Comparative Phylogenetic Analysis of Blcap/Nnat Reveals Eutherian-Specific Imprinted Gene

Heather K. Evans,*† Jennifer R. Weidman,*† Dale O. Cowley,* and Randy L. Jirtle*†

*Department of Radiation Oncology, Duke University Medical Center; †University Program in Genetics and Genomics, Duke University Medical Center; and Department of Genetics and Lineberger Comprehensive Cancer Center, University of North Carolina

Imprinted genes are parent-of-origin dependent, monoallelically expressed genes present in marsupials and eutherian mammals. Altered expression of imprinted genes plays a significant role in the etiology of a variety of human disorders and diseases. Nevertheless, the regulatory mechanisms of imprinting remain poorly defined. The imprinted gene Neuronatin (Nnat) is an excellent candidate for studying imprinting because it resides within the 8.5-kb intron of the nonimprinted gene Bladder Cancer-Associated Protein (Blcap) and is the only imprinted gene within the region. A phylogenetic comparison of this micro-imprinted domain in mouse, human, and rat revealed several candidates for imprint control, including tandem repeats and putative binding sites for trans-acting factors known to be involved in chromatin remodeling. Genome-wide phylogenetic comparisons of species from the three major extant mammalian clades failed, however, to show any evidence of Nnat outside the eutherian lineage. Thus, Nnat is the first identified eutherian-specific imprinted gene, demonstrating that imprinted genes did not arise at a single point during evolution. This finding also suggests that the complexity of imprinting regulation observed at other loci may, in part, be directly related to the amount of time they have been imprinted.

Introduction

Genomic imprinting refers to parent-of-origin specific, monoallelic expression of a gene. The expression from a singular parental allele occurs despite identical sequence information and cellular environment. Imprinting regulation is complex with layers of control involving multiple cis- and trans-acting elements. The presence of tandem repeats, differential methylation of CpG islands, boundary elements, nontranslated RNA, and temporal differences in DNA replication have all been associated with imprinting (Reik and Walter 2001b; Delaval and Feil 2004).

Three general transcriptional regulatory motifs for imprint control have been proposed (Ferguson-Smith and Surani 2001; Reik, Dean, and Walter 2001). The most complex method uses long-range cis-acting elements to control multiple, clustered imprinted genes. Examples in humans of imprinted gene clusters regulated in this manner include the Beckwith-Wiedemann locus at 11p15.5 and the Prader-Willi/Angelman syndrome locus at 15q11-q13 (Paulsen et al. 1998; Nicholls and Knepper 2001). Secondly, some genes, such as Igf2r, are regulated in cis by the production of an imprinted antisense RNA that represses expression (Smilich et al. 1999; Chamberlain and Brannan 2001; Sleutels, Zwart, and Barlow 2002). Thirdly, allele-specific promoter hypermethylation represses expression and leads to imprinting at some loci, including Neuronatin (Nnat) (Evans et al. 2001; John et al. 2001) and U2af1-rs1 (Nabetai et al. 1997). This last regulatory method is postulated to be the primordial imprint because it is the simplest in form (Reik and Walter 2001a).

Despite differences in the methods of imprinting regulation, all imprinted genes should share some imprinting characteristics. Inherited, differentially marked motifs that identify an imprinted allele must be present in the genome. Furthermore, sequences required for imprint regulation should be subject to the forces of natural selection, and thus be preferentially conserved both within and between species. Previous reports have shown phylogenetic comparisons between the three major extant mammalian clades Eutheria (placental mammals), Metatheria (marsupials), and Prototheria (monotremes) to be useful in identifying such conserved sequences (Killian et al. 2000; Weidman et al. 2004). These comparative phylogenetic studies were conducted with Igf2 and Igf2r, models for two of the distinct classes of imprint control—imprinted clusters and antisense RNA. A similar phylogenetic comparison has not been conducted, however, with a representative example for the third imprint motif—promoter hypermethylation. Consequently, we conducted a phylogenetic analysis of the imprinted gene Nnat, a candidate for this class of imprint regulation (Reik and Walter 2001a), in order to identify sequences potentially involved in its imprint control.

