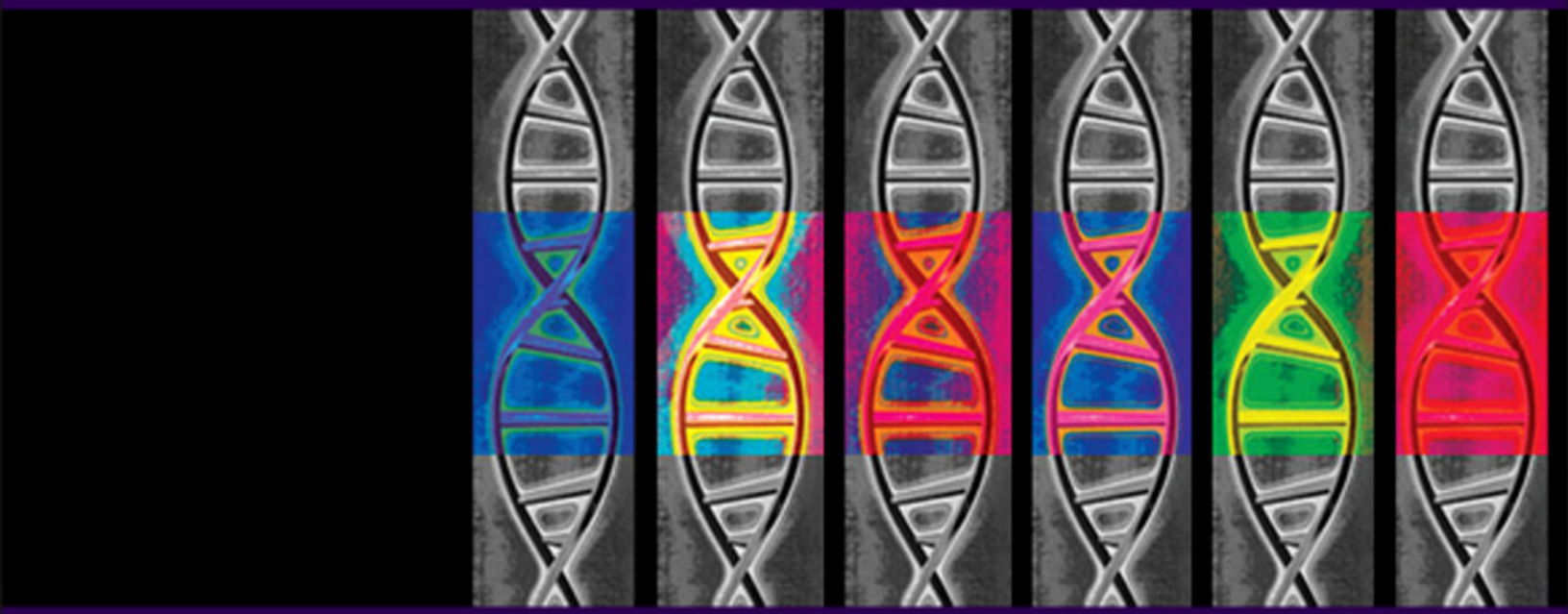
HANDBOOK OF EPIGENETICS

The New Molecular and Medical Genetics



Edited by

Trygve Tollefsbol

Handbook of Epigenetics

The New Molecular and Medical Genetics

Edited by

Trygve Tollefsbol

Department of Biology
University of Alabama at Birmingham
Birmingham, AL, 35294-1170, USA

Amsterdam • Boston • Heidelberg • London • New York • Oxford
Paris • San Diego • San Francisco • Singapore • Sydney • Tokyo

Academic Press is an imprint of Elsevier

PREFACE

Epigenetics is considered by many to be the "new genetics" because many biological processes are controlled not through gene mutations, but rather through reversible and heritable epigenetic phenomena ranging from DNA methylation to histone modifications to prions. Epigenetic processes occur in diverse organisms and control a vast array of biological functions, such as tissue/organ regeneration, X-chromosome inactivation, stem cell differentiation, genomic imprinting, and aging. Epigenetic aberrations underlie many diseases, including cancer and disorders of the immune, endocrine, and nervous systems; clinical intervention is already in place for some of these disorders and many novel epigenetic therapies are likely on the horizon.

Handbook of Epigenetics: The New Molecular and Medical Genetics is the first comprehensive analysis of epigenetics, and summarizes recent advances in this intriguing field of study. This book will interest students and researchers in both academics and industry by illuminating the evolution of epigenetics, the epigenetic basis of normal and pathological processes, and the practical applications of epigenetics in research and therapeutics.

CHAPTER 22

Genomic Imprinting

Wendy Chao
Department of Ophthalmology, Schepens Eye Research Institute, Harvard Medical School, 20 Staniford Street, Boston, MA 02114, USA

INTRODUCTION

When eukaryotic organisms reproduce sexually, each parent contributes a haploid set of chromosomes to create diploid progeny. Two copies, or *alleles*, exist for most gene loci. According to classical Mendelian genetics, both copies are expressed, and certain variations in DNA sequence may allow the phenotype of a dominant allele to prevail over a recessive one. However, some genes carry epigenetic marks that distinguish between maternally and paternally inherited alleles. These genomic "imprints" can dramatically alter gene expression depending on parent of origin – even if the two alleles are otherwise identical.

In mammals, genomic imprinting manifests in monoallelic silencing according to parental lineage. Because a second allele may provide genetic diversity and mask undesirable traits [1], it is somewhat counterintuitive to find functionally haploid genes in complex diploid species. However, genomic imprinting arose in mammalian evolution over 150 million years ago [2], which implies that monoallelic expression is not necessarily detrimental to genetic fitness. Rather, it is imperative for several loci to maintain imprinted monoallelic expression. In humans, aberrant imprinting underlies numerous developmental and neurological disorders [reviewed in Ref. 3], and loss of imprinting is common in cancer [reviewed in Ref. 4].

Even though it is highly relevant to human development and disease, genomic imprinting went undiscovered in mammals until relatively recently. Imprinted (as opposed to random) X chromosome inactivation has been a known phenomenon since the early 1970s [5,6], but it was twenty more years before imprinted autosomal genes were discovered in mammals [7–9]. The existence of these genes was predicted by earlier nuclear transplantation experiments, that produced mouse embryos with both sets of chromosomes derived from one parent. Not only were these uniparental embryos abnormal, but gynogenetic (female-derived) and androgenetic (male-derived) embryos displayed contrasting phenotypes [10–13]. These studies demonstrated the nonequivalence of maternal and paternal genomes – even after accounting for sex chromosome differences. Subsequent complementation studies narrowed these parental effects to discrete autosomal regions [14].

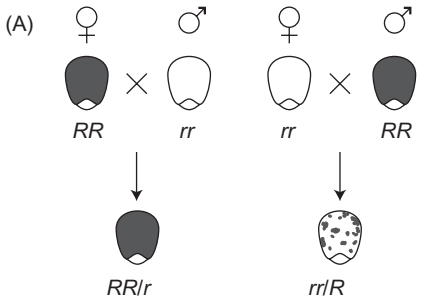
In 1991, three imprinted genes in mice were characterized: insulin-like growth factor 2 receptor (*Igf2r*), which is maternally expressed [7]; its ligand, insulin-like growth factor 2 (*Igf2*), a paternally-expressed regulator of growth and development [9]; and *H19*, a maternally-expressed noncoding RNA [8] that is physically linked to *Igf2* [15] and regulated by shared elements [16,17]. These archetypes of genomic imprinting have yielded much insight into various epigenetic regulatory mechanisms. The *Igf2* gene is particularly

Handbook of Epigenetics: The New Molecular and Medical Genetics. DOI: 10.1016/B978-0-12-375709-8.00022-8

interesting because its complex regulation involves both the *H19* gene at the transcriptional level and the Igf2r protein at the post-translational level [reviewed in Ref. 18]. *Igf2* is also highly conserved among vertebrates [2,19], but its imprinting status is not. The divergence of *Igf2* imprinting in the phylogenetic tree has fueled many theories on vertebrate evolution and the origin of genomic imprinting.

Genomic imprinting is often described as an exclusively mammalian phenomenon, yet parental effects on gene expression were documented in insects and plants long before the discovery of imprinted mammalian genes. The term "imprint" was actually used as early as 1960 to describe epigenetic parental effects in fungus gnats of the genus *Sciara*. At various stages of sciarid development, certain paternally-derived chromosomes are heterochromatized and eliminated from cells independently of genomic constitution and "determined only by the sex of the germ line through which the chromosome has been inherited" [20]. Allele-specific silencing in *Drosophila* was recorded in the mid-1930s, when vague reports noted the preferential silencing of the X-linked *scute-8* gene when paternally inherited [21,22]. The earliest (and perhaps most extreme) example of genomic imprinting in insects can be traced to a 1931 report, which described sex determination in the family Pseudococcidae [23]. Coccids (commonly known as mealybugs) represent a striking example of haplodiploidy, a system of sex determination commonly employed by insects, in which females are diploid but males are haploid [reviewed in Ref. 24]. In males, all paternally derived chromosomes are either silenced by heterochromatin or completely eliminated; thus, all male coccids are functionally haploid [reviewed in Ref. 25].

Epigenetic parent-specific effects were demonstrated even earlier in plants. In 1918 and 1919, two independent studies demonstrated parent-specific effects at the maize *R* locus, which controls anthocyanin pigment expression in the aleurone endosperm [26,27]. When the female gamete transmits the dominant *R* allele in *RR* (pigmented) × *rr* (colorless) crosses, the aleurone seed covering is solidly pigmented; conversely, if *R* originates from the paternal (pollen) parent in a reciprocal cross, the endosperm is lightly pigmented, and mottled or spotted in appearance (Fig. 22.1A). Although the endosperm of flowering plants is usually



(B)

Solid (M/P)	Mottled (M/P)
RR/*r*	*rr*/*R*
RR/*rr*	*rr*/*RR*
RRRR/*r*	*rrrr*/*R*
–R/*r*	*–*/*R*

FIGURE 22.1

Influence of parental lineage on the *R*-mottled phenotype in maize *RR* (pigmented) × *rr* (colorless) crosses. (A) When *R* is transmitted by the maternal gamete (left), the *RR*/*r* triploid aleurone of the progeny is solidly pigmented. In the reciprocal cross (right), the resulting *rr*/*R* aleurone is lightly pigmented and mottled [26,27]. (B) Maize genotypes that yield solid or mottled phenotypes. M and P designate maternal and paternal origin, respectively. The solid phenotype occurs only when *R* is maternally derived, irrespective of gene dosage. A paternally transmitted *R* allele yields a mottled phenotype if the female gamete is homozygous for the *r* allele, and even when the maternal gamete is deficient (–) for the *R* region [28].

polyploid, later experiments [28] confirmed that the mottling effect is indeed dependent on parental origin rather than differences in gene dosage (Fig. 22.1B).

Definitions of "genomic imprinting" do not always include parent-of-origin effects in insects and plants, but are instead limited to those observed in mammals. However, syntactic differences aside, these processes use conserved regulatory mechanisms to achieve a common purpose: to epigenetically distinguish maternal and paternal genomes. Though genomic imprinting may not be a uniquely mammalian phenomenon, its discovery in mice did uncover an epigenetic basis of human disease, and catalyzed a field of research devoted to parent-specific gene expression. As with many genetic processes, the epigenetic phenomena in other organisms have helped elucidate the mechanisms of mammalian imprinting and its evolutionary origin.

MECHANISMS OF GENOMIC IMPRINTING

What distinguishes an imprinted gene from its non-imprinted counterparts, and destines it for allele-specific expression? This question has puzzled scientists for decades, and is now beginning to be understood. Genomic imprinting is currently known to involve numerous epigenetic processes – many of which are conserved among diverse species. While its exact catalyst is not entirely clear, the primary imprint in many organisms involves the classic epigenetic mark: DNA methylation. This primary mark is propagated by *cis* and *trans* factors that trigger additional modifications and culminate in allele-specific gene expression.

DNA Methylation

Many organisms – from primitive bacteria to complex eukaryotes – use methyl groups to distinguish DNAs of different origins [reviewed in Ref. 29]. In higher eukaryotes, methylation primarily occurs on the cytosine residue of CpG dinucleotides, which tend to cluster around promoter regions as "CpG islands" but appear sparsely in the rest of the genome due to spontaneous deamination of 5-methylcytosine to thymine [reviewed in Ref. 30]. In the mammalian genome, much of the DNA methylation targets transposable elements, which illustrates a role in host defense mechanisms that silence invasive DNAs [31]. Methylation is also a common denominator of differential DNA regulation; before the first imprinted genes were discovered in mammals, studies demonstrated that transgenes could acquire allele-specific methylation patterns depending on the transmitting parent [32,33]. Differential methylation was thus identified as a heritable epigenetic feature that distinguishes maternal and paternal alleles – and a central mechanism in genomic imprinting.

IMPRINT ESTABLISHMENT AND MAINTENANCE IN MAMMALS

Mammalian imprinted genes are often (if not always) situated near differentially-methylated regions (DMRs), also known as differentially-methylated domains (DMDs), which are believed to be the primary targets of epigenetic modifications. DMRs may in turn direct other *cis* and *trans* elements to achieve stable allele-specific gene expression [reviewed in Ref. 34]. Thus, DMRs may serve as imprinting control regions (ICRs), also known as imprinting control elements (ICEs) or imprinting centers (ICs). Aside from the high frequency of CpG dinucleotides, DMRs share little sequence homology; instead, they are characterized by tandemly repeated elements [35–37]. These repetitive structures are believed to trigger *de novo* differential methylation [35–37], similarly to how retrotransposon-derived repetitive sequences – such as short interspersed nuclear elements (SINEs) and CpG-rich Alu repeats – may acquire germline-specific differential methylation [38,39]. Primary imprints occur in the germline, where the prospective parent's existing imprints can be erased and reestablished in the haploid gametes. This occurs through global demethylation in germ cells, followed by differential methylation by the *de novo* DNA methyltransferase Dnmt3a [40] and its cofactor Dnmt3L [41].

It is not entirely clear how the *de novo* methyltransferases differentiate between maternal and paternal DMRs in the germline, though this seems to be partially determined by DMR location. Maternally-methylated DMRs coincide with transcription units, whereas the few known paternally-methylated DMRs occur within intergenic regions (Table 22.1). Paternal-specific germline methylation appears to target tandem repeat sequences, as evidenced by the *H19* [42] and *Rasgrf1* [43] loci, which contain two of the three known paternal germline DMRs. On the other hand, the act of transcription may dictate maternal-specific methylation in oocytes, as demonstrated by the *Gnas*/*Nesp* locus (Fig. 22.2). This complex imprinted domain [reviewed in Ref. 44] includes *Gnas*, which exhibits maternal-specific expression in some tissues. *Gnas* encodes the highly conserved signaling protein Gsaα; alternative promoters give rise to two paternal-specific transcripts, *Gnasxl* (which encodes the variant protein XLas) and the noncoding *1A*. The upstream *Nesp* locus encodes the Nesp55 protein, which is involved in the secretory pathway. This maternally expressed *Nesp* transcript also appears to have a major functional role in imprinting the entire locus. There are two DMRs in this domain: one encompassing the promoters for the paternally-expressed *Gnasxl* and the noncoding *Nespas*, and another at the *1A* promoter. Truncating *Nesp* transcription disrupts methylation of both DMRs in the female germline, suggesting that the act of transcription facilitates *de novo* methylation [45]. This model is supported by the fact that all known maternal DMRs occur in transcribed regions – either in introns or near promoters that are downstream of alternate transcription start sites (Table 22.1). It is hypothesized that oocyte-specific transcription facilitates germline DMR methylation by favorably altering chromatin structure; alternatively, the RNA itself might recruit *de novo* methyltransferases or other *trans* regulatory factors that promote germline methylation [45].

Allele-specific methylation also involves germline-specific timing of Dnmt3L expression. This protein lacks *in vitro* methyltransferase activity, but is required for *de novo* methylation by Dnmt3a [40,41]. Additionally, alternate germline-specific promoters lead to differential Dnmt3L expression in oocytes and spermatocytes [46]. Oocytes express Dnmt3L for only

TABLE 22.1 Known Germline DMRs and Their Locations

DMR/location	Locus		
Maternal/intron	*Gnas (1A)*	*Inpp5f*	*Peg3*
	Gnas (Nespas/Gnasxl)	*Kcnq1 (KvDMR)*	*Peg13*
	Grb10	*Mcts2*	*Snrpn*
	Igf2r (Air)	*Nap1l5*	*U2af1-rs1*
	Impact	*Peg1*	*Zac1*
Maternal/promoter	*Peg10*	*Slc38a4*	
Paternal/intergenic	*Dlk1-Gtl2 (IG-DMR)*	*H19*	*Rasgrf1*

Adapted from Ref. 45

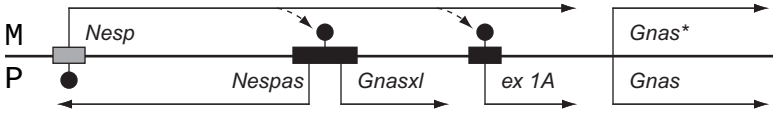


FIGURE 22.2

The *Gnas* imprinted domain (not drawn to scale), including the protein-coding transcripts (*Nesp*, *Gnasxl*, and *Gnas*) and noncoding transcripts (*Nespas* and *1A*). Maternal (M) transcripts are indicated by arrows above the line, while paternal (P) transcripts are below the line. Arrowheads indicate the direction of transcription. *Gnas* shows maternal-specific expression in some tissues, as indicated by the asterisk (*). Two germline DMRs (black boxes) acquire maternal methylation imprints (black circles) in oocytes. *Nesp* transcription is believed to facilitate germline methylation, as indicated by dashed arrows. A somatic DMR (gray box), which covers the *Nesp* promoter, is methylated on the paternal allele after fertilization. Adapted from Ref.45.

a few days before ovulation, and primary methylation imprints are established during this short time frame [41]. On the other hand, Dnmt3L expression begins prenatally in embryonic prospermatogonia, and continues until a few days after birth; paternal-specific DMR methylation then persists in the male germline well into adulthood [47]. Because methylated cytosine residues spontaneously deaminate to thymine, the prolonged methylation time may explain why paternal DMRs tend to have far fewer CpGs than their maternal counterparts; moreover, this gradual sequence degeneration may explain why so few paternal DMRs have been identified [reviewed in Ref. 48]. Coincidentally, retrotransposon silencing depends heavily on Dnmt3L in male germ cells [49] but to a lesser extent in the female germline [41,50]. Since paternal-specific DMR methylation requires tandemly-repeated elements [42,43], the primary paternal imprint likely targets the retrotransposon-like nature of DMRs.

