

# A Resource for the Allele-Specific Analysis of DNA Methylation at Multiple Genomically Imprinted Loci in Mice

Jadiel A. Wasson,\* Onur Birol,<sup>†</sup> and David J. Katz<sup>\*,1</sup>

\*Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138 and <sup>†</sup>Department of Cell Biology, Emory University School of Medicine, Atlanta, Georgia 30322

**ABSTRACT** Genomically imprinted loci are expressed mono-allelically, dependent upon the parent of origin. Their regulation not only illuminates how chromatin regulates gene expression but also how chromatin can be reprogrammed every generation. Because of their distinct parent-of-origin regulation, analysis of imprinted loci can be difficult. Single nucleotide polymorphisms (SNPs) are required to accurately assess these elements allele specifically. However, publicly available SNP databases lack robust verification, making analysis of imprinting difficult. In addition, the allele-specific imprinting assays that have been developed employ different mouse strains, making it difficult to systemically analyze these loci. Here, we have generated a resource that will allow the allele-specific analysis of many significant imprinted loci in a single hybrid strain of *Mus musculus*. This resource includes verification of SNPs present within 10 of the most widely used imprinting control regions and allele-specific DNA methylation assays for each gene in a C57BL/6J and CAST/EiJ hybrid strain background.

## KEYWORDS

bisulfite analysis  
DNA methylation  
imprinting  
mouse  
single nucleotide  
polymorphism

Genomically imprinted loci, which are expressed mono-allelically dependent upon their parent of origin, highlight how DNA methylation and chromatin structure can regulate gene expression (Bartolomei and Ferguson-Smith 2011). For example, many of the chromatin mechanisms that regulate imprinted loci are involved in other contexts, including cancer biology and stem cell reprogramming. In addition, alterations at multiple imprinted loci can be used as a readout of global epigenetic misregulation. As a result, there is an increasing need to assay multiple imprinted loci in different mouse models. In this resource article, we provide a streamlined resource for assaying the methylation status of a number of the most studied imprinted genes in a single hybrid strain background.

To date, ~150 imprinted genes have been identified in mice and ~100 in humans (Gregg *et al.* 2010; DeVeale *et al.* 2012; Kelsey and Bartolomei 2012). These genes tend to be organized on chromosomes in clusters (Wan and Bartolomei 2008; Bartolomei 2009). This clustering allows multiple imprinted loci to be regulated together, under the control of *cis*-regulatory domains termed imprinting control regions (ICRs) (Wan and Bartolomei 2008; Bartolomei 2009). ICRs are typically between 100 and 3700 bp long and are rich in CpG dinucleotides (Bartolomei and Tilghman 1997; Barlow 2011; Ferguson-Smith 2011). In mammals, DNA methylation occurs mainly in the context of CpG dinucleotides, and within ICRs these CpG dinucleotides are differentially methylated, dependent upon the parent of origin (Reik and Dean 2001; Reik and Walter 2001). This differential methylation determines the expression status of the multiple imprinted genes located within the imprinting cluster (Reik and Walter 2001). Therefore, to globally interrogate the epigenetic control of genomically imprinted loci in a particular mouse model, it is necessary to be able to assay the DNA methylation status of multiple ICRs allele specifically.

Assessing ICRs allele specifically requires taking advantage of single nucleotide polymorphisms (SNPs). C57BL/6J (hereafter referred to as B6) mice are the most commonly used strain of *Mus musculus domesticus* and were the first mouse strain to be fully sequenced (Beck *et al.* 2000). To generate hybrids with SNPs on each allele, B6 mice can be crossed to *M. musculus castaneus* (hereafter referred to as CAST) mice, which originate from a well-defined subgroup of wild mice (Beck *et al.*

Copyright © 2018 Wasson *et al.*

doi: <https://doi.org/10.1534/g3.117.300417>

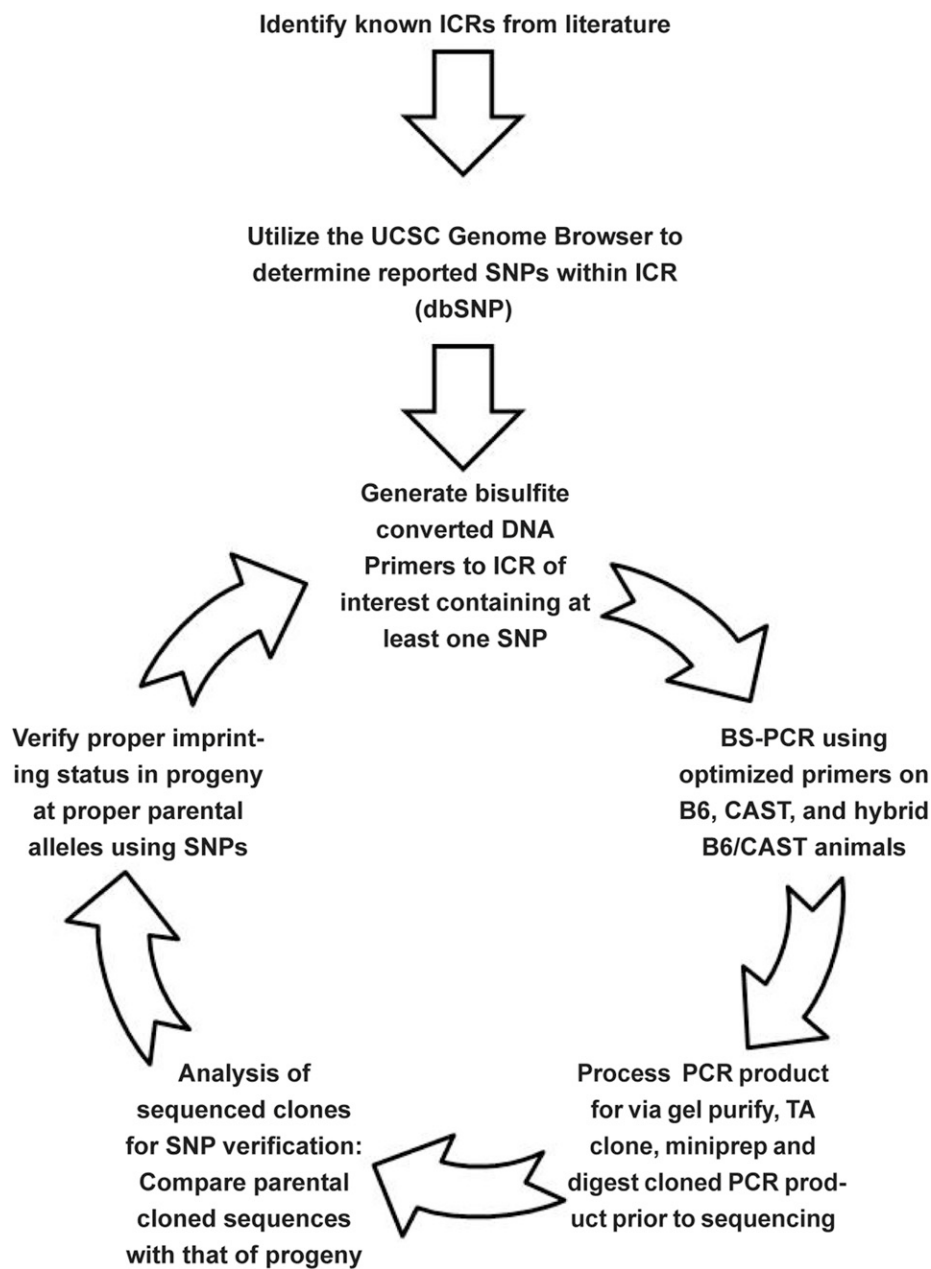
Manuscript received July 14, 2017; accepted for publication October 30, 2017; published Early Online November 14, 2018.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplemental material is available online at [www.g3journal.org/lookup/suppl/doi:10.1534/g3.117.300417/-/DC1](http://www.g3journal.org/lookup/suppl/doi:10.1534/g3.117.300417/-/DC1).

<sup>1</sup>Corresponding author: Department of Cell Biology, Emory University School of Medicine, 615 Michael St., Rm. 443, Whitehead Medical Research Bldg., Atlanta, GA 30322. E-mail: [djkatz@emory.edu](mailto:djkatz@emory.edu)

101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161



**Figure 1** Workflow for SNP verification within ICRs. Known ICRs were first pulled from literature followed by identification of putative SNPs present within each region. These SNPs then underwent a verification process through bisulfite analysis of both parental and hybrid progeny strains. SNPs that fail to verify were fed back to the verification process.

162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222

2000). Genome-wide DNA sequence analysis between different strains of *M. musculus* revealed a 50% allelic difference between B6 and CAST at potential SNPs (Frazer *et al.* 2007). This makes these hybrid progeny especially useful for analyzing imprinted loci.

SNPs between B6 and CAST are cataloged in the database of SNPs (dbSNP) (<https://www.ncbi.nlm.nih.gov/projects/SNP/>) (Smigielski *et al.* 2000; Sherry *et al.* 2001). This database reports SNPs that have been observed in various assays performed by individual researchers, consortiums, and genome sequencing centers, for the purpose of facilitating genome-wide association studies (Smigielski *et al.* 2000; Sherry *et al.* 2001). Unfortunately, this database is phasing out all nonhuman organism data by September 2017. However, very similar information will still be housed in the European variation archive (<http://www.ebi.ac.uk/eva/?Home>). This database overlaps with the dbSNP database and also the Sanger SNP viewer database ([https://www.sanger.ac.uk/sanger/Mouse\\_SnpViewer/rel-1505](https://www.sanger.ac.uk/sanger/Mouse_SnpViewer/rel-1505)) (Keane *et al.* 2011; Yalcin *et al.*

2011), which provides SNP information in multiple different strain backgrounds.