Nnat expresses highly in the central nervous system from midgestation through early postnatal development, correlating with the onset and termination of brain development in mice and humans (Joseph, Dou, and Tsang 1994; Wijnholds et al. 1995). Its expression appears to be limited, however, to the anterior pituitary in the adult brain. Nnat is also expressed in the developing pancreas, becoming restricted in adulthood to beta cells (Chu and Tsai 2005). Nnat is imprinted in both mouse and human, and all three exons of Nnat lie within the singular intron of the nonimprinted gene, Bladder Cancer-Associated Protein (Blcap) (Kagitani et al. 1997; Evans et al. 2001; John et al. 2001). This finding is especially surprising given that the intron of Blcap is only 8.5 kb, and imprinted genes can be coordinate regulated in clusters spanning distances up to 2 Mb (Buiting et al. 1995; Nicholls, Saitoh, and Horsthemke 1998; Thorvaldsen, Duran, and Bartolomei 1998). Therefore, we postulated that cis-acting sequences important for maintaining the transcriptional status of this microimprinted domain should be present within a relatively short distance of Nnat. Such sequences would allow for the differential allelic expression observed for Nnat while preventing the spread of imprinting to Blcap.

Key words: Blcap, eutherian, evolution, imprinting, Nnat, phylogenetics.

E-mail: jirtle@radonc.duke.edu.

doi:10.1093/molbev/msi165
Advance Access publication May 18, 2005

© The Author 2005. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved. For permissions, please e-mail: journals.permissions@oupjournals.org
Our phylogenetic analysis of the orthologous imprinted domain of Nnat in human, mouse, and rat identified several highly conserved, noncoding sequences predicted to be important in regulating the imprinting of this gene. We also demonstrated that Nnat does not reside within the opossum Blcap gene (oBlcap) as it does in humans and mice. Moreover, we provide evidence that Nnat does not exist outside the eutherian lineage, thereby defining Nnat as the first known eutherian-specific imprinted gene.

Materials and Methods
Tissue and DNA Samples, Bacterial Artificial Chromosome Library

Tissues and DNA were obtained from the following resources—chicken: Department of Poultry Science, North Carolina State University, Raleigh, N.C.; didelphis opossum: Michael Stoskopf, North Carolina State University and the North Carolina Wildlife Commission, Raleigh, N.C.; elephant: North Carolina Zoo, Asheboro, N.C.; genomic yeast DNA: Invitrogen, Carlsbad, Calif.; human: Birth Defects Research Laboratory, University of Washington, Seattle, Wash.; lemur: Duke University Primate Center, Durham, N.C.; monodelphis opossum: Kathleen Smith, Duke University, Durham, N.C.; mouse kidney DNA: BD Biosciences, Palo Alto, Calif.; sheep: U.S. Department of Agriculture Meat Animal Research Center, Clay Center, Nebr.; and wallaby, echidna, and platypus: Barry Munday, University of Tasmania, Launceston, Tasmania. All tissues were stored at −80°C. Amplexic Express (Pullman, Wash.) constructed the 5× coverage opossum (Didelphis virginiana) genomic bacterial artificial chromosome (BAC) library. This study was performed in accordance with current regulations and standards of the U.S. Department of Agriculture, Department of Health and Human Services, and National Institutes of Health.

Probing BAC Membranes

Probes were generated using the RadPrime DNA Labeling System (Invitrogen) according to manufacturer’s directions. Primers for Blcap polymerase chain reaction (PCR) template were 5′-CAG CTC CTG GAG AGA GAG TCG-3′ and 5′-TGG GGA ATC GGA GCA GTG-3′, and the product was amplified using Platinum Taq DNA Polymerase (Invitrogen). For Nnat, the product was amplified using the GC-RICH PCR System (Roche, Indianapolis, Ind.); primers were 5′-GGA CTC CGA GAC CAG TAG ACC-3′ and 5′-TGG TGC CTA CGC CCA TAT C-3′. Cycling conditions for both reactions were 63°C annealing, 30 s extension, and 35 cycles. Membranes were prehbed in 10 ml of ExpressHyb (BD Biosciences) for 2 h at 65°C. Approximately 10 ml of denatured probe was added to 10 ml of fresh solution, and each membrane was hybridized overnight at 65°C. Membranes were washed as follows: twice 2× standard saline citrate (SSC) for 5 min at room temperature, twice 0.5× SSC/1% sodium dodecyl sulfate (SDS) for 30 min at 65°C, and twice 0.5× SSC for 5 min at room temperature. Films were exposed for 5 days and then developed. Positive clones were selected from Amplexic Express, grown, and prepped using a modified Qiagen midi protocol (Valencia, Calif.) in which the P1, P2, and P3 buffer volumes were increased to 10 ml. Elution was performed by five applications of 1 ml of elution buffer prewarmed to 65°C.

