OPEN ACCESS Freely available online

PLoS one

A Model for Transgenerational Imprinting Variation in Complex Traits

Chenguang Wang[1], Zhong Wang[2], Jiangtao Luo[2], Qin Li[1], Yao Li[3], Kwangmi Ahn[2], Daniel R. Prows[4], Rongling Wu[2]*

1 Department of Statistics, University of Florida, Gainesville, Florida, United States of America, **2** Center for Statistical Genetics, Pennsylvania State University, Hershey, Pennsylvania, United States of America, **3** Department of Statistics, West Virginia University, Morgantown, West Virginia, United States of America, **4** Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, Ohio, United States of America

Abstract

Despite the fact that genetic imprinting, i.e., differential expression of the same allele due to its different parental origins, plays a pivotal role in controlling complex traits or diseases, the origin, action and transmission mode of imprinted genes have still remained largely unexplored. We present a new strategy for studying these properties of genetic imprinting with a two-stage reciprocal F_2 mating design, initiated with two contrasting inbred lines. This strategy maps quantitative trait loci that are imprinted (i.e., iQTLs) based on their segregation and transmission across different generations. By incorporating the allelic configuration of an iQTL genotype into a mixture model framework, this strategy provides a path to trace the parental origin of alleles from previous generations. The imprinting effects of iQTLs and their interactions with other traditionally defined genetic effects, expressed in different generations, are estimated and tested by implementing the EM algorithm. The strategy was used to map iQTLs responsible for survival time with four reciprocal F_2 populations and test whether and how the detected iQTLs inherit their imprinting effects into the next generation. The new strategy will provide a tool for quantifying the role of imprinting effects in the creation and maintenance of phenotypic diversity and elucidating a comprehensive picture of the genetic architecture of complex traits and diseases.

Citation: Wang C, Wang Z, Luo J, Li Q, Li Y, et al. (2010) A Model for Transgenerational Imprinting Variation in Complex Traits. PLoS ONE 5(7): e11396. doi:10.1371/journal.pone.0011396

Editor: Amanda Ewart Toland, Ohio State University Medical Center, United States of America

Received February 17, 2010; **Accepted** May 6, 2010; **Published** July 14, 2010

Funding: National Science Foundation (NSF) grant DMS/NIGMS-0540745 and Changjiang Scholarship Award. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: rwu@hes.hmc.psu.edu

Introduction

Many traits important to agriculture, biology, and human health are complex in terms of the genetic machineries that determine trait formation and development. Broadly speaking, these machineries are equipped with a web of actions and interactions of numerous DNA sequence polymorphisms, modified or altered by environmental factors. To elucidate a detailed picture of the genetic architecture of complex traits, various molecular, statistical, and computational tools have been developed and used in the mapping and identification of specific genes underlying the traits [1–8]. The biological basis for developing these tools is that variation in phenotypic traits is due to the changes of DNA sequences in particular regions of the genome and, thus, by analyzing the linkage or association between the genotype and phenotype, significant genes can be detected. More recently, a growing body of new evidence has indicated that chromatin variation, such as differential DNA methylation, independent of DNA sequence changes, may play an important role in regulating the phenotypic formation and progression of complex traits [9–12]. Examples of these findings include a spontaneous epigenetic change in the SBP-box promoter leading to the inhibition of fruit ripening in tomatoes [13], the imprinted expression of the axin-fused ($Axin^{Fu}$) allele resulting in kinked tails in mice [14], and a global loss of cytosine methylation during aging in mice, rats, and humans [15].

To describe variation among individuals in the number or distribution of methylated nucleotides at specific gene sequences, a new term, called epialleles, has been coined [16]. Because epiallele phenotypes can have identical underlying DNA sequences, the genetic control mechanisms of these phenotypes are likely to differ from those estimated from traditional models of quantitative genetics. Thus, it is crucial to screen for epiallelic variants within a population and disentangle epigenetic from more standard genetic sources of phenotypic variance, such as additive genetic variance, dominance variance, epistasis and maternal genetic effects [17]. More recently, Johannes et al. [12] developed a panel of epigenetic Recombinant Inbred Lines (epiRILs) in the reference plant *Arabidopsis thaliana* to identify the genetic variation due to epiallelic variants in flowering time and plant height. Epiallelic variation can also be studied by tracing parent-dependent differences of the same allele. If the same allele functions differently, depending on which parent the allele is derived from, a phenomenon known as genetic imprinting or parent-of-origin effect, this allele may be epigenetic. Previous studies have suggested that genetic imprinting results from an epigenetic mark of differential methylation set during gametogenesis [18–20], forming part of the genetic architecture involved in the formation, development, function, and evolution of complex traits and diseases [21–25].

The past several years have witnessed an intense interest in mapping and identifying the regions of the genome that contain

imprinted sequence variants with genome-wide linkage and association studies. Cheverud et al. [26] and Wolf et al. [27] used a three-generation F_2 design to map genome-wide imprinted quantitative trait loci (iQTLs) that affect body weight and growth in mice, and they found that these traits may be controlled by QTLs with more complex and diverse effect patterns than previously assumed. Li et al. [28] proposed a reciprocal backcross design to estimate the distribution of iQTLs and quantify their effects on physiological traits related to endosperm development in maize. By modeling alleles identical-by-descent in a multi-generational pedigree of canines, Liu et al. [29] derived a linkage-based random effect to genome-wide scan for the existence of iQTLs that affect canine hip dysplasia. However, there is limited knowledge about whether imprinted effects are inherited over generations and, if yes, how imprinting inheritance takes place [19,30–37]. An understanding of these question will help to characterize the impacts of imprinting loci on the genetic diversity of a biological trait or process [38–40].