Using SNPs from all of these databases, we sought to develop allele-specific DNA methylation assays at multiple ICRs in a B6/CAST hybrid background. However, we encountered two significant hurdles. First, since the dbSNP database and the European variation archive are public repositories, many reported SNPs have not been additionally verified (Mitchell *et al.* 2004; Nekrutenko and Taylor 2012). Moreover, they currently have no minimum requirements for allelic frequencies (Mitchell *et al.* 2004; Nekrutenko and Taylor 2012). This further contributes to the lack of verification for many SNPs. As a result, false positives have been reported at a rate of between 15 and 17% (Mitchell *et al.* 2004; Nekrutenko and Taylor 2012). In addition, these two databases pool sequence differences from different strains into one combined output. Thus, we discovered that relying solely on the dbSNP database or European variation archive leads to an even higher rate of false positives within ICRs. These hurdles can partially be overcome by



223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283

■ Table 1 Primer sequences

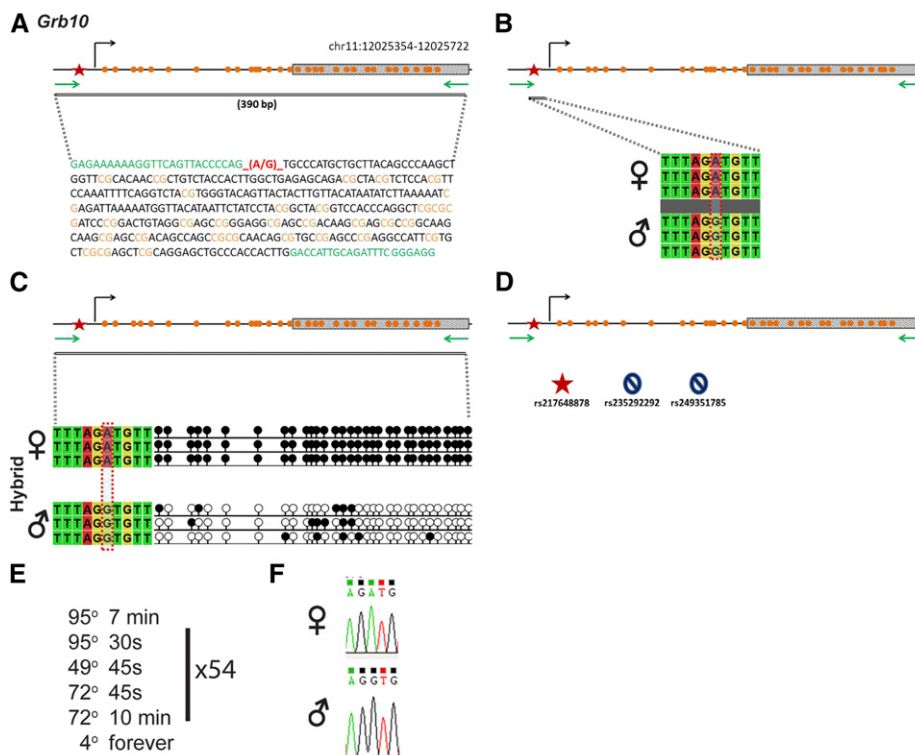
Gene	DNA Sequence 5' → 3'	Bisulfite Converted Sequence 5' → 3'
<i>Grb10</i>	F-GAGAAAAAGGTTTCAGTTACCCAG(A/G) R-CCTCCCAGAAATCTGCAATGGTC	F-GAGAAAAAGGTTTAGTTATTTTAG(A/G) R-CCTCCCAGAAATCTACAATAATC
<i>H19</i>	F-ATTCACAAATGGCAATGCTGTGG R-CCTCATGAAGCCCATGACTAT	F-ATTTATAAATGGTAATGTTGTGG R-CCTCATAAAACCCATAACTAT
<i>Igf2r</i>	F-CAGAGGATTTTAGCACAACCTCAA R-CACTTTTGAGCTTGCTCTCTTGC	F-TAGAGGATTTTAGTATAATTTTAA R-CACTTTTGAGCTTGCTCTCTTGC
<i>Impact</i>	F-CTGCATAGTTTTGCTCTCATAAGTG R-GGCCTGCTCATGTGACAATGCGGC	F-TTGTATAGTTTTGTTTTTATAAGTG R-AACCTACTCATATAACAATACAAC
<i>Lit1</i>	F-CAAGGTGAGTGGCCTAGGAC R-AATCCCCACACCTGAATTC	F-TAAGGTGAGTGGTTTAGGAT R-AATCCCCACACCTAAATTC
<i>Mest</i>	F-GGGTGTTTTATGTTTCCAGGG(T/G) R-CCCAGATTCTAGTGAAGAAAGCCTTCCCAT	F-GGGTGTTTTATGTTTTTAGGG(T/G) R-CCCAATTCTAATAAAAAAACCTTCCCAT
<i>Peg3</i>	F-GGTGCATCTTTACTGCCAACTAGCAAAG R-CAGGTTTGCTGCACAGGCTTATCC	F-GGTGTATTTTTATTGTTAATTAGTAAAG R-CAAATTTACTACACAACTTATCC
<i>Peg10</i>	F-GCAAAGTGAAGTGGCTCTGCACTCTTAAGTG R-TTGTTACTCTCTGCACTTCCAAATT	F-GTAAAGTGAAGTGGTTTGTATTTTAAAGTG R-TTAATTACTCTCTACAACTTCCAAATT
<i>Snrpn</i>	F-GCAATTATATCCATTATCCAGATTGACAGTGA(T/G) R-ATAGGATGCACCTTCACTACTAGAATCC	F-GTAATTATATTTATTTTATGATTGATAGTGA(T/G) R-ATAAAATACACTTCTACTACTAAAATCC
<i>Zac1</i>	F-GGGTAGGTAAGTAGTGACAA R-CCTAAACACCAAAGTAGCA	F-GGGTAGGTAAGTAGTGATAA R-CCTAAACACCAAATAACA

also incorporating the Sanger database, which contains information from individual strain backgrounds. However, a drawback of the Sanger database is that it contains much less information on intergenic regions, where many ICRs are found. For example, it contains no information on three of the ICRs that we sought to interrogate. In the end, we assessed 93 B6/CAST SNPs from the three databases at 10 of the most commonly studied mouse ICRs, and were able to validate only 18 of them (19%).

The second hurdle that we encountered is the generation of bisulfite PCR assays within ICRs. The gold standard in probing the DNA methylation status of any locus is bisulfite analysis (Hayatsu *et al.*

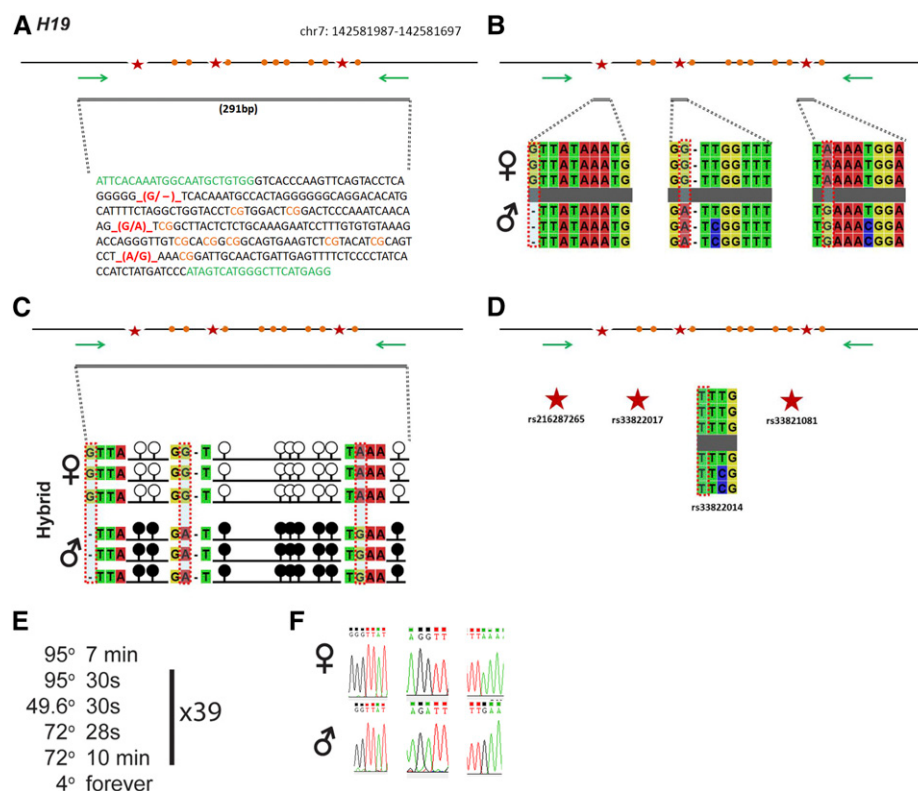
2008; Laird 2010). As bisulfite analysis relies on detecting base pair changes at CpG dinucleotides, primer sets used for bisulfite PCR cannot contain any CpG dinucleotides because of the uncertainty of whether a cytosine base in the primer-annealing sequence may be methylated. As a result, generating bisulfite-specific primer sets in these highly CpG-rich ICR regions can be difficult. In addition, because the CpG-rich ICRs tend to be repetitive, finding primer sets that amplify a unique product can also be challenging.

Based on the significant hurdles we encountered, we identified a need for optimized protocols for allele-specific DNA methylation analysis of ICRs in a B6/CAST hybrid mouse background. As a result, we developed



**Figure 2** SNP verification within *Grb10* ICR. (A) Schematic of *Grb10* ICR. Probed region is highlighted by double-dashed line with number of base pairs covered reported. CpG island indicated by dotted box. Green indicates primer sequences; orange indicates CpG dinucleotides; red star and bases indicate verified SNP. (B) Verified SNP presented as sequences from B6 female and CAST male. A-to-G SNP is highlighted by red dotted rectangle. (C) Verification of proper imprinted status in hybrid B6/CAST progeny. SNP highlighted by red dotted rectangle. DNA methylation presented as lollipop diagram; White circles indicate unmethylated cytosines; black circles indicate methylated cytosines. (D) Other SNPs reported in all three databases within the probed region with the SNP highlighted by red dotted rectangle. dbSNP identification number indicated under each SNP. Red star indicates validated SNP and blue closed circle indicates C-to-T polymorphism that cannot be assayed in bisulfite analysis. (E) Optimal PCR conditions for probed region with the given primers. (F) The electropherogram indicating A-to-G polymorphism for the SNP region. ♀, maternal; ♂, paternal.