DNA Isolation

Tissue was homogenized and incubated with 30 µl of 20 mg/ml proteinase K (Qiagen) and 600 µl HCl proK buffer (50 mM Tris-HCl pH 8.0, 100 mM EDTA, 100 mM NaCl, 1% SDS) at 55°C. After overnight incubation, 10 µl of 10 mg/ml RNase A was added and incubated for 1 h at 37°C. DNA was subsequently extracted with phenol-chloroform followed by chloroform, then precipitated with an equal volume of isopropanol. The resulting pellet was suspended in 10 mM Tris-HCl, pH 8.0, and heated at 65°C to ensure complete resuspension.

Genomic and BAC Southern Blots

Approximately 10 µg of genomic DNA or 2 µg of BAC DNA were digested with 100 units of EcoRI and 100 units of BglII (New England Biolabs, Beverly, Mass.). Digested DNA was run on a 1% agarose gel. Transfer was accomplished by soaking the gel for 5 min in 250 mM HCl, twice for 15 min in 0.5 M NaOH/1.5 M NaCl, twice for 15 min in 0.5 M Tris-HCl pH 7.5/1.5 M NaCl, and for 10 min in 20 × SSC. DNA was transferred to a nylon membrane (Roche) with Turboblotter (Schleicher & Schuell, Keene, N.H.) according to manufacturer’s directions. DNA was bound to the membrane by UV cross-linking.

For hybridization, probes were generated using the PCR DIG Probe Synthesis Kit (Roche). Blcap primers were 5′-CAG CTC CTG GAG AGA GAG TCG-3′ and 5′-GGA CTC CGA GAC CAG TAG ACC-3′ and 5′-TGG TGC CTA CGC CCA TAT C-3′. PCR amplification conditions were 60°C annealing, 30 s extension, and 35 cycles. For Nnat, 1 µl of the PCR product used to generate the BAC probes was used as template with primers 5′-GGA CTC CGA GAC CAG TAG ACC-3′ and 5′-TCT TCA TCA TTG TCG GTG-3′; PCR amplification conditions were 63°C, 45 s extension, and 35 cycles. Approximately 20 µl of labeled probe was added to a prehybed membrane and incubated overnight at appropriate calculated temperatures. The membranes were washed twice with 2× SSC/0.1% SDS at room temperature for 5 min, twice with 0.5× SSC/0.1% SDS at 65°C for 15 min, blocked and developed using DIG Wash and Block Buffer Set (Roche) according to manufacturer’s directions. oBlcap Sequencing

Plasmids containing the subcloned 3-kb oBlcap fragment were sequenced at Duke University’s Automated Sequencing Facility using 0.5 µg DNA and 1.67 mM primer. Direct sequencing of BACs was performed similarly using 2 µg DNA. Primer sequences are available upon request.

5′ Rapid Amplification of cDNA Ends and Reverse Transcriptase–Polymerase Chain Reaction

The 5′ rapid amplification of cDNA ends (RACE) was conducted using 5′/3′ RACE Kit (Roche) according
to manufacturer’s directions. The cDNA synthesis primer was 5’-GAG GAT CCT AAG TCC CCA CTA TC-3’. RACE products were obtained using seminested PCR with Platinum Taq DNA Polymerase (Invitrogen). Round 1: oligo dT anchor primer (provided) and 5’-TTT CCC CAA CAG CTG TAA CAG-3’. PCR amplification conditions were 63°C annealing, 1 min extension, and 35 cycles. Round 2: 5’-AAG GCT GCC AGG AAG ACC-3’ and oligo dT anchor primer; PCR amplification conditions were 60°C annealing, 1 min extension, and 35 cycles. PCR products were purified using the High Pure PCR Product Purification Kit (Roche), ligated, and transformed with pGem T-Easy Vector System II (Promega, Madison, Wis.). Clones were prepped using Qiagen mini-prep, Purification Kit (Roche), ligated, and transformed with pGem T-Easy Vector System II (Promega, Madison, Ohio.). The sequences obtained were confirmed using reverse transcriptase (RT)-PCR and the following primer sets: (RE2) 5’-CGC TCC AGG AGG AAG AAC CTC-3’, with (FE2) 5’-TTC AGA ATA CGG GAC AAC TTG-3’ (positive control), (FI) 5’-TGT GGT ACC AAT ATG TGA CAC C-3’ (intron), (F1E1) 5’-TGT CCA ACA CCC TCC TCA TC-3’ (exon 1), and (F2E1) 5’-GCC TTC AGC CTT CCA GAC-3’ (exon 1); PCR amplification conditions were 60°C annealing, 1 min extension, and 35 cycles.