In this article, we develop a novel strategy for identifying imprinted genes and understanding the transgenerational changes of their effects with a three-generation pedigree. This pedigree is initiated by reciprocally crossing two contrasting inbred lines, leading to two different F_1 families. The F_1 males and females from the same and different families are further crossed to generate four F_2 families. Thus, the inheritance of alleles at a gene from a male or female parent can be traced by observing the segregation of the gene in different families. A joint likelihood model is constructed to formulate the effect of imprinted genes on a complex trait. Traditional quantitative genetic theory is integrated to define the effects of imprinting genes (due to the parent-dependent expression of an allele), their interactions with other genetic effect sources (such as additive, dominant, and epistatic), and their generation-dependent actions. We implement the EM algorithm to estimate different genetic effects of imprinted genes and their changes across generations. A testing procedure is proposed to study the pattern of transgenerational imprinting inheritance. The statistical behavior of the model is examined through simulation studies and its usefulness validated from a real data analysis in a three-generation pedigree of mice.

Methods

Mating Design

Suppose there are two inbred lines that are sharply contrasting in a complex trait. Each line can serve as a maternal and paternal parent, thus allowing a reciprocal cross. An F_1 family is produced by mating a dam from one parental line with a sire from the other line, while a reciprocal F_1 family produced by using the dam and sire from the opposing lines. According to traditional Mendel's first law, these two F_1 families should be genetically identical. However, if there is an imprinting effect, the two families will be different. Here, we assume that these two F_1 families are epigenetically different. The females and males from the same F_1 families are crossed to produce two epigenetically "inbred" F_2 families, whereas those from the opposing F_1 families are crossed to produce two epigenetically "outbred" F_2 families. Using a quantitative trait locus (QTL) with two alleles A and a, the mating design involving the original parents, reciprocal F_1 families, and reciprocal F_2 families is illustrated in Figure S1.

Assume that each F_2 family is typed for the same panel of molecular markers and phenotyped for the same trait of interest. Linkage analysis with these markers allows the construction of an integrative linkage map that covers the genome by combining the four F_2 families. The map is then used to identify imprinted quantitative trait loci (iQTLs) that control the trait. The model presented in this article enables geneticists to map iQTLs by combining the segregation pattern of an iQTL in the four different F_2 populations.

Quantitative Genetic Model

Using the iQTL demonstrated in Figure S1, we formulate quantitative genetic models of an iQTL that affects a complex trait. Two inbred lines are reciprocally crossed to generate two F_1 configurations, Aa and aA, with the same allele inherited from different parents. These two F_1 configurations will perform differently if this iQTL shows a significant imprinted effect in the F_1 generation. Reciprocal crosses with these F_1 configurations lead to four F_2 combinations, $Aa \times Aa$, $Aa \times aA$, $aA \times Aa$, and $aA \times aA$, each of which will have the same group of segregating QTL genotypes/configurations, AA, Aa, aA, and aa. The imprinted effect of the iQTL is inherited into the next generation if two F_2 configurations, Aa and aA, are still different. To test whether this imprinted effect is inheritable and how much it is inherited, we will need to quantify the difference of the imprinted effect of the iQTL expressed in the F_1 and F_2 generations. To do that, we attributed the differences among the F_2 genotypes to two different sources:

(1) The same QTL genotype is different from different mating types due to the genetic imprinting of the F_1 generation. For example, F_2 genotype AA from $Aa \times Aa$ is different than F_2 genotype AA from $Aa \times aA$ because of the imprinting effect of the F_1 male parent formed in the cross of original inbred lines;

(2) F_2 configurations Aa and aA from the same mating type are different because of genetic imprinting formed in the cross of F_1 individuals (F_2).

Thus, a final genotypic value of an F_2 genotype is determined by the imprinting effects of the iQTL in the F_1 and F_2 generations, additive and dominance effects, and their interactions. Genotypic values of four F_2 configurations at the iQTL from different mating types are decomposed into different components expressed in Table 1. The component parameters are sorted into seven different groups:

(1) μ is the overall mean of all the F_2 populations,

(2) i_m and i_p are the imprinting effects of iQTL expressed by the F_1 maternal and paternal parents, respectively,

(3) e_{mp} is the interaction between i_m and i_p,

(4) a_o, d_o, and i_o are the additive, dominant, and imprinting effects of the iQTL formed in the F_2,

(5) e_{ma}, e_{md}, and e_{mi} are the interaction effects between the imprinting effects of the F_1 maternal parent and additive, dominant, and imprinting effects expressed in the F_2, respectively,

(6) e_{pa}, e_{pd}, and e_{pi} are the interaction effects between the imprinting effects of the F_1 paternal parent and additive, dominant, and imprinting effects expressed in the F_2, respectively,

(7) e_{mpa}, e_{mpd}, and e_{mpi} are the interactions between e_{mp} and a_o, d_o, and i_o, respectively.

Mixture Likelihood

The four epigenetically different F_2 families (Table 1) are observed for a complex trait with respective sample sizes $n_1, \ldots, n_4$. Let $\mathbf{y}_1, \ldots, \mathbf{y}_4$ denote the phenotypic values of the trait for different families. An iQTL for the trait that is segregating in four F_2

Table 1. Genetic components of 16 F_2 configurations derived from two successive reciprocal crosses.