284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344



**Figure 3** SNP verification within H19 ICR. (A) Schematic of H19 ICR. Probed region is highlighted by double-dashed line with number of base pairs covered reported. CpG island indicated by dotted box. Green indicates primer sequences; orange indicates CpG dinucleotides; red star and bases indicate verified SNPs. The chromosome location is from high to low, see Materials and Methods for more details. (B) Verified SNPs presented as sequences from B6 female and CAST male. G-to-del, G-to-A, and A-to-G SNPs are highlighted by red dotted rectangle. (C) Verification of proper imprinted status in hybrid B6/CAST progeny. SNPs highlighted by red dotted rectangle. DNA methylation presented as lollipop diagram; white circles indicate unmethylated cytosines; black circles indicate methylated cytosines. (D) Other SNPs reported in all three databases within the probed region with the SNP highlighted by red dotted rectangle. dbSNP identification number indicated under each SNP. Red star indicates validated SNP and blue closed circle indicates C-to-T polymorphism that cannot be assayed in bisulfite analysis. (E) Optimal PCR conditions for probed region with the given primers. (F) The electropherograms indicating G-to-del, G-to-A, and A-to-G polymorphisms for the SNPs. ♀, maternal; ♂, paternal.

a resource that includes verification of SNPs present in ICRs, primer information, and optimal PCR conditions. This resource will enable the systematic interrogation of many significant imprinted genes in different mouse models.

## MATERIALS AND METHODS

### Bisulfite analysis and bisulfite PCR optimization

Mouse tail DNA from single C57BL/6J and CAST/EiJ animals was used for the original identification of SNPs. Subsequently, DNA from sagittal sections of perinatal pups was used for allele-specific DNA methylation analysis. Bisulfite conversion was done according to the Zymo EZ DNA Methylation Kit (Zymo D5001) protocol from 400 ng of DNA. PCR products were amplified in a 15- $\mu$ l reaction and 3  $\mu$ l was saved for subsequent TA cloning using the standard TOPO TA cloning protocol (K4500J10; ThermoFisher). The remaining volume was run on a 1% agarose gel to confirm that there is a single PCR product. Bisulfite primers were optimized on bisulfite-converted DNA using 12 different conditions, including four different concentrations of  $MgCl_2$  (1.5, 2.5, 3.5, and 4.5 mM) paired with three different concentrations of DMSO (0, 1.5, and 5%). In addition, primers were optimized across a temperature gradient. Primer sets, polymorphisms, and optimal PCR conditions for each gene are listed in the individual figures. Of note, because of the difficulty in finding primer sequences in highly CpG-rich regions that do not contain a CpG dinucleotide, many of the primers contained suboptimal base composition and/or did not match the annealing temperature of the other primer used in the reaction. As a result, several of the optimized PCR protocols contain relatively large numbers of cycles to enable the amplification of a product. The BiQ Analyzer program was used for the analysis of bisulfite-converted sequences. During the

bisulfite analysis, depending on the choice of primers, two different DNA strands will lead to two different sequencing results. Some of the genes we report here were surveyed on the opposite strand of the gene assembly and therefore have a reversed order of their SNPs compared to the databases. These genes are shown with their chromosome location number in reverse order, from high to low, and this is noted in the corresponding figure legend.

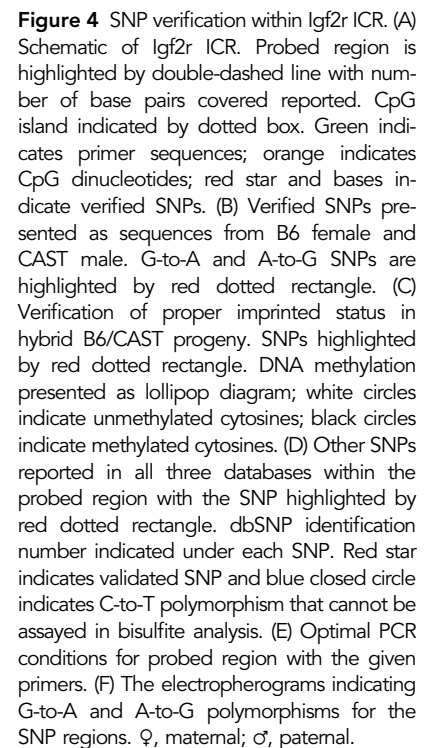
### Data availability

The authors affirm that all data necessary for confirming the results in the article are present in the article. Reagents are available upon request.

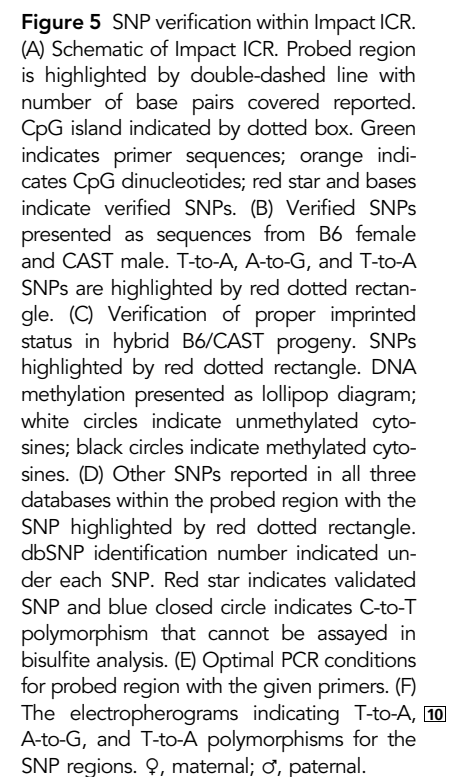
## RESULTS

To begin the process of interrogating specific imprinted loci, we generated a workflow to streamline the process (Figure 1). Our first criterion was to identify well-defined ICRs that have been extensively studied. We focused on the following ICRs due to their prevalence in the literature: *Grb10*, *H19*, *Igf2r*, *Impact*, *Lit1/Kcnq1ot1*, *Mest/Peg1*, *Peg3*, *Peg10*, *Snrpn*, and *Zac1/Plagl1*. These ICRs also had well-defined locations in the genome and are associated with differentially methylated regions that allowed us to probe their methylation status via bisulfite analysis.

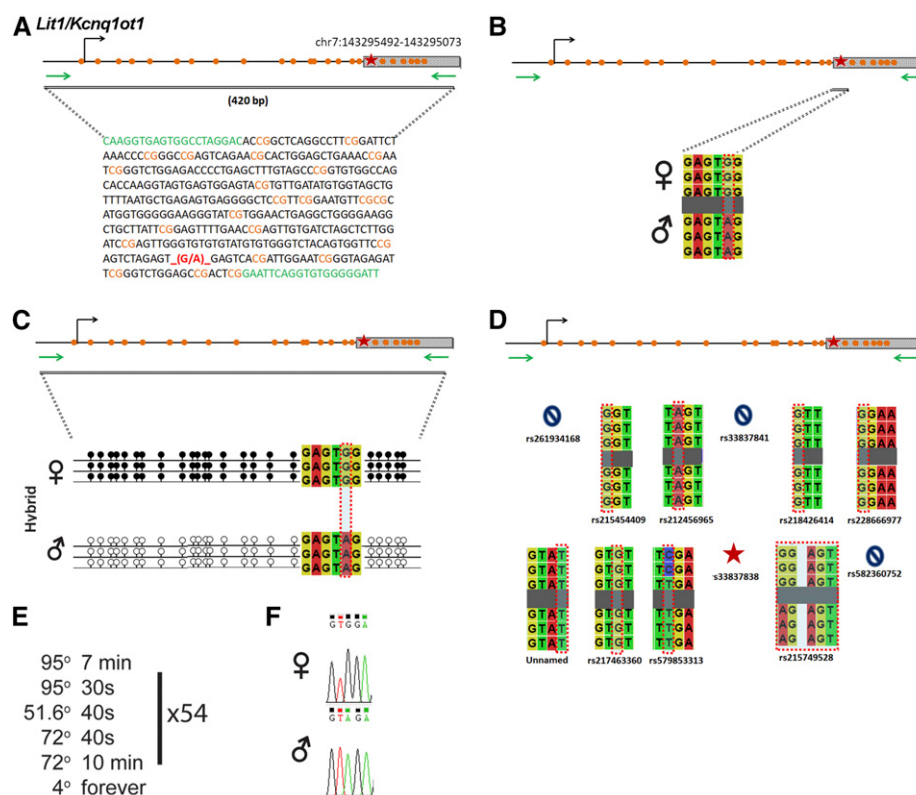
We then used the UCSC Genome Browser in conjunction with dbSNP to determine reported SNPs within a 10-kb window surrounding and including the ICRs, and these SNPs were then cross-checked against the European database as well as the Sanger database to determine their presence in specific strain backgrounds. Following this *in silico* analysis, we designed bisulfite-specific primers to the regions of interest (Table 1). These regions were <1 kb and were within our 10-kb defined window, including a significant portion of the ICR and at least one SNP.



Using this workflow, we validated SNPs in all 10 ICRs and identified PCR conditions for the analysis of each. The relevant details are reported







for each gene below. In addition, we have shown each single band amplicon run in an agarose gel (Supplemental Material, [Figure S1](#)).