RNA Isolation

Opossum liver tissue was homogenized and incubated in 1 ml RNA Stat-60 (TEL-TEST, Friendswood, Tex.) for 15 min at room temperature. RNA was extracted using 200 μl of chloroform and precipitated with 500 μl of isopropanol. The pellets were washed with 1 ml of 75% ethanol and dissolved in appropriate volumes of water.

### Table 1

<table>
<thead>
<tr>
<th>Species</th>
<th>CDS (%)</th>
<th>Protein (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blacap</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chimpanzee</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Mouse</td>
<td>95</td>
<td>100</td>
</tr>
<tr>
<td>Rat</td>
<td>94</td>
<td>100</td>
</tr>
<tr>
<td>Dog</td>
<td>96</td>
<td>100</td>
</tr>
<tr>
<td>Pig</td>
<td>99</td>
<td>100</td>
</tr>
<tr>
<td>Cow</td>
<td>98</td>
<td>100</td>
</tr>
<tr>
<td>Opossum</td>
<td>84</td>
<td>94</td>
</tr>
<tr>
<td>Zebraslah</td>
<td>84</td>
<td>94</td>
</tr>
<tr>
<td>Drosophilia</td>
<td>58</td>
<td>64</td>
</tr>
<tr>
<td>Caenorhabditis elegans</td>
<td>51</td>
<td>54</td>
</tr>
</tbody>
</table>

*Reference sequence for all species except opossum was compiled from multiple GenBank entries to ensure correct sequence; sequence for opossum (Didelphis virginiana) was obtained in our laboratory as described herein. Homologies were calculated based on number of identities relative to human Blacap coding sequence (264 nt) and protein (87 aa) and human Nnat coding sequence (246 nt) and protein (81 aa).*

Bioinformatics

The following Web sites were used for data analyses: Blast (www.ncbi.nlm.nih.gov/BLAST), VISTA (genomes.tbi.ox.ac.uk/vista/index.shtml), RepeatMasker (www.repeatmasker.org), WebGene (www.itb.cnr.it/sun/webgene), Grail (compbio.ornl.gov/grailexp), and Consite (mordor.cgb.ksu.edu/cgi-bin/CONSITE/consite).

### Results

**Opossum Blacap Isolation**

The opossum Nnat/Blacap domain was obtained by probing our genomic opossum BAC library using the highly conserved coding sequence of mouse Blacap (table 1). We isolated two BAC clones containing oBlacap. To guard against sequence resulting from a clone containing deletions or rearrangements, both BAC clones were sequenced simultaneously to obtain genomic oBlacap sequence.

Interestingly, after obtaining 14.5 kb of sequence that included oBlacap, no sequence homologous to Nnat was identified (GenBank accession number AF166161). To confirm this finding, we conducted 5’ RACE on opossum liver RNA to obtain full-length oBlacap mRNA (figs. 1 and 2). The resulting RACE product revealed that only 2 kb of intronic sequence resides between exon 1 and exon 2 of

---

**Fig. 1.** Opossum Blacap RACE RT-PCR. (A) Genomic structure of opossum Blacap. Open boxes denote exons, long arrow indicates direction of transcription, and small arrows indicate positions of primers for RT-PCR. (B) oBlcap RACE RT-PCRs. RACE products were confirmed by RT-PCR. All PCRs were conducted with one primer residing within the known exon 2 of oBlacap (RE2) and with another 5’ primer. Lanes 1–4 contain RT, lanes 5–8 do not contain reverse transcriptase, and lane 9 is a negative PCR control. Lanes 1 and 5 show a positive control using a second primer within exon 2 (FE2—250 bp); lanes 2 and 6 show a negative control using a primer within the intronic region of oBlacap (FI); lanes 2 and 6 show a primer within exon 1 of oBlacap (F1E1—300 bp), and lanes 4 and 8 show a second, distinct primer within exon 1 (F2E1—450 bp). All band sizes are those expected for cDNA products, and no contaminating DNA is present as shown by the absence of product in all negative controls.
**Nnat—A Eutherian-Specific Gene**