No.	Mating Type	F_2 Generation	
		Configuration	Genotypic Value
1	$Aa \times Aa$	AA	$\mu_{11}=\mu+i_m+i_p+e_{mp}+a_o+e_{ma}+e_{pa}+e_{mpa}$
		Aa	$\mu_{12}=\mu+i_m+i_p+e_{mp}+d_o+i_o+e_{md}+e_{mi}+e_{pd}+e_{pi}+e_{pmd}+e_{pmi}$
		aA	$\mu_{13}=\mu+i_m+i_p+e_{mp}+d_o-i_o+e_{md}-e_{mi}+e_{pd}-e_{pi}+e_{pmd}-e_{pmi}$
		aa	$\mu_{14}=\mu+i_m+i_p+e_{mp}-a_o-e_{ma}-e_{pa}-e_{mpa}$
2	$Aa \times aA$	Aa	$\mu_{21}=\mu+i_m-i_p-e_{mp}+d_o+i_o+e_{md}+e_{mi}-e_{pd}-e_{pi}-e_{mpd}-e_{mpi}$
		AA	$\mu_{22}=\mu+i_m-i_p-e_{mp}+a_o+e_{ma}-e_{pa}+e_{mpa}$
		aa	$\mu_{23}=\mu+i_m-i_p-e_{mp}-a_o-e_{ma}+e_{pa}+e_{mpa}$
		aA	$\mu_{24}=\mu+i_m-i_p-e_{mp}+d_o-i_o+e_{md}+e_{mi}-e_{pd}+e_{pi}-e_{mpd}+e_{mpi}$
3	$aA \times Aa$	aA	$\mu_{31}=\mu-i_m+i_p-e_{mp}+d_o-i_o-e_{md}+e_{mi}+e_{pd}-e_{pi}-e_{mpd}+e_{mpi}$
		aa	$\mu_{32}=\mu-i_m+i_p-e_{mp}-a_o+e_{ma}-e_{pa}+e_{mpa}$
		AA	$\mu_{33}=\mu-i_m+i_p-e_{mp}+a_o-e_{ma}+e_{pa}-e_{mpa}$
		Aa	$\mu_{34}=\mu-i_m+i_p-e_{mp}+d_o+i_o+e_{md}-e_{mi}-e_{pd}-e_{pi}-e_{mpd}-e_{mpi}$
4	$aA \times aA$	aa	$\mu_{41}=\mu-i_m-i_p+e_{mp}-a_o+e_{ma}+e_{pa}-e_{mpa}$
		aA	$\mu_{42}=\mu-i_m-i_p+e_{mp}+d_o-i_o-e_{md}-e_{mi}-e_{pd}+e_{pi}+e_{mpd}-e_{mpi}$
		Aa	$\mu_{43}=\mu-i_m-i_p+e_{mp}+d_o+i_o-e_{md}+e_{mi}-e_{pd}-e_{pi}+e_{mpd}+e_{mpi}$
		AA	$\mu_{44}=\mu-i_m-i_p+e_{mp}+a_o-e_{ma}-e_{pa}+e_{mpa}$

doi:10.1371/journal.pone.0011396.t001

populations can be mapped with interval mapping. Consider a pair of markers between which the iQTL for the trait is hypothesized to be located. The configurations of the iQTL are unobserved, but can be inferred from the genotypes of the markers that bracket the QTL. This inference needs the construction of a likelihood based on a mixture model. Such a likelihood combines the information from four F_2 families, expressed as

$$\log L(\theta;\mathbf{y})=\sum_{i=1}^{n_1}\log\left\{\sum_{j=1}^{4}\omega_{j|1i}f_{j1}(y_{1i})\right\}+\sum_{i=1}^{n_2}\log\left\{\sum_{j=1}^{4}\omega_{j|2i}f_{j2}(y_{2i})\right\}+\sum_{i=1}^{n_3}\log\left\{\sum_{j=1}^{4}\omega_{j|3i}f_{j3}(y_{3i})\right\}+\sum_{i=1}^{n_4}\log\left\{\sum_{j=1}^{4}\omega_{j|4i}f_{j4}(y_{4i})\right\} \quad (1)$$

where $\omega_{j|ik}$ is the conditional probability of an iQTL configuration j ($j=1$ for AA, 2 for Aa, 3 for aA, and 4 for aa) given the marker genotype of individual i from F_2 family k ($k=1,2,3,4$), and $f_{kj}(y_{ik})$ is the normal distribution function of the trait with iQTL configuration-specific mean (μ_{kj}) and variance (σ_k^2). In Wu et al. [41], the procedure for deriving these conditional probabilities are given in terms of the recombination fractions between the left marker and QTL, QTL between the right marker, and the two markers. The EM algorithm was implemented to estimate the genotypic means and variance from the mixture model (1) (see Methods S1).

Hypothesis Tests

To determine whether there is an iQTL for the complex trait can be tested with log-likelihood ratio approaches. We first tested whether a significant QTL exists in the four F_2 populations using the following null hypothesis,

$$H_0: i_m=i_p=e_{mp}=a_o=d_o=i_o=e_{ma}=e_{md}=e_{mi}=e_{pa}=e_{pd}=e_{pi}=e_{mpa}=e_{mpd}=e_{mpi}=0. \quad (2)$$

The log-likelihood ratio calculated under the null and alternative hypotheses is compared with the critical threshold determined from permutation tests [42].

After a significant QTL is determined, then the imprinting effect of the QTL can be tested using the following null hypothesis,

$$H_0: i_m=i_p=e_{mp}=i_o=e_{ma}=e_{md}=e_{mi}=e_{pa}=e_{pd}=e_{pi}=e_{mpa}=e_{mpd}=e_{mpi}=0. \quad (3)$$

The rejection of null hypothesis (3) implies that the QTL has an accumulative imprinting effect expressed in different generations, which includes main and interaction effects related with genetic imprinting. The imprinting effects expressed in the F_1 and F_2 are tested by the null hypotheses, respectively,

$$H_0: i_m=i_p=e_{mp}=0, \quad (4)$$

$$H_0: i_o=e_{ma}=e_{md}=e_{mi}=e_{pa}=e_{pd}=e_{pi}=e_{mpa}=e_{mpd}=e_{mpi}]=0. \quad (5)$$

The interactions between the imprinting effect expressed in the F_1 maternal or paternal parents and the additive, dominant, and imprinting genetic effects in the F_2 can also be tested, respectively, by

$$H_0: e_{ma}=e_{md}=e_{mi}=0, \quad (6)$$

$$H_0: e_{pa}=e_{pd}=e_{pi}=0. \quad (7)$$

The higher-order interactions among the maternally- and paternally-expressed genetic imprinting in the F_1 and the additive, dominant, and imprinting genetic effects in the F_2 are tested by the null hypothesis,

$$H_0 : e_{mpa} = e_{mpd} = e_{mpi} = 0. \quad (8)$$

All the genetic effects in equations (3)–(8) can be tested individually. The log-likelihood ratios for hypothesis tests related with genetic imprinting can be thought of being asymptotically χ^2-distributed.