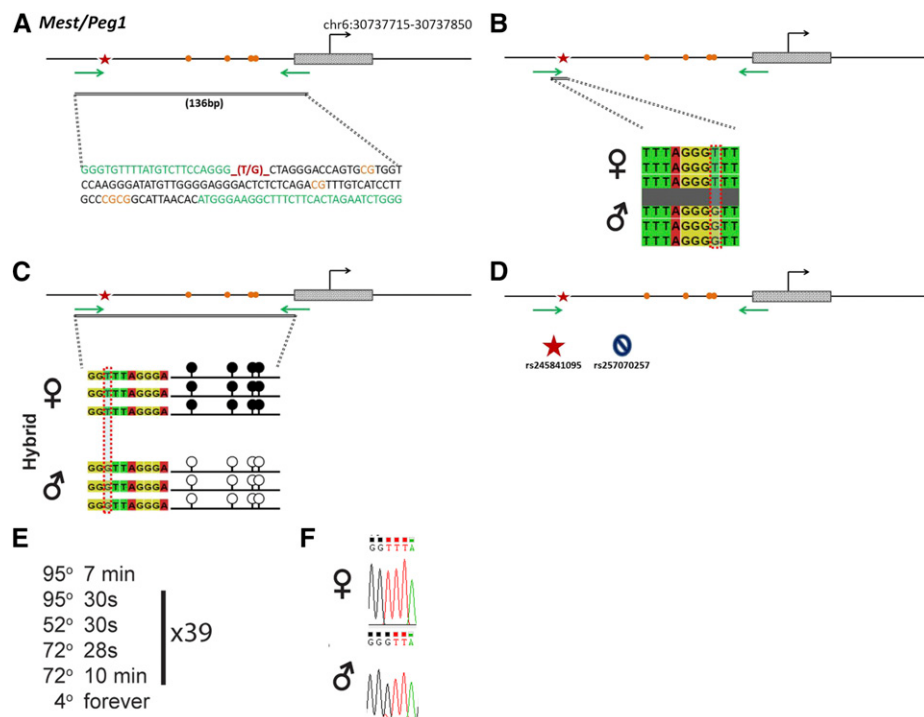
## Grb10

*Grb10* is regulated by an ICR that is ~1.4 kb and located on chromosome 11 in mouse (Figure 2A). Within our probed region, we validated one SNP out of three reported SNPs from the dbSNP database (Figure 2D). The validated SNP is within a 390-bp region containing 31 CpG residues (Figure 2A), with the polymorphic base being an A in the B6 background and a G in the CAST background (Figure 2B). *Grb10* is methylated on the maternal allele and unmethylated on the paternal allele. This methylation pattern was correctly observed in the hybrid progeny using our optimized assay (Figure 2, C and E).

## H19

**Igf2r**



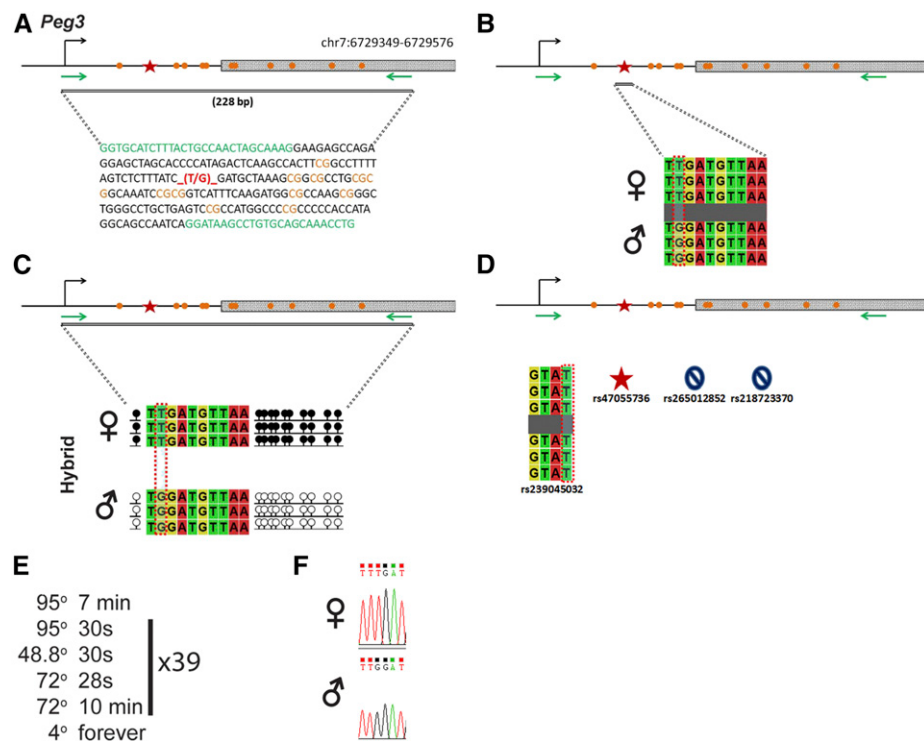


**Figure 7** SNP verification within Mest/Peg1 ICR. (A) Schematic of Mest/Peg1 ICR. Probed region is highlighted by double-dashed line with number of base pairs covered reported. CpG island indicated by dotted box. Green indicates primer sequences; orange indicates CpG dinucleotides; red star and bases indicate verified SNP. (B) Verified SNP presented as sequences from B6 female and CAST male. T-to-G SNP is highlighted by red dotted rectangle. (C) Verification of proper imprinted status in hybrid B6/CAST progeny. SNP highlighted by red dotted rectangle. DNA methylation presented as lollipop diagram; white circles indicate unmethylated cytosines; black circles indicate methylated cytosines. (D) Other SNPs reported in all three databases within the probed region with the SNP highlighted by red dotted rectangle. dbSNP identification number indicated under each SNP. Red star indicates validated SNP and blue closed circle indicates C-to-T polymorphism that cannot be assayed in bisulfite analysis. (E) Optimal PCR conditions for probed region with the given primers. (F) The electropherogram indicating T-to-G polymorphism for the SNP region. ♀, maternal; ♂, paternal.

17 CpG residues (Figure 6A). The polymorphic base is a G in the B6 background and an A in the CAST background (Figure 6B). *Lit1* is methylated on the maternal allele and unmethylated on the paternal allele. This methylation pattern was correctly observed in the hybrid progeny using our optimized assay (Figure 6, C and E).

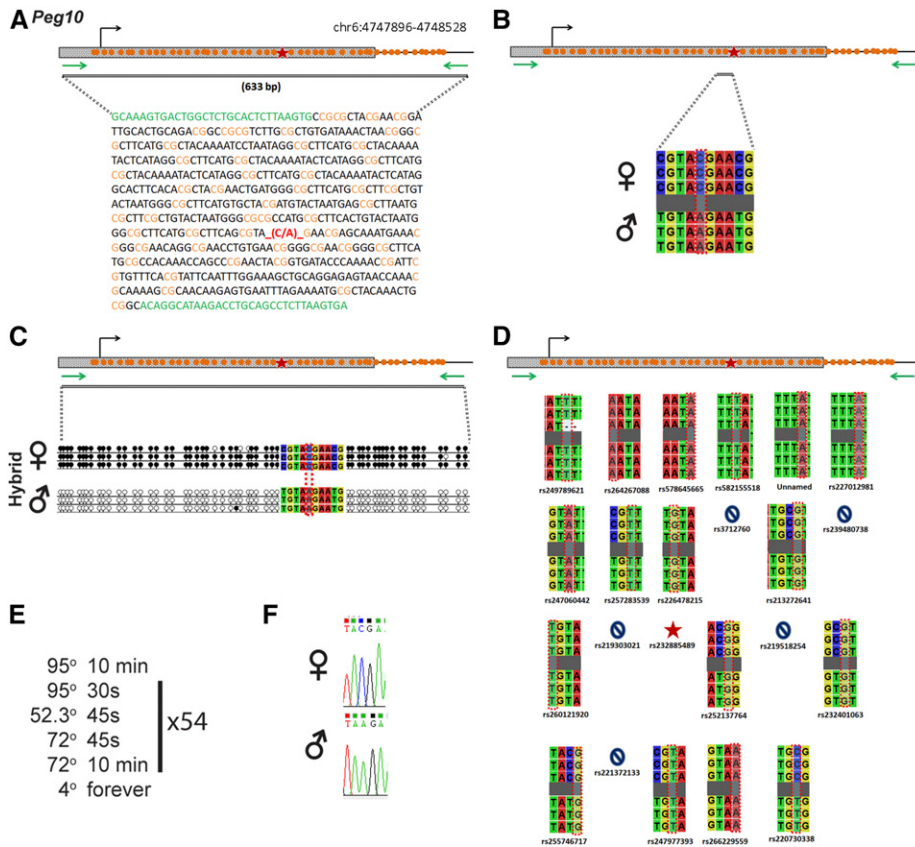
### Mest/Peg1

*Mest/Peg1* is regulated by an ICR on chromosome 6 (Figure 7A). Within our probed region, we validated one SNP out of two reported SNPs from the dbSNP database (Figure 7D). This validated SNP is within a 136-bp region that contains four CpG residues (Figure 7A).



**Figure 8** SNP verification within Peg3 ICR. (A) Schematic of Peg3 ICR. Probed region is highlighted by double-dashed line with number of base pairs covered reported. CpG island indicated by dotted box. Green indicates primer sequences; orange indicates CpG dinucleotides; red star and bases indicate verified SNP. (B) Verified SNP presented as sequences from B6 female and CAST male. T-to-G SNP is highlighted by red dotted rectangle. (C) Verification of proper imprinted status in hybrid B6/CAST progeny. SNP highlighted by red dotted rectangle. DNA methylation presented as lollipop diagram; white circles indicate unmethylated cytosines; black circles indicate methylated cytosines. (D) Other SNPs reported in all three databases within the probed region with the SNP highlighted by red dotted rectangle. dbSNP identification number indicated under each SNP. Red star indicates validated SNP and blue closed circle indicates C-to-T polymorphism that cannot be assayed in bisulfite analysis. (E) Optimal PCR conditions for probed region with the given primers. (F) The electropherogram indicating the T-to-G polymorphism for the SNP region. ♀, maternal; ♂, paternal.