Because Nnat was not present within the oBcap intron, we determined if Nnat was positioned elsewhere in the opossum genome or if this gene arose in a lineage specific to eutherian mammals. To resolve this issue, we first performed a Southern blot analysis of the opossum BACs containing oBcap to confirm that Nnat was not located at a distant site on the BACs, either endogenously or from a BAC-specific rearrangement. Southern blot analyses indicated that Nnat did not exist in the two opossum BAC clones containing oBcap (data not shown). We then hybridized our opossum genomic BAC library with a mouse Nnat probe. No positive clones were detected, indicating that Nnat does not exist within the opossum genome (data not shown).

Finally, we performed a Southern blot analysis using DNA from a number of species in the three major mammalian lineages—eutherians (human, lemur, mouse, sheep, and elephant), marsupials (monodelphis opossum, didelphis opossum, and wallaby), and monotremes (echidna and platypus); outlier groups used were chicken and yeast. The eutherian group included species from three of the four major mammalian clades: Euarchontoglires, Laurasiatheria, and Afrotheria. Nnat-specific bands were identified in human, lemur, mouse, sheep, and elephant, but bands were not detected in any of the marsupial or monotreme species (fig. 3). To check the integrity of the DNA, we also hybridized our blots with a probe for opossum Igf2r. Using this probe we detected bands in all three of the marsupial species examined (data not shown), demonstrating that our DNA was intact. For the Nnat Southern blot analysis, available genome sequence for mouse and human correlates with the observed band sizes for these two species. The lack of Nnat in any species outside placental mammals, and the presence of Nnat in the elephant, a member of the most ancestral eutherian superordinal clade (Afrotheria), indicates that Nnat originated in eutherians after their divergence from marsupials.

**Blast Searches**

To further our search for Nnat and its origins, we performed several Blast searches. We performed a Blast search of Nnat full-length mouse mRNA and protein against the nonredundant, EST, month, chromosome, and all species-specific databases presently available. Nnat ESTs were identified in human, mouse, rat, hamster, chimpanzee, cow, pig, and dog. No Nnat ESTs were identified from mRNA or protein searches against fungi, eukaryotes, chicken, frog, drosophila, malaria, nematode, plant, viral, sea urchin, zebrafish, or monodelphis opossum databases. These results support our finding that Nnat is indeed a eutherian-specific
gene. Moreover, these Blast searches revealed no other gene homologous to *Nnat*, and it has not been classified into any known family.

**VISTA Analysis**

In order to identify *cis*-acting sequences that may be important for the imprinting of *Nnat*, we performed a VISTA analysis of the *Nnat*/*Blcap* domain using available sequence for all species known to be imprinted at this locus: human, mouse, and rat, all of which belong to the Euarchontoglires mammalian clade. Didelphis opossum was used as an outlier for comparison because this species represents the closest relation to the Euarchontoglires for which *Nnat* is definitively not imprinted. This examination revealed a highly conserved syntenic region between exon 1 and exon 2 of *Blcap* in human, mouse, and rat (fig. 4A). Conversely, the only sequence to show homology between human and opossum is the coding sequence and the very 3' untranslated region of *Blcap*. Because *Nnat* is absent in the opossum and imprinted in every eutherian mammal investigated, these highly conserved regions may