Results

Worked Example

The newly developed model was used to analyze a data set from a large-scale QTL analysis project in which mice serve as a model system to study survival time to hyperoxic acute lung injury (HALI) [43]. In a screen of 18 inbred mouse strains, C57BL/6J (B) mice were selected as sensitive and 129X1/SvJ (S) mice resistant, based on total survival time in >95% oxygen (hyperoxia). Reciprocal F_1 ($B \times S$ and $S \times B$) mice demonstrated a significant difference in acute lung injury survival time, suggesting possible occurrence of parent-of-origin effects. To further identify specific loci displaying a imprinting effect, both pairs of reciprocal F_1 crosses were bred to generate 840 F_2 mice, including 213 for $(B \times S) \times (B \times S)$, 221 for $(B \times S) \times (S \times B)$, 197 for $(S \times B) \times (B \times S)$, and 209 for $(S \times B) \times (S \times B)$. A genome-wide linkage map was constructed by typing 93 microsatellite markers located on the 19 autosomes and X-chromosome for four F_2 populations of mice derived from sensitive B and resistant S strains.

Phenotype differences between the F_2 crosses further support possible existence of imprinted genes that affect HALI. By scanning over the linkage map with the log-likelihood ratio test statistics calculated from hypothesis (2), the number and distribution of QTLs for HALI are detected (Figure S2), which is consistent with the discoveries by traditional interval mapping [43]. Five significant QTLs were located between Mit236 and Mit478 on chromosome 1, Mit196 and Mit17 on chromosome 4, Mit116 and Mit145 on chromosome 4, Mit289 and Mit355 on chromosome 9, and Mit175 and Mit5 on chromosome 15. Given their long genetic distance, two significant peaks on chromosome 4 were thought to carry different QTLs. At each of the detected QTLs, the 15 genetic effect parameters including the imprinting, additive, and dominant effects and their interactions across generations, as defined in Table 1, were estimated (Table 2). All these estimated parameters were tested for imprinting effects at different levels. The first test was made for the overall imprinting effects and their interactions expressed in both generations F_1 and F_2, including i_m, i_p, e_{mp}, i_o, e_{ma}, e_{md}, e_{mi}, e_{pa}, e_{pd}, e_{pi}, e_{mpa}, e_{mpd}, and e_{mpi}. It is found that all the detected QTLs are highly significant for the overall imprinting effects, with the p-values ranging from 4.86×10^{-10} to 2.22×10^{-16} (Table 3). Therefore, these QTLs are regarded as iQTLs.

The second test concerns the imprinting effects expressed in the F_1 generation by testing whether the paternally- (i_m) and maternally-imprinted effects (i_p) and their interaction (e_{mp}) during the cross of the original inbred lines are equal to zero (Table 3). Except for the QTL on chromosome 1 and one QTL on chromosome 4, which are significant at $p = 0.0036 - -0.0006$, all others display highly significant imprinting effects in the F_1 generation ($p = 1.60 \times 10^{-5} - -8.90 \times 10^{-8}$). The third test was conducted to see whether there is an imprinting effect in the F_2 generation by jointly testing the significance of i_o, e_{ma}, e_{md}, e_{mi}, e_{pa}, e_{pd}, e_{pi}, e_{mpa}, e_{mpd}, and e_{mpi}. It appears that all the QTLs are highly significant, except for one on chromosome 9 displaying a

Table 2. Maximum likelihood estimates of genetic effect parameters for each iQTL detected on different chromosomes.

Para-meters	Chromosome				
	1	4	4	9	15
	(Mit236-Mit478)	(Mit196-Mit17)	(Mit116-Mit145)	(Mit289-Mit355)	(Mit175-Mit5)
Genetic imprinting expressed in the F_1					
i_m	−6.5207	−9.0352	−7.5241	−8.6968	−14.6362
i_p	1.6017	0.9479	1.6623	2.7283	11.4645
e_{mp}	0.6448	0.8244	−0.9077	−1.9431	−7.8337
Genetic effects expressed in the F_2					
a_o	−1.1171	−4.0853	−0.1759	2.1381	11.4756
d_o	5.5179	−2.7714	1.7809	6.8690	−9.0239
i_o	−1.2043	−4.1082	−4.7386	−4.3694	6.0973
Two-way interactions between genetic effects expressed over generations					
e_{ma}	−3.9038	8.7865	3.8376	2.3743	−6.3636
e_{md}	−2.2013	0.5825	−2.2923	0.0604	7.5975
e_{mi}	−4.4437	2.4542	2.3924	3.1781	1.3964
e_{pa}	−3.9038	2.6049	8.7361	6.2536	16.7322
e_{pd}	3.5226	4.7608	4.2266	1.1876	−11.7537
e_{pi}	10.6457	−5.4277	−7.1118	−4.2282	−5.0311
Three-way interactions between genetic effects expressed over generations					
e_{mpa}	−1.1171	2.0963	−5.0744	−1.7413	−11.6203
e_{mpd}	−2.6972	−4.4730	−1.1203	1.4102	8.4446
e_{mpi}	−4.9976	1.1347	0.0192	−0.8235	4.6168

doi:10.1371/journal.pone.0011396.t002

Table 3. *P*-values for testing the imprinting effects of iQTLs expressed at different levels.