833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893



**Figure 9** SNP verification within *Peg10* ICR. (A) Schematic of *Peg10* ICR. Probed region is highlighted by double-dashed line with number of base pairs covered reported. CpG island indicated by dotted box. Green indicates primer sequences; orange indicates CpG dinucleotides; red star and bases indicate verified SNP. (B) Verified SNP presented as sequences from B6 female and CAST male. C-to-A SNP is highlighted by red dotted rectangle. (C) Verification of proper imprinted status in hybrid B6/CAST progeny. SNP highlighted by red dotted rectangle. DNA methylation presented as lollipop diagram; white circles indicate unmethylated cytosines; black circles indicate methylated cytosines. (D) Other SNPs reported in all three databases within the probed region with the SNP highlighted by red dotted rectangle. dbSNP identification number indicated under each SNP. Red star indicates validated SNP and blue closed circle indicates C-to-T polymorphism that cannot be assayed in bisulfite analysis. (E) Optimal PCR conditions for probed region with the given primers. (F) The electropherogram indicating C-to-A polymorphism for the SNP region. ♀, maternal; ♂, paternal.

894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954

This polymorphic base is a T in the B6 background and a G in the CAST background (Figure 7B). *Mest* is methylated on the maternal allele and unmethylated on the paternal allele. This methylation pattern was correctly observed in the hybrid progeny using our optimized assay (Figure 7, C and E).

### Peg3

*Peg3* is regulated by an ICR on chromosome 7 (Figure 8A). Within our probed region, we validated one SNP out of four reported SNPs from the dbSNP database (Figure 8D). This validated SNP is within a 228-bp region that contains 11 CpG residues (Figure 8A). This polymorphic base is a T in the B6 background and a G in the CAST background (Figure 8B). *Peg3* is methylated on the maternal allele and unmethylated on the paternal allele. This methylation pattern was correctly observed in the hybrid progeny using our optimized assay (Figure 8, C and E).

### Peg10

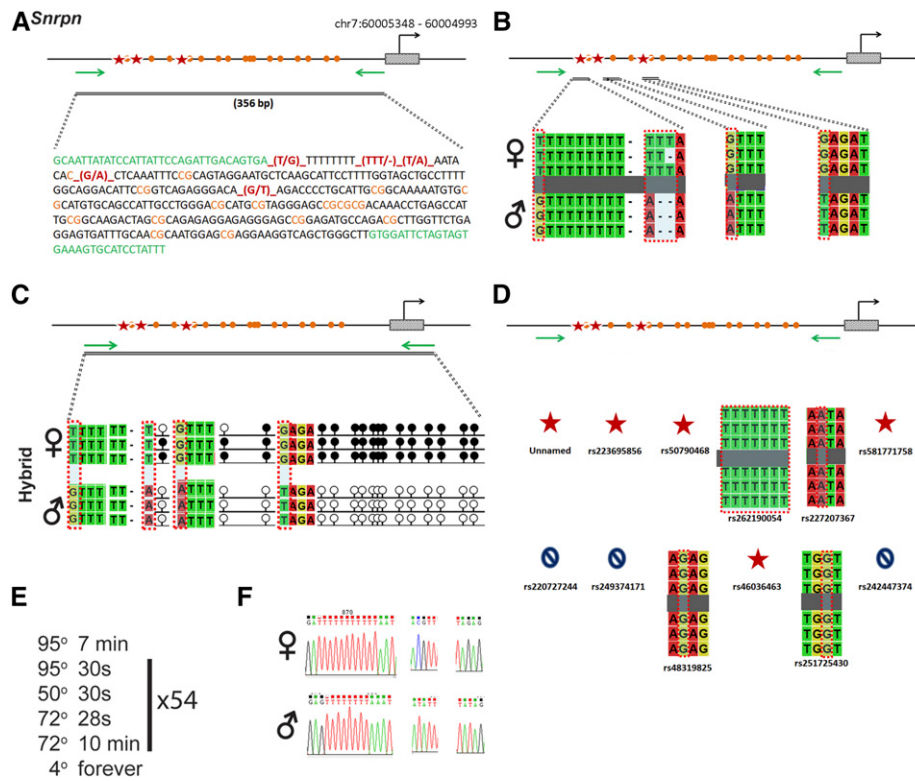
*Peg10* is regulated by an ICR on chromosome 6 (Figure 9A). Within our probed region, we validated one SNP out of 23 reported SNPs from the dbSNP and European databases (Figure 9D). One of the SNPs that was not validated was an unnamed SNP from the European database. The validated SNP is within a 663-bp region that contains 54 CpG residues (Figure 9A). This polymorphic base is a C in the B6 background and an A in the CAST background (Figure 9B). *Peg10* is methylated on the maternal allele and unmethylated on the paternal allele. This methylation pattern was correctly observed in the hybrid progeny using our optimized assay (Figure 9, C and E).

### Snrpn

*Snrpn* is regulated by an ICR on chromosome 7 (Figure 10A). Within our probed region, we validated four SNPs out of 11 reported SNPs from the dbSNP database (Figure 10D). We also identified a novel SNP that is not present in any of the three databases. All five of the validated SNPs are within a 356-bp region that contains 16 CpG residues (Figure 10A). These polymorphic bases include (1) a T in the B6 background and a G in the CAST background, this is the novel SNP that we identified; (2) a TTT in the B6 background and a deletion in the CAST background; (3) a T in the B6 background and an A in the CAST background; (4) a G in the B6 background and an A in the CAST background; and (5) a G in the B6 background and a T in the CAST background (Figure 10B). *Snrpn* is methylated on the maternal allele and unmethylated on the paternal allele. This methylation pattern was correctly observed in the hybrid progeny using our optimized assay (Figure 10, C and E).

### Zac1/Plagl1

*Zac1/Plagl1* is regulated by an ICR on chromosome 10 (Figure 11A). Within our probed region, we validated one SNP out of 11 reported SNPs from the dbSNP and European databases (Figure 11D). The unnamed SNPs are not found in the dbSNP. The validated SNP is within a 578-bp region that contains 33 CpG residues (Figure 11A). This polymorphic base is an A in the B6 background and a G in the CAST background (Figure 11B). *Zac1* is methylated on the maternal allele and unmethylated on the paternal allele. This methylation pattern was correctly observed in the hybrid progeny using our optimized assay (Figure 11, C and E).



**Figure 10** SNP verification within *Snrpn* ICR. (A) Schematic of *Snrpn* ICR. Probed region is highlighted by double-dashed line with number of base pairs covered reported. CpG island indicated by dotted box. Green indicates primer sequences; orange indicates CpG dinucleotides; red star and bases indicate verified SNPs. The chromosome location is from high to low, see *Materials and Methods* for more details. (B) Verified SNPs presented as sequences from B6 female and CAST male. T-to-G, TTT-to-del, T-to-A, G-to-A, and G-to-T SNPs are highlighted by red dotted rectangle. (C) Verification of proper imprint status in hybrid B6/CAST progeny. SNP highlighted by red dotted rectangle. DNA methylation presented as lollipop diagram; white circles indicate unmethylated cytosines; black circles indicate methylated cytosines. (D) Other SNPs reported in all three databases within the probed region with the SNP highlighted by red dotted rectangle. dbSNP identification number indicated under each SNP. Red star indicates validated SNP and blue closed circle indicates C-to-T polymorphism that cannot be assayed in bisulfite analysis. (E) Optimal PCR conditions for probed region with the given primers. (F) The electropherograms indicating T-to-G, TTT-to-del, T-to-A, G-to-A, and G-to-T polymorphisms for the SNP regions. ♀, maternal; ♂, paternal.

## DISCUSSION

Of the SNPs that we analyzed, we were able to validate 18, while we failed to validate 75 SNPs within those same regions (Table 2, red and black). In addition, of those 75 SNPs, 28 of them were C/T polymorphisms that bisulfite analysis was unable to differentiate (Table 2, blue). We also identified a SNP in the *Snrpn* ICR, which was not present in any of the three databases (Table 2, orange). Furthermore, during our optimization we failed to validate multiple SNPs that lie outside of our bisulfite primers. These SNPs are reported in Figure S2. Among the many SNPs reported in the dbSNP database that we failed to verify, most were identified as SNPs between strains other than CAST in the Sanger database. In the end, we could only find one SNP that was supposed to show a polymorphism based on the reported data but did not in our experiments (Table 2, purple). Thus, in general, we recommend using the Sanger database. However, it is important to note that since the Sanger database primarily contains SNPs located close to or within genes, certain ICR SNPs had to be identified in the dbSNP database.

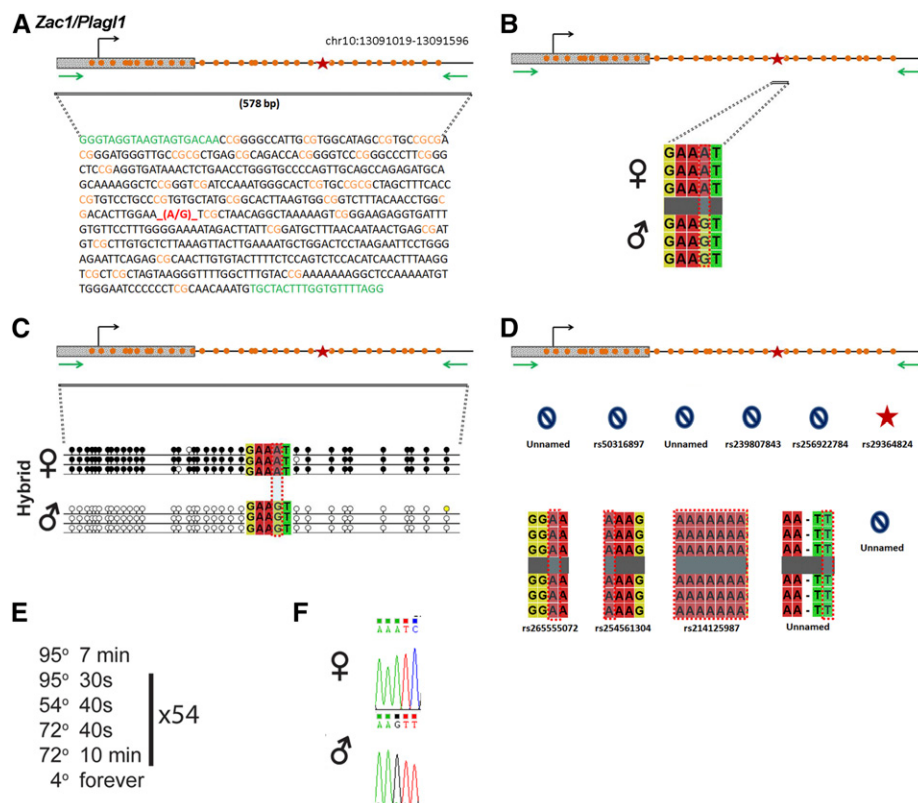
In this resource, we have validated a number of SNPs within the ICRs of the most commonly imprinted loci. In addition, we have demonstrated a high frequency of invalid SNPs within ICRs when the pooled SNPs from the dbSNP (European variation archive) are used alone, highlighting the drawbacks of the mixed strain databases compared to the Sanger strain-specific polymorphism database. Using the validated SNPs, we have optimized allele-specific DNA methylation assays that will allow for the rapid analysis of multiple imprinted loci in a variety of contexts, including at several ICRs that are not contained within the Sanger database. This resource will enable the systematic analysis of multiple imprinted genes in a number of potential applications.