---

Fig. 4.—Multispecies comparison of *Blcap/Nnat* domain. (A) VISTA analysis of human (GenBank, AL109614), mouse (GenBank, AL672259), rat (GenBank, NW_047659), and didelphis opossum (GenBank, AY821914). Genomic sequences were compared using VISTA to determine evolutionary conserved regions among all four species using human as the reference sequence. Exons are shown in blue and are conserved, conserved nonexonic regions are shown in pink. We used 75% homology and a 50-bp sliding window as our parameters for determining conservation. This region shows high homology in human, mouse, and rat with virtually no homology between human and opossum. The position of the small L2 repeat potentially responsible for transposition of *Nnat* is displayed as a small red bar at the 3' end of *Nnat* exon 3. Yellow bars indicate locations of tandem repeats. Dashed box represents the region enlarged in figure 4B. (B) Conserved region of interest in human, mouse, and rat that lies immediately 5' to *Blcap* exon 2 (blue box). It contains a CAGA simple repeat followed immediately by an A residue simple repeat (yellow box), a putative SUH-binding site (teal box), and a putative CTCF-binding site (purple box). The human sequence of the putative SUH- and CTCF-binding sites are shown (top lines) in comparison with their relative consensus binding sites (bottom line); mismatches are shown in red.
be involved in transcriptional rather than imprinting regulation of Nnat. Nevertheless, some interesting characteristics worthy of note do appear from our phylogenetic comparative analysis.

Firstly, VISTA plots show the presence of several repeats within the Bicap domain. A tandem repeat exists at both the 5' and 3' ends of the Bicap intron. These repeats abut the Bicap exons in human, mouse, and rat, but are absent in opossum. The VISTA plot also shows the presence of two L2 repeats within the Nnat gene. The largest L2 repeat resides within the first intron of Nnat, but other bioinformatics programs (RepeatMasker, WebGene, Grail) failed to recognize this sequence as a repeat. This region is highly A rich, a characteristic often observed for matrix attachment regions (MARs) (Cockerill and Garrard 1986). The second L2 is located at the 3' end of Nnat and is conserved in human, mouse, and rat.

Interestingly, a Blast search identified a mouse Nnat pseudogene on chromosome 7. The 6-kb band in the mouse lane of our Southern blot analysis correlates with the expected band size for this pseudogene. The Nnat pseudogene contains only exons 1 and 3 and lacks both introns and an open reading frame; however, the small L2 repeat was conserved. This finding indicated that the Nnat gene may derive from a retroelement and that the small L2 repeat may be responsible for the transposition of Nnat into the Bicap locus in an eutherian ancestor.

Finally, the conserved noncoding sequences delineated by VISTA were further screened for the presence of known DNA-binding motifs (Consite). Most of the binding sites identified were for transcription factors necessary for the regulation of gene expression. Nevertheless, we did detect a binding site for the human homolog of Suppressor of Hairless (SUH) in the region immediately 5' to Bicap exon 2 and adjacent to one of the tandem repeats previously mentioned (fig. 4B).

This region was also analyzed for the presence of putative CTCF-binding sites. Although CTCF binds to a wide variety of DNA sequences (Mukhopadhyay et al. 2004), a consensus imprinted gene-binding motif for CTCF has been identified from studies of the Igf2/H19 and Dlk1/Gtl2 JRIG domains (Wylie et al. 2000). We utilized this specific consensus sequence to screen for putative CTCF-binding sites in the Bicap/Nnat domain. Interestingly, we found the presence of one possible CTCF-binding site that is conserved in human, mouse, and rat, but is absent in the opossum. This putative binding site differs from the previous consensus site by a single residue and resides in the same location as the putative SUH-binding site and tandem repeat immediately upstream of Bicap exon 2 (fig. 4B).

Discussion

Comparative phylogenetics not only provides a powerful tool to identify crucial imprint regulatory sequences, but this technique can also offer insight into the process of imprinting evolution. The Nnat/Bicap micro-imprinted domain presents a unique region for study because the imprinted gene Nnat must maintain an expression pattern that is distinct from that of its close, nonimprinted neighbor, Bicap. We provide evidence herein by Southern blot and genomic sequence analysis that Nnat arose in a lineage specific to placental mammals, providing the first evidence of an eutherian-specific imprinted gene.

Comparative phylogenomic analyses of the Bicap/Nnat region allows for the visualization of the Bicap intronic expansion that accompanied the apparent transposition of Nnat into a common ancestor to eutherians. We have identified a small L2 repeat that may be responsible for the transposition of Nnat because this repeat is conserved in the human, mouse, and rat Nnat gene, as well as in the mouse Nnat pseudogene on chromosome 7. Similarly, U2af1-rs1 is imprinted in mouse and sits within intron 1 of the biallelically expressed Murf1 (Nabetani et al. 1997). This gene has also been proposed to have arisen through retrotransposition. Therefore, genes that transpose into the genome may trigger imprinting either through the transposition event itself or through the presence of sequence common to both transposition and imprinting regulation. Further evidence for the linkage between transposition and epigenetic controls comes from plants and mouse cells in which altered methylation levels enhance transposition (Miura et al. 2001; Ros and Kunze 2001; Singer, Yordan, and Martienssen 2001; Yusa, Takeda, and Horie 2004).