QTL		Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
Chrom.	**Marker Interval**						
1	Mit236-Mit478	2.22×10^{-16}	0.0036	3.64×10^{-8}	0.2240	0.1406	0.7263
4	Mit196-Mit17	7.93×10^{-14}	2.26×10^{-5}	2.63×10^{5}	0.6955	0.2073	0.3244
4	Mit116-Mit145	3.30×10^{-14}	0.0006	4.62×10^{-6}	0.4300	0.2143	0.9806
9	Mit289-Mit355	4.86×10^{-10}	1.60×10^{-5}	0.0163	0.8872	0.8447	0.9396
15	Mit175-Mit5	1.00×10^{-13}	8.90×10^{-8}	2.21×10^{-5}	0.1072	0.0213	0.0016

Note: The null hypotheses used are
H_0: $i_m=i_p=e_{mp}=i_o=e_{ma}=e_{md}=e_{mi}=e_{pa}=e_{pd}=e_{pi}=e_{mpa}=e_{mpd}=e_{mpi}=0$ for Test 1.
H_0: $i_m=i_p=e_{mp}=0$ for Test 2.
H_0: $i_o=e_{ma}=e_{md}=e_{mi}=e_{pa}=e_{pd}=e_{pi}=e_{mpa}=e_{mpd}=e_{mpi}=0$ for Test 3.
H_0: $e_{ma}=e_{md}=e_{mi}=0$ for Test 4.
H_0: $e_{pa}=e_{pd}=e_{pi}=0$ for Test 5.
H_0: $e_{mpa}=e_{mpd}=e_{mpi}=0$ for Test 6.
doi:10.1371/journal.pone.0011396.t003

marginally significant effect. The last three tests focus on the interactions of the imprinting effect in the F_1 with the additive, dominant and imprinting effects in the F_2. We did not detect many significant interactions between the imprinted effect in the F_1 and the overall genetic effects in the F_2, but with two exceptions (Table 3). One is the interaction between the paternally-imprinted effect in the F_1 and the overall genetic effects in the F_2 for the QTL on chromosome 15 ($p=0.0213$), and the other is the three-way interaction among the maternally- and paternally-imprinted effects in the F_1 and the overall genetic effects in the F_2 for the same QTL ($p=0.0016$).

In sum, all the detected iQTLs show a similar pattern of genetic effect on HALI in the F_1 generation, with the maternally-imprinted effect (negative) larger than with the paternally-imprinted effect (positive) (Table 2). Pronounced diversity was observed in the additive and dominant effects among the QTLs when they inherit into the F_2 generation. Main imprinting effects in the F_2 generation were largely reduced, but there is some evidence that imprinted effects are preserved into the F_2 through their interactions with other genetic effects such as additive and dominant.

Computer Simulation

To examine the statistical behavior of the new model, we performed Monte Carlo simulation studies by mimicking the example of the F_2 mice. The simulation includes two different parts. In part 1, we simulated 10 evenly-spaced markers in a linkage group of 200 cM. An iQTL is located 35 cM from the first marker at the left. The markers and iQTL are segregating in four reciprocal F_2 families (Figure S1), initiated with two contrasting inbred lines. The 15 parameters of genetic effects were given and the genotypic values of 16 F_2 configurations were then calculated. The phenotypic values were then simulated by summing the genotypic values and residual errors assumed to follow a normal distribution with mean zero and variance scaled for different heritabilities 0.10 and 0.40. Two different sample sizes were assumed, i.e., 300 and 500 progeny, for each F_2 family. All the parameters can be reasonably well estimated with the new model (Table 4). At the modest heritability (0.10), the main imprinting effects and their interactions in the F_1 and the main additive, dominant, and imprinting effects in the F_2 can reasonably well be estimated, even with a smaller sample size (Table 4). To better estimate interactions between imprinting effects of the F_1 generation and genetic effects of the F_2, a larger sample size is needed. All parameters can be more precisely estimated when the heritability increases from 0.1 to 0.4. The precise estimation of three-way interactions of imprinting effects between different generations requires a large sample size (2000 in total) and large heritability (0.4).

In part 2, the simulation was used to test the power of the new model and its false positive rates. The conditions used for power calculation were the same as described above. Table 5 tabulates the results from three different simulation scenarios. There is full power for the detection of overall genetic imprinting effects even when the heritability and sample size are modest (Test 1, Scenario I). Also, great power (>0.86) was detected for the overall genetic imprinting effects expressed in the F_1 generations (Test 2, Scenarios I and II). Yet, to detect the genetic imprinting expressed in the F_2, a larger sample size (2000 in total) is needed to achieve a power of 0.99 (Test 3, Scenario II). Much larger heritabilities and/or sample sizes are needed for detecting the interactions between the imprinting effects in the F_1 and genetic effects in the F_2, especially when the values of these interactions are small (Tests 4–6, Scenario I). The false positive rates of the estimation for genetic effects by the new model were calculated by simulating the data assuming the absence of those effects (see Scenarios II and III). In general, false positive rates are low for overall genetic imprinting effects (<0.08) (Test 1, Scenario III), regardless of different heritabilities and sample sizes. Also, false positive rates for overall genetic imprinting effects expressed in the F_1 are reasonably low (Test 2, Scenario III). Genetic imprinting effects expressed in the F_2 generation, as well as interactions between the imprinting effects of the F_1 and genetic effects of the F_2, all have very low false positive rates.

Discussion

According to traditional Mendelian genetic theory, the maternally and paternally derived alleles of a gene should have a similar amount of expression because they carry the same DNA sequence. However, a growing number of studies suggest that alleles may be expressed from only one of the two parental chromosomes [18,44] due to the difference of DNA methylation. Such genetic imprinting or parent-of-origin effects provide a possible source of phenotypic variation for complex traits in the absence of DNA sequence variants [21–25]. Thus, to better elucidate the genetic architecture of complex traits and diseases for

Table 4. Maximum likelihood estimates (and their standard errors) of genetic effect parameters from simulated data under different sample sizes (300 and 500) and heritabilities (0.1 and 0.4).