## Potential Applications

As this resource offers extensive and straightforward assays to interrogate the most commonly studied imprinted loci, it can be used across a number of fields. There are two major instances where we envision the utility of this resource: (1) cases where a regulatory mechanism directly interacts with multiple imprinted loci and (2) cases where a mechanism either indirectly regulates many imprinted loci or affects multiple imprinted loci by generally disrupting the epigenetic landscape.

Recently, a number of proteins have been demonstrated to directly regulate multiple imprinted loci. These include, but are not limited to, *Dnmt3l*, *Dnmt1*, *Lsd2*, *Trim28*, *Zfp57*, and *Tet1/2*, each with a different mechanism of action (Bourc'his *et al.* 2001; Howell *et al.* 2001; Reik *et al.* 2003; Li *et al.* 2008; Karytinis *et al.* 2009; Fang *et al.* 2010; Messerschmidt *et al.* 2012; Yamaguchi *et al.* 2013; Canovas and Ross 2016). For example, deletion of the regulatory subunit of the *de novo* DNA methyltransferase *Dnmt3L* results in the failure to establish maternal DNA methylation at a number of maternally imprinted loci, including *Peg3*, *Lit1/Kcnq1ot1*, and *Snrpn* (Bourc'his *et al.* 2001; Hata *et al.* 2002). Another maternal effect enzyme required for the establishment of DNA methylation at maternally imprinted loci is the histone demethylase *Lsd2*. Mechanistically, *Lsd2* is required to remove H3K4 methylation to get proper DNA methylation at imprinted loci including *Mest*, *Grb10*, and *Zac1* (Ciccone *et al.* 2009; Fang *et al.* 2010; Zhang *et al.* 2012; Stewart *et al.* 2015). Furthermore, *Zfp57*, a KRAB domain zinc-finger protein, is required both maternally and zygotically to maintain the imprinting status of various imprinted loci including *Snrpn* (Li *et al.* 2008; Strogantsev and Ferguson-Smith 2012; Strogantsev *et al.* 2015). This protein is thought to bind directly to





**Figure 11** SNP verification within *Zac1/Plagl1* ICR. (A) Schematic of *Zac1/Plagl1* ICR. Probed region is highlighted by double-dashed line with number of base pairs covered reported. CpG island indicated by dotted box. Green indicates primer sequences; orange indicates CpG dinucleotides; red star and bases indicate verified SNP. (B) Verified SNP presented as sequences from B6 female and CAST male. A-to-G SNP is highlighted by red dotted rectangle. (C) Verification of proper imprinted status in hybrid B6/CAST progeny. SNP highlighted by red dotted rectangle. DNA methylation presented as lollipop diagram; white circles indicate unmethylated cytosines; black circles indicate methylated cytosines. (D) Other SNPs reported in all three databases within the probed region with the SNP highlighted by red dotted rectangle. dbSNP identification number indicated under each SNP. Red star indicates validated SNP and blue closed circle indicates C-to-T polymorphism that cannot be assayed in bisulfite analysis. (E) Optimal PCR conditions for probed region with the given primers. (F) The electropherogram indicating A-to-G polymorphism for the SNP region. ♀, maternal; ♂, paternal.

DNA with its zinc fingers and subsequently recruit factors that repress transcription (Li *et al.* 2008; Quenneville *et al.* 2011; Strogantsev *et al.* 2015). These studies demonstrate how disruptions in mechanistically distinct regulatory mechanisms can affect multiple imprinted loci.

Alternatively, a number of mechanisms have been demonstrated to indirectly affect imprinted loci via general epigenetic disruptions. For example, mutations in human *NLRP* genes, which are required maternally for the transition to zygotic gene expression, result in hydatidiform moles and loss of imprinting (Docherty *et al.* 2015). Another maternal effect gene, *Lsd1*, the homolog of *Lsd2*, is also maternally required at fertilization for the maternal to zygotic transition (Ancelin *et al.* 2016; Wasson *et al.* 2016). Loss of maternal *Lsd1* leads to a general disruption of DNA methylation in the resulting progeny at both maternally and paternally imprinted loci (Ancelin *et al.* 2016; Wasson *et al.* 2016). These studies demonstrate how maternal factors, deposited into the zygote from the mother, are required for proper imprinting and development of the embryo.

As ICRs are inherently asymmetric in their epigenetic modifications and opposing mechanisms are required at each parental ICR, even slight disturbances in the epigenetic landscape can lead to significant changes in expression at these loci. For example, disruptions in the maternal expression of *Grb10* results in developmental defects in mice, while disruption of the paternal allele of *Grb10* leads to changes in behavior, including increased social dominance (Garfield *et al.* 2011; Dent and Isles 2014). This highlights differences in the roles of imprinted parental alleles in mice. Another study that highlights the relative contributions of each parental allele describes parental-specific duplications of the 15q11.2-q13.3 region of human chromosome 15 (Isles *et al.* 2016). Paternal duplications were more associated with autism spectrum disorder and developmental delay, while maternal duplications were more

associated with psychiatric disorders (Isles *et al.* 2016). These studies demonstrate the complexity of outcomes associated with maternal vs. paternal inheritance.

Finally, mechanisms that affect imprinted genes indirectly through general epigenetic disruptions highlight how the methylation status of ICRs can act as a proxy for global epigenetic alterations. For example, studies have demonstrated hypomethylation of a differentially methylated region in the *Igf2-H19* locus in Wilms tumor patients (Scharnhorst *et al.* 2001). In addition, embryos conceived using artificial reproductive technologies have higher incidences of Prader-Willi and Angelman syndromes (Horsthemke and Wagstaff 2008; Buiting 2010; Butler 2011). These syndromes are caused by large-scale chromosomal abnormalities that affect multiple imprinted loci (Horsthemke and Wagstaff 2008; Buiting 2010; Butler 2011). It is also possible that imprinting may be disrupted by environmental factors. For example, Bisphenol A, an environmental toxin, as well as various endocrine disruptors, have been revealed to significantly alter the epigenetic landscape (Kang *et al.* 2011; Susiarjo *et al.* 2013). Also, Vinclozolin exposure in mice leads to infertility due to sperm defects in mice, which correlates with global alterations in the DNA methylation landscape (Anway *et al.* 2005; Kang *et al.* 2011). These studies demonstrate additional mechanisms that may lead to broad imprinting disruptions.

## Conclusion

Due to various mechanisms that can disrupt the epigenetic landscape, we anticipate a growing need to assay imprinted loci in different mouse models. The resource provided here will facilitate the future analysis of multiple imprinted loci in a single hybrid genetic background.



1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259

■ Table 2 The complete list of all the SNPs from 3 databases within surveyed regions

Gene	Chromosome	Position	dbSNP	Reference	129P2/OlaHsd	129S1/SvImJ	129SSvEvBrd	A/J	AKR/J	BALB/cJ	C3H/HeJ	C57BL/6NJ	CAST/EiJ	CBA/J	DBA/2J	FVB/NJ	LP/J	NOD/ShiLtJ	NZO/HILtJ	PWK/PhJ	SPRET/EiJ	WSB/EiJ	Present in SANGER
Grb10	11	12,025,379	rs217648878	A	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	No
Grb10	11	12,025,628	rs235292292	C/T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Grb10	11	12,025,688	rs249351785	T/C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
H19	7	142,581,765	rs33821081	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	C	C	-	-	Yes
H19	7	142,581,783	rs33822014	G	-	-	-	-	-	-	-	-	C	-	-	-	-	-	C	-	-	-	Yes
H19	7	142,581,852	rs33822017	C	-	-	-	-	-	-	-	-	T	-	-	-	-	-	T	T	-	-	Yes
H19	7	142,581,933	rs216287265	C	-	-	-	-	-	-	-	-	DEL	-	-	-	-	-	-	-	-	-	No
Igf2r	17	12,742,167	rs222297088	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	Yes
Igf2r	17	12,742,203	rs242482749	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	Yes
Igf2r	17	12,742,239	rs578459511	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	Yes
Igf2r	17	12,742,253	rs211862027	T/-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Igf2r	17	12,742,283	rs229760939	G	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	Yes
Igf2r	17	12,742,373	rs108681933	C	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	Yes
Igf2r	17	12,742,426	rs250523644	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	Yes
Igf2r	17	12,742,469	rs265144059	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	Yes
Igf2r	17	12,742,474	rs107811421	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	Yes
Igf2r	17	12,742,517	rs245573738	A	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	Yes
Igf2r	17	12,742,538	rs216289274	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	Yes
Igf2r	17	12,742,554	rs234366750	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	Yes
Igf2r	17	12,742,579	rs107601903	A	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	G	G	-	Yes
Impact	18	12,972,845	rs29925054	T/C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Impact	18	12,972,852	rs235629089	A/G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Impact	18	12,972,910	rs223891728	T/-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Impact	18	12,972,953	rs261124226	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	No
Impact	18	12,972,960	UNNAMED	C/T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Impact	18	12,972,965	rs31057356	A	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	No
Impact	18	12,972,968	rs240274686	CAG/-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Impact	18	12,973,031	rs29558070	C/T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Impact	18	12,973,055	rs251991535	C/T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Impact	18	12,973,080	rs220788023	C	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	No
Lit1	7	143,295,133	rs582360752	G/A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Lit1	7	143,295,136	rs215749528	ACTCC/-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Lit1	7	143,295,140	rs33837838	C	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	No
Lit1	7	143,295,152	rs579853313	G/C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Lit1	7	143,295,180	rs217463360	C/T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Lit1	7	143,295,277	UNNAMED	A/T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Lit1	7	143,295,291	rs228666977	C/A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Lit1	7	143,295,295	rs218426414	C/-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Lit1	7	143,295,335	rs33837841	A/G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Lit1	7	143,295,366	rs212456965	T/C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Lit1	7	143,295,375	rs215454409	C/G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Lit1	7	143,295,438	rs261934168	G/A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Mest	6	30,737,737	rs245841095	T	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	G	G	-	Yes
Mest	6	30,737,801	rs257070257	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	Yes
Peg3	7	6,729,398	rs239045032	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	Yes
Peg3	7	6,729,440	rs47057736	T	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	G	G	-	Yes
Peg3	7	6,729,451	rs265012852	C/T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Peg3	7	6,729,539	rs218723370	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	Yes
Peg10	6	4,748,007	rs249789621	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	Yes
Peg10	6	4,748,010	rs264267088	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	Yes
Peg10	6	4,748,036	rs578645665	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	Yes
Peg10	6	4,748,039	rs582155518	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	Yes
Peg10	6	4,748,069	UNNAMED	A/G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No

(continued)

1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381

■ Table 2, continued

Gene	Chromosome	Position	dbSNP	Reference	129p2/OlaHsd	129s1/SvimJ	129S5SvEvBrd	A/J	AKR/J	BALB/cJ	C3H/HeJ	C57BL/6NJ	CAST/EiJ	CBA/J	DBA/2J	FVB/NJ	LP/J	NOD/ShiLtJ	NZO/HILtJ	PWK/PHJ	SPRET/EiJ	WSB/EiJ	Present in SANGER
Peg10	6	4,748,127	rs227012981	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	Yes
Peg10	6	4,748,133	rs247060442	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	Yes
Peg10	6	4,748,174	rs257283539	A/C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Peg10	6	4,748,176	rs226478215	C/G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Peg10	6	4,748,196	rs3712760	T	C	C	C	C	C	C	C	-	C	C	C	C	C	-	C	C	-	-	Yes
Peg10	6	4,748,197	rs213272641	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	Yes
Peg10	6	4,748,249	rs239480738	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	Yes
Peg10	6	4,748,262	rs260121920	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	Yes
Peg10	6	4,748,291	rs219303021	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	Yes
Peg10	6	4,748,295	rs232885489	C	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	Yes
Peg10	6	4,748,313	rs252137764	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	Yes
Peg10	6	4,748,341	rs219518254	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	Yes
Peg10	6	4,748,351	rs232401063	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	Yes
Peg10	6	4,748,382	rs255746717	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	Yes
Peg10	6	4,748,384	rs221372133	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	C	C	-	Yes
Peg10	6	4,748,413	rs247977393	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	C	-	Yes
Peg10	6	4,748,442	rs266229559	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	Yes
Peg10	6	4,748,482	rs220730338	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	Yes
Snrpn	7	60,005,033	rs242447374	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	Yes
Snrpn	7	60,005,074	rs251725430	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	Yes
Snrpn	7	60,005,215	rs46036463	C	A	-	-	-	-	-	-	-	A	-	A	-	-	-	-	A	A	-	Yes
Snrpn	7	60,005,223	rs48319825	C	T	-	-	-	-	-	-	-	-	-	T	-	-	-	-	T	-	-	Yes
Snrpn	7	60,005,271	rs249374171	G	A	-	-	-	-	-	-	-	A	-	A	-	-	-	-	A	-	-	Yes
Snrpn	7	60,005,282	rs220727244	G	A	-	-	-	-	-	-	-	A	-	A	-	-	-	-	A	A	-	Yes
Snrpn	7	60,005,295	rs581771758	C	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	Yes
Snrpn	7	60,005,301	rs227207367	T/-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Snrpn	7	60,005,303	rs262190054	AAAAAAA/-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Snrpn	7	60,005,303	rs50790468	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	No
Snrpn	7	60,005,303	rs223695856	AAA	-	-	-	-	-	-	-	-	DEL	-	-	-	-	-	-	-	-	-	No
Snrpn	7	60,005,316	UNNAMED	A	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	No
Plagl1	10	13,091,157	UNNAMED	C/T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Plagl1	10	13,091,167	rs50316897	C	-	-	-	-	-	-	-	-	T	-	T	-	-	-	-	T	-	-	Yes
Plagl1	10	13,091,224	UNNAMED	T/C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Plagl1	10	13,091,272	rs239807843	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	Yes
Plagl1	10	13,091,284	rs256922784	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	Yes
Plagl1	10	13,091,296	rs29364824	A	-	-	-	-	-	-	-	-	G	-	G	-	-	-	-	G	G	G	Yes
Plagl1	10	13,091,347	rs265555072	A/-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Plagl1	10	13,091,403	rs254561304	A/-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Plagl1	10	13,091,531	rs214125987	AAAAAAAAA/-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Plagl1	10	13,091,559	UNNAMED	C/-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No
Plagl1	10	13,091,566	UNNAMED	C/T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	No

Orange indicates that the SNP is the SNP we have found; it is not present in any database. Red indicates that SNPs are validated polymorphisms. Blue indicates that SNPs are C/T (or G/A) variations that bisulfite sequencing assay can't detect. Purple indicates that the SNP is the only inconsistency between our sequencing result (C on B6 background) and the reported Sanger data (G on B6 background). Green nucleotides indicate the present polymorphism on both assayed backgrounds (B6 and CAST) at reported SNP locations

## ACKNOWLEDGMENTS

We thank the epigenetic community at Emory University for their feedback, T. Lee for help in editing this article, and A. Ferguson-Smith and M. Bartolomei for feedback on the article. In addition, we thank M. Bartolomei for providing the *H19* assay and D. Cutler for bioinformatics assistance. J.A.W. was supported by the Biochemistry, Cell and Molecular Biology Training Grant (5T32GM008367). The work was supported by a grant to D.J.K. from the National Science Foundation (IOS1354998).

## LITERATURE CITED

Ancelin, K., L. Syx, M. Borensztein, N. Ranisavljevic, I. Vassilev *et al.*, 2016 Maternal LSD1/KDM1A is an essential regulator of chromatin

and transcription landscapes during zygotic genome activation. *Elife* 5: e08851.  
Anway, M. D., A. S. Cupp, M. Uzumcu, and M. K. Skinner, 2005 Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* 308: 1466–1469 (erratum: *Science* 328: 690).  
Barlow, D. P., 2011 Genomic imprinting: a mammalian epigenetic discovery model. *Annu. Rev. Genet.* 45: 379–403.  
Bartolomei, M. S., 2009 Genomic imprinting: employing and avoiding epigenetic processes. *Genes Dev.* 23: 2124–2133.  
Bartolomei, M. S., and A. C. Ferguson-Smith, 2011 Mammalian genomic imprinting. *Cold Spring Harb. Perspect. Biol.* 3: 1–17.  
Bartolomei, M. S., and S. M. Tilghman, 1997 Genomic imprinting in mammals. *Annu. Rev. Genet.* 31: 493–525.