After the maternal and paternal genomes join during fertilization, the primary germline imprints persist while the rest of the zygotic genome is demethylated [reviewed in Ref. 51]. It is not entirely clear how these primary parental imprints survive this early embryonic demethylation; however, methyl-CpG-binding proteins (MBDs), which regulate transcription by binding methylated DNA and recruiting additional silencing factors [52], are required to maintain differential methylation of imprinted mouse genes in somatic cells [53]. The murine maintenance methyltransferase Dnmt1 is also pivotal for preserving parental methylation patterns during zygotic demethylation and subsequent somatic cell divisions [54]. PGC7/Stella and ZFP57 are additional factors that appear to protect imprinted DMRs from zygotic demethylation. PCG7/Stella is a nuclear protein that is highly expressed in both male and female primordial germ cells (PGCs); maternal-specific expression of PCG7/Stella continues in the early embryo, and is essential for maintaining methylation at maternal DMRs [55]. ZFP57 is a Kruppel-associated box (KRAB) zinc finger protein that is expressed in oocytes and in certain somatic tissues; it is required for maternal imprint establishment (and for both maternal and paternal imprint maintenance) at some DMRs [56].

Many imprinted genes occur in clusters in the genome, which can be several megabases (Mb) in length and contain multiple differentially-expressed genes under the control of one or two DMRs. Such is the case with an approximately 3-Mb region on human chromosome 15q11-13 that is implicated in Prader–Willi syndrome (PWS) and Angelman syndrome (AS) [reviewed in Ref. 57]. At least five paternally expressed genes (*MKRN3*, *MAGEL2*, *NDN*, *SNURF*, and *SNRPN*) and two maternally expressed genes (*UBE3A* and *ATP-10A*) lie in this region (Fig. 22.3). On the homologous region on mouse chromosome 7c, *Atp10a* is not imprinted [58,59]; however, an additional paternally-expressed gene, *Peg12/Frat3*, lies distal to the *Mkrn3/Magel2/Ndn* cluster [60,61]. A bipartite ICR for this region was defined by observed microdeletions in PWS and AS patients; it encompasses the maternally-methylated *SNURF/SNRPN* promoter and a region 35 kb upstream that also has ICR function. Deletions in this upstream region (known as AS-IC) cause AS when maternally inherited, whereas

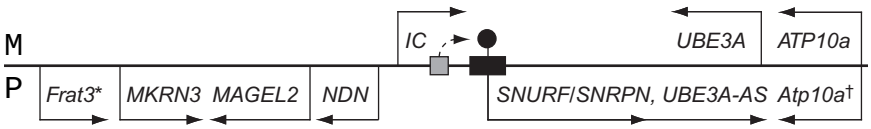


FIGURE 22.3

The PWS/AS imprinted domain on human 15q11-13 (not drawn to scale). Maternal transcripts are above the line; paternal transcripts are below the line. An additional paternal-specific gene, *Peg12/Frat3*, lies distal to *Mkrn3* on mouse chromosome 7c (*); however, murine *Atp10a* is not imprinted (†). The ICR near the *SNURF/SNRPN* promoter (PWS-IC, black box) is methylated in the female germline (black circle) and implicated in PWS. The upstream ICR (AS-IC, gray box) is implicated in Angelman syndrome, and facilitates germline methylation of PWS-IC (indicated by the curved dashed arrow). *IC* transcripts that arise from oocyte-specific alternate promoters may facilitate methylation of PWC-IC [64].

deletions in the *SNURF/SNRPN* promoter (called PWS-IC) cause PWS when paternally inherited [reviewed in Ref. 62]. The upstream ICR is believed to help establish the primary maternal imprint on PWS-IC [reviewed in Ref. 63]. In mouse and human oocytes, alternative promoters upstream of AS-IC give rise to maternal-specific *IC* transcripts that may facilitate PWS-IC methylation [64]. This is consistent with the hypothesis that maternal germline imprints are dependent on transcription [45].

The 5′ DMR of *H19* is one of three known paternal germline DMRs (Table 22.1) and displays persistent paternal methylation in both mice [65] and humans [66]. It was initially identified as an ICR when its deletion in mice abolished imprinting for both *Igf2* and *H19* [67]. CpG mutations that prevent methylation also disrupt imprinting in this region [68]. In humans, *IGF2* overexpression is implicated in Wilms' tumor and in the overgrowth disease Beckwith-Wiedemann syndrome (BWS) [69]. Loss of imprinting (resulting in biallelic expression) is one of several known mechanisms that cause *IGF2* overexpression [70,71]. Not surprisingly, many BWS or Wilms' tumor cases involve mutations or deletions in the ICR [72–76]. Because *IGF2* overexpression is a common denominator in carcinogenesis, it is also not surprising that ICR dysfunction has been noted in numerous cancers [reviewed in Refs 77,78].

METHYLATION AND GENOMIC IMPRINTING IN PLANTS AND INVERTEBRATES

Some plants and insects also use DNA methylation to achieve parent-specific gene expression, albeit with several distinct features. In plants, the major role of methylation appears to be in maintaining differential expression [79,80], whereas the primary imprint is established by a DNA glycosylase rather than a *de novo* methyltransferase [81]. In mealybugs, parent-specific genome silencing also involves differential methylation, though the heterochromatic paternal DNA is associated with hypomethylation rather than hypermethylation [82]. In *Drosophila*, parent-specific genomic imprints appear to involve chromatin-modifying proteins rather than DNA methylation [83,84]. It was once believed that DNA methylation does not occur at all in *Drosophila* [85,86]; however, it is now known that methylation does in fact occur – though only at early stages of development at non-CpG dinucleotides (most often CpT and CpA) [87,88]. It remains unclear whether methylation is involved in *Drosophila* imprinting, yet homologs of *Drosophila* Polycomb group (PcG) proteins appear to coordinate differential methylation of imprinted genes in mice [89] and plants [reviewed in Ref. 90]. Another model organism, *Caenorhabditis elegans*, exhibits imprinted X chromosome inactivation [91] and also imprints exogenous transgenes [92], despite the long-held notion that *C. elegans* does not genomically imprint [93]. However, these parent-specific effects in *C. elegans* apparently do not involve methylation, as attempts to detect 5-methylcytosine have failed [94]. Because the *C. elegans* genome is considerably less complex than higher eukaryotic genomes, constitutive silencing mechanisms such as DNA methylation are thought to be less critical [95].

Replication Timing

Asynchronous DNA replication is a curious hallmark of imprinted alleles and other monoallelically-expressed genes, including those on the active and inactive X chromosomes [5]; actively transcribed genes tend to replicate early, while late replication is characteristic of repressed genes and transcriptionally silent heterochromatin [reviewed in Ref. 96]. Differential methylation correlates with replication asynchrony on human chromosomal region 15q11-q13, which contains multiple imprinted genes and is associated with PWS and AS (see Fig. 22.3). This implies that replication and methylation are coordinately regulated [97]. While differential replication could simply be a consequence of genomic imprinting, there is some evidence that it may occur independently of methylation imprints and might even play a regulatory role. In some cases of aberrant human *IGF2/H19* imprinting, loss of differential methylation does not disrupt asynchronous replication [98]. Moreover,

in mouse embryonic stem (ES) cells that lack *de novo* methylation machinery (Dnmt1 and Dnmt3L), the *Igf2*/*H19* locus continues to replicate asynchronously despite loss of imprinting [99]. On the other hand, methylation imprints may also be established at the *Igf2*/*H19* locus without affecting replication timing [100]. Nonetheless, these results suggest that replication asynchrony does not necessarily occur secondary to genomic imprinting. Of note, asynchronous replication is reset during gametogenesis and maintained throughout zygotic development, which coincides temporally with imprint erasure, reestablishment, and maintenance [101]. Thus, a component of the primary epigenetic imprint may indeed involve replication timing; however, this remains uncertain.

Chromatin Modifications

Gene expression is not only affected by covalent DNA modifications such as methylation, but also by higher-order changes in chromatin structure that involve DNA-protein interactions. Conformational changes in chromatin (revealed by differential DNAseI hypersensitivity) may determine whether genes are accessible to transcription factors and other regulatory proteins [reviewed in Ref. 102]. Thus, differential modifications of key chromatin structures, such as core histones, are epigenetic events that can contribute to allele-specific gene expression.

HISTONE MODIFICATIONS

Differential histone modifications occur in many examples of genomic imprinting, including paternal genome silencing/elimination in insects [reviewed in Ref. 103] and imprinted X chromosome inactivation in mammals [104]. Modifications (generally to the lysine residues in histone N-terminal tails) may result in either transcriptional activation or silence, and often coordinate with DNA methylation status [reviewed in Ref. 105]. Acetylation is a well-known histone modification that generally associates with active transcription; deacetylation may require DNA methylation, as certain methyl-CpG binding proteins (such as MECP2) can recruit histone deacetylases that repress transcription in mice [106] and frogs [107].

Methylation not only takes place on CpG residues of DNA, but also on lysine residues of histones; these may be either transcriptionally repressing or activating, depending on lysine position and level of methylation [reviewed in Ref. 108]. Methylation on lysine 9 of histone H3 (H3K9) is required for DNA methylation in *Neurospora* [109] and *Arabidopsis* [110]; it also coincides with methylation of pericentric heterochromatin [111] and the inactive X chromosome in mammals [reviewed in Ref. 108]. Differential H3K9 methylation corresponds with imprinting in the PWS/AS domain (see Fig. 22.3). In human cells, both H3K9 methylation and CpG methylation occur at the maternal PWS-IC [112]. In mouse ES cells, deleting the gene that encodes G9a (the H3K9 methyltransferase) reduces PWS-IC methylation and disrupts imprinting at the PWS/AC locus, suggesting that H3K9 methylation regulates allele-specific ICR methylation [113].

While H3K9 methylation is associated with transcriptional repression, methylation of histone H3 lysine 4 (H3K4) is associated with transcriptional activation [reviewed in Ref. 108]. On the unmethylated and transcriptionally active paternal PWS-IC (see Fig. 22.3), H3K4 is methylated and H3K9 is unmethylated [112]. H3K4 has also emerged as an important variable in primary imprint establishment. Dnmt3L, the critical cofactor of the *de novo* DNA methylase Dnmt3a [41], functions in part by binding H3 and recruiting Dnmt3a to DNA; this complex cannot occur with methylated H3K4, which essentially prevents *de novo* CpG methylation [114]. H3K4 demethylation by the lysine demethylase KDM1B, which is highly expressed during late oogenesis, is required for *de novo* methylation of some (but not all) maternal DMRs [115]. Because different maternal DMRs acquire methylation imprints at specific stages of oocyte development [116], and only DMRs imprinted during late oogenesis are associated with H3K4 demethylation, KDM1B expression timing is likely to be a factor in maternal DMR specificity [115]. At the three known paternal DMRs

(see Table 22.1), histone methylation may also direct primary imprint establishment. In sperm, H3K4 methylation occurs specifically at unmethylated maternal DMRs; in somatic tissues, the methylated paternal DMRs coincide with H3K9 and H4K20 methylation [117]. This concurs with the DNA/histone methylation patterns of the heterochromatin in pericentric satellite repeats [111]. Thus, histone methylation appears to be a critical precursor to primary imprinting of both maternal and paternal DMRs.

POLYCOMB AND TRITHORAX GROUP PROTEINS

Polycomb group (PcG) and trithorax group (TrxG) proteins, which have reciprocal functions in maintaining chromatin stability, may also regulate differential expression at imprinted regions. Both PcG and TrxG proteins were originally identified as modifiers of *Drosophila* position effect variegation, but are now known to control gene expression in mammals and many other species [reviewed in Ref. 118]. In *Drosophila*, TrxG and Su(var) (suppressor of variegation) proteins mediate imprinting of the mini-X chromosome [119]. *Drosophila* PcG proteins are also able to recognize the murine *Igf2/H19* ICR [120], and the mouse PcG protein Eed apparently mediates histone methylation and initiation of imprinted (but not random) X chromosome inactivation [104,121]. Murine Eed may also regulate autosomal imprinted loci, where it appears to modulate differential methylation of ICRs [89] and histones [122]. PcG proteins appear to coordinately regulate DNA and histone methylation to achieve tissue-specific gene expression, as evidenced by the differential methylation patterns associated with tissue-specific *Grb10* imprinting [123]. Interestingly, genes that encode PcG proteins may themselves be imprinted; in mice, the PcG gene *Sfmbt2* has paternal-specific expression in extraembryonic tissues [124], and several PcG genes that control endosperm development are imprinted in *Arabidopsis* and maize [reviewed in Ref. 90].

Chromosomal Position Effects

When inserted into imprinted chromosomal regions, transgenes acquire allele-specific methylation patterns that are determined by the transmitting parent [32,33]. This phenomenon bears resemblance to *Drosophila* position effect variegation, in which gene expression patterns may change upon transposition to other chromosomal locations – either by juxtaposition to enhancing elements, or insertion into heterochromatic DNA [reviewed in Ref. 125]. Thus, it has been proposed that position effects, similar to those observed in *Drosophila* variegation, may occur at imprinted domains. In this model, primary imprints established at ICRs lead to secondary methylation and heterochromatization, which can in turn spread to surrounding regions and silence distant genes [126]. In *Drosophila*, classical modifiers of position-effect variegation (such as chemicals and PcG proteins) mediate paternal-specific silencing of three closely linked genes on a mini-X chromosome [84,119]. The site of the primary imprint is a *cis* regulatory ICR; its effects extend to silence a distal gene within a 1.2-Mb region, as well as the entire 1.5-Mb mini-X chromosome, which is distinguished by reduced transcription and late replication [126]. Imprinting of a novel gene in the mouse PWS/AS cluster, *Peg12/Frat3* (Fig. 22.3), is believed to be a product of analogous position effects; this gene lacks a human homolog, and was likely retrotransposed during species divergence into the mouse PWS/AS locus, where it acquired the imprinted status of the surrounding genes [60,61]. Interestingly, ICRs themselves remain faithfully imprinted when inserted into non-imprinted regions, and can imprint hybrid transgenes [127,128]. Even more remarkably, mammalian ICRs function as silencers in *Drosophila*; although imprinting is not established, this demonstrates that genomic imprinting involves highly conserved silencing mechanisms [129,130].

Chromatin Insulators

Chromatin insulators establish boundaries between different DNA regulatory domains, and demarcate transcriptionally inactive heterochromatin from euchromatin that is conducive

to transcription. Thus, insulators can protect against chromosomal positional effects and *cis* regulatory elements such as enhancing and silencing elements [reviewed in Ref. 131]. Insulators feature at several imprinted domains, most notably at the *Igf2*/*H19* locus (Fig. 22.4), which has become a paradigm of imprinted gene regulation. The ICR between *Igf2* and *H19* contains binding sites for CCCTC binding factor (CTCF), a multifunctional transcription factor; at the *Igf2*/*H19* locus, CTCF acts as a methylation-sensitive insulator protein that dictates whether a shared downstream enhancer can activate either promoter [132,133]. On the maternal chromosome, CTCF binds the ICR and forms a chromatin barrier between *Igf2* and the downstream enhancer, resulting in *Igf2* silence and *H19* expression. On the paternal chromosome, methylation of the ICR prevents CTCF binding – thus allowing the enhancer to activate the *Igf2* promoter (Fig. 22.4). This likely occurs through CTCF-mediated chromatin arrangements, that allow the enhancer to differentially activate the *Igf2* and *H19* promoters [134]. ICR methylation then leads to secondary *H19* methylation and transcriptional repression on the paternal allele [135].

Chromatin insulators have been proposed to regulate several other imprinted domains. At the *Xist*/*Tsix* locus, the identification of functional CTCF binding sites [136,137] has led to a similar model for X chromosome inactivation (Fig. 22.5). The ICR for this region exhibits differential methylation [138] and bidirectional promoter activity [139]; it is also part of a bipartite enhancer that is believed to activate *Tsix* expression on the active X chromosome [140,141]. CTCF also binds at the ICRs of several other imprinted loci [142], and is strongly correlated with imprinting in cross-species comparisons [143]. Another multifunctional

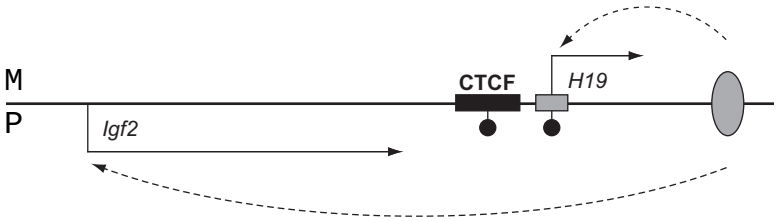


FIGURE 22.4
The insulator model of genomic imprinting at the *Igf2/H19* locus. This region (not drawn to scale) contains a downstream enhancer (gray oval) that may activate either *Igf2* or *H19* (curved dashed arrows), likely by chromatin looping events that allow the enhancer to contact either promoter [134]. The intergenic ICR (black box) is one of three known paternal germline DMRs (Table 22.1). When this DMR is paternally methylated (black circle), the enhancer activates *Igf2* expression on that allele, as indicated by the curved dashed arrow. The paternal *H19* promoter is then silenced by secondary methylation (gray box with black circle). On the maternal allele, the unmethylated DMR binds CTCF, which prevents the enhancer from accessing *Igf2*. The downstream enhancer thus activates maternal-specific *H19* expression (dashed curved arrow).

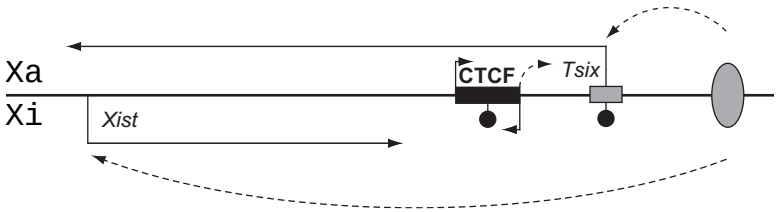


FIGURE 22.5
The *Xist/Tsix* region (not drawn to scale), which controls both random and imprinted X chromosome inactivation. The *Xist* ncRNA mediates silencing of the inactive X chromosome (Xi), and is antagonized by the *Tsix* antisense ncRNA on the active X chromosome (Xa) [reviewed in Ref. 152]. The ICR for this region exhibits paternal-specific methylation (black circles) and contains multiple CTCF binding sites; it also shows bidirectional promoter activity (arrows). The ICR and a downstream element (gray oval) both have enhancing activity on the *Tsix* promoter (dashed curved arrows). The *Tsix* promoter is believed to acquire biallelic methylation (gray box) as a secondary event after X inactivation takes place [138].

transcription factor, yin yang 1 (YY1), similarly functions as a methylation-sensitive insulator that mediates parent-specific expression at several imprinted loci [144–146]. Interestingly, YY1 associates with CTCF through protein-protein interactions and serves as a cofactor in X chromosome inactivation; because both proteins are ubiquitously expressed, they are hypothesized to alternately regulate (or co-regulate) developmentally or tissue-specific imprinting [147].