Parameters	True Value	300		500	
		$H^2=0.1$	$H^2=0.4$	$H^2=0.1$	$H^2=0.4$
Genetic imprinting expressed in the F_1					
i_m	0.15	0.152±0.0664	0.1519±0.0278	0.1466±0.0472	0.1501±0.0233
i_p	0.15	0.154±0.0678	0.1534±0.0273	0.1450±0.0478	0.1482±0.0203
e_{mp}	0.1	0.090±0.0694	0.0990±0.0291	0.0954±0.0505	0.1005±0.0234
Genetic effects expressed in the F_2					
a_o	0.3	0.334±0.1173	0.2947±0.0479	0.3199±0.0960	0.2943±0.0370
d_o	0.6	0.612±0.0934	0.5982±0.0394	0.5940±0.0804	0.5992±0.0303
i_o	0.2	0.244±0.1106	0.19660±0.0450	0.2300±0.0986	0.19820±0.0357
Two-way interactions between genetic effects expressed over generations					
e_{ma}	0.04	0.041±0.1106	0.04081±0.0402	0.0425±0.0947	0.04201±0.0349
e_{md}	0.04	0.038±0.1041	0.03758±0.0430	0.0441±0.0775	0.03828±0.0351
e_{mi}	0.04	0.022±0.1086	0.04262±0.0397	0.0153±0.0871	0.04022±0.0324
e_{pa}	0.04	0.048±0.1026	0.03688±0.0429	0.0415±0.0884	0.04118±0.0349
e_{pd}	0.04	0.034±0.0969	0.03574±0.0409	0.0463±0.0741	0.04274±0.0290
e_{pi}	0.04	0.020±0.1153	0.04290±0.0421	0.0193±0.0832	0.04130±0.0347
Three-way interactions between genetic effects expressed over generations					
e_{mpa}	0.04	0.005±0.1130	0.0461±0.0451	0.0295±0.0995	0.0467±0.0373
e_{mpd}	0.04	0.059±0.0952	0.0406±0.0416	0.0470±0.0760	0.0385±0.0335
e_{mpi}	0.04	0.092±0.1102	0.0353±0.0464	0.0753±0.0923	0.0342±0.0378

doi:10.1371/journal.pone.0011396.t004

Table 5. Power and Type I error rates of the model for detecting genetic imprinting effects at different levels.

Scenario	Sample Size	H^2	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
I	300	0.1	100	86	20	3	2	2
		0.4	100	98	33	6	3	4
	500	0.1	100	100	99	22	20	18
		0.4	100	100	100	40	37	32
II	300	0.1	99	98	5	3	2	2
		0.4	100	100	3	4	1	1
	500	0.1	100	100	4	4	1	5
		0.4	100	100	4	2	2	3
III	300	0.1	3	5	2	2	3	1
		0.4	6	6	3	1	3	4
	500	0.1	8	12	7	4	3	4
		0.4	4	7	1	2	2	1

The null hypotheses used are
H_0:$i_m=i_p=e_{mp}=i_o=e_{ma}=e_{md}=e_{mi}=e_{pa}=e_{pd}=e_{pi}=e_{mpa}=e_{mpd}=e_{mpi}=0$ for Test 1.
H_0: $i_m=i_p=e_{mp}=0$ for Test 2.
H_0: $i_o=e_{ma}=e_{md}=e_{mi}=e_{pa}=e_{pd}=e_{pi}=e_{mpa}=e_{mpd}=e_{mpi}=0$ for Test 3.
H_0: $e_{ma}=e_{md}=e_{mi}=0$ for Test 4.
H_0: $e_{pa}=e_{pd}=e_{pi}=0$ for Test 5.
H_0: $e_{mpa}=e_{mpd}=e_{mpi}=0$ for Test 6.
Three scenarios used are
I. $i_m=i_p=0.15, e_{mp}=0.1, a_o=0.3, d_o=0.6, i_o=0.2, e_{ma}=e_{md}=e_{mi}=e_{pa}=e_{pd}=e_{pi}=e_{mpa}=e_{mpd}=e_{mpi}=0.04$,
II. $i_m=i_p=0.15, e_{mp}=0.1, a_o=0.3, d_o=0.6, i_o=0, e_{ma}=e_{md}=e_{mi}=e_{pa}=e_{pd}=e_{pi}=e_{mpa}=e_{mpd}=e_{mpi}=0$,
III. $i_m=i_p=e_{mp}=0, a_o=0.3, d_o=0.6, i_o=e_{ma}=e_{md}=e_{mi}=e_{pa}=e_{pd}=e_{pi}=e_{mpa}=e_{mpd}=e_{mpi}=0$.
doi:10.1371/journal.pone.0011396.t005

various organisms including humans, the magnitude and pattern of imprinting effects should be estimated and their impact on quantitative variation quantified.

The attempts to characterize imprinting effects are affected by our incapacity to discern the effect of DNA methylation variants from that of DNA sequence variants using a mapping study. This issue was, however, resolved by comparing two reciprocal crosses in which the maternally- or paternally-derived version of the same allele at a gene can be identified [28,45]. Liu et al. [29] incorporated identical-by-descent (IBD) sharing into a random-effect mapping model, allowing the characterization of the discrepancy of allelic transmission through different parents. Linkage mapping using controlled crosses or pedigrees with known parents has led to the genome-wide identification of imprinted quantitative trait loci (iQTLs) that affect body weight and growth in mice [26,27], physiological traits related to endosperm development in maize [28], and hip dysplasia in canines [29].

However, to study the precise genetic mechanisms through which chromatin dynamics alter quantitative variation, a simple test of imprinting effects of iQTLs is not adequate. Rather, a detailed understanding of whether and how imprinting effects are transmitted across generations is crucial for determining the contribution of epigenetic modification to heritable phenotypic variation for a complex trait. In this article, we present a new strategy for estimating and testing imprinting effects of iQTLs and their transgenerational transmission through two-generation reciprocal crosses leading to four epigenetically different F_2 families (Figure S1). The new strategy displays two advantages compared with previous models. First, it provides a comprehensive elucidation of the genetic control mechanisms for a complex trait or disease in terms of traditionally defined additive and dominant effects, newly defined imprinted effects, and their interactions. Second, the strategy has power to detect the changes of imprinting effects from generation to generation, thus facilitating the modeling of transgenerational epigenetic variation and inheritance.