1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442

1443 Beck, J. A., S. Lloyd, M. Hafezparast, M. Lennon-Pierce, J. T. Eppig *et al.*,  
1444 2000 Genealogies of mouse inbred strains. *Nat. Genet.* 24: 23–25.  
1445 Bourc'his, D., G. L. Xu, C. S. Lin, B. Bollman, and T. H. Bestor,  
1446 2001 Dnmt3L and the establishment of maternal genomic imprints.  
1447 *Science* 294: 2536–2539.  
1448 Buiting, K., 2010 Prader-Willi syndrome and Angelman syndrome. *Am.*  
1449 *J. Med. Genet. C Semin. Med. Genet.* 154C: 365–376.  
1450 Butler, M. G., 2011 Prader-Willi syndrome: obesity due to genomic im-  
1451 printing. *Curr. Genomics* 12: 204–215.  
1452 Canovas, S., and P. J. Ross, 2016 Epigenetics in preimplantation mamma-  
1453 lian development. *Theriogenology* 86: 69–79.  
1454 Ciccone, D. N., H. Su, S. Hevi, F. Gay, H. Lei *et al.*, 2009 KDM1B is a  
1455 histone H3K4 demethylase required to establish maternal genomic im-  
1456 prints. *Nature* 461: 415–418.  
1457 Dent, C. L., and A. R. Isles, 2014 Brain-expressed imprinted genes and adult  
1458 behaviour: the example of Nesp and Grb10. *Mamm. Genome* 25: 87–93.  
1459 DeVeale, B., D. van der Kooy, and T. Babak, 2012 Critical evaluation of  
1460 imprinted gene expression by RNA-seq: a new perspective. *PLoS Genet.*  
1461 8: e1002600.  
1462 Docherty, L. E., F. I. Rezwan, R. L. Poole, C. L. S. Turner, E. Kivuva *et al.*,  
1463 2015 Mutations in NLRP5 are associated with reproductive wastage and  
1464 multilocus imprinting disorders in humans. *Nat. Commun.* 6: 8086.  
1465 Fang, R., A. J. Barbera, Y. Xu, M. Rutenber, T. Leonor *et al.*, 2010 Human  
1466 LSD2/KDM1b/AOF1 regulates gene transcription by modulating intra-  
1467 genic H3K4me2 methylation. *Mol. Cell* 39: 222–233.  
1468 Ferguson-Smith, A. C., 2011 Genomic imprinting: the emergence of an  
1469 epigenetic paradigm. *Nat. Rev. Genet.* 12: 565–575.  
1470 Frazer, K. A., E. Eskin, H. M. Kang, M. A. Bogue, D. A. Hinds *et al.*, 2007 A  
1471 sequence-based variation map of 8.27 million SNPs in inbred mouse  
1472 strains. *Nature* 448: 1050–1053.  
1473 Garfield, A. S., M. Cowley, F. M. Smith, K. Moorwood, J. E. Stewart-Cox  
1474 *et al.*, 2011 Distinct physiological and behavioural functions for pa-  
1475 rental alleles of imprinted Grb10. *Nature* 469: 534–538.  
1476 Gregg, C., J. Zhang, B. Weissbourd, S. Luo, G. P. Schroth *et al.*, 2010 High  
1477 resolution analysis of parent-of-origin allelic expression in the mouse  
1478 brain. *Science* 329: 643–648.  
1479 Hata, K., M. Okano, H. Lei, and E. Li, 2002 Dnmt3L cooperates with the  
1480 Dnmt3 family of de novo DNA methyltransferases to establish maternal  
1481 imprints in mice. *Development* 129: 1983–1993.  
1482 Hayatsu, H., M. Shiraishi, and K. Negishi, 2008 Bisulfite modification for  
1483 analysis of DNA methylation. *Curr. Protoc. Nucleic Acid Chem.* Chapter  
1484 6: Unit 6.10.  
1485 Horsthemke, B., and J. Wagstaff, 2008 Mechanisms of imprinting of the  
1486 Prader-Willi/Angelman region. *Am. J. Med. Genet. A.* 146: 2041–2052.  
1487 Howell, C. Y., T. H. Bestor, F. Ding, K. E. Latham, C. Mertineit *et al.*,  
1488 2001 Genomic imprinting disrupted by a maternal effect mutation in  
1489 the *Dnmt1* gene. *Cell* 104: 829–838.  
1490 Isles, A. R., A. Ingason, C. Lowther, J. Walters, M. Gawlick *et al.*,  
1491 2016 Parental origin of interstitial duplications at 15q11.2-q13.3 in  
1492 schizophrenia and neurodevelopmental disorders. *PLoS Genet.* 12:  
1493 e1005993.  
1494 Kang, E.-R., K. Iqbal, D. A. Tran, G. E. Rivas, P. Singh *et al.*, 2011 Effects of  
1495 endocrine disruptors on imprinted gene expression in the mouse embryo.  
1496 *Epigenetics* 6: 937–950.  
1497 Karytinis, A., F. Forneris, A. Profumo, G. Ciossani, E. Battaglioli *et al.*,  
1498 2009 A novel mammalian flavin-dependent histone demethylase.  
1499 *J. Biol. Chem.* 284: 17775–17782.  
1500 Keane, T. M., L. Goodstadt, P. Danecek, M. A. White, K. Wong *et al.*,  
1501 2011 Mouse genomic variation and its effect on phenotypes and gene  
1502 regulation. *Nature* 477: 289–294.  
1503 Kelsey, G., and M. S. Bartolomei, 2012 Imprinted genes... and the number  
1504 is? *PLoS Genet.* 8: e1002601.  
1505 Laird, P. W., 2010 Principles and challenges of genome-wide DNA meth-  
1506 ylation analysis. *Nat. Rev. Genet.* 11: 191–203.  
1507 Li, X., M. Ito, F. Zhou, N. Youngson, X. Zuo *et al.*, 2008 A maternal-zygotic  
1508 effect gene, *Zfp57*, maintains both maternal and paternal imprints. *Dev.*  
1509 *Cell* 15: 547–557.  
1510 Messerschmidt, D. M., W. de Vries, M. Ito, D. Solter, A. Ferguson-Smith  
1511 *et al.*, 2012 Trim28 is required for epigenetic stability during mouse  
1512 oocyte to embryo transition. *Science* 335: 1499–1502.  
1513 Mitchell, A. A., M. E. Zwick, A. Chakravarti, and D. J. Cutler,  
1514 2004 Discrepancies in dbSNP confirmation rates and allele frequency  
1515 distributions from varying genotyping error rates and patterns. *Bioin-*  
1516 *formatics* 20: 1022–1032.  
1517 Nekrutenko, A., and J. Taylor, 2012 Next-generation sequencing data in-  
1518 terpretation: enhancing reproducibility and accessibility. *Nat. Rev. Genet.*  
1519 13: 667–672.  
1520 Quenneville, S., G. Verde, A. Corsinotti, A. Kapopoulou, J. Jakobsson *et al.*,  
1521 2011 In embryonic stem cells, ZFP57/KAP1 recognize a methylated  
1522 hexanucleotide to affect chromatin and DNA methylation of imprinting  
1523 control regions. *Mol. Cell* 44: 361–372.  
1524 Reik, W., and W. Dean, 2001 DNA methylation and mammalian epige-  
1525 netics. *Electrophoresis* 22: 2838–2843.  
1526 Reik, W., and J. Walter, 2001 Genomic imprinting: parental influence on  
1527 the genome. *Nat. Rev. Genet.* 2: 21–32.  
1528 Reik, W., F. Santos, and W. Dean, 2003 Mammalian epigenomics: re-  
1529 programming the genome for development and therapy. *Theriogenology*  
1530 59: 21–32.  
1531 Scharnhorst, V., A. J. Van Der Eb, and A. G. Jochemsen, 2001 WT1 pro-  
1532 teins: functions in growth and differentiation. *Gene* 273: 141–161.  
1533 Sherry, S. T., M. H. Ward, M. Kholodov, J. Baker, L. Phan *et al.*,  
1534 2001 dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.*  
1535 29: 308–311.  
1536 Smigielski, E. M., K. Sirotkin, M. Ward, and S. T. Sherry, 2000 dbSNP: a  
1537 database of single nucleotide polymorphisms. *Nucleic Acids Res.* 28: 352–  
1538 355.  
1539 Stewart, K. R., L. Veselovska, J. Kim, J. Huang, H. Saadeh *et al.*,  
1540 2015 Dynamic changes in histone modifications precede de novo DNA  
1541 methylation in oocytes. *Genes Dev.* 29: 2449–2462.  
1542 Strogantsev, R., and A. C. Ferguson-Smith, 2012 Proteins involved in es-  
1543 tablishment and maintenance of imprinted methylation marks. *Brief.*  
1544 *Funct. Genomics* 11: 227–239.  
1545 Strogantsev, R., F. Krueger, K. Yamazawa, H. Shi, P. Gould *et al.*,  
1546 2015 Allele-specific binding of ZFP57 in the epigenetic regulation of  
1547 imprinted and non-imprinted monoallelic expression. *Genome Biol.* 16:  
1548 112.  
1549 Susiarjo, M., I. Sasson, C. Mesaros, and M. S. Bartolomei, 2013 Bisphenol A  
1550 exposure disrupts genomic imprinting in the mouse. *PLoS Genet.* 9:  
1551 e1003401.  
1552 Wan, L. B., and M. S. Bartolomei, 2008 Regulation of imprinting in clusters:  
1553 noncoding RNAs versus insulators. *Adv. Genet.* 61: 207–223.  
1554 Wasson, J. A., A. K. Simon, D. A. Myrick, G. Wolf, S. Driscoll *et al.*,  
1555 2016 Maternally provided LSD1/KDM1A enables the maternal-to-zy-  
1556 gotic transition and prevents defects that manifest postnatally. *Elife* 5:  
1557 e08848.  
1558 Yalcin, B., K. Wong, A. Agam, M. Goodson, T. M. Keane *et al.*,  
1559 2011 Sequence-based characterization of structural variation in the  
1560 mouse genome. *Nature* 477: 326–329.  
1561 Yamaguchi, S., L. Shen, Y. Liu, D. Sendler, and Y. Zhang, 2013 Role of Tet1  
1562 in erasure of genomic imprinting. *Nature* 504: 460–464.  
1563 Zhang, Q., S. Qi, M. Xu, L. Yu, Y. Tao *et al.*, 2012 Structure-function  
1564 analysis reveals a novel mechanism for regulation of histone demethylase  
1565 LSD2/AOF1/KDM1b. *Cell Res.* 2: 1–17.

Communicating editor: D. Threadgill

**GGG January (2018)**  
**Author query sheet Wasson (GGG\_300417)**

**QAI** If you or your coauthors would like to include an ORCID ID in this article, please provide your respective ORCID IDs along with your corrections.

Note: If you do not yet have an ORCID ID and would like one, you may register for this unique digital identifier at <https://orcid.org/register>.

- 1** Please verify corresponding author address.
- 2** Please check all figure legends carefully to confirm that any and all labels, designators, directionals, colors, *etc.* are represented accurately in comparison with the figure images.
- 3** Please check the Supplemental Material links in your article.
- 4** Please confirm or update any and all URLs in your article.
- 5** Please check use of italics throughout your article, including all taxonomic and genetic nomenclature. Uppercase Greek letters should remain roman per journal style even when appearing in a term where the overall style is italic (*e.g.*, a gene name such as *kap108Δ*). Note that headings are set all Roman or all italics based on journal style and should not be changed.
- 6** Please verify the keywords are correct as listed.
- 7** Any alternations between capitalization and/or italics in genetic nomenclature have been retained per the original manuscript. Please confirm that all genetic nomenclature has been formatted properly throughout.
- 8** Throughout the article, instances of “Castaneus background” have been edited to “CAST background” to match the terminology defined in the sentence “To generate hybrids with SNPs on each allele. . .,” please verify or amend this edit.
- 9** In the legends for Figures 2–11, “♀, maternal; ♂, paternal” has been inserted to explain the symbols on the diagrams. Please check if this is correct.
- 10** Figure panel 5F is poor quality. Please provide better quality artwork.