Where Nnat transposed from initially is still unknown. Genetic theory states that all genes arose from a rearrangement, deletion, mutation, or duplication of genes within a single ancestral genome (Cooper 1999). The intronless U2af1-rs1 most likely derived from transposition of its mouse parologue U2af1-rs2. While Nnat has limited structural similarity to the proteolipid class of proteins (Dou and Joseph 1996), Blast searches with Nnat mRNA, exons, and protein did not reveal sequence homologies to any other gene. Thus, we found no evidence for an ancestral origin of this gene.

The eutherian-specific lineage and imprinting of Nnat is particularly interesting because expression profiles indicate that this gene plays an important role in brain development. In addition, mouse embryos carrying paternal duplications of the region containing Nnat (i.e., two functional copies of Nnat) have a notable increase in the depth of the cerebellar folds (Kikyo et al. 1997). Conversely, a reduction of cerebellar fold depth is present in embryos with a maternal duplication of this area (i.e., no functional copies of Nnat) (Kikyo et al. 1997). Therefore, it is tempting to speculate that the transposition of this small imprinted gene provided early eutherian mammals with a selective advantage in responding to cognitive pressures.

Moreover, our finding of Nnat as an eutherian-specific imprinted gene sheds new light on the evolution of imprinting. Previous studies demonstrated that Igf2r and Igf2 imprinting is present in therian mammals but not in monotremes (Killian et al. 2000; O’Neill et al. 2000; Killian et al. 2001). These findings suggest that genomic imprinting originated approximately 180 MYA in an ancestor common to both marsupials and eutherian mammals sometime after their divergence from monotremes (Penny et al. 1999; Hasegawa, Thorne, and Kishino 2003; Woodburne, Rich, and Springer 2003). Because Nnat does not exist in the marsupial lineage, the phenomenon of imprinting had already evolved by the time Nnat transposed into the eutherian lineage and became imprinted. Thus, our studies demonstrate
for the first time that not all imprinted genes evolved at a single evolutionary time point because $\textit{Nnat}$, and possibly other genes, acquired this unique form of regulation after the initial development of imprinting.

Our findings also lend weight to an evolutionary hypo-
thesis of Kaneko-Ishino, Kohda, and Ishino (2003) that postulates that imprinted genes arose with the origination of placental tissues. Consequently, imprinting was vital to the speciation of therian mammals and therefore cannot be easily reversed. Once in place, other genes could usurp the imprinting machinery either by chance or by selective force. Interestingly, of the six genes examined within the placenta in their study, $\textit{Nnat}$ did not follow the extraembryonic expression patterns of the other five genes. Because we have shown that $\textit{Nnat}$ does not exist within the marsupial lineage, it was not present in the genome at the time of placenta
tion. Consequently, $\textit{Nnat}$ would not have been involved in the initial establishment of imprinting, but would have acquired this mechanism of gene regulation after imprinting was established within the genome of therian mammals.

Our data, however, does not preclude other theories of imprinting evolution including the widely discussed kinship theory (Wilkins and Haig 2003). This theory states that imprinting of genes evolved due to a parental conflict between the genomes to control the extraction of resources from the mother by her offspring. Moreover, an imprinted gene’s expression will confer a fitness benefit to a parental lineage. For example, genes expressed from the paternal allele often encourage fetal growth, while maternally expressed genes often suppress growth in an effort to con
derve resources for the mother (Wilkins and Haig 2003). Accordingly, a role for $\textit{Nnat}$ in metabolism has recently been demonstrated because downregulation of this im
tprinted gene decreases insulin production in response to glucose stimulation (Chu and Tsai 2005). In addition, embryos carrying a maternal duplication of the region contain
ing $\textit{Nnat}$ were 60%–80% smaller than embryos with normal $\textit{Nnat}$ genotypes (Kikyo et al. 1997). Alternatively, $\textit{Nnat}$ may provide a benefit to the patriarchal lineage by ensuring appropriate nurturing behavior in offspring as seen for other imprinted genes such as $\textit{Peg1}$ and $\textit{Peg3}$ (Lefeuvre et al. 1998, Li et al. 1999).

The absence of $\textit{Nnat}$ in marsupials also supports the postulate that promoter hypermethylation represents the pri
mordial imprint mark (Ferguson-Smith and Surani 2001; Reik, Dean, and Walter 2001). Because $\textit{Nnat}$ has clearly had less time to develop the layers of transcriptional control observed in genes known to be imprinted more ancestrally such as $\textit{Igf2r}$, $\textit{Igf2}$, and $\textit{Peg1}/\textit{Mest}$ (Killian et al. 2000; O’Neill et al. 2000; Killian et al. 2001; Suzuki et al. 2005), imprinting at this locus may be mostly or even entirely regulated by the differential promoter methylation identified in eutherian mammals (Kikyo et al. 1997; Evans et al. 2001). Given enough evolutionary time, genes that utilize simple promoter hypermethylation may acquire additional levels of control and achieve the same degree of complexity as imprinted genes that arose earlier in the course of imprinting evolution, a hypothesis that will be further tested as more mammalian genomes are sequenced and compared.

Finally, our phylogenetic analysis of the $\textit{Bcap}/\textit{Nnat}$ micro-imprinted domain reveals several intriguing, con
served structures including a possible MAR, putative binding sites for $\textit{CTCF}$ and $\textit{SUH}$, and a tandem repeat both 5’ and 3’ of $\textit{Nnat}$. With the exception of $\textit{SUH}$, all of these factors have been implicated in establishing and/or maintaining imprinting (Neumann, Kubicka, and Barlow 1995; Greally et al. 1999; Bell and Felsenfeld 2000; Yoon et al. 2002). Interestingly, although we cannot define the boundaries of the ancestral $\textit{Bcap}$ intron versus that of eutherian mammals, neither tandem repeat is conserved in the opossum despite its virtual juxtaposition to the $\textit{Bcap}$ exons in eutherians. Furthermore, the conserved noncoding region containing one of these repeats also contains putative binding sites for the human homolog of $\textit{SUH}$ and $\textit{CTCF}$. While $\textit{SUH}$ has not been implicated previously in imprinting, this finding is nonetheless intriguing because $\textit{SUH}$ acts as a transcriptional repressor when complexed with other factors such as histone deacetylase to form a chromatin remodeling complex (Hsieh et al. 1999). It is conceivable that such a complex would play a role in the imprinting process as histone deacetylases are known to be critical components of the imprinting machinery (Nan et al. 1998).

Our phylogenetic examination of the $\textit{Nnat}/\textit{Bcap}$ domain has uncovered several genomic motifs whose role in the imprint regulation of $\textit{Nnat}$ would be interesting to determine experimentally. Moreover, even though $\textit{Nnat}$ is highly conserved in eutherian mammals and imprinted in mouse and human, it does not exist in species outside the eutherian lineage. These findings demonstrate that after imprinting evolved in an initial subset of genes in viviparous mammals during the Jurassic era (Killian et al. 2000; O’Neill et al. 2000; Killian et al. 2001), other genes continued to acquire imprinting, at least in eutherians. Indeed, approximately a 70 Myr span exists between the divergence of the Therian and Eutherian base species (Hasegawa, Thorne, and Kishino 2003; Woodburne, Rich, and Springer 2003), providing ample time for the formation of eutherian-specific imprinted genes such as $\textit{Nnat}$. Continuing phy
genetic analyses of imprinted domains may reveal additional examples of the diversity of imprinting between species and thereby enhance our understanding of the mechanisms of imprinting, the history of imprint evolution, and the role these genes play in speciation and brain development.

Acknowledgments

We wish to thank individuals and institutions that aided our efforts in collecting tissues for this study, includ
ing Guy Liechty and the North Carolina Zoo. We are also grateful to Scott Langdon and the Duke Sequencing Facility for their assistance in obtaining sequence described in this work. This study was supported by National Institutes of Health grants CA25951, ES08823, and ES13053.

Literature Cited

Bell, A. C., and G. Felsenfeld. 2000. Methylation of a CTCF


Mark Springer, Associate Editor

Accepted May 13, 2005