We formulated a mixture model-based likelihood for the imprinting effects of iQTLs flanked by markers in four epigenetically different F_2 families. A closed form of the EM algorithm was derived to estimate a high-dimensional set of genetic parameters that define the maternally- and paternally-imprinted genetic effects and their interactions in the F_1, the additive, dominant, and imprinting effects in the F_2, and the interactions of different orders between these effects expressed in different generations. The algorithm was tested through simulation studies from which the minimum heritability and sample size for reasonable estimates of each parameter are determined. Additional simulation studies were performed to test the power for the detection of imprinting effects at different levels. In general, the model shows reasonably low false positive rates for the data in which no imprinting effects exist. In an application of the new model for genetic mapping of iQTL in mice, we identified five significant QTLs on chromosomes 1, 4, 9, and 15 for the overall survival time to hyperoxic acute lung injury (HALI). Each of these QTLs displays remarked imprinting effects on HALI. The model was further used to test when and how these imprinting effects are activated to affect the expression of HALI. In general, all the iQTLs trigger marked imprinting effects in the F_1 (see i_m and i_p estimates in Table 2). During transmission into the next generation, these imprinting effects were observed to be shrunk (see i_o estimates in Table 2). But highly significant imprinting effects in the F_2 generation can still be detected (Table 3; see also [46]) when the interactions between the imprinting effects of the F_1 and main effects of the F_2 are jointly tested. This result suggests that imprinting effects detected from pure F_2 generations, as conducted in [46], may have confounded their interactions with other effects formed during transmission. The results from reanalyzing the mouse data with the new model shed light on the new inheritance and aetiology of HALI.

The model developed in this article will provide a useful tool for studying transgenerational imprinting inheritance and its impact on the variation in complex traits and diseases. As a first attempt of its kind, the model will need to be modified so as to broaden the scope of its application. Given its ubiquitousness in trait control, epistasis between different genes should be incorporated into the current model, helping to draw a comprehensive atlas of the genetic architecture for complex traits. Also, the expression of any genetic effects cannot be isolated from the environment in which organisms are reared [47,48]. The interactions between different genetic effects and environmental factors should be modeled when a powerful imprinting model is developed. Genetic imprinting may be expressed at the DNA sequence level [49–51]. Thus, the integration of haplotype diversity into the model will gain new insights into the genetic control mechanisms of complex traits. All these extensions, although straightforward in theory, will face with an increasing number of parameters being estimated. Statistical explorations for enhancing the efficiency of parameter estimation will be largely demanded. In sum, the development of the new strategy will facilitate our efforts to address many biological questions of fundamental importance in elucidating the genetic architecture of complex traits.

Supporting Information

Figure S1 A mating design generating four reciprocal F_2 families, initiated with two inbred lines *AA* and *aa*. The two inbred lines that serve as female (red) and male parents (blue) are crossed reciprocally to generate two F_1 families. From each of these two families, two progeny, one being a female (red) and the other being a male (blue), are selected to make all possible crosses, leading to four different F_2 families (with four genotype configurations *AA*, *Aa*, *aA*, and *aa* listed in the box).
Found at: doi:10.1371/journal.pone.0011396.s001 (0.04 MB EPS)

Figure S2 The plot of log-likelihood ratio across the mouse genome composed of 19 autosomes and one sex chromosome. Ticks on the x-axis are molecular markers. The peaks of the profile, at which significant QTLs on chromosomes 1, 4, 9, and 15 are detected by the new model, are indicated by arrowed vertical lines. The critical threshold for claiming the existence of significant QTLs is indicated by a horizontal line.
Found at: doi:10.1371/journal.pone.0011396.s002 (0.03 MB EPS)

Methods S1 Supporting Methods.
Found at: doi:10.1371/journal.pone.0011396.s003 (0.04 MB PDF)

Author Contributions

Conceived and designed the experiments: DP RW. Performed the experiments: DP. Analyzed the data: CW ZW JL QL YL KA. Contributed reagents/materials/analysis tools: CW ZW JL QL YL KA. Wrote the paper: RW.

References

1. Lander E, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121: 185–199.

2. Zeng Z (1994) Precision mapping of quantitative trait loci. Genetics 136: 1457–1468.

3. Ma C, Casella G, Wu R (2002) Functional mapping of quantitative trait loci underlying the character process: a theoretical framework. Genetics 161: 1751–1762.
4. Xu S (2003) Estimating polygenic effects using markers of the entire genome. Genetics 163: 789–801.
5. Li C, Zhou A, Sang T (2006) Rice domestication by reducing shattering. Science 311: 1936–1939.
6. Anholt R, Mackay T (2004) Genetic analysis of complex behaviors in Drosophila. Nat Rev Genet 5: 838–849.
7. Paterson A (2006) Leafing through the genomes of our major crop plants: strategies for capturing unique information. Nature Reviews Genetics 7: 174–184.
8. Veyrieras J, Kudaravalli S, Kim S, Dermitzakis E, Gilad Y, et al. (2008) High-resolution mapping of expression-QTLs yields insight into human gene regulation. PLoS genetics 4: e1000214. doi:10.1371/journal.pgen.1000214.
9. Martienssen R, Colot V (2001) DNA methylation and epigenetic inheritance in plants and filamentous fungi. Science 293: 1070–1074.
10. Richards E (2006) Inherited epigenetic variation—revisiting soft inheritance. Nature Reviews Genetics 7: 395–401.
11. Richards E (2008) Population epigenetics. Current Opinion in Genetics & Development 18: 221–226.
12. Johannes F, Porcher E, Teixeira F, Saliba-Colombani V, Simon M, et al. (2009) Assessing the impact of transgenerational epigenetic variation on complex traits. PLoS Genetics 5: e1000530. doi:10.1371/journal.pgen.1000530.
13. Manning K, Tor M, Poole M, Hong Y, Thompson A, et al. (2006) A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. Nature genetics 38: 948–952.
14. Rakyan V, Chong S, Champ M, Cuthbert P, Morgan H, et al. (2003) Transgenerational inheritance of epigenetic states at the murine AxinFu allele occurs after maternal and paternal transmission. Proceedings of the National Academy of Sciences 100: 2538–2543.
15. Kahn A, Fraga M (2009) Epigenetics and aging: status, challenges, and needs for the future. The Journals of Gerontology: Series A 64: 195–198.
16. Kakutani T (2002) Epi-alleles in plants: inheritance of epigenetic information over generations. Plant and cell physiology 43: 1106–1111.
17. Kalisz S, Purugganan M (2004) Epialleles via DNA methylation: consequences for plant evolution. Trends in Ecology & Evolution 19: 309–314.
18. Reik W, Walter J (2001) Genomic imprinting: parental influence on the genome. Nature Reviews Genetics 2: 21–32.
19. Sasaki H, Matsui Y (2008) Epigenetic events in mammalian germ-cell development: reprogramming and beyond. Nature Reviews Genetics 9: 129–140.
20. Sha K (2008) A mechanistic view of genomic imprinting. Annu Res Genomics Hum Genet 9: 197–216.
21. Itier J, Tremp G, Léonard J, Multon M, Ret G, et al. (1998) Imprinted gene in postnatal growth role. Nature 393: 125–126.
22. Li L, Keverne E, Aparicio S, Ishino F, Barton S, et al. (1999) Regulation of maternal behavior and offspring growth by paternally expressed Peg3. Science 284: 330–333.
23. Isles A, Wilkinson L (2000) Imprinted genes, cognition and behaviour. Trends in Cognitive Sciences 4: 309–318.
24. Constância M, Kelsey G, Reik W (2004) Resourceful imprinting. Nature 432: 53–57.
25. Wilkinson L, Davies W, Isles A (2007) Genomic imprinting effects on brain development and function. Nature Reviews Neuroscience 8: 832–843.
26. Cheverud J, Hager R, Roseman C, Fawcett G, Wang B, et al. (2008) Genomic imprinting effects on adult body composition in mice. Proceedings of the National Academy of Sciences 105: 4253–4258.
27. Wolf J, Cheverud J, Roseman C, Hager R (2008) Genome-wide analysis reveals a complex pattern of genomic imprinting in mice. PLoS Genetics 4: e1000091.
28. Li Y, Coelho C, Liu T, Wu S, Wu J, et al. (2008) A statistical model for estimating maternal-zygotic interactions and parent-of-origin effects of QTLs for seed development. PLoS One 3: e3131.
29. Liu T, Todhunter R, Wu S, Hou W, Mateescu R, et al. (2007) A random model for mapping imprinted quantitative trait loci in a structured pedigree: An implication for mapping canine hip dysplasia. Genomics 90: 276–284.
30. Morgan H, Santos F, Green K, Dean W, Reik W (2005) Epigenetic reprogramming in mammals. Human molecular genetics 14: R47–R58.
31. Li E (2002) Chromatin modification and epigenetic reprogramming in mammalian development. Nature Reviews Genetics 3: 662–673.
32. Surani M, Barton S, Norris M (1984) Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. Nature 308: 548–550.
33. McGrath J, Solter D (1984) Inability of mouse blastomere nuclei transferred to enucleated zygotes to support development in vitro. Science 226: 1317–1319.
34. Morgan H, Sutherland H, Martin D, Whitelaw E (1999) Epigenetic inheritance at the agouti locus in the mouse. Nature Genetics 23: 314–318.
35. Cropley J, Suter C, Beckman K, Martin D (2006) Germ-line epigenetic modification of the murine Avy allele by nutritional supplementation. Proceedings of the National Academy of Sciences 103: 17308–17312.
36. Dolinoy D, Weidman J, Waterland R, Jirtle R (2006) Maternal genistein alters coat color and protects Avy mouse offspring from obesity by modifying the fetal epigenome. Environmental health perspectives 114: 567–572.
37. Skinner M (2008) What is an epigenetic transgenerational phenotype?:: F3 or F2. Reproductive toxicology 25: 2–6.
38. Youngson N, Whitelaw E (2008) Transgenerational epigenetic effects. Annu Rev Genomics Hum Genet 9: 233–257.
39. Whitelaw N, Whitelaw E (2008) Transgenerational epigenetic inheritance in health and disease. Current opinion in genetics & development 18: 273–279.
40. Pembrey M, Bygren L, Kaati G, Edvinsson S, Northstone K, et al. (2005) Sex-specific, male-line transgenerational responses in humans. European Journal of Human Genetics 14: 159–166.
41. Wu R, Ma C, Casella G (2007) Statistical genetics of quantitative traits: linkage, maps, and QTL Springer Verlag.
42. Churchill G, Doerge R (1994) Empirical threshold values for quantitative trait mapping. Genetics 138: 963–971.
43. Prows D, Hafertepen A, Winterberg A, Gibbons Jr W, Liu C, et al. (2007) Genetic analysis of hyperoxic acute lung injury survival in reciprocal intercross mice. Physiological genomics 30: 271–281.
44. Wilkins J, Haig D (2003) What good is genomic imprinting: the function of parent-specific gene expression. Nature Reviews Genetics 4: 359–368.
45. Cui Y (2007) A statistical framework for genome-wide scanning and testing of imprinted quantitative trait loci. Journal of theoretical biology 244: 115–126.
46. Yang R, Wang X, Wu Z, Prows D, Lin M (2010) Bayesian model selection for characterizing genomic imprinting effects and patterns. Bioinformatics 26: 235–241.
47. Jirtle R, Skinner M (2007) Environmental epigenomics and disease susceptibility. Nature Reviews Genetics 8: 253–262.
48. Kaati G, Bygren L, Pembrey M, Sjostrom M (2007) Transgenerational response to nutrition, early life circumstances and longevity. European Journal of Human Genetics 15: 784–790.
49. Cheng Y, Berg A, Wu S, Li Y, Wu R (2009) Computing genetic imprinting expressed by haplotypes. Methods in molecular biology (Clifton, NJ) 573: 189–212.
50. Wen S, Wang C, Berg A, Li Y, Chang M, et al. (2009) Modeling genetic imprinting effects of DNA sequences with multilocus polymorphism data. Algorithms for Molecular Biology 4: 11.
51. Schalkwyk L, Meaburn E, Smith R, Dempster E, Jeffries A, et al. (2010) Allelic Skewing of DNA Methylation Is Widespread across the Genome. The American Journal of Human Genetics 86: 196–212.