Noncoding RNAs

While only a fraction of the mammalian genome is actively transcribed in differentiated cells [148], much of the transcriptional activity produces non-coding RNAs (ncRNAs) [reviewed in Ref. 149]. Several of these ncRNAs – ranging from ~21 nucleotides to several kilobases in length – are now known to serve regulatory functions [reviewed in Refs150,151]. Perhaps the most well known ncRNA in mammalian gene regulation is *Xist*, which mediates long-range silencing of the X chromosome to achieve dosage compensation [reviewed in Ref. 152]. Many imprinted loci studied to date contain ncRNAs [reviewed in Ref. 153], which form the basis of another genomic imprinting paradigm.

The maternally-expressed *Igf2r* was the first of three imprinted mouse genes identified in 1991 [7]. It encodes a receptor for the Igf2 protein, and primarily serves as a negative regulator by internalizing Igf2 and targeting it for degradation [reviewed in Ref. 154]. The *Igf2r* domain includes several nonimprinted genes (*Slc22a1*, *Mas1*, and *Plg*), two additional maternally-expressed genes (*Slc22a2* and *Slc22a3*), and a single paternally-expressed gene that encodes the antisense *Igf2r* RNA (*Air*) (Fig. 22.6). The first intron of the *Igf2r* gene contains a DMR that coordinates the maternal-specific expression of *Igf2r* [128], *Slc22a2*, and *Slc22a3* [155]. *Air* transcription originates within this DMR, and proceeds in an antisense orientation to *Igf2r* [156]. When *Air* transcription is prematurely terminated, the remaining promoter retains its imprint; however, all of the paternally-silenced genes (*Igf2r*, *Slc22a2*,

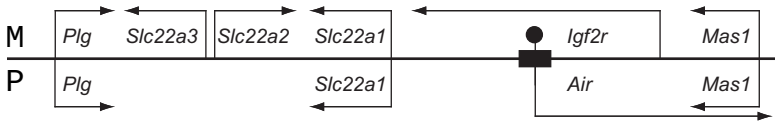


FIGURE 22.6
The *Igf2r/Air* region (not drawn to scale), which demonstrates ncRNA-dependent imprinting. Maternally expressed genes (*Slc22a2* and *Slc22a3*) are represented by solid arrows above the line, and the single paternally expressed antisense *Igf2r* RNA (*Air*) is indicated below the line. Biallelic genes (*Plg*, *Slc22a2*, and *Mas1*) are also shown. The DMR for this region resides in the first intron of the *Igf2r* gene, and is methylated on the maternal allele (black circle); it also serves as the origin of *Air* transcription on the paternal allele. The *Air* ncRNA overlaps the reciprocally imprinted *Igf2r*, and also mediates silencing at *Slc22a2* and *Slc22a3* [157]. The *Air* transcript overlaps (but does not silence) *Mas1* [156].

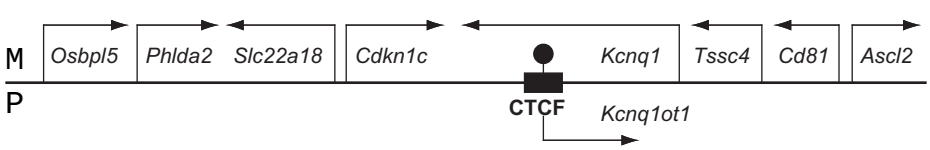


FIGURE 22.7
The *Kcnq1* locus (not drawn to scale) contains several maternally expressed genes (indicated above the line); a single paternally expressed *Kcnq1ot1* ncRNA is transcribed antisense to the *Kcnq1*. The ICR lies within a *Kcnq1* intron and is methylated on the maternal allele; on the paternal allele, the unmethylated ICR binds CTCF and also serves as the origin of *Kcnq1ot1* transcription. While the *Kcnq1ot1* ncRNA is required to imprint the maternally expressed genes [159], CTCF has been proposed to regulate imprinting of *Cdkn1c* in some tissues [161].

and *Slc22a3*) become active, which indicates that the full-length *Air* transcript is required to silence the maternal-specific genes in this region [157].

The *Kcnq1* locus also supports an ncRNA-dependent model of genomic imprinting. It contains several maternally expressed genes and one paternally expressed ncRNA, *Kcnq1ot1*, which is transcribed antisense to the *Kcnq1* potassium channel gene (Fig. 22.7). The ICR for this domain, which is also the origin of *Kcnq1ot1* transcription, lies within a *Kcnq1* intron (Table 22.1) and is methylated on the maternal allele [158]. Premature termination of the *Kcnq1ot1* ncRNA disrupts imprinting in the locus, but the truncated *Kcnq1ot1* retains its imprint [159]. This is consistent with the observations at the *Igf2r-Air* locus [157].

Interestingly, the *Kcnq1* locus may also utilize the insulator model of genomic imprinting, as the unmethylated paternal ICR not only serves as the origin of *Kcnq1ot1* transcription but also binds CTCF [160]. Truncating *Kcnq1ot1* transcription does not affect *Cdkn1c* imprinting in some tissues, which implicates ncRNA-independent mechanisms that are perhaps mediated by CTCF [161]. Likewise, the *Igf2/H19* domain (Fig. 22.4), which is believed to follow a strict CTCF-dependent insulator model [162], may also use multiple imprinting mechanisms – some of which may require the *H19* ncRNA. In a targeted disruption of the *H19* transcriptional unit, the DNA cassette inserted in its place becomes imprinted – yet *Igf2* becomes biallelic; this indicates that full-length *H19* is required for imprinting *Igf2* [163], though its precise function remains unclear.

These ncRNAs are distinct from the protein-coding transcripts that contribute to maternal germline imprints [45], and the antisense orientations of *Air* and *Kcnq1ot1* evoke a possible dsRNA-based mechanism for silencing *Igf2r* and *Kcnq1*, respectively. This model would emulate RNAi, a system that likely evolved to silence transposable elements, viral DNAs, and other parasitic nucleic acids [reviewed in Ref. 150]. However, these antisense ncRNAs do not overlap all oppositely imprinted genes in their domains; moreover, non-imprinted genes may be overlapped, such as *Mas1* in the *Igf2r/Air* locus [156]. These observations argue against the likelihood of a homology-dependent silencing mechanism.

Imprinted silencing by ncRNAs may be similar to *Xist*-mediated X chromosome inactivation. In this scenario, transcripts coat the DNA and recruit chromatin modifying proteins and silencing factors [reviewed in Ref. 152]. Consistent with this hypothesis, *Kcnq1ot1* associates with the PcG proteins at the chromatin level [164], and both *Air* and *Kcnq1ot1* recruit repressive histone methyltransferases to their target promoters – an effect that requires the full-length ncRNAs [165,166]. The silent genes in both the *Kcnq1* and *Igf2r* loci also become contracted into repressive nuclear compartments that exclude RNA polymerase II [164], which mirrors the transcriptionally silent nuclear compartment formed by the repressive *Xist* RNA [167]. Thus, imprinted silencing by ncRNA has been proposed to be mechanistically similar to *Xist*-mediated gene silencing [164].

However, there is some indication that *Xist* antagonism by the complementary *Tsix* does involve RNAi, whereas RNAi does not appear to be sufficient for silencing at the *Kcnq1* locus. The *Xist* and *Tsix* ncRNAs have been shown to form double-stranded duplexes processed by Dicer, a central protein in the RNAi pathway [168]. Though this study showed that Dicer deficiency abolishes *Tsix*-dependent *Xist* repression, other studies have shown Dicer-deficient ES cells to have normal *Xist* expression patterns and X chromosome inactivation [169,170]. Likewise, abolishing Dicer function does not affect *Kcnq1ot1*-mediated gene silencing, which suggests that RNAi pathways are not involved [171]. Interestingly, in the Dicer-deficient embryos that exhibit normal *Xist* expression, the *Xist* promoter is hypomethylated – suggesting that Dicer affects promoter methylation indirectly by regulating Dnmt3a [170], the *de novo* methylase involved in primary imprint establishment [40]. These studies demonstrate an interesting intersection of various epigenetic mechanisms, which may all contribute to genomic imprinting.

Transposable Elements

As evidenced by the incorporation of *Peg12/Frat3* into the murine PWS/AS domain (Fig. 22.3), previously non-imprinted genes may acquire allele-specific expression when transposed next to other imprinted genes [60,61]. However, there is increasing evidence that transposition not only adds genes to existing imprinted domains, but may also establish genomic imprinting *de novo*. This is conceptually similar to position effect variegation in *Drosophila*, in which genes are silenced when placed adjacent to heterochromatin [reviewed in Ref. 125]. Interestingly, genomic imprinting in *Drosophila* is usually confined to heterochromatin [reviewed in Ref. 172], which is characterized by repetitive sequences and transposable elements [reviewed in Ref. 173]. Due to its ability to silence genes by juxtaposition and its correlation with genomic imprinting in *Drosophila*, heterochromatin has been hypothesized to establish imprinted domains when transposed to euchromatin [172].

The concept of transposable "controlling elements" originated with Barbara McClintock's seminal discovery of transposons in maize [174], and transposable elements are now believed to serve major regulatory functions in genomic imprinting. Retrotransposons, which are transposons that replicate via RNA intermediates, are especially abundant in eukaryotes; much of the eukaryotic 5-methylcytosine is targeted to these elements [reviewed in Ref. 31]. In plants, transposable elements are differentially methylated and maintained by DDM1 [175], a homolog of the yeast SWI2/SNF2 chromatin-remodeling complex [176]. Differential methylation of the imprinted *Arabidopsis FWA* promoter is targeted to a retrotransposon-derived SINE element, which in itself is sufficient for imprinted silencing of *FWA* [177]. This SINE element of *FWA* also corresponds to small interfering RNAs (siRNAs), which supports a role for RNAi in genomic imprinting in plants [175].

Retrotransposons compose about 50% of the human genome [reviewed in Ref. 178], which also contains several hundred coding sequences for reverse transcriptases that facilitate retrotransposon replication [31]. While mammalian imprinted domains tend to have an overall lower frequency of retrotransposon-derived SINEs [143,179], DMRs and ICRs themselves are highly enriched with repeats [35–37], which may be remnants of transposition. Indeed, the repeat-rich ICR of the *Xist/Tsix* locus (Fig. 22.5) bears a striking resemblance to the ERV family of endogenous retrovirus-like transposons [139]. Known paternal DMRs are also correlated with tandemly repeated sequences [42,43], which are likely to be vestiges of retrotransposons that are targeted for allele-specific methylation by Dnmt3L [49].

Retrotransposons have been shown to act as novel promoters [180], which may also be imprinted. In mice, a retrotransposon inserted upstream of the *agouti* gene drives ectopic expression in a parent-specific manner [181,182]. Several imprinted retrotransposons have been found within introns, which are sometimes called "microimprinted" domains; these may serve as maternal germline DMRs that give rise to paternal-specific transcripts on the opposite allele [183]. Retrotransposons also serve as oocyte-specific promoters, which produce abundant transcripts (over 10% of the mRNA pool) that persist into the early embryo [184]. It has been proposed that transcripts from oocyte-specific alternative promoters facilitate maternal DMR methylation [45]; given the abundance of retrotransposon-derived transcripts in the oocyte [184], it is very likely that these alternative promoters are of retrotransposon origin.

ON THE ORIGIN OF GENOMIC IMPRINTING

The underlying mechanisms of genomic imprinting – such as DNA methylation, chromatin modification, and ncRNAs – are well conserved across diverse taxa and may theoretically be traced to common origins. It is likely that diverging species independently recruited these mechanisms, and evolved modes of imprinting that fundamentally differ while retaining

striking similarities. For example, entire chromosomes are often silenced or eliminated in insects [reviewed in Refs 25,185] whereas in mammals (with the exception of the X chromosome), imprinting generally targets discrete loci [14]. However, when transgenically introduced into *Drosophila* DNA, mammalian ICRs can silence neighboring genes [129,130], which demonstrates remarkable mechanistic conservation. The most highly conserved silencing mechanism (at least among higher eukaryotes) is methylation, which is used by both animals and plants to maintain differential gene expression [reviewed in Ref. 186]. In mammals, *de novo* methylases establish primary imprints [reviewed in Ref. 187]; in contrast, imprinted alleles in plants are selectively activated from a default silent state by DNA glycosylases [81]. Yet in both phyla, imprinted genes feature prominently in tissues that support embryonic development – such as the mammalian placenta and the seed endosperm of flowering plants [reviewed in Ref. 186]. Selective forces in these tissues form the basis of prominent evolutionary models of genomic imprinting.

Though ongoing research is steadily increasing our understanding of genomic imprinting, the phenomenon remains puzzling in many ways. Opinions differ as to how genomic imprinting originated, how parent-specific expression is selected, why imprinting affects only some genes, and why only certain taxa exhibit genomic imprinting. As expected, the field is rife with theories and models – the most prominent of which consider the phylogenetic distribution of genomic imprinting in vertebrates. The *Igf2* gene is of particular interest because it is well conserved among vertebrates but not universally imprinted [reviewed in Refs 188,189]; thus, studies have examined *Igf2* in various vertebrate classes [2,19,143, 190–194] in an effort to pinpoint the evolutionary origin of genomic imprinting. Only mammals in the subclass Theria, which includes metatherians (marsupials) and eutherians (mice, humans, and most other contemporary mammals), are known to genomically imprint; therefore, established theories are largely based on both physiological and genomic differences between therian and prototherian (monotreme) mammals [reviewed in Ref. 195].

According to philosophical tradition, scientific theories may be classified as *organismic*, *mechanistic*, or *reductionist* [196]. Organismic theories center on interactions between individual organisms; in terms of genomic imprinting, these interactions may favor allele-specific expression, rather than biallelic or stochastic (random) monoallelic expression, for certain genes. These models consider the physiological functions of imprinted genes, and attribute their parent-specific expression to genetic conflicts or co-adaptive interactions within populations. In contrast, mechanistic models focus mainly on the fundamental processes of genomic imprinting, such as silencing mechanisms that target certain genomic elements. According to the established mechanistic theories, genomic imprinting evolved from host defense mechanisms against invasive genetic elements, which reflect increasing genome complexity. Finally, reductionist theories dissect complex systems into interactions between individual parts; in the case of genomic imprinting, this may focus on the interaction between two alleles. Reductionist theories of genomic imprinting examine the possible benefits of monoallelic gene expression in diploid organisms. Various organismic, mechanistic, and reductionist theories have been presented as opposing viewpoints; however, they address different levels of hierarchy in imprinted gene regulation, and may actually form complementary models of genomic imprinting.

Organismic Models of Genomic Imprinting

The majority of known imprinted genes have established roles in growth, metabolism, or behavior during mammalian development [reviewed in Refs 188,189]. Because resources must be carefully allocated between the mother, offspring, and siblings during this time, it has been proposed that competition for resources imposes selective pressure on these genes. Prominent evolutionary models of genomic imprinting are based on these organismic interactions between related individuals. The *parent–offspring conflict theory* was originally

formulated by Trivers [197], before imprinted genes were characterized, and proposed that genes (such as those governing altruistic behavior and maternal instinct) would be positively or negatively selected to optimize resource allocation during maternal care. A variation on this concept, known as the *kinship theory*, includes adaptations driven by additional familial or social interactions – such as in litters or social insects [24]. Thus, these selective pressures optimize survival not just for individuals, but also for populations. After the discovery of genomic imprinting, genes were increasingly viewed as modular alleles as opposed to singular units; thus, these concepts were adopted by Haig to explain parent-specific monoallelic expression in angiosperm (flowering) plants [198] and mammals [199].

Therian mammals are distinguished from the extant egg-laying prototherians by viviparity (live birth) and the presence of a highly-developed placenta, which is the site of nutrient transfer between mother and fetus [reviewed in Ref. 200]. Therefore, many organismic models of genomic imprinting correlate the known functions of imprinted genes with prolonged gestation and postnatal care. The parent–offspring conflict theory proposes that growth-promoting genes, such as *Igf2*, favor paternal expression – particularly in the placenta – to maximize resource transfer to the embryo, which represents the paternal genetic contribution; in contrast, maternal expression of growth-suppressing genes (such as *H19* and *Igf2r*, which both negatively regulate *Igf2*) would optimize maternal health [reviewed in Ref. 188]. Imprinted genes are not only common in the placenta but also in the brain [reviewed in Ref. 201], where they may contribute to cognitive processes [202], postnatal adaptation to feeding and novel environments [203,204], and other neurological processes that may also be relevant to parent-offspring interactions.

Because imprinted genes were first identified based on developmental phenotypes, this may have created a sampling bias for genes involved in embryonic or postnatal development; thus, it is not surprising that most imprinted genes conform to the parent-offspring conflict theory. However, not all imprinted expression patterns are so easily predicted by this theory. *Mash2*, a gene required for trophoblast development in mice [205], is a notable example [reviewed in Ref. 206]. The trophoblast is one of the more critical placental tissues for embryonic growth, as it promotes nutrient transfer to the embryo [207]. The conflict theory predicts a paternal expression pattern, yet *Mash2* is biallelic in the early embryo then maternally expressed by 8.5 days past coitum [208]. Complex trophoblast-mediated processes involving placental hormones may justify this paradoxical expression pattern [reviewed in Ref. 206]. Imprinted genes, such as *Rasgrf1*, may also indirectly control growth by regulating non-imprinted growth factors such as insulin-like growth factor I (*Igf1*) [209]. Other genes may display complex imprinting patterns that manifest more strongly in adulthood, after maternal contribution has ceased [210]; these expression patterns may not be obviously consistent with the simple parent-offspring conflict theory, but the more intricate aspects of the kinship theory may apply.

In social animals, survival is not restricted to maternal-offspring conflict, but also involves interactions between related individuals and other members of society [reviewed in Refs 206,211]. Hence, the more inclusive kinship theory may apply to imprinted genes without obvious relevance to simple maternal-offspring conflict. For animals that group together for warmth, such as emperor penguins and species with large litters, this might include genes that govern "huddling" behavior in addition to metabolic processes, such as thermogenesis [reviewed in Ref. 212]. Furthermore, parental investment in higher mammals is not limited to perinatal development, but includes courtship and mating; these complex interactions may explain why some genes that affect sexual behavior may also be imprinted. One such gene is the paternally expressed *Peg3*, which not only regulates suckling behavior in mouse pups but also olfactory-dependent maternal instincts (such as licking and grooming of pups) and male sexual behavior [213–216]. Because *Peg3* is involved in complex co-adaptive interactions that are beyond the scope of maternal–offspring conflict, the kinship theory

has been challenged with the *co-adaptation theory* [217,218]. The co-adaptation theory has also been proposed for unexpected patterns of imprinted expression, such as the apparent preponderance of maternally-expressed genes [219]. However, since the kinship theory is not limited to maternal–offspring interactions but involves all interactions between related individuals [206], the co-adaptation theory may not be that dissimilar to the kinship theory.

Theories based on organismic interactions also predict genomic imprinting to occur in species that make significant maternal contributions to their young during gestation and postnatal development, such as the placental (therian) mammals. Indeed, genomic imprinting is characteristic of viviparous therian mammals [reviewed in Ref. 200] and apparently lacking in egg-laying (oviparous) animals, including monotreme mammals [191] and birds [19]. However, viviparity is also characteristic of certain placental fish and reptiles [reviewed in Ref. 220], which so far have not been associated with genomic imprinting [reviewed in Refs 188,189]. Comparative analyses in teleost fishes (which include both oviparous and viviparous species) have linked placental development with positive selection at the *Igf2* locus, which supports the notion that genomic imprinting is coincident with placental evolution [192]. However, the apparent lack of genomic imprinting in monotreme mammals – which have primordial placentas yet lay eggs – suggests that a stronger link may exist between genomic imprinting and viviparity [reviewed in Ref. 200].

Oviparity has also been linked to primordial imprinting mechanisms. Orthologs to imprinted mammalian genes (complete with CpG islands) occur in many egg-laying fish, such as zebrafish [221], puffer fish [222,223], and goldfish [194]; however, conservation is rather poor in terms of synteny (chromosomal position) [190,222] and differential methylation [194,223]. The biallelic methylation patterns in fish orthologs may be explained in part by Dnmt3L, the cofactor in primary imprint establishment in mouse [41]. Dnmt3L is conserved among therian mammals [224] but lacking in animals that do not genomically imprint, such as monotremes [225], fish, and birds [224]. Interestingly, though zebrafish lack Dnmt3L, they can differentially methylate exogenous transgenes according to parent of origin [226]. Furthermore, the CpG island near the goldfish *Igf2* gene is hypermethylated in goldfish sperm but not in eggs; this mirrors the methylation patterns of mammalian orthologs, though differential methylation is not maintained zygotically in goldfish [194]. Together with the reduced viability of uniparental zebrafish [227] and goldfish [228], both of which are oviparous, these data suggest that a primordial form of genomic imprinting exists in fish – and is not strongly correlated with viviparity or placentation. In addition to fish, other invertebrate genomes (such as chicken and frog) contain orthologous arrays with varying degrees of synteny with mammalian imprinted loci; this suggests that primordial imprinting mechanisms existed in a common vertebrate ancestor prior to mammalian divergence [190].

Mechanistic Models of Genomic Imprinting

It is important to note that the organismic conflict-based theories seek to explain the parent-specific expression patterns of imprinted genes, as well as to justify their natural selection during evolution; they do not explain how or why the mechanisms of genomic imprinting arose. This aspect of the kinship theory is sometimes interpreted as a weakness. However, the kinship theory assumes preexisting mechanisms and proposes selective pressures to impart allele specificity on these mechanisms; it does not actually attempt to explain their origins [206]. Mechanistic models of genomic imprinting address these fundamental processes.

One such model proposes that genomic imprinting evolved from a primitive host-defense mechanism [229,230]. In many species, this manifests as methylation and silencing of foreign nucleic acids, endogenous transposed DNAs, and repetitive elements [229]. Unlike conflict-based theories, the host defense hypothesis does not consider the biological functions of imprinted genes or species-specific reproductive features. Rather, it correlates imprinting with distinguishing genomic features and mammalian divergence [reviewed in Ref. 195].

Molecular and genomic evidence for this hypothesis again centers on Dnmt3L. This component of the *de novo* methylation machinery not only establishes germline-specific imprints in mammals [41], but also silences retrotransposons and repetitive DNA sequences [49]. Cross-species genome comparisons also support the link between retrotransposon silencing and imprinting, as Dnmt3L is present in therian mammals [224] but lacking in monotremes, fish, and birds [224,225]. Interestingly, most *de novo* methyltransferases are highly conserved between mice and humans (at least 80% identical), yet the Dnmt3L protein sequence is highly divergent (<60% identical); this rapid rate of evolution is consistent with a role in host defense [231].

Recent comparative genome analyses also support the notion that genomic imprinting is an incarnation of a primordial host defense mechanism against retrotransposons. *Paternally expressed 10* (*Peg10*) belongs to the sushi class of retrotransposon-derived genes that have lost

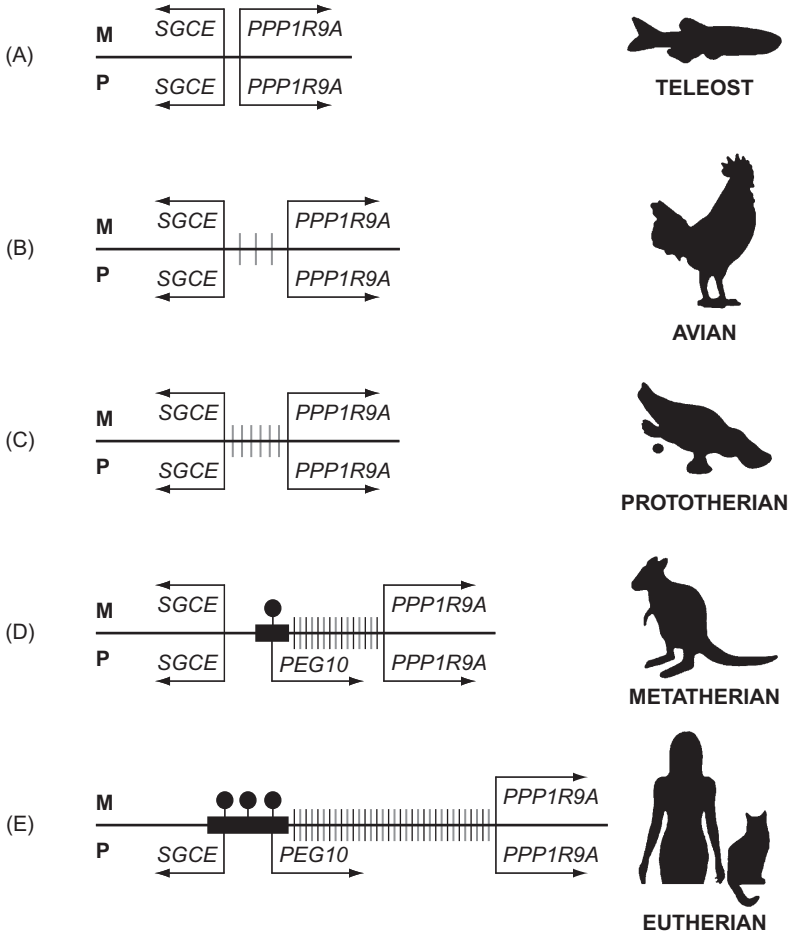


FIGURE 22.8

The *Peg10* locus (not drawn to scale), which correlates genomic imprinting with retrotransposition and genome expansion during vertebrate evolution. (A) In teleost fish, *SGCE* lies adjacent to the *PPP1R9A* gene. (B,C) This intergenic region is expanded in avian (bird) and prototherian (monotreme) genomes by retrotransposon-derived short and long interspersed elements (SINES and LINES), indicated by the short gray vertical lines. (D,E) In metatherian (marsupial) and eutherian mammals, the acquisition of the *Peg10* gene is coincident with imprinted paternal expression (solid arrow below the line) and maternal methylation (black circle). This novel imprinted gene is associated with increased long terminal repeat (LTR)-like sequences (short black lines), which are also indicative of retrotransposition. (E) In eutherians (humans and most contemporary mammals), *SGCE* is also imprinted; this is coincident with further genome expansion. [Adapted from Ref. 234].

the ability to transpose in mammals, which presents a link between genomic imprinting and retrotransposon silencing [232,233]. *Peg10* and the neighboring sarcoglycan epsilon (*Sgce*) gene are imprinted with paternal-specific expression in eutherians (mice and humans); however, the *Peg10* gene is conspicuously absent in the non-imprinted homologous region in platypus (Fig. 22.8), which correlates its retrotransposition into the therian genome with genomic imprinting [234]. Moreover, *Peg10* is imprinted in a marsupial (the tammar wallaby) but *Scge* is not (Fig. 22.8). This suggests that *Peg10* acquired differential methylation and primordial imprinting upon retrotransposition [234]. Compared to imprinted domains in therians, orthologous monotreme regions also have fewer repeat elements; thus, genomic imprinting coincides evolutionarily with retrotransposition and the accumulation of repetitive sequences in therian genomes [235].

The genomic evidence strongly correlates imprinting with increased retrotransposition during therian divergence. This implicates a host-defense origin for genomic imprinting mechanisms, yet does not address why certain alleles are preferentially expressed based on parent of origin. This may be better resolved with organismic models, such as the kinship theory; however, the host defense hypothesis may also provide underlying mechanisms for allele-specific silencing. The primary imprint in the male germline requires repeat elements in paternal DMRs [42,43] and is coincident with retrotransposon silencing by Dnmt3L [49]; thus, it is possible that Dnmt3L expression in the male germline targets the retrotransposon-like characteristics of DMRs. Conversely, maternal germline imprints appear to require transcription from oocyte-specific alternative promoters [45], which may also be of retrotransposon origin. This is supported by the regulatory role of retrotransposons in oocyte-specific transcription [184]. Since Dnmt3L-mediated retrotransposon silencing appears to be limited to the male germline [50], this may explain the abundance of retrotransposon-derived transcripts in the oocyte. Nonetheless, it remains unanswered why genes should be monoallelically expressed at all – whether they are silenced randomly or in a parent-dependent fashion. This question is addressed with reductionist theories of genomic imprinting.

Reductionist Models of Genomic Imprinting

Considering the purported genetic advantages of diploidy [1], it is perplexing for autosomal genes to be hemizygous – or expressed from just one allele – in diploid organisms. Reductionist theories of genomic imprinting address this paradoxical nature of genomic imprinting. These hypotheses not only apply to imprinted genes, which number around 125 according to several online databases (Table 22.2), but also to autosomal genes that are subject to stochastic (random) monoallelic expression. This includes genes in the immune and odorant systems that are randomly silenced via allelic exclusion [reviewed in Refs 96,236,237], as opposed to genes on the X chromosome that are silenced to achieve dosage compensation. Random monoallelic expression is a widespread phenomenon; according to a recent genome survey, 5–10% of autosomal genes – or well over 1000 – may be randomly expressed from only one allele at any given time [238]. While genomic imprints are set in the germline, random monoallelic expression is a zygotic process; however, similar mechanisms are used in either imprinted or random allelic silencing, which suggests that the two processes are related [reviewed in Ref. 96]. The prevalence of monoallelic expression – imprinted or random – implies that it may serve some sort of evolutionary purpose.

It is generally agreed that hemizygosity increases the evolvability of a particular locus, and hence the adaptability of the overall population [217,239,240]. Because diploidy may mask both deleterious and beneficial mutations, functional haploidy may quickly eliminate undesirable recessive traits while simultaneously promoting beneficial mutations. In complex multicellular organisms, monoallelic expression of multiple loci may combinatorially increase phenotypic variability and facilitate adaptive responses.

TABLE 22.2 Select Online Databases of Imprinted Genes

Database	Taxonomic Group(s)
Brain Imprinted Source Tables | Cardiff University http://www.bgg.cf.ac.uk/imprinted_tables	Mouse (brain)
Catalog of Parent of Origin Effects | Otago University http://igc.otago.ac.nz	Human, mouse, rat, cow, pig, sheep, marsupial, monotreme
Geneimprint | Duke University http://www.geneimprint.com	Human, mouse, rat
Genomic Imprinting | MRC Harwell http://www.har.mrc.ac.uk/research/genomic_imprinting	Mouse
WAMIDEX | King's College London https://atlas.genetics.kcl.ac.uk/atlas.php	Mouse

This is plainly illustrated by the broad range of highly specific receptors generated via allelic exclusion in the immune and olfactory systems [reviewed in Ref. 241]. Monoallelic expression is even characteristic of *Drosophila* odorant receptor genes, which are unrelated to vertebrate olfactory genes [reviewed in Ref. 242]. Thus, monoallelic expression is fundamental to adaptive processes, and may in fact be evolutionarily advantageous.

Because hemizygosity is so widespread, one might speculate that it reflects the economic costs associated with maintaining diploid genomes. Heterozygosity may be genetically advantageous; however, for a multicellular organism, the benefits of diploidy must outweigh the costs of replicating or expressing two copies of a locus in every single cell. An analogy can be made using one of the simplest model organisms, the bacterium *Escherichia coli*. It is a widely known fact that *E. coli* cultures only retain plasmids if they confer some selective advantage, such as antibiotic resistance; if there is no selective pressure, then cells without plasmids are at a growth advantage because they do not expend resources replicating extraneous plasmid DNA [243]. More primitive forms of genomic imprinting, such as paternal genome elimination in mealybugs and fungus gnats [reviewed in Refs 25,185], might illustrate this extraneous nature of duplicate genomes. In higher eukaryotes, the need to repress superfluous DNA is demonstrated by the targeted methylation and silencing of duplicated genes [244,245]. Gene silencing is also a fundamental property of complex multicellular organisms [95], and it is estimated that less than 10% of mammalian genes are transcribed at any given time in differentiated cell lineages [148]. Thus, it should not come as a surprise that monoallelic silencing is a widespread phenomenon. In the most basic model of genomic imprinting, monoallelic expression may simply represent an economical means of maintaining a diploid genome. Whatever the rationale, monoallelic expression may be subject to additional selective pressures that determine random or allele-specific silence.

One common limitation of reductionism is that it may not adequately accommodate complex systems or concepts [246]. Indeed, neither the mechanistic nor reductionist models presented here can fully account for the parent-specific expression that is central to genomic imprinting, which is best explained by the organismic kinship theory. Nonetheless, the mechanistic theories provide insight into the origin of genomic imprinting processes, and, as with many reductionist theories, those presented here may serve to facilitate the understanding of this complex phenomenon. Most importantly, none of these theories are mutually exclusive, and may serve complementary functions in deciphering the complex phenomenon of genomic imprinting.

CONCLUSION

Genomic imprinting was once perceived as a bizarre characteristic of plants, insects, and a handful of mammalian genes; it has since become the focus of intensive research, which

has produced numerous implications for development, disease, and evolution. Genomic imprinting may also explain the long-standing mystery of reciprocal interspecies hybrids, which produce progeny with dramatically different phenotypes. For example, crossing a male tiger with a female lion produces a tigon, which is about the same size as either parent; however, the reciprocal cross produces a liger, which is known for its great size [247]. Another notable example is the mule, a hybrid of a male donkey and a female horse. It has been known for millennia that the hinny, which is the lesser-known reciprocal hybrid, differs remarkably in appearance from the mule [248]. Genomic imprinting is now known to underlie the inequality of reciprocal hybrid crosses [reviewed in Ref. 249].

In just three decades, the number of known imprinted genes in mice and humans has grown from three to over 100. In the past, most imprinted genes were recognized based on measurable phenotypes or by association with previously known imprinted genes; however, newer genome-wide analyses may predict novel imprinted genes based on expression profiles or even DNA sequence characteristics. Conservative estimates suggest that up to 150 or so additional imprinted genes exist [250,251]. Recent genome-wide analyses have also revealed complex patterns of imprinting that manifest over multiple generations, and depend not just on parent of origin but also on other imprinted alleles [210]. Thus, more finely tuned approaches may identify additional candidates – particularly those with expression profiles that are more complicated than the traditional binary (on-or-off) definition of genomic imprinting.

References

1. Otto SP, Goldstein DB. Recombination and the evolution of diploidy. *Genetics* 1992;**131**:745–51.
2. Smits G, Mungall AJ, Griffiths-Jones S, Smith P, Beury D, Matthews L, et al. Conservation of the H19 noncoding RNA and H19-IGF2 imprinting mechanism in therians. *Nat Genet* 2008;**40**:971–6.
3. Amor DJ, Halliday J. A review of known imprinting syndromes and their association with assisted reproduction technologies. *Hum Reprod* 2008;**23**:2826–34.
4. Feinberg AP, Ohlsson R, Henikoff S. The epigenetic progenitor origin of human cancer. *Nat Rev Genet* 2006;**7**:21–33.
5. Sharman GB. Late DNA replication in the paternally derived X chromosome of female kangaroos. *Nature* 1971;**230**:231–2.
6. Takagi N, Sasaki M. Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* 1975;**256**:640–2.
7. Barlow DP, Stoger R, Herrmann BG, Saito K, Schweifer N. The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. *Nature* 1991;**349**:84–7.
8. Bartolomei MS, Zemel S, Tilghman SM. Parental imprinting of the mouse H19 gene. *Nature* 1991;**351**:153–5.
9. DeChiara TM, Robertson EJ, Efstratiadis A. Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* 1991;**64**:849–59.
10. McGrath J, Solter D. Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* 1984;**37**:179–83.
11. Surani MA, Barton SC, Norris ML. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature* 1984;**308**:548–50.
12. Mann JR, Lovell-Badge RH. Inviability of parthenogenones is determined by pronuclei, not egg cytoplasm. *Nature* 1984;**310**:66–7.
13. Barton SC, Surani MA, Norris ML. Role of paternal and maternal genomes in mouse development. *Nature* 1984;**311**:374–6.
14. Cattanach BM, Kirk M. Differential activity of maternally and paternally derived chromosome regions in mice. *Nature* 1985;**315**:496–8.
15. Zemel S, Bartolomei MS, Tilghman SM. Physical linkage of two mammalian imprinted genes, H19 and insulin-like growth factor 2. *Nat Genet* 1992;**2**:61–5.
16. Leighton PA, Saam JR, Ingram RS, Stewart CL, Tilghman SM. An enhancer deletion affects both H19 and Igf2 expression. *Genes Dev* 1995;**9**:2079–89.
17. Leighton PA, Ingram RS, Eggenschwiler J, Efstratiadis A, Tilghman SM. Disruption of imprinting caused by deletion of the H19 gene region in mice. *Nature* 1995;**375**:34–9.
18. Chao W, D'Amore PA. IGF2: epigenetic regulation and role in development and disease. *Cytokine Growth Factor Rev* 2008;**19**:111–20.

19. O'Neill MJ, Ingram RS, Vrana PB, Tilghman SM. Allelic expression of IGF2 in marsupials and birds. *Dev Genes Evol* 2000;**210**:18–20.
20. Crouse HV. The controlling element in sex chromosome behavior in sciara. *Genetics* 1960;**45**:1429–43.
21. Noujdin NI. An investigation of an unstable chromosome in *Drosophila melanogaste*r and the mosaicism caused by it. *Zoologishcheski Zh* 1935;**14**:317–52.
22. Noujdin NI. Influence of the Y chromosome and of the homologous region of the X on mosaicism in *Drosophila*. *Nature* 1936;**137**:319–20.
23. Schrader FaH-SS. Haploidy in metazoa. *Q Rev Biol* 1931;**6**:411–38.
24. Trivers RL, Hare H. Haploidploidy and the evolution of the social insect. *Science* 1976;**191**:249–63.
25. Khosla S, Mendiratta G, Brahmachari V. Genomic imprinting in the mealybugs. *Cytogenet Genome Res* 2006;**113**:41–52.
26. Emerson RA. A fifth pair of factors, *Aa*, for aleurone color in maize, and its relation to the *Cc* and *Rr* pairs. *Cornell Univ Agric Esp Station Memoir* 1918;**16**:231–89.
27. Kempton JH. Inheritance of spotted aleurone color in hybrids of chinese maize. *Genetics* 1919;**4**:261–74.
28. Kermicle JL. Dependence of the R-mottled aleurone phenotype in maize on mode of sexual transmission. *Genetics* 1970;**66**:69–85.
29. Doerfler W. In pursuit of the first recognized epigenetic signal–DNA methylation: a 1976 to 2008 synopsis. *Epigenetics* 2008;**3**:125–33.
30. Illingworth RS, Bird AP. CpG islands–"a rough guide". *FEBS Lett* 2009;**583**:1713–20.
31. Yoder JA, Walsh CP, Bestor TH. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet* 1997;**13**:335–40.
32. Swain JL, Stewart TA, Leder P. Parental legacy determines methylation and expression of an autosomal transgene: a molecular mechanism for parental imprinting. *Cell* 1987;**50**:719–27.
33. Reik W, Collick A, Norris ML, Barton SC, Surani MA. Genomic imprinting determines methylation of parental alleles in transgenic mice. *Nature* 1987;**328**:248–51.
34. Lewis A, Reik W. How imprinting centres work. *Cytogenet Genome Res* 2006;**113**:81–9.
35. Reinhart B, Paoloni-Giacobino A, Chaillet JR. Specific differentially methylated domain sequences direct the maintenance of methylation at imprinted genes. *Mol Cell Biol* 2006;**26**:8347–56.
36. Paoloni-Giacobino A, D'Aiuto L, Cirio MC, Reinhart B, Chaillet JR. Conserved features of imprinted differentially methylated domains. *Gene* 2007;**399**:33–45.
37. Hutter B, Helms V, Paulsen M. Tandem repeats in the CpG islands of imprinted genes. *Genomics* 2006;**88**:323–32.
38. Hellmann-Blumberg U, Hintz MF, Gatewood JM, Schmid CW. Developmental differences in methylation of human Alu repeats. *Mol Cell Biol* 1993;**13**:4523–30.
39. Rubin CM, VandeVoort CA, Teplitz RL, Schmid CW. Alu repeated DNAs are differentially methylated in primate germ cells. *Nucleic Acids Res* 1994;**22**:5121–7.
40. Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E, et al. Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* 2004;**429**:900–3.
41. Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH. Dnmt3L and the establishment of maternal genomic imprints. *Science* 2001;**294**:2536–9.
42. Lewis A, Mitsuya K, Constancia M, Reik W. Tandem repeat hypothesis in imprinting: deletion of a conserved direct repeat element upstream of H19 has no effect on imprinting in the Igf2-H19 region. *Mol Cell Biol* 2004;**24**:5650–6.
43. Pearsall RS, Plass C, Romano MA, Garrick MD, Shibata H, Hayashizaki Y, et al. A direct repeat sequence at the Rasgrf1 locus and imprinted expression. *Genomics* 1999;**55**:194–201.
44. Plagge A, Kelsey G. Imprinting the Gnas locus. *Cytogenet Genome Res* 2006;**113**:178–87.
45. Chotalia M, Smallwood SA, Ruf N, Dawson C, Lucifero D, Frontera M, et al. Transcription is required for establishment of germline methylation marks at imprinted genes. *Genes Dev* 2009;**23**:105–17.
46. Shovlin TC, Bourc'his D, La Salle S, O'Doherty A, Trasler JM, Bestor TH, et al. Sex-specific promoters regulate Dnmt3L expression in mouse germ cells. *Hum Reprod* 2007;**22**:457–67.
47. Li JY, Lees-Murdock DJ, Xu GL, Walsh CP. Timing of establishment of paternal methylation imprints in the mouse. *Genomics* 2004;**84**:952–60.
48. Bourc'his D, Proudhon C. Sexual dimorphism in parental imprint ontogeny and contribution to embryonic development. *Mol Cell Endocrinol* 2008;**282**:87–94.
49. Bourc'his D, Bestor TH. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* 2004;**431**:96–9.
50. Lucifero D, La Salle S, Bourc'his D, Martel J, Bestor TH, Trasler JM. Coordinate regulation of DNA methyltransferase expression during oogenesis. *BMC Dev Biol* 2007;**7**:36.

51. Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. *Science* 2001;**293**:1089–93.
52. Ohki I, Shimotake N, Fujita N, Jee J, Ikegami T, Nakao M, et al. Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA. *Cell* 2001;**105**:487–97.
53. Reese KJ, Lin S, Verona RI, Schultz RM, Bartolomei MS. Maintenance of paternal methylation and repression of the imprinted H19 gene requires MBD3. *PLoS Genet* 2007;**3**:e137.
54. Hirasawa R, Chiba H, Kaneda M, Tajima S, Li E, Jaenisch R, et al. Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. *Genes Dev* 2008;**22**:1607–16.
55. Nakamura T, Arai Y, Umehara H, Masuhara M, Kimura T, Taniguchi H, et al. PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat Cell Biol* 2007;**9**:64–71.
56. Li X, Ito M, Zhou F, Youngson N, Zuo X, Leder P, et al. A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. *Dev Cell* 2008;**15**:547–57.
57. Horsthemke B, Wagstaff J. Mechanisms of imprinting of the Prader–Willi/Angelman region. *Am J Med Genet A* 2008;**146A**:2041–52.
58. Kayashima T, Yamasaki K, Joh K, Yamada T, Ohta T, Yoshiura K, et al. Atp10a, the mouse ortholog of the human imprinted ATP10A gene, escapes genomic imprinting. *Genomics* 2003;**81**:644–7.
59. Dubose AJ, Johnstone KA, Smith EY, Hallett RA, Resnick JL. Atp10a, a gene adjacent to the PWS/AS gene cluster, is not imprinted in mouse and is insensitive to the PWS-IC. *Neurogenetics* 2009.
60. Kobayashi S, Kohda T, Ichikawa H, Ogura A, Ohki M, Kaneko-Ishino T, et al. Paternal expression of a novel imprinted gene, Peg12/Frat3, in the mouse 7C region homologous to the Prader–Willi syndrome region. *Biochem Biophys Res Commun* 2002;**290**:403–8.
61. Chai JH, Locke DP, Ohta T, Greally JM, Nicholls RD. Retrotransposed genes such as Frat3 in the mouse Chromosome 7C Prader–Willi syndrome region acquire the imprinted status of their insertion site. *Mamm Genome* 2001;**12**:813–21.
62. Nicholls RD, Knepper JL. Genome organization, function, and imprinting in Prader–Willi and Angelman syndromes. *Annu Rev Genomics Hum Genet* 2001;**2**:153–75.
63. Kantor B, Shemer R, Razin A. The Prader–Willi/Angelman imprinted domain and its control center. *Cytogenet Genome Res* 2006;**113**:300–5.
64. Mapendano CK, Kishino T, Miyazaki K, Kondo S, Yoshiura K, Hishikawa Y, et al. Expression of the Snurf-Snrpn IC transcript in the oocyte and its putative role in the imprinting establishment of the mouse 7C imprinting domain. *J Hum Genet* 2006;**51**:236–43.
65. Tremblay KD, Duran KL, Bartolomei MS. A 5′ 2-kilobase-pair region of the imprinted mouse H19 gene exhibits exclusive paternal methylation throughout development. *Mol Cell Biol* 1997;**17**:4322–9.
66. Jinno Y, Sengoku K, Nakao M, Tamate K, Miyamoto T, Matsuzaka T, et al. Mouse/human sequence divergence in a region with a paternal-specific methylation imprint at the human H19 locus. *Hum Mol Genet* 1996;**5**:1155–61.
67. Thorvaldsen JL, Duran KL, Bartolomei MS. Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2. *Genes Dev* 1998;**12**:3693–702.
68. Engel N, West AG, Felsenfeld G, Bartolomei MS. Antagonism between DNA hypermethylation and enhancer-blocking activity at the H19 DMD is uncovered by CpG mutations. *Nat Genet* 2004;**36**:883–8.
69. Scott J, Cowell J, Robertson ME, Priestley LM, Wadey R, Hopkins B, et al. Insulin-like growth factor-II gene expression in Wilms′ tumour and embryonic tissues. *Nature* 1985;**317**:260–2.
70. Ogawa O, Eccles MR, Szeto J, McNoe LA, Yun K, Maw MA, et al. Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms′ tumour. *Nature* 1993;**362**:749–51.
71. Weksberg R, Shen DR, Fei YL, Song QL, Squire J. Disruption of insulin-like growth factor 2 imprinting in Beckwith-Wiedemann syndrome. *Nat Genet* 1993;**5**:143–50.
72. Sparago A, Cerrato F, Vernucci M, Ferrero GB, Silengo MC, Riccio A. Microdeletions in the human H19 DMR result in loss of IGF2 imprinting and Beckwith-Wiedemann syndrome. *Nat Genet* 2004;**36**:958–60.
73. Prawitt D, Enklaar T, Gartner-Rupprecht B, Spangenberg C, Lausch E, Reutzel D, et al. Microdeletion and IGF2 loss of imprinting in a cascade causing Beckwith-Wiedemann syndrome with Wilms′ tumor. *Nat Genet* 2005;**37**:786–7.
74. Prawitt D, Enklaar T, Gartner-Rupprecht B, Spangenberg C, Oswald M, Lausch E, et al. Microdeletion of target sites for insulator protein CTCF in a chromosome 11p15 imprinting center in Beckwith-Wiedemann syndrome and Wilms′ tumor. *Proc Natl Acad Sci USA* 2005;**102**:4085–90.
75. Ravenel JD, Broman KW, Perlman EJ, Niemitz EL, Jayawardena TM, Bell DW, et al. Loss of imprinting of insulin-like growth factor-II (IGF2) gene in distinguishing specific biologic subtypes of Wilms tumor. *J Natl Cancer Inst* 2001;**93**:1698–703.
76. Cui H, Niemitz EL, Ravenel JD, Onyango P, Brandenburg SA, Lobanenkov VV, et al. Loss of imprinting of insulin-like growth factor-II in Wilms′ tumor commonly involves altered methylation but not mutations of CTCF or its binding site. *Cancer Res* 2001;**61**:4947–50.

77. Yun K. Genomic imprinting and carcinogenesis. *Histol Histopathol* 1998;**13**:425–35.

78. Ohlsson R. Loss of IGF2 imprinting: mechanisms and consequences. *Novartis Found Symp* 2004;**262**:108–21. discussion 121-104, 265-108.

79. Vielle-Calzada JP, Thomas J, Spillane C, Coluccio A, Hoeppner MA, Grossniklaus U. Maintenance of genomic imprinting at the *Arabidopsis medea* locus requires zygotic DDM1 activity. *Genes Dev* 1999;**13**:2971–82.

80. Jullien PE, Kinoshita T, Ohad N, Berger F. Maintenance of DNA methylation during the *Arabidopsis* life cycle is essential for parental imprinting. *Plant Cell* 2006;**18**:1360–72.

81. Kinoshita T, Miura A, Choi Y, Kinoshita Y, Cao X, Jacobsen SE, et al. One-way control of FWA imprinting in Arabidopsis endosperm by DNA methylation. *Science* 2004;**303**:521–3.

82. Bongiorni S, Cintio O, Prantera G. The relationship between DNA methylation and chromosome imprinting in the coccid *Planococcus citri*. *Genetics* 1999;**151**:1471–8.

83. Golic KG, Golic MM, Pimpinelli S. Imprinted control of gene activity in *Drosophila*. *Curr Biol* 1998;**8**:1273–6.

84. Lloyd VK, Sinclair DA, Grigliatti TA. Genomic imprinting and position-effect variegation in *Drosophila melanogaster*. *Genetics* 1999;**151**:1503–16.

85. Patel CV, Gopinathan KP. Determination of trace amounts of 5-methylcytosine in DNA by reverse-phase high-performance liquid chromatography. *Anal Biochem* 1987;**164**:164–9.

86. Urieli-Shoval S, Gruenbaum Y, Sedat J, Razin A. The absence of detectable methylated bases in *Drosophila melanogaster* DNA. *FEBS Lett* 1982;**146**:148–52.

87. Lyko F, Ramsahoye BH, Jaenisch R. DNA methylation in *Drosophila melanogaster*. *Nature* 2000;**408**:538–40.

88. Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP, Jaenisch R. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci USA* 2000;**97**:5237–42.

89. Mager J, Montgomery ND, de Villena FP, Magnuson T. Genome imprinting regulated by the mouse Polycomb group protein Eed. *Nat Genet* 2003;**33**:502–7.

90. Huh JH, Bauer MJ, Hsieh TF, Fischer R. Endosperm gene imprinting and seed development. *Curr Opin Genet Dev* 2007;**17**:480–5.

91. Bean CJ, Schaner CE, Kelly WG. Meiotic pairing and imprinted X chromatin assembly in *Caenorhabditis elegans*. *Nat Genet* 2004;**36**:100–5.

92. Sha K, Fire A. Imprinting capacity of gamete lineages in *Caenorhabditis elegans*. *Genetics* 2005;**170**:1633–52.

93. Hodgkin J. Epigenetics and the maintenance of gene activity states in *Caenorhabditis elegans*. *Dev Genet* 1994;**15**:471–7.

94. Simpson VJ, Johnson TE, Hammen RF. *Caenorhabditis elegans* DNA does not contain 5-methylcytosine at any time during development or aging. *Nucleic Acids Res* 1986;**14**:6711–19.

95. Bird AP. Gene number, noise reduction and biological complexity. *Trends Genet* 1995;**11**:94–100.

96. Goldmit M, Bergman Y. Monoallelic gene expression: a repertoire of recurrent themes. *Immunol Rev* 2004;**200**:197–214.

97. Shemer R, Hershko AY, Perk J, Mostoslavsky R, Tsuberi B, Cedar H, et al. The imprinting box of the Prader–Willi/Angelman syndrome domain. *Nat Genet* 2000;**26**:440–3.

98. Reik W, Brown KW, Schneid H, Le Bouc Y, Bickmore W, Maher ER. Imprinting mutations in the Beckwith-Wiedemann syndrome suggested by altered imprinting pattern in the IGF2-H19 domain. *Hum Mol Genet* 1995;**4**:2379–85.

99. Gribnau J, Hochedlinger K, Hata K, Li E, Jaenisch R. Asynchronous replication timing of imprinted loci is independent of DNA methylation, but consistent with differential subnuclear localization. *Genes Dev* 2003;**17**:759–73.

100. Cerrato F, Dean W, Davies K, Kagotani K, Mitsuya K, Okumura K, et al. Paternal imprints can be established on the maternal Igf2-H19 locus without altering replication timing of DNA. *Hum Mol Genet* 2003;**12**:3123–32.

101. Simon I, Tenzen T, Reubinoff BE, Hillman D, McCarrey JR, Cedar H. Asynchronous replication of imprinted genes is established in the gametes and maintained during development. *Nature* 1999;**401**:929–32.

102. Feil R, Khosla S. Genomic imprinting in mammals: an interplay between chromatin and DNA methylation? *Trends Genet* 1999;**15**:431–5.

103. Sha K. A mechanistic view of genomic imprinting. *Annu Rev Genomics Hum Genet* 2008;**9**:197–216.

104. Plath K, Fang J, Mlynarczyk-Evans SK, Cao R, Worringer KA, Wang H, et al. Role of histone H3 lysine 27 methylation in X inactivation. *Science* 2003;**300**:131–5.

105. Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev Genet* 2009;**10**:295–304.

106. Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, et al. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 1998;**393**:386–9.

107. Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, et al. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat Genet* 1998;**19**:187–91.

108. Sims RJ 3rd, Nishioka K, Reinberg D. Histone lysine methylation: a signature for chromatin function. *Trends Genet* 2003;**19**:629–39.

109. Tamaru H, Selker EU. A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature* 2001;**414**:277–83.

110. Jackson JP, Lindroth AM, Cao X, Jacobsen SE. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* 2002;**416**:556–60.

111. Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, et al. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* 2003;**13**:1192–200.

112. Xin Z, Allis CD, Wagstaff J. Parent-specific complementary patterns of histone H3 lysine 9 and H3 lysine 4 methylation at the Prader–Willi syndrome imprinting center. *Am J Hum Genet* 2001;**69**:1389–94.

113. Xin Z, Tachibana M, Guggiari M, Heard E, Shinkai Y, Wagstaff J. Role of histone methyltransferase G9a in CpG methylation of the Prader–Willi syndrome imprinting center. *J Biol Chem* 2003;**278**:14996–5000.

114. Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* 2007;**448**:714–17.

115. Ciccone DN, Su H, Hevi S, Gay F, Lei H, Bajko J, et al. KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. *Nature* 2009;**461**:415–18.

116. Obata Y, Kono T. Maternal primary imprinting is established at a specific time for each gene throughout oocyte growth. *J Biol Chem* 2002;**277**:5285–9.

117. Delaval K, Govin J, Cerqueira F, Rousseaux S, Khochbin S, Feil R. Differential histone modifications mark mouse imprinting control regions during spermatogenesis. *EMBO J* 2007;**26**:720–9.

118. Schwartz YB, Pirrotta V. Polycomb silencing mechanisms and the management of genomic programmes. *Nat Rev Genet* 2007;**8**:9–22.

119. Joanis V, Lloyd VK. Genomic imprinting in *Drosophila* is maintained by the products of Suppressor of variegation and Trithorax group, but not Polycomb group, genes. *Mol Genet Genomics* 2002;**268**:103–12.

120. Erhardt S, Lyko F, Ainscough JF, Surani MA, Paro R. Polycomb-group proteins are involved in silencing processes caused by a transgenic element from the murine imprinted H19/Igf2 region in *Drosophila*. *Dev Genes Evol* 2003;**213**:336–44.

121. Kalantry S, Magnuson T. The Polycomb group protein EED is dispensable for the initiation of random X-chromosome inactivation. *PLoS Genet* 2006;**2**:e66.

122. Umlauf D, Goto Y, Cao R, Cerqueira F, Wagschal A, Zhang Y, et al. Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat Genet* 2004;**36**:1296–300.

123. Yamasaki-Ishizaki Y, Kayashima T, Mapendano CK, Soejima H, Ohta T, Masuzaki H, et al. Role of DNA methylation and histone H3 lysine 27 methylation in tissue-specific imprinting of mouse Grb10. *Mol Cell Biol* 2007;**27**:732–42.

124. Kuzmin A, Han Z, Golding MC, Mann MR, Latham KE, Varmuza S. The PcG gene Sfmbt2 is paternally expressed in extraembryonic tissues. *Gene Expr Patterns* 2008;**8**:107–16.

125. Weiler KS, Wakimoto BT. Heterochromatin and gene expression in *Drosophila*. *Annu Rev Genet* 1995;**29**:577–605.

126. Anaka M, Lynn A, McGinn P, Lloyd VK. Genomic imprinting in *Drosophila* has properties of both mammalian and insect imprinting. *Dev Genes Evol* 2009;**219**:59–66.

127. Reinhart B, Eljanne M, Chaillet JR. Shared role for differentially methylated domains of imprinted genes. *Mol Cell Biol* 2002;**22**:2089–98.

128. Wutz A, Smrzka OW, Schweifer N, Schellander K, Wagner EF, Barlow DP. Imprinted expression of the Igf2r gene depends on an intronic CpG island. *Nature* 1997;**389**:745–9.

129. Lyko F, Brenton JD, Surani MA, Paro R. An imprinting element from the mouse H19 locus functions as a silencer in *Drosophila*. *Nat Genet* 1997;**16**:171–3.

130. Lyko F, Buiting K, Horsthemke B, Paro R. Identification of a silencing element in the human 15q11-q13 imprinting center by using transgenic *Drosophila*. *Proc Natl Acad Sci USA* 1998;**95**:1698–702.

131. Geyer PK. The role of insulator elements in defining domains of gene expression. *Curr Opin Genet Dev* 1997;**7**:242–8.

132. Bell AC, Felsenfeld G. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* 2000;**405**:482–5.

133. Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 2000;**405**:486–9.

134. Murrell A, Heeson S, Reik W. Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. *Nat Genet* 2004;**36**:889–93.

135. Srivastava M, Frolova E, Rottinghaus B, Boe SP, Grinberg A, Lee E, et al. Imprint control element-mediated secondary methylation imprints at the Igf2/H19 locus. *J Biol Chem* 2003;**278**:5977–83.

136. Chao W, Huynh KD, Spencer RJ, Davidow LS, Lee JT. CTCF, a candidate trans-acting factor for X-inactivation choice. *Science* 2002;**295**:345–7.

137. Xu N, Donohoe ME, Silva SS, Lee JT. Evidence that homologous X-chromosome pairing requires transcription and Ctcf protein. *Nat Genet* 2007;**39**:1390–6.

138. Boumil RM, Ogawa Y, Sun BK, Huynh KD, Lee JT. Differential methylation of Xite and CTCF sites in Tsix mirrors the pattern of X-inactivation choice in mice. *Mol Cell Biol* 2006;**26**:2109–17.

139. Cohen DE, Davidow LS, Erwin JA, Xu N, Warshawsky D, Lee JT. The DXPas34 repeat regulates random and imprinted X inactivation. *Dev Cell* 2007;**12**:57–71.

140. Stavropoulos N, Rowntree RK, Lee JT. Identification of developmentally specific enhancers for Tsix in the regulation of X chromosome inactivation. *Mol Cell Biol* 2005;**25**:2757–69.

141. Ogawa Y, Lee JT. Xite, X-inactivation intergenic transcription elements that regulate the probability of choice. *Mol Cell* 2003;**11**:731–43.

142. Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RD, et al. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* 2007;**128**:1231–45.

143. Weidman JR, Murphy SK, Nolan CM, Dietrich FS, Jirtle RL. Phylogenetic footprint analysis of IGF2 in extant mammals. *Genome Res* 2004;**14**:1726–32.

144. Kim J, Kim JD. In vivo YY1 knockdown effects on genomic imprinting. *Hum Mol Genet* 2008;**17**:391–401.

145. Kim J, Kollhoff A, Bergmann A, Stubbs L. Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, Peg3. *Hum Mol Genet* 2003;**12**:233–45.

146. Kim JD, Hinz AK, Choo JH, Stubbs L, Kim J. YY1 as a controlling factor for the Peg3 and Gnas imprinted domains. *Genomics* 2007;**89**:262–9.

147. Donohoe ME, Zhang LF, Xu N, Shi Y, Lee JT. Identification of a Ctcf cofactor, Yy1, for the X chromosome binary switch. *Mol Cell* 2007;**25**:43–56.

148. Allis CD, Gasser SM. Chromosomes and expression mechanisms New excitement over an old word: 'chromatin'. *Curr Opin Genet Dev* 1998;**8**:137–9.

149. Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, et al. The transcriptional landscape of the mammalian genome. *Science* 2005;**309**:1559–63.

150. Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. *Nat Rev Genet* 2009;**10**:94–108.

151. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. *Nat Rev Genet* 2009;**10**:155–9.

152. Lee JT. Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome. *Genes Dev* 2009;**23**:1831–42.

153. Koerner MV, Pauler FM, Huang R, Barlow DP. The function of non-coding RNAs in genomic imprinting. *Development* 2009;**136**:1771–83.

154. Hebert E. Mannose-6-phosphate/insulin-like growth factor II receptor expression and tumor development. *Biosci Rep* 2006;**26**:7–17.

155. Zwart R, Sleutels F, Wutz A, Schinkel AH, Barlow DP. Bidirectional action of the Igf2r imprint control element on upstream and downstream imprinted genes. *Genes Dev* 2001;**15**:2361–6.

156. Lyle R, Watanabe D, te Vruchte D, Lerchner W, Smrzka OW, Wutz A, et al. The imprinted antisense RNA at the Igf2r locus overlaps but does not imprint Mas1. *Nat Genet* 2000;**25**:19–21.

157. Sleutels F, Zwart R, Barlow DP. The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* 2002;**415**:810–13.

158. Fitzpatrick GV, Soloway PD, Higgins MJ. Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1. *Nat Genet* 2002;**32**:426–31.

159. Mancini-Dinardo D, Steele SJ, Levorse JM, Ingram RS, Tilghman SM. Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. *Genes Dev* 2006;**20**:1268–82.

160. Fitzpatrick GV, Pugacheva EM, Shin JY, Abdullaev Z, Yang Y, Khatod K, et al. Allele-specific binding of CTCF to the multipartite imprinting control region KvDMR1. *Mol Cell Biol* 2007;**27**:2636–47.

161. Shin JY, Fitzpatrick GV, Higgins MJ. Two distinct mechanisms of silencing by the KvDMR1 imprinting control region. *EMBO J* 2008;**27**:168–78.

162. Ideraabdullah FY, Vigneau S, Bartolomei MS. Genomic imprinting mechanisms in mammals. *Mutat Res* 2008;**647**:77–85.

163. Ripoche MA, Kress C, Poirier F, Dandolo L. Deletion of the H19 transcription unit reveals the existence of a putative imprinting control element. *Genes Dev* 1997;**11**:1596–604.

164. Terranova R, Yokobayashi S, Stadler MB, Otte AP, van Lohuizen M, Orkin SH, et al. Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos. *Dev Cell* 2008;**15**:668–79.

165. Pandey RR, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, et al. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol Cell* 2008;**32**:232–46.

166. Nagano T, Mitchell JA, Sanz LA, Pauler FM, Ferguson-Smith AC, Feil R, et al. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* 2008;**322**:1717–20.

167. Chaumeil J, Le Baccon P, Wutz A, Heard E. A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes Dev* 2006;**20**:2223–37.

168. Ogawa Y, Sun BK, Lee JT. Intersection of the RNA interference and X-inactivation pathways. *Science* 2008;**320**:1336–41.

169. Kanellopoulou C, Muljo SA, Dimitrov SD, Chen X, Colin C, Plath K, et al. X chromosome inactivation in the absence of Dicer. *Proc Natl Acad Sci USA* 2009;**106**:1122–7.

170. Nesterova TB, Popova BC, Cobb BS, Norton S, Senner CE, Tang YA, et al. Dicer regulates Xist promoter methylation in ES cells indirectly through transcriptional control of Dnmt3a. *Epigenetics Chromatin* 2008;**1**:2.

171. Redrup L, Branco MR, Perdeaux ER, Krueger C, Lewis A, Santos F, et al. The long noncoding RNA Kcnq1ot1 organises a lineage-specific nuclear domain for epigenetic gene silencing. *Development* 2009;**136**:525–30.

172. Lloyd V. Parental imprinting in *Drosophila*. *Genetica* 2000;**109**:35–44.

173. Dimitri P. Constitutive heterochromatin and transposable elements in *Drosophila melanogaster*. *Genetica* 1997;**100**:85–93.

174. McClintock B. The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci USA* 1950;**36**:344–55.

175. Lippman Z, Gendrel AV, Black M, Vaughn MW, Dedhia N, McCombie WR, et al. Role of transposable elements in heterochromatin and epigenetic control. *Nature* 2004;**430**:471–6.

176. Jeddeloh JA, Stokes TL, Richards EJ. Maintenance of genomic methylation requires a SWI2/SNF2-like protein. *Nat Genet* 1999;**22**:94–7.

177. Fujimoto R, Kinoshita Y, Kawabe A, Kinoshita T, Takashima K, Nordborg M, et al. Evolution and control of imprinted FWA genes in the genus *Arabidopsis*. *PLoS Genet* 2008;**4**:e1000048.

178. Rafalski A, Morgante M. Corn and humans: recombination and linkage disequilibrium in two genomes of similar size. *Trends Genet* 2004;**20**:103–11.

179. Greally JM. Short interspersed transposable elements (SINEs) are excluded from imprinted regions in the human genome. *Proc Natl Acad Sci USA* 2002;**99**:327–32.

180. Ferrigno O, Virolle T, Djabari Z, Ortonne JP, White RJ, Aberdam D. Transposable B2 SINE elements can provide mobile RNA polymerase II promoters. *Nat Genet* 2001;**28**:77–81.

181. Morgan HD, Sutherland HG, Martin DI, Whitelaw E. Epigenetic inheritance at the agouti locus in the mouse. *Nat Genet* 1999;**23**:314–18.

182. Michaud EJ, van Vugt MJ, Bultman SJ, Sweet HO, Davisson MT, Woychik RP. Differential expression of a new dominant agouti allele (Aiapy) is correlated with methylation state and is influenced by parental lineage. *Genes Dev* 1994;**8**:1463–72.

183. Wood AJ, Roberts RG, Monk D, Moore GE, Schulz R, Oakey RJ. A screen for retrotransposed imprinted genes reveals an association between X chromosome homology and maternal germ-line methylation. *PLoS Genet* 2007;**3**:e20.

184. Peaston AE, Evsikov AV, Graber JH, de Vries WN, Holbrook AE, Solter D, et al. Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev Cell* 2004;**7**:597–606.

185. Goday C, Esteban MR. Chromosome elimination in sciarid flies. *Bioessays* 2001;**23**:242–50.

186. Feil R, Berger F. Convergent evolution of genomic imprinting in plants and mammals. *Trends Genet* 2007;**23**:192–9.

187. Scott RJ, Spielman M. Epigenetics: imprinting in plants and mammals–the same but different? *Curr Biol* 2004;**14**:R201–R3.

188. Wilkins JF, Haig D. What good is genomic imprinting: the function of parent-specific gene expression. *Nat Rev Genet* 2003;**4**:359–68.

189. Reik W, Lewis A. Co-evolution of X-chromosome inactivation and imprinting in mammals. *Nat Rev Genet* 2005;**6**:403–10.

190. Dunzinger U, Haaf T, Zechner U. Conserved synteny of mammalian imprinted genes in chicken, frog, and fish genomes. *Cytogenet Genome Res* 2007;**117**:78–85.

191. Killian JK, Nolan CM, Stewart N, Munday BL, Andersen NA, Nicol S, et al. Monotreme IGF2 expression and ancestral origin of genomic imprinting. *J Exp Zool* 2001;**291**:205–12.

192. O'Neill MJ, Lawton BR, Mateos M, Carone DM, Ferreri GC, Hrbek T, et al. Ancient and continuing Darwinian selection on insulin-like growth factor II in placental fishes. *Proc Natl Acad Sci USA* 2007;**104**:12404–9.

193. Suzuki S, Renfree MB, Pask AJ, Shaw G, Kobayashi S, Kohda T, et al. Genomic imprinting of IGF2, p57(KIP2) and PEG1/MEST in a marsupial, the tammar wallaby. *Mech Dev* 2005;**122**:213–22.

194. Xie B, Zhang L, Zheng K, Luo CT. The evolutionary foundation of genomic imprinting in lower vertebrates. *Chinese Sci Bull* 2009;**54**:1354–60.

195. Renfree MB, Hore TA, Shaw G, Marshall Graves JA, Pask AJ. Evolution of genomic imprinting: Insights from marsupials and monotremes. *Annu Rev Genomics Hum Genet* 2009;**10**:241–62.

196. Nagel E. *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York: Harcourt, Brace & World;1961.

197. Trivers RL. Parent-offspring conflict. *Am Zool* 1974;**14**:249–64.

198. Haig D, Westoby M. Parent-specific gene expression and the triploid endosperm. *Am Nat* 1989;**134**:147–55.

199. Haig D. The kinship theory of genomic imprinting. *Annu Rev Ecol Sys* 2000;**31**:9–32.

200. Renfree MB, Ager EI, Shaw G, Pask AJ. Genomic imprinting in marsupial placentation. *Reproduction* 2008;**136**:523–31.

201. Davies W, Isles AR, Humby T, Wilkinson LS. What are imprinted genes doing in the brain? *Adv Exp Med Biol* 2008;**626**:62–70.

202. Davies W, Isles A, Smith R, Karunadasa D, Burrmann D, Humby T, et al. Xlr3b is a new imprinted candidate for X-linked parent-of-origin effects on cognitive function in mice. *Nat Genet* 2005;**37**:625–9.

203. Plagge A, Gordon E, Dean W, Boiani R, Cinti S, Peters J, et al. The imprinted signaling protein XL alpha s is required for postnatal adaptation to feeding. *Nat Genet* 2004;**36**:818–26.

204. Plagge A, Isles AR, Gordon E, Humby T, Dean W, Gritsch S, et al. Imprinted Nesp55 influences behavioral reactivity to novel environments. *Mol Cell Biol* 2005;**25**:3019–26.

205. Guillemot F, Nagy A, Auerbach A, Rossant J, Joyner AL. Essential role of Mash-2 in extraembryonic development. *Nature* 1994;**371**:333–6.

206. Haig D. Genomic imprinting and kinship: how good is the evidence? *Annu Rev Genet* 2004;**38**:553–85.

207. Constancia M, Hemberger M, Hughes J, Dean W, Ferguson-Smith A, Fundele R, et al. Placental-specific IGF-II is a major modulator of placental and fetal growth. *Nature* 2002;**417**:945–8.

208. Guillemot F, Caspary T, Tilghman SM, Copeland NG, Gilbert DJ, Jenkins NA, et al. Genomic imprinting of Mash2, a mouse gene required for trophoblast development. *Nat Genet* 1995;**9**:235–42.

209. Drake NM, Park YJ, Shirali AS, Cleland TA, Soloway PD. Imprint switch mutations at Rasgrf1 support conflict hypothesis of imprinting and define a growth control mechanism upstream of IGF1. *Mamm Genome* 2009;**20**:654–63.

210. Wolf JB, Cheverud JM, Roseman C, Hager R. Genome-wide analysis reveals a complex pattern of genomic imprinting in mice. *PLoS Genet* 2008;**4**:e1000091.

211. Constancia M, Kelsey G, Reik W. Resourceful imprinting. *Nature* 2004;**432**:53–7.

212. Haig D. Huddling: brown fat, genomic imprinting and the warm inner glow. *Curr Biol* 2008;**18**:R172–R4.

213. Champagne FA, Curley JP, Swaney WT, Hasen NS, Keverne EB. Paternal influence on female behavior: the role of Peg3 in exploration, olfaction, and neuroendocrine regulation of maternal behavior of female mice. *Behav Neurosci* 2009;**123**:469–80.

214. Swaney WT, Curley JP, Champagne FA, Keverne EB. Genomic imprinting mediates sexual experience-dependent olfactory learning in male mice. *Proc Natl Acad Sci USA* 2007;**104**:6084–9.

215. Curley JP, Barton S, Surani A, Keverne EB. Coadaptation in mother and infant regulated by a paternally expressed imprinted gene. *Proc Biol Sci* 2004;**271**:1303–9.

216. Curley JP, Pinnock SB, Dickson SL, Thresher R, Miyoshi N, Surani MA, et al. Increased body fat in mice with a targeted mutation of the paternally expressed imprinted gene Peg3. *FASEB J* 2005;**19**:1302–4.

217. Keverne EB, Curley JP. Epigenetics, brain evolution and behaviour. *Front Neuroendocrinol* 2008;**29**:398–412.

218. Hager R, Johnstone RA. The genetic basis of family conflict resolution in mice. *Nature* 2003;**421**:533–5.

219. Wolf JB, Hager R. A maternal-offspring coadaptation theory for the evolution of genomic imprinting. *PLoS Biol* 2006;**4**:e380.

220. Zeh JA, Zeh DW. Viviparity-driven conflict: more to speciation than meets the fly. *Ann N Y Acad Sci* 2008;**1133**:126–48.

221. Hahn Y, Yang SK, Chung JH. Structure and expression of the zebrafish mest gene, an ortholog of mammalian imprinted gene PEG1/MEST. *Biochim Biophys Acta* 2005;**1731**:125–32.

222. Paulsen M, Khare T, Burgard C, Tierling S, Walter J. Evolution of the Beckwith-Wiedemann syndrome region in vertebrates. *Genome Res* 2005;**15**:146–53.

223. Brunner B, Grutzner F, Yaspo ML, Ropers HH, Haaf T, Kalscheue VM. Molecular cloning and characterization of the *Fugu rubripes* MEST/COPG2 imprinting cluster and chromosomal localization in *Fugu* and *Tetraodon nigroviridis*. *Chromosome Res* 2000;**8**:465–76.

224. Yokomine T, Hata K, Tsudzuki M, Sasaki H. Evolution of the vertebrate DNMT3 gene family: a possible link between existence of DNMT3L and genomic imprinting. *Cytogenet Genome Res* 2006;**113**:75–80.

225. Warren WC, Hillier LW, Marshall Graves JA, Birney E, Ponting CP, Grutzner F, et al. Genome analysis of the platypus reveals unique signatures of evolution. *Nature* 2008;**453**:175–83.

226. Martin CC, McGowan R. Genotype-specific modifiers of transgene methylation and expression in the zebrafish, *Danio rerio*. *Genet Res* 1995;**65**:21–8.

227. Corley-Smith GE, Lim CJ, Brandhorst BP. Production of androgenetic zebrafish (*Danio rerio*). *Genetics* 1996;**142**:1265–76.

228. Luo C, Li B. Diploid-dependent regulation of gene expression: a genetic cause of abnormal development in fish haploid embryos. *Heredity* 2003;**90**:405–9.

229. Barlow DP. Methylation and imprinting: from host defense to gene regulation? *Science* 1993;**260**:309–10.

230. McDonald JF, Matzke MA, Matzke AJ. Host defenses to transposable elements and the evolution of genomic imprinting. *Cytogenet Genome Res* 2005;**110**:242–9.

231. Bestor TH, Bourc'his D. Transposon silencing and imprint establishment in mammalian germ cells. *Cold Spring Harb Symp Quant Biol* 2004;**69**:381–7.

232. Youngson NA, Kocialkowski S, Peel N, Ferguson-Smith AC. A small family of sushi-class retrotransposon-derived genes in mammals and their relation to genomic imprinting. *J Mol Evol* 2005;**61**:481–90.

233. Ono R, Nakamura K, Inoue K, Naruse M, Usami T, Wakisaka-Saito N, et al. Deletion of Peg10, an imprinted gene acquired from a retrotransposon, causes early embryonic lethality. *Nat Genet* 2006;**38**:101–6.

234. Suzuki S, Ono R, Narita T, Pask AJ, Shaw G, Wang C, et al. Retrotransposon silencing by DNA methylation can drive mammalian genomic imprinting. *PLoS Genet* 2007;**3**:e55.

235. Pask AJ, Papenfuss AT, Ager EI, McColl KA, Speed TP, Renfree MB. Analysis of the platypus genome suggests a transposon origin for mammalian imprinting. *Genome Biol* 2009;**10**:R1.

236. Zakharova IS, Shevchenko AI, Zakian SM. Monoallelic gene expression in mammals. *Chromosoma* 2009;**118**:279–90.

237. Olender T, Lancet D, Nebert DW. Update on the olfactory receptor (OR) gene superfamily. *Hum Genomics* 2008;**3**:87–97.

238. Gimelbrant A, Hutchinson JN, Thompson BR, Chess A. Widespread monoallelic expression on human autosomes. *Science* 2007;**318**:1136–40.

239. Beaudet AL, Jiang YH. A rheostat model for a rapid and reversible form of imprinting-dependent evolution. *Am J Hum Genet* 2002;**70**:1389–97.

240. McGowan RA, Martin CC. DNA methylation and genome imprinting in the zebrafish, *Danio rerio*: some evolutionary ramifications. *Biochem Cell Biol* 1997;**75**:499–506.

241. Cedar H, Bergman Y. Choreography of Ig allelic exclusion. *Curr Opin Immunol* 2008;**20**:308–17.

242. Fuss SH, Ray A. Mechanisms of odorant receptor gene choice in *Drosophila* and vertebrates. *Mol Cell Neurosci* 2009;**41**:101–12.

243. Casali N, Preston AE. coli plasmid vectors: methods and applications. Totowa, NJ: Humana Press; 2003.

244. Kricker MC, Drake JW, Radman M. Duplication-targeted DNA methylation and mutagenesis in the evolution of eukaryotic chromosomes. *Proc Natl Acad Sci USA* 1992;**89**:1075–9.

245. Lynch M, Conery JS. The evolutionary fate and consequences of duplicate genes. *Science* 2000;**290**:1151–5.

246. Anderson PW. More is different. *Science* 1972;**177**:393–6.

247. Gray AP. Mammalian hybrids: a checklist with bibliography. Farnham Royal, Bucks: Commonwealth Agricultural Bureaux; 1972.

248. Savory TH. The mule. *Sci Am* 1970;**223**:102–9.

249. Glaser RL, Morison IM. Equality of the sexes? parent-of-origin effects on transcription and de novo mutations. *Bioinform Syst Biol* 2009:485–513.

250. Luedi PP, Dietrich FS, Weidman JR, Bosko JM, Jirtle RL, Hartemink AJ. Computational and experimental identification of novel human imprinted genes. *Genome Res* 2007;**17**:1723–30.

251. Pollard KS, Serre D, Wang X, Tao H, Grundberg E, Hudson TJ, et al. A genome-wide approach to identifying novel-imprinted genes. *Hum Genet* 2008;**122**:625–34.

Available online at www.sciencedirect.com
ScienceDirect
Cytokine & Growth Factor Reviews 19 (2008) 111–120
Cytokine & Growth Factor Reviews
www.elsevier.com/locate/cytogfr

Survey

IGF2: Epigenetic regulation and role in development and disease

Wendy Chao [a], Patricia A. D'Amore [b,*]

[a] *Schepens Eye Research Institute, 20 Staniford Street, Boston, MA 02114, United States*
[b] *Departments of Pathology and Ophthalmology, Harvard Medical School, Schepens Eye Research Institute, 20 Staniford Street, Boston, MA 02114, United States*

Abstract

Insulin-like growth factor II (IGF2) is perhaps the most intricately regulated of all growth factors characterized to date. Its gene is imprinted – only one allele is active, depending on parental origin – and this pattern of expression is maintained epigenetically in almost all tissues. IGF2 activity is further controlled through differential expression of receptors and IGF-binding proteins (IGFBPs) that determine protein availability. This complex and multifaceted regulation emphasizes the importance of accurate IGF2 expression and activity. This review will examine the regulation of the *IGF2* gene and what it has revealed about the phenomenon of imprinting, which is frequently disrupted in cancer. IGF2 protein function will be discussed, along with diseases that involve IGF2 overexpression. Roles for IGF2 in sonic hedgehog (Shh) signaling and angiogenesis will also be explored.

Keywords: Insulin-like growth factor; Imprinting; Angiogenesis; Vascular endothelial growth factor; Sonic hedgehog

Contents

1. Introduction

In the early 1900s, the innovative surgeon Alexis Carrel experimented with maintaining tissues and whole organs *in vitro*, hoping to advance techniques in organ transplantation. Carrel observed that certain tissue extracts could induce cell proliferation, and he published his findings with this disclaimer:

"Possibly the finding of the activating power of tissue extracts will have no immediate practical application. Nevertheless, it may be indirectly useful by leading to the discovery of some of the factors determining the growth of tissues and of the unknown laws of cell dynamics . . . [1]."

* Corresponding author. Tel.: +1 617 912 2559; fax: +1 617 912 0128.
E-mail addresses: wendychao@post.harvard.edu (W. Chao), patricia.damore@schepens.harvard.edu (P.A. D'Amore).

1359-6101/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cytogfr.2008.01.005

Table 1
A brief history of the insulin-like growth factors (IGFs)

1912	Albert Schaefer coins the term "insulin" for a substance in blood that controls glucose metabolism [2]
1913	The landmark article "Artificial activation of the growth *in vitro* of connective tissues" is published [1]
1957	Serum is found to contain "sulfation factor activity" (SFA), which mediates the effect of growth hormone (GH) on sulfate uptake by cartilage [3]
1963	The insulin-like factor in human serum, that is not neutralized by anti-insulin antibodies, is given the term "non-suppressible insulin-like activity" (NSILA) [4]
1972	Because SFA and NSILA have similar (if not identical) activities, the term "somatomedin" is proposed to denote the ability to promote somatic growth [5]
1973	"Multiplication-stimulating activity" (MSA), which induces proliferation of chicken embryo fibroblasts, is found in rat liver cell-conditioned media [6]
1976	NSILA is sequenced and found to be two distinct proteins, similar to human and tuna fish insulin. They are named IGF1 and IGF2 [7]
1981	MSA is purified, sequenced, and found to differ from human IGF2 by only five amino acids. Thus, MSA is designated as rat IGF2 [8]
1987	The IGF nomenclature is adopted to denote SFA, NSILA, somatomedin, and MSA [9]

Carrel was mistaken that this finding would have no practical application—rather, it pioneered the discipline of tissue culture and the widespread use of serum to support *in vitro* cell growth. He was right, however, that this "activating power" would eventually lead to the discovery of growth factors, many of which were isolated and characterized in the decades that followed. Two of these factors, which were structurally similar to insulin, had many effects on cell growth and differentiation. In 1987, after 30 years of confusing nomenclature, these proteins were designated as insulin-like growth factor I (IGF1) and insulin-like growth factor II (IGF2) (Table 1).

The IGFs regulate cell growth and differentiation in many species. The anabolic functions of growth hormone are largely mediated by IGF1, which designates IGF1 as a major determinant of somatic growth [10]. Rare mutations in the human *IGF1* gene lead to severe growth inhibition and mental retardation [11]. *Igf1*-null mice are born at 60% of normal birth weight, and the few that survive to adulthood are less than one-third the size of normal mice [12,13]. On the other hand, IGF2 is virtually dispensable for post-natal development in mice, since *Igf2* expression is almost entirely limited to the embryo in rodents [14]. At birth, *Igf2*-null mice are also growth-impaired but are otherwise normal, and subsequent growth proceeds at normal rates [13].

These studies support a somewhat redundant role for IGF2; furthermore, its designation as the "second" IGF seems to have relegated it to a lesser role than IGF1. However, IGF2 is the predominant IGF in adult humans (reviewed in Ref. [15]), and inappropriate IGF2 expression is implicated in a growing number of diseases (reviewed in Ref. [16]). The importance of IGF2 is highlighted by its complex and multifaceted regulation. The gene that codes for IGF2 is imprinted such that only one allele is expressed, depending on parental origin [14]. Besides the intriguing mechanisms that surround its imprinted expression, IGF2 is further modulated by a concert of differentially expressed proteins and receptors that determine IGF availability (reviewed in Ref. [17]). This review will examine the complex epigenetic regulation of the *IGF2* gene and provide a broad introduction to IGF2 signaling. The ability of IGF2 to stimulate cell proliferation and differentiation will be reviewed, which will lead to a discussion on its involvement in various cancers and other diseases. The angiogenic functions of IGF2 will be addressed, and conclude with a proposal that IGF2 is a key mediator facilitating the angiogenic activity of sonic hedgehog (Shh).

2. The *IGF2* gene

2.1. Epigenetic regulation of Igf2

Igf2 is widely expressed during murine embryonic development, and is particularly important in placental growth [18]. As with many genes that regulate placental development, *Igf2* is imprinted, or expressed monoallelically, and active only on the paternally inherited allele. *Igf2* is highly expressed in the mouse embryo, but levels decline dramatically after birth; in adult mice, *Igf2* transcripts are detectable only in the choroid plexus and leptomeninges, where expression is biallelic [14]. *IGF2* is also imprinted in humans, but is expressed biallelically in the choroid plexus, leptomeninges, and perhaps the developing retina [19]. However, human *IGF2* is also expressed in the adult, with transcripts arising from an adult-specific promoter [20]. The corresponding region in the mouse *Igf2* gene contains two pseudo-exons and what appears to be a remnant of this adult-specific promoter which may explain why *Igf2* expression ceases after birth in mice but not in humans [21].

Almost all known imprinted genes occur in clusters with one or more reciprocally imprinted genes (reviewed in Ref. [22]). The mouse *Igf2* gene lies on the distal region of chromosome 7 with the oppositely imprinted, non-coding gene *H19*. *Igf2* and *H19* share a set of enhancers that act on either gene, depending on parental origin. In eukaryotic DNA, promoters generally harbor regions dense with CpG dinucleotides, which are targets of methylation. These "CpG islands" are often methylated in inactive promoters. On the paternal chromosome, the *H19* promoter region is methylated and inactive; this methylation and expression

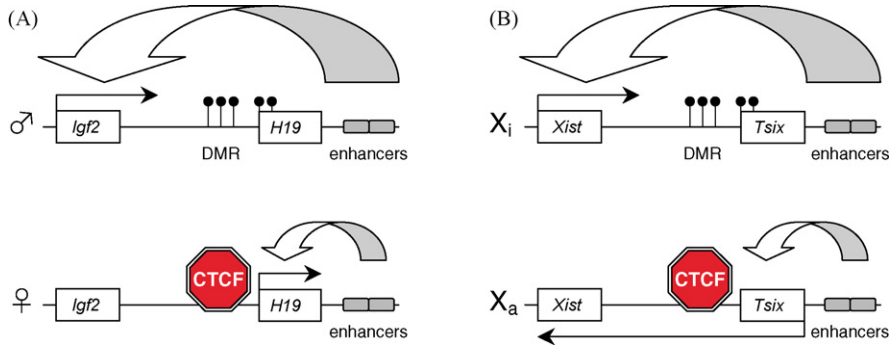


Fig. 1. (A) Model of imprinted regulation at the *Igf2-H19* locus. Adapted from Ref. [25]. (B) Model of allele-specific repression in X chromosome inactivation by CTCF. Adapted from Refs. [26–28]. DMR: differentially methylated region. Lollipops: methylated CpGs. X_i and X_a: inactive and active X chromosomes, respectively.

pattern is passed on when cells divide. Because this inheritance of gene expression patterns is achieved without altering the DNA sequence, it is called *epigenetic*.

The *Igf2* promoter is not methylated on the maternal chromosome, so another mechanism must account for silencing *Igf2*. Several kilobases (kb) upstream of the H19 promoter is a differentially methylated region (DMR) that, when deleted, reactivates *Igf2* on the maternal chromosome [23]. This region, also called the imprinting control region (ICR), was found to harbor binding sites for CCCTC binding factor (CTCF), an insulator protein that demarcates active and inactive chromatin domains (reviewed in Ref. [24]). Methylation of the CG-rich CTCF binding sequence prevents CTCF binding. Thus, on the paternal chromosome, the DMR/ICR is methylated, CTCF is excluded, and the enhancers act on the *Igf2* promoter. Conversely, on the maternal chromosome, CTCF forms a chromatin insulator that blocks the enhancers from activating *Igf2* (Fig. 1A).

2.2. *Igf2* imprinting as a model of allele-specific repression

Murine *Igf2* was the first gene found to be imprinted, and has served as a model of allele-specific gene repression—the most extreme example being X chromosome inactivation, where one X is silenced in each somatic cell of XX female mammals to equalize gene dosage with XY males [29]. *Igf2* imprinting and X chromosome inactivation are the most well-studied mechanisms of epigenetic regulation, and the parallels between these mechanisms give insight into the epigenetic alterations that are abundant in cancer.

X chromosome inactivation generally occurs in a random fashion and silences either X; however, in some mammals and in certain tissues of others, the paternal X is always silenced. In either random or imprinted X chromosome inactivation, the X that is destined to be silenced expresses the non-coding *Xist* RNA, which covers the chromosome and mediates silencing (reviewed in Ref. [30]).

Xist lies in a region called the X inactivation center (XIC) along with another non-coding gene, *Tsix,* which is transcribed antisense to *Xist* and expressed on the active X chromosome [31]. Not long after CTCF was shown to regulate imprinting at the *Igf2/H19* locus, a similar mechanism was found at the *Xist/Tsix* locus. In a region implicated in controlling both random and imprinted X chromosome inactivation, functional methylation-sensitive CTCF binding sites were identified (Fig. 1B). This region was later found to contain developmentally specific enhancers [28] and to be differentially methylated *in vivo* [27]. CTCF has since been demonstrated to control imprinting at several other gene domains, and putative binding sites have been discovered in several other imprinted loci [32]. However, not all imprinted genes contain functional CTCF binding sites. It is proposed that another multifunctional transcription factor, yin yang 1 (YY1), functions as a methylation-sensitive insulator that mediates allele-specific gene activation or silencing at some loci. YY1 has been found to control imprinting at the human *SNURF-SNRPN* locus within the Prader–Willi syndrome and Angelman syndrome locus, and the *PEG3*, *Gnas*, and *Nespas* genes ([33] and references therein). Interestingly, it was reported recently that YY1 is a cofactor for CTCF in X chromosome inactivation [34]. Because both CTCF and YY1 are ubiquitously expressed, it is possible that tissue- and developmentally specific imprinting of *Igf2* is accomplished through a combination of these factors.

The similarities between *Igf2/H19* and *Xist/Tsix* regulation have additional implications for other regulatory mechanisms that may be aberrant in cancer. The X chromosomes initiate silencing after forming a transient interchromosomal complex (reviewed in Ref. [30]). This pairing phenomenon has also been observed with the *Igf2/H19* region, in which CTCF mediates interchromosomal colocalization and induces *trans* effects on a non-homologous chromosome [35]. Interchromosomal pairing may increase the frequency of mitotic recombination, which can account for both heritable and sporadic mutations [36]. Because CTCF mediates interchromosomal pairing of the *IGF2/H19* region, it may very well facilitate such mitotic recombination events. X chromosome inactivation has also drawn attention in the field of cancer research with the recent discovery of X-linked tumor suppressor genes; when mutated, these can lead to hemizygosity in males and skewed X inactivation in females

(reviewed in Ref. [37]). One gene, *FOXP3*, codes for a forkhead family transcription factor that represses the *HER-2/ErbB2* oncogene [38]. Interestingly, the forkhead transcription factors are targets of the PI3-kinase pathway, which is activated by IGF signaling (reviewed in Ref. [39]). The other X-linked tumor suppressor, *WTX*, is frequently inactivated in Wilms' tumor, a disease also associated with disrupted *IGF2* imprinting [40].

X chromosome inactivation can have other implications for *Igf2/H19* regulation as well. There is mounting evidence that non-coding (especially antisense) RNAs regulate allele-specific gene expression (reviewed in Ref. [30]). Multiple sense and antisense transcripts have been detected in the mouse *Igf2* 5′ region, and the major antisense transcript, *Igf2AS,* is paternally expressed and non-coding [41]. An antisense message transcribed from a homologous region near human *IGF2* encodes a putative 273-amino acid protein of unknown function [42]. It remains unclear whether *IGF2AS* regulates *IGF2* or *H19* imprinting; nonetheless, it may have biological importance. In Wilms' tumor, *IGF2AS* is highly expressed and demonstrates sporadic loss of imprinting [42,43]. As stated before, disrupted *IGF2* imprinting is implicated in a number of diseases, and may be attributed to increased gene dosage and subsequent increases in IGF2 signaling, which will be discussed in the following section.

3. The IGF2 protein

3.1. IGF system overview

The IGFs signal primarily through the type I IGF receptor (IGF1R), but there is significant crosstalk between the IGF and insulin systems as certain variants of the insulin receptor (IR) have been shown to bind IGFs (Fig. 2). The alternatively spliced IR-A isoform, which is expressed predominantly during embryogenesis [44], binds insulin and IGF2 (but not IGF1) with high affinity [45]. IGF2 can also stimulate insulin-like metabolic responses by binding the classical IR-B isoform; furthermore, functional heterodimers can form between IGF1R and the IR isoforms (reviewed in Ref. [46]). Thus, tissue-specific effects of insulin and the IGFs may be accomplished through differential expression of the receptors and receptor hybrids. Though IGF1R is activated more efficiently by IGF1 [47], the ability to signal through IR potentially gives IGF2 a broader range of biological functions than IGF1.

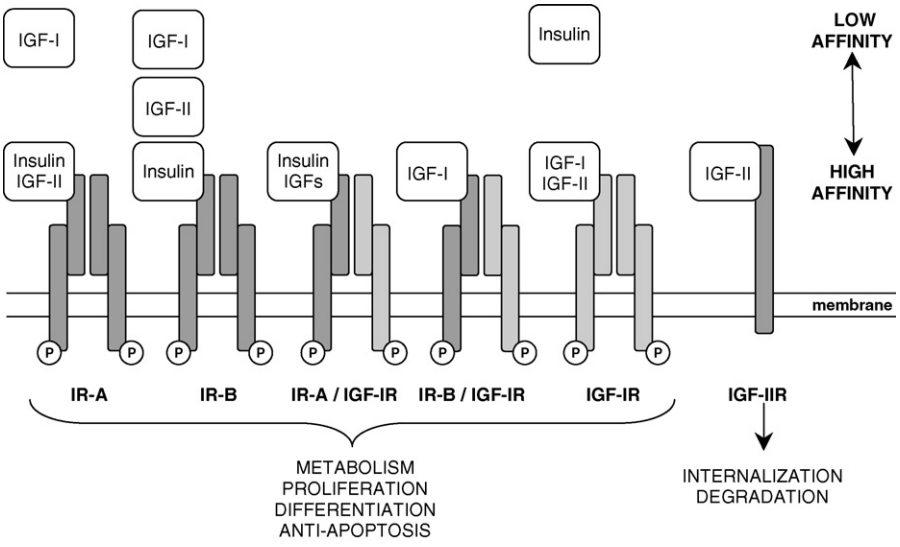


Fig. 2. Overview of the insulin/IGF system. IR exists in two isoforms: IR-A and IR-B. IR-B is responsible for the classic metabolic responses induced by insulin, and also binds IGF1 and IGF2 with low and intermediate affinity, respectively. IR-A has high affinity for insulin and IGF2, and binds IGF1 with low affinity. IGF1R binds the IGFs to stimulate anabolic activity, and also binds insulin at high concentrations. IR-A/IGF1R heterodimers bind insulin and IGFs with similar affinity, whereas IR-B/IGF1R heterodimers bind IGF1 exclusively. IGF2R exclusively binds IGF2 and targets it for degradation. Adapted from Refs. [47,46].

Table 2
IGFBP functions (adapted from Ref. [17])

IGFBP-1	Physiological levels stimulate IGF1 action; molar excess inhibits mitogenic and insulin-like activities of IGF1 and IGF2
IGFBP-2	Inhibitor of IGF-induced DNA synthesis; stimulatory effects on IGF function have also been observed
IGFBP-3	Major carrier of IGFs in serum and modulator of IGF endocrine action; potentiates IGF activity; excessive levels are inhibitory
IGFBP-4	Only IGFBP shown to consistently inhibit IGF action; serum concentration generally low; expression appears to be tissue-specific
IGFBP-5	Inhibitory; association with extracellular matrix (ECM) lowers its affinity for the IGFs, resulting in increased IGF activity
IGFBP-6	Specifically binds IGF2; generally thought to be inhibitory

IGF2 has high affinity for another receptor, IGF2R, and is its principal ligand (Fig. 2). However, IGF2R does not transduce a signal; rather, it serves mainly to limit IGF2 bioavailability by targeting IGF2 for degradation (reviewed in Ref. [48]). Interestingly, the *IGF2R* gene is also imprinted but it is maternally expressed (reviewed in Ref. [16]).

Whereas insulin circulates freely in the bloodstream, the IGFs are found in complexes with the IGF-binding proteins (IGFBPs). Six different IGFBPs have been identified, and each binds the IGFs with significantly higher affinity than IGF1R. The expression patterns of the various IGFBPs differ both spatially and temporally, and they have distinct activities (Table 2). Thus, IGFBPs are important modulators of IGF action, availability, and tissue distribution (reviewed in Ref. [17]). Differential expression of IGFBPs, as well as differential expression of IGF receptors and receptor hybrids, may govern the cell- and tissue-specific actions of IGFs.

3.2. IGF2 in cell growth and differentiation

IGF1 and IGF2 are well known for their mitogenic activities. Almost all cell types express IGF1R, so the IGFs can stimulate growth and differentiation in many tissues (reviewed in Ref. [49]). Upon binding to IGF1R, the IGFs trigger the receptor tyrosine kinase activity, which leads to phosphorylation of itself and its major substrate, the insulin receptor substrate 1 (IRS-1). Phosphorylated IRS-1 can activate the Ras/Raf/MAPK and PI3-kinase/Akt cascades, and depending on the cell type, stimulate proliferation, differentiation, or both (reviewed in Ref. [50]). PI3-kinase activation can lead to anti-apoptotic signals, and components of this pathway are frequently amplified or mutated in cancers (reviewed in Ref. [51]).

The role of IGF2 in muscle development has been studied extensively. IGF2 is upregulated early in MyoD-induced in myocyte differentiation, and signals in an autocrine loop to activate PI3-kinase and Akt [52]. IGF2 inhibition leads to reduced expression of MyoD target genes, which suggests that IGF2 is essential for amplifying and maintaining MyoD efficacy [53]. IGF2 is also essential in bone development, where it promotes proliferation and differentiation of bone cells. Down-regulation of IGF2 most likely accounts for the decrease in bone mass observed with cortisol use [54]. Thus, IGF2 has great therapeutic potential in wound and fracture healing.

Growth in the developing mouse embryo is largely governed by IGF2. When a targeted *Igf2* deletion is transmitted paternally, mouse embryos inherit only the inactive maternal allele and are born runted [14]. Conversely, IGF2 overexpression, achieved by disrupting the inhibitory *Igf2r* [55], by deleting *H19* [56], or by transactivating *Igf2* [57], leads to fetal overgrowth and malformations with characteristics that resemble Beckwith–Wiedemann syndrome (BWS, discussed below).

4. IGF2 and disease

4.1. Loss of IGF2 imprinting

IGF2 is regulated precisely to ensure monoallelic expression in most tissues [19], which emphasizes the importance of gene dosage. Normal development requires accurate expression, and many disorders can be attributed to an abnormally high dose of *IGF2* caused by loss of imprinting (LOI). BWS is one such disease, characterized by fetal and neonatal overgrowth, and is often accompanied by an increased risk of childhood cancers (reviewed in Ref. [58]). BWS patients almost always have mutations in the chromosome 11p15.5 region, a large cluster of imprinted genes that includes *IGF2* and $p57^{KIP2}$ (Fig. 3). Most of these mutations affect imprinting; quite often, biallelic *IGF2* expression and *H19* methylation are observed (reviewed in Ref. [16]). BWS usually occurs sporadically, but in rare familial cases *IGF2* LOI may be caused by deletions of the CTCF binding sites in the maternal *IGF2/H19* ICR [59,60].

Disrupted imprinting is perhaps the most common observation in cancer (reviewed in Ref. [61]), and *IGF2* overexpression is a recurring theme. Wilms' tumor, a childhood cancer of the kidney, is often associated with defects in the WT1 gene, which encodes a transcriptional repressor of *IGF2* [62]. Wilms' tumor is also associated with mutations in the 11p15.5 region that affect *IGF2* imprinting: altered *IGF2* expression accounts for nearly 50% of all cases of Wilms' tumor, and *IGF2* LOI is found in the vast majority (90%) of pathological cases [63]. *IGF2* LOI has also been

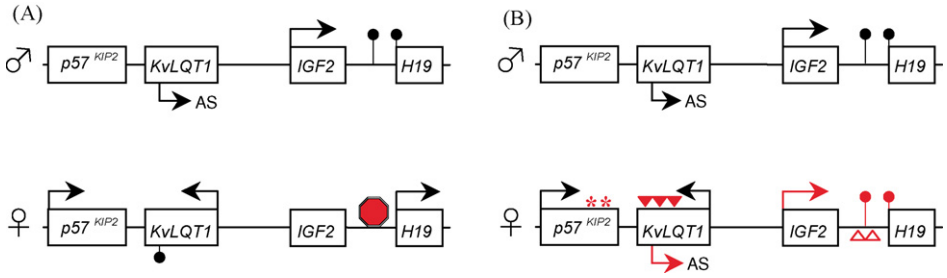


Fig. 3. (A) Normal and (B) BWS gene expression patterns on chromosome 11p15.5. Arrows represent active genes. Lollipops: methylated CpGs. Red octagon: CTCF. Asterisks: point mutations. Filled triangles: translocation breakpoints. Open triangles: deletions. Adapted from Refs. [16,59,60].

observed in many other cancers. Both benign and malignant breast lesions show biallelic *IGF2* expression, and altered imprinting of *IGF2* has been identified in hepatoblastoma, lung cancer, cervical carcinoma, rhabdomyosarcoma, choriocarcinoma, and testicular cancer ([64] and references therein).

The epigenetic mutations associated with cancer, such as aberrant methylation or LOI, may magnify the effects of genetic mutations or even have causal roles. In either case, epigenetic changes have potential value for assessing disease risk and prognosis. In a mouse model of intestinal cancer, where the *adenomatous polyposis coli* (*Apc*) gene is mutated, supplementary *Igf2* LOI increases the incidence of intestinal hyperplasia. The clinical relevance of this is corroborated by the fact that patients with *IGF2* LOI also have an increased risk of developing colorectal cancer [65]. Alterations involving CTCF may also be informative. Elevated CTCF expression levels have been reported in breast cancer, where it is postulated to have anti-apoptotic actions [66]. Gene activation by a CTCF homolog is observed in lung cancer [67,68], and methylation changes in CTCF binding sites have also been reported in osteosarcoma [69]. Because epigenetic changes such as LOI and demethylation are among the earliest events in cancer progression (reviewed in Ref. [70]), assays for epigenetic biomarkers may allow for early detection, prevention, and treatment of cancer.

4.2. IGF2 and other signaling pathways in disease pathogenesis

Igf2 overexpression sometimes occurs without apparent LOI or gene duplication. Other factors, such as sonic hedgehog (Shh), can also transcriptionally activate *Igf2*. Shh is a developmental morphogen involved with patterning and organ specification, and its signaling pathway is mutated in several diseases (reviewed in Ref. [71]). The Shh cascade culminates in the activation of Gli, a transcription factor that induces several target genes (Fig. 4).

Shh has been demonstrated to upregulate *Igf2* both *in vitro* and *in vivo*. When mouse mesenchymal cells are treated with Shh or transfected with *Gli1*, *Igf2* mRNA is upregulated [72]. A *Ptc*-deficient mutation in mice, which results in constitutive Gli activation, increases IGF2 protein levels and also the formation of medulloblastomas and rhabdomyosarcomas [73]. It is not entirely clear how Shh induces *Igf2* expression. Though putative Gli-binding sites have been identified in the mouse *Igf2* promoter [72], it is not known whether these sites are functional, or if they exist in the human VEGF promoter. However, functional Gli sites have been documented in the human *IGFBP-6* promoter [74]. IGFBP-6 specifically binds IGF2 and is generally thought to have anti-proliferative properties. Nonetheless, like most of the IGFBPs (Table 2), IGFBP-6 can have contrasting activities, and has also been shown to be anti-apoptotic and tumorigenic (reviewed in Ref. [75]).

IGF2 itself may provide an oncogenic signal in some systems, such as the mouse mammary gland, where transgenic *Igf2* overexpression induces adenocarcinomas [76]. In mouse models of rhabdomyosarcoma and medulloblastoma, *Igf2* alone is insufficient to generate tumors; however, it can enhance the tumorigenic potential of Shh [73,77]. Interestingly, tumors often overexpress the IR-A variant, which binds IGF2 with high affinity; thus, concomitant IGF2 and IR-A overexpression can potentially generate an autoproliferative loop [30]. Taken together, these observations substantiate the hypothesis that IGF2 can supply the ''second hit'' necessary for oncogene-induced tumors [78].

4.3. IGF2 and angiogenesis

Angiogenesis, or blood vessel growth, is another critical element of tumor progression that may involve IGF2. Oxygen, nutrients, and metabolic wastes can simply diffuse in and out of small tumors, but growth beyond a critical size (1 mm^3) requires a vascular network (reviewed in Ref. [79]). Areas of hypoxia within tumors induce the expression of

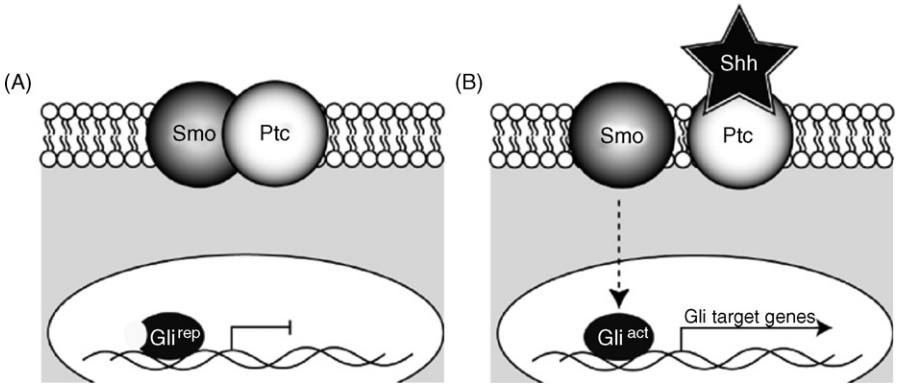


Fig. 4. The Shh signaling pathway. (A) In the absence of signal, the receptor patched (Ptc) is complexed with smoothened (Smo), and Gli exists in a truncated form that acts as a transcriptional repressor [99,100]. (B) When bound by Shh, Ptc releases Smo, which signals to produce a full-length Gli activator protein. Gli target genes include *Gli*, *Ptc*, and genes involved in proliferation and morphogenesis.

angiogenic factors, which prompt an influx of vessels from surrounding tissues. Neovascularization also facilitates the spread of cancer cells to other tissues; thus, there is a correlation between high metastatic potential and tumor vascularity (reviewed in Ref. [80]).

Vascular endothelial growth factor (VEGF) has a central function in both normal and pathological neovascularization, and its expression is upregulated in tumors (reviewed in Ref. [81]). Hypoxia-inducible factors (HIFs) are principle mediators of VEGF upregulation, though VEGF mRNA levels are also increased via message stabilization [82]. Transcriptional regulation also occurs through other *cis* elements in the VEGF promoter, and can be instigated by various growth factors, hormones, and oncogenes (reviewed in Ref. [83]).

Though studies of the IGFs in vascular development are limited, IGF2 may participate in angiogenesis through its ability to upregulate VEGF. In hepatocellular carcinomas cells, hypoxia-induced VEGF expression is increased by IGF2, which is itself upregulated by HIFs [84]**.** Other studies have suggested that IGF2 signaling upregulates VEGF in part by increasing HIF levels [85,86]. Because reciprocal upregulation of IGF2 and HIF has been demonstrated [87], they may act in synergy to induce VEGF expression. Though the mechanisms remain unclear, the ability to induce VEGF accentuates the importance of IGF2 in tumor development.

IGF2 may also be involved in the pathological neovascularization that characterizes proliferative diabetic retinopathy (PDR) and retinopathy of prematurity (ROP). Several studies have implicated IGF1 in retinopathy (reviewed in Ref. [88]), but IGF2 has been largely overlooked—despite reports of 10- to 30-fold more IGF2 in the vitreous of diabetic patients than IGF1 ([89] and references therein). A recent study showed that IGFBP-3 suppressed retinal neovascularization irrespective of IGF1 levels [90], which supported the long-standing notion that IGFBPs can act independently of IGF signaling through IGF1R (reviewed in Ref. [75]). However, the potential contribution of IGF2 needs to be examined—specifically, its interactions with other receptors (such as IR-A variant) and whether these interactions are subject to IGFBP regulation. Clearly, the likely role of IGF2 in retinopathy calls for further exploration.

4.4. IGF2: the missing link between Shh and angiogenesis?

In recent years, Shh has been identified as an angiogenic factor. Studies in zebrafish reveal vascular defects in Shh-mutant embryos [91,92], and place Shh upstream of VEGF signaling during arterial differentiation [93]. The cascades induced by Shh also appear to regulate vessel formation in mammals. In the mouse embryo, Indian hedgehog (Ihh), a Shh homolog, has been suggested to be critical for early vasculogenesis [94,95]. In *Shh*-deficient mice, the developing lung is poorly vascularized [96]; conversely, *Shh* overexpression in the neural tube results in hypervascularization [97]. Shh can also induce angiogenic factors (including VEGF) and promote neovascularization in adult mice [98]. Thus, vessel formation may depend on the ability of Shh to induce VEGF. Though the exact mechanism remains elusive, it may very well involve IGF2, which is a downstream target of the Shh cascade [72] and has a demonstrated ability to synergize with Shh [73,77]. Moreover, IGF2 has also been shown to induce VEGF [84–86]. Thus, IGF2 may mediate the angiogenic effects of Shh, and provide the critical link between Shh and VEGF.

5. Conclusions

Though interest in IGF2 has been somewhat skewed towards the study of gene regulation and imprinting, it is likely to attract attention from other fields as studies implicate IGF2 in an increasing number of diseases. The complexity of IGF2 regulation indicates that overexpression can occur at multiple levels. Since IGF2 is pivotal in many developmental and pathological processes, its multifaceted regulation presents a number of potential therapeutic targets.

Because imprinting defects are now recognized as common in the pathogenesis of cancer, the mechanisms surrounding *IGF2* imprinting are likely to gain interest as well. Perhaps the most thoroughly studied of known imprinted genes, *IGF2* has yielded valuable insight into other epigenetic gene regulatory mechanisms—namely X chromosome inactivation, which also gained significance with the discovery of X-linked tumor suppressors (reviewed in Ref. [37]). These studies highlight the multifactorial nature of cancer, in which IGF2 may have a pivotal role. More importantly, they suggest that imprinting and X inactivation are not just interesting epigenetic phenomena, but have considerable functional relevance.

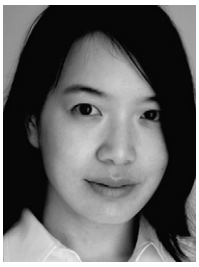
References

[1] Carrel A. Artificial activation of the growth in vitro of connective tissue. J Exp Med 1913;17:14–9.
[2] Westwood M, Jabbour H, Bloom S, Whitehead S, Barber T, Chapman J, et al. Ten hot hormones. Endocrinologist)2005;(75) [Spring 2005, Feature].
[3] Salmon Jr WD, Daughaday WH. A hormonally controlled serum factor which stimulates sulfate incorporation by cartilage in vitro. J Lab Clin Med 1957;49(6):825–36.
[4] Froesch ER, Buergi H, Ramseier EB, Bally P, Labhart A. Antibody-suppressible and nonsuppressible insulin-like activities in human serum and their physiologic significance. an insulin assay with adipose tissue of increased precision and specificity. J Clin Invest 1963;42:1816–34.
[5] Daughaday WH, Hall K, Raben MS, Salmon Jr WD, van den Brande JL, van Wyk JJ. Somatomedin: proposed designation for sulphation factor. Nature 1972;235(5333):107.
[6] Dulak NC, Temin HM. Multiplication-stimulating activity for chicken embryo fibroblasts from rat liver cell conditioned medium: a family of small polypeptides. J Cell Physiol 1973;81(2):161–70.

[7] Rinderknecht E, Humbel RE. Amino-terminal sequences of two polypeptides from human serum with nonsuppressible insulin-like and cell-growth-promoting activities: evidence for structural homology with insulin B chain. Proc Natl Acad Sci USA 1976;73(12):4379–81.
[8] Marquardt H, Todaro GJ, Henderson LE, Oroszlan S. Purification and primary structure of a polypeptide with multiplication-stimulating activity from rat liver cell cultures. Homology with human insulin-like growth factor II. J Biol Chem 1981;256(13):6859–65.
[9] Daughaday WH, Hall K, Salmon Jr WD, Van den Brande JL, Van Wyk JJ. On the nomenclature of the somatomedins and insulin-like growth factors. Endocrinology 1987;121(5):1911–2.
[10] Zhou Y, Xu BC, Maheshwari HG, He L, Reed M, Lozykowