea that effects of parental competence are expressed via SES formation, we next tested the hypothesis that the PC_EA PRS effect would be removed by controlling for parental SES. This hypothesis was confirmed: the effect size of non-transmitted alleles dropped to zero, $\beta = 0.01$ (95% CI [-0.04, 0.07]), and was no longer significant, $F(1, 594) = 0.011$, $p = .917$.

Flynn (2016) recently formulated a G × SES model in which home environments superior to a child's own genetic tendency tend to raise attainment, while for children with higher genetic potential, home environments are typically lower than their potential, and in most cases therefore retard their development relative to their potential. We examined evidence for this effect in the present sample by testing the significance of an interaction between EA PRS and PC_EA PRS as an index of disparity between parent and child potential. No support was found, however, for this predicted interaction, $F(1, 605) = 0.339$, $p = .561$, which was estimated at near zero, $\beta = 0.003$ (95% CI [-0.052, 0.057]).
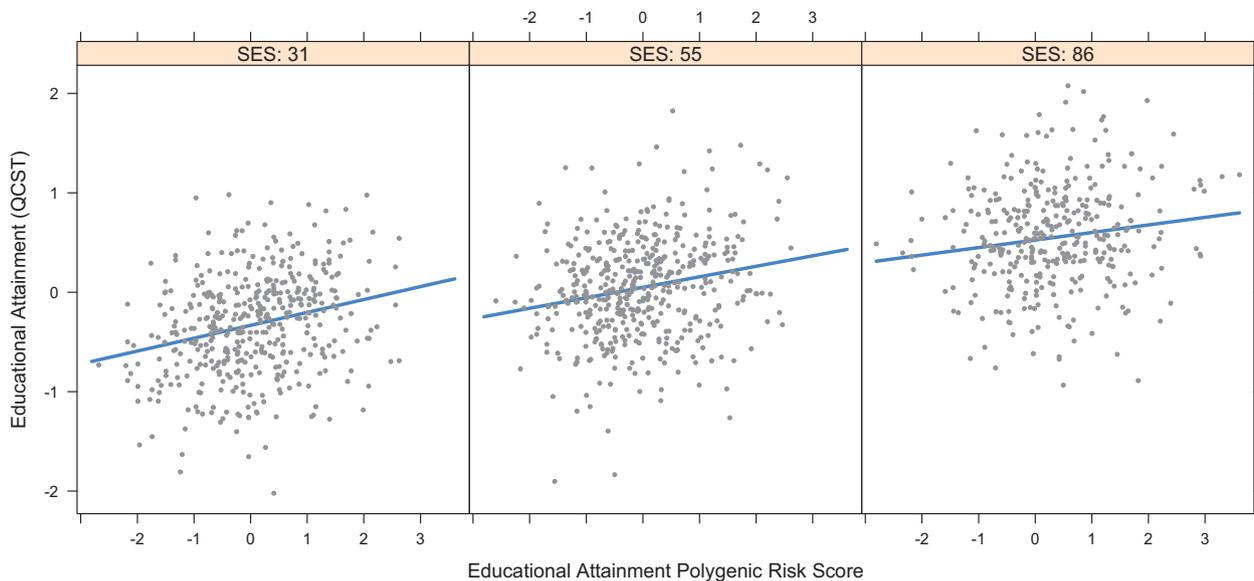
FIGURE 4

(Colour online) Plot of PRS × SES interaction for predicted educational attainment for low (left), medium (middle), and high (right) SES.

Structural Equation Model-Based Testing

The results of the multilevel models suggested that while no predicted G × E effects emerged, there was support for effects of non-transmitted parental polygenic scores on offspring EA. These models controlled for nesting in family and zygosity in the data and, empirically, the non-transmitted and transmitted scores were shown to be independent (see Figure 1). These models do not, however, capture the full, complex familial structure of genetic relatedness for the transmitted and non-transmitted scores among the different classes of twins and their sibs. To do this, we next reproduced the final results in a structural model capturing these sources of covariance including the expectation of a 0.5 correlation of parents' non-transmitted PRS for a child with the transmitted PRS for their sibs. We also included the known complete sharing of transmitted and non-transmitted scores for MZ twin members. Finally, the residual QCST covariance was included in the model (see Figure 5). This model yielded a similar estimate of the transmitted EA PRS effect (0.167) and a similarly ratio of non-transmitted effects (0.062, or approximately 1/3 the size). The latter effect was, however, non-significant in this model, with its many more degrees of freedom expended in modeling the covariance of family structure.

This work was begun in August of 2016 (with ethics permission) completed in June of 2017. Given the value of increased power, we decided to await the imminent release of EA3 and publish the results of the identical analyses with an enhanced genetic predictor. However, delays in the publication and release of EA3 and the publication of a test of the same idea by independent researchers (Kong et al., 2018)

leads us to publish this version now, with analyses to be re-run when EA3 results are available.

Discussion

By using a method capable of separating genetic transmission from purely environmental influence, we untangled otherwise confounded G × E covariation. This revealed a parenting genotype causing higher EA in offspring independent of offspring genetic propensity. The results show that, in part, what molecular genetic research into EA has captured is behavior present in parents that forms an environmental inheritance for children, operating in addition to their genetic inheritance and impacting their attainment. The results thus demonstrate that, in part, offspring cultural and cognitive development results from genes operating in parents.

Our four hypotheses and the results were as follows. We hypothesized that parent SES would associate with offspring attainment and that, in this family-based sample, children's EA PRS would remain associated with their measured EA. These initial hypotheses were confirmed. We then hypothesized that parental EA PRSs would be significantly positively associated with the SES parents' attainment and which therefore is available to their children. This was supported also (see Figure 3). Finally, we hypothesized that children's EA would reflect in part the effects of a heritable phenotype for parenting, reflected in significant effects of PC_EA PRS. This effect of non-transmitted alleles on offspring EA was also supported. Finally, supporting the idea that this association is realized
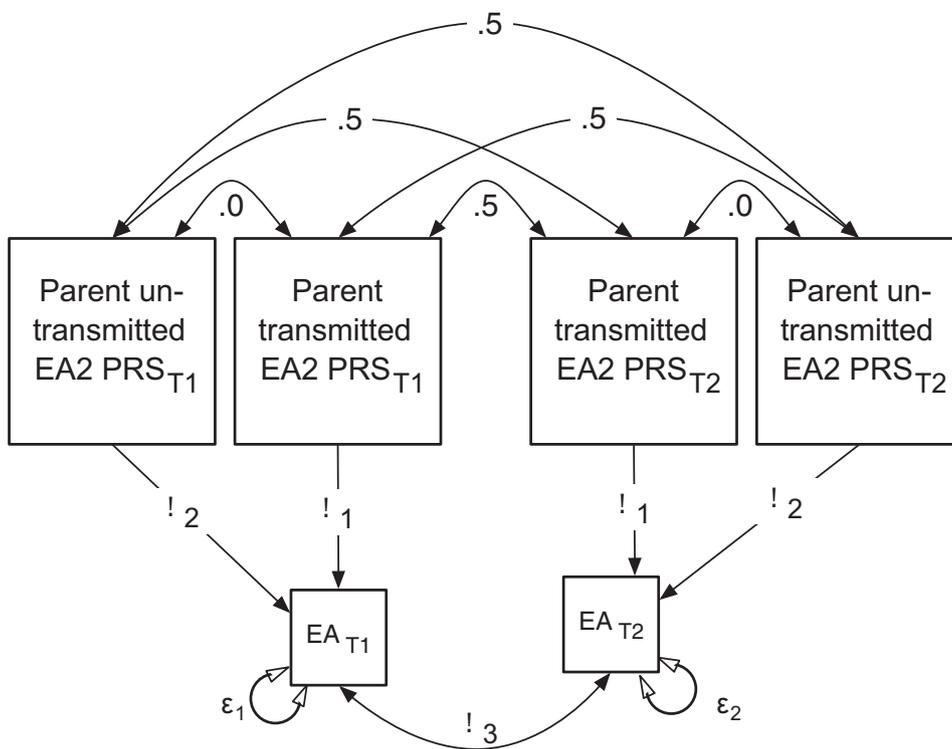
FIGURE 5

Structural model testing the relationship of non-transmitted and transmitted EA2 PRS to attained EA. Shown for DZ twins. β_2 is the key parameter tested: the effect of EA2 SNPs not transmitted to offspring. Full model includes additional sibs and MZ twin-family group.

in parental SES, this effect was abolished when controlling for SES.

The confirmation that the association of EA PRS with EA is replicable in a family design further bolsters the validity of this indicator. The finding that parents' EA PRS was linked to their own attained SES (see Figure 3) extends the finding that offspring EA2 scores are associated with parental SES (Selzam et al., 2017) to show that parental genotypes themselves are associated with SES. It thus significantly broadens the phenotype associated with EA PRS. As education raises SES, one plausible route for this association would be via EA PRS impacting their own EAs (unmeasured here). Alternatively, the EA PRS may be an SES niche-creation phenotype distinct from competencies required for education per se.

The finding that parent SES was significantly associated with offspring EA confirms that even in a cultural context where G × SES interactions for IQ are non-significant (Bates et al., 2016), SES is still linked to important social outcomes such as children's EA. The finding that the association of the EA PRS with measured EA was not moderated by family SES (see Figure 4) confirms, via molecular instruments, this finding of non-association from the twin literature in Australia (see also Grasby et al., 2017).

We found an effect of parents' non-transmitted alleles approximately 38% of the effect size of children's own EA PRS. This supports the idea of a parental competence phenotype supporting child cognitive and social development. The finding that the effect of non-transmitted alleles could be reduced to zero by controlling for parental SES suggests that the proximal outcome is the SES-niche created by parents. We did not find support for the idea (Flynn, 2016) that differences between parent and child genetic potentials impact children's attainment, that is, evidence was not found for high-ability parents raising low-ability children's outcomes or vice versa. This lack of interaction is in keeping with a lack of G × SES effects in general in this population (Bates et al., 2016) and in Australia more broadly (see also Grasby et al., 2017). Home environments appeared to exert additive effects, independent of offspring potential.

Questions Raised by the Results

The finding from this 'virtual parent' design indicating that non-transmitted alleles in the EA PRS nevertheless cause higher EA in offspring prompts several additional questions. One concerns how close to optimal our PRS is for detecting the parenting phenotype. The rationale prompting this paper was based on reverse engineering the social

hypothesis of parenting: If children did not differ in their genetic potential for education, then any signal in studies such as those included in generating the EA PRS (Okbay et al., 2016) would be carried on the back of social stratification. In the worst case, we risk detecting genes merely confounded with cultures in which parents are socialized to deliver high-competence parenting, an effect known as genetic stratification (Pritchard & Rosenberg, 1999). This worst case can be adequately ruled out, not only because methods were in place to control such artifacts in these samples (Price et al., 2006; Price et al., 2010), but because family designs are robust against stratification by virtue of providing a unit of analysis with common ancestry. That we detected associations for children's EA PRS and their own attainment in this family design argues against artifact as a cause of association.

Swagerman et al. (2017) examined reading fluency in an extended twin design, finding a non-significant and near-zero ($\beta = 0.006$) cultural transmission effect. In the United States, Wadsworth et al. (2002) reported a null correlation between adoptees' reading skill and that of their adoptive parents. Both these papers suggest zero or near-zero cultural transmission for reading. These results are compatible with our zero estimate of cultural transmission for EA in this Australian sample.

Replicating Selzam et al. (2017), we found a higher associated of EA PRS with the SES of children's parents than with children's attainment. This may reflect factors such as the increasing heritability with age associated with attainment outcomes or may indicate that the EA PRS targets latent SES more closely than it does the EA target it was originally designed to assess.

The mechanism of this effect of non-transmitted alleles on offspring attainment requires clarification. We observed that effect was consistent with complete mediation via parental SES. The precise components of social competence accounting for this effect; however, remain unclear. Behavioral differences between low- and high-SES parents may form one mechanism — for instance, increased responsiveness toward offspring needs (Bradley & Corwyn, 2002). Alternatively, social competence is manifested in physical factors such as high-quality shelter and reliable health and food provision, which have been argued to account for the positive outcomes associated with higher SES (Duncan et al., 2011). These extended aspects of the phenotype (Dawkins, 1983) may form the mechanism. It is possible that the selection pressure for the cognitive and personality traits underpinning SES attainment may in part or whole have arisen due to selection for these benefits, including child rearing of the created niche. Additional evidence demonstrating that SES behavior was selected for its effects on parent or infant survival (rather than arising incidentally) is required.

It is important to note that while much of EA reflects cognitive ability (Roth et al., 2015), these parental differences need not involve IQ. The question of the genetic correlation between genes for intelligence and genes for parenting that encourages intelligence in one's children is unknown but is unlikely to be unity. Parental attentiveness, warmth, and so on may form important targets for genetic research, contributing to raising smart kids in a manner not strongly related to how smart the parents are. This also raises the question of whether a much better PRS could be developed, based on specific measures of parenting. Most of the EA PRS must consist of alleles operating in the offspring to support cognitive and other educationally relevant skills and abilities. If the parenting-competence genome is more aligned with factors such as investment or attentiveness, a large-scale GWAS for parenting behaviors could cast valuable light on the parenting phenotype for nurture in general.

Summary

Our results provide support for a heritable parenting-phenotype that causes higher offspring EA, independent of offspring genetic propensity. They call for further investigation of this parenting phenotype, including the extent to which it may be environmentally malleable. They open doors to research into nurturing and caregiving behavior beyond the parental bond: for instance, differences in effective teaching (Taylor et al., 2010). It is possible our technique may be extended also to parental influences on additional traits, such as psychopathology, stimulating research into the role of genetically influenced parental behavior on non-intellectual traits in offspring.

Acknowledgments

References

Bates, T. C., Hansell, N. K., Martin, N. G., & Wright, M. J. (2016). When does socioeconomic status (SES) moderate the heritability of IQ? No evidence for g × SES interaction for IQ in a representative sample of 1176 Australian adolescent twin pairs. *Intelligence, 56*, 10–15.

Belsky, D. W., Moffitt, T. E., Corcoran, D. L., Domingue, B., Harrington, H., Hogan, S., … Caspi, A. (2016). The genetics of success: How single-nucleotide polymorphisms

associated with educational attainment relate to life-course development. *Psychological Science, 27*, 957–972.

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53*, 371–399.

Bronfenbrenner, U., & Ceci, S. J. (1994). Nature-nurture reconceptualized in developmental perspective: A bioecological model. *Psychological Review, 101*, 568–586.

Cesarini, D., & Visscher, P. M. (2017). Genetics and educational attainment. *npj Science of Learning, 2*, Article 4.

Cordell, H. J., Barratt, B. J., & Clayton, D. G. (2004). Case/pseudocontrol analysis in genetic association studies: A unified framework for detection of genotype and haplotype associations, gene-gene and gene-environment interactions, and parent-of-origin effects. *Genetic Epidemiology, 26*, 167–185.

Dawkins, R. (1983). *The extended phenotype: The long reach of the gene*. Oxford, UK: Oxford University Press.

Duncan, G. J., Morris, P. A., & Rodrigues, C. (2011). Does money really matter? Estimating impacts of family income on young children's achievement with data from random-assignment experiments. *Developmental Psychology, 47*, 1263–1279.

Eaves, L. J., Pourcain, B. S., Smith, G. D., York, T. P., & Evans, D. M. (2014). Resolving the effects of maternal and offspring genotype on dyadic outcomes in genome wide complex trait analysis ('M-GCTA'). *Behavior Genetics, 44*, 445–455.

Flynn, J. R. (2016). *Does your family make you smarter? Nature, nurture, and human autonomy*. Cambridge, UK: Cambridge University Press.

Franke, B., Stein, J. L., Ripke, S., Anttila, V., Hibar, D. P., van Hulzen, K. J. E., … Sullivan, P. F. (2016). Genetic influences on schizophrenia and subcortical brain volumes: Large-scale proof of concept. *Nature Neuroscience, 19*, 420–431.

Galobardes, B., Shaw, M., Lawlor, D. A., Lynch, J. W., & Davey Smith, G. (2006). Indicators of socioeconomic position (Part 1). *Journal of Epidemiology and Community Health, 60*, 7–12.

Grasby, K. L., Coventry, W. L., Byrne, B., & Olson, R. K. (2017). Little evidence that socioeconomic status modifies heritability of literacy and numeracy in Australia. *Child Development*. Advance online publication.

Howey, R. (2014). PseudoCons [Computer software]. Retrieved from http://www.staff.ncl.ac.uk/richard.howey/pseudocons/.

International Schizophrenia Consortium, Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O'Donovan, M. C., … Sklar, P. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature, 460*, 748–752.

Kendler, K. S. (2016). The schizophrenia polygenic risk score: To what does it predispose in adolescence? *JAMA Psychiatry, 73*, 193–194.

Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsson, B. J., Young, A. I., Thorgeirsson, T. E., … Stefansson, K. (2018).

The nature of nurture: Effects of parental genotypes. *Science, 359*, 424–428.

Krapohl, E., Rimfeld, K., Shakeshaft, N. G., Trzaskowski, M., McMillan, A., Pingault, J. B., … Plomin, R. (2014). The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proceedings of the National Academy of Sciences of the United States of America, 111*, 15273–15278.

Loehlin, J. (2000). Group differences in intelligence. In R. J. Sternberg (Ed.), *The handbook of intelligence* (pp. 176–195). Cambridge, UK: Cambridge University Press.

Martin, N. G., & Martin, P. G. (1975). The inheritance of scholastic abilities in a sample of twins. I. Ascertainments of the sample and diagnosis of zygosity. *Annals of Human Genetics, 39*, 213–218.

Martin, N. W., Hansell, N. K., Wainwright, M. A., Shekar, S. N., Medland, S. E., Bates, T. C., … Wright, M. J. (2009). Genetic covariation between the Author Recognition Test and reading and verbal abilities: What can we learn from the analysis of high performance? *Behavior Genetics, 39*, 417–426.

McMillan, J., Beavis, A., & Jones, F. L. (2009). A new socioeconomic index for Australia. *Journal of Sociology, 45*, 123–149.

Medland, S. E., Nyholt, D. R., Painter, J. N., McEvoy, B. P., McRae, A. F., Zhu, G., … Martin, N. G. (2009). Common variants in the trichohyalin gene are associated with straight hair in Europeans. *American Journal of Human Genetics, 85*, 750–755.

Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families. Vol. 67*. Dordrecht: Kluwer Academic Publishers.

Okbay, A., Beauchamp, J. P., Fontana, M. A., Lee, J. J., Pers, T. H., Rietveld, C. A., … Benjamin, D. J. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. *Nature, 533*, 539–542.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2017). Nlme: Linear and nonlinear mixed effects models. R package version 3.1-131. [Computer software] Retrieved from https://CRAN.R-project.org/package=nlme.

Plomin, R. (1994). *Genetics and experience: The interplay between nature and nurture*. Thousand Oaks, CA: Sage.

Plomin, R., & Loehlin, J. C. (1989). Direct and indirect IQ heritability estimates: A puzzle. *Behavior Genetics, 19*, 331–342.

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics, 38*, 904–909.

Price, A. L., Zaitlen, N. A., Reich, D., & Patterson, N. (2010). New approaches to population stratification in genome-wide association studies. *Nature Reviews Genetics, 11*, 459–463.

Pritchard, J. K., & Rosenberg, N. A. (1999). Use of unlinked genetic markers to detect population stratification in association studies. *American Journal of Human Genetics, 65*, 220–228.

Puglisi, M. L., Hulme, C., Hamilton, L. G., & Snowling, M. J. (2017). The home literacy environment is a correlate, but

perhaps not a cause, of variations in children's language and literacy development. *Scientific Studies of Reading, 21*, 498–514.

Queensland Studies Authority. (2003). *The 2003 Queensland core skills test retrospective*. Brisbane, Australia: Author.

Rietveld, C. A., Medland, S. E., Derringer, J., Yang, J., Esko, T., Martin, N. W., … Koellinger, P. D. (2013). GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science, 340*, 1467–1471.

Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence, 53*, 118–137.

Rowe, D. C. (2001). Do people make environments or do environments make people? *Annals of the New York Academy of Sciences, 935*, 62–74.

Scarr, S., & McCartney, K. (1983). How people make their own environments - A theory of genotype-environment effects. *Child Development, 54*, 424–435.

Selzam, S., Krapohl, E., von Stumm, S., O'Reilly, P. F., Rimfeld, K., Kovas, Y., … Plomin, R. (2017). Predicting educational achievement from DNA. *Molecular Psychiatry, 22*, 267–272.

Sniekers, S., Stringer, S., Watanabe, K., Jansen, P. R., Coleman, J. R. I., Krapohl, E., … Posthuma, D. (2017). Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nature Genetics, 49*, 1107–1112.

Swagerman, S. C., van Bergen, E., Dolan, C., de Geus, E. J. C., Koenis, M. M. G., Hulshoff Pol, H. E., & Boomsma, D. I. (2017). Genetic transmission of reading ability. *Brain and Language, 172*, 3–8.

Taylor, J., Roehrig, A. D., Soden Hensler, B., Connor, C. M., & Schatschneider, C. (2010). Teacher quality moderates the genetic effects on early reading. *Science, 328*, 512–514.

Tucker-Drob, E. M., & Bates, T. C. (2016). Large cross-national differences in gene x socioeconomic status interaction on intelligence. *Psychological Science, 27*, 138–149.

Vilhjalmsson, B. J., Yang, J., Finucane, H. K., Gusev, A., Lindstrom, S., Ripke, S., … Price, A. L. (2015). Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *American Journal of Human Genetics, 97*, 576–592.

Wadsworth, S. J., Corley, R. P., Hewitt, J. K., Plomin, R., & DeFries, J. C. (2002). Parent-offspring resemblance for reading performance at 7, 12 and 16 years of age in the Colorado adoption project. *Journal of Child Psychology and Psychiatry, 43*, 769–774.

Wainwright, M. A., Wright, M. J., Geffen, G. M., Luciano, M., & Martin, N. G. (2005). The genetic basis of academic achievement on the Queensland Core Skills Test and its shared genetic variance with IQ. *Behavior Genetics, 35*, 133–145.

Wright, M. J., & Martin, N. G. (2004). The Brisbane adolescent twin study: Outline of study methods and research projects. *Australian Journal of Psychology, 56*, 65–78.

Zhang, Y., Qi, G., Park, J.-H., & Chatterjee, N. (2017). Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits and implications for the future. *bioRxiv*, 175406. doi:https://doi.org/10.1101/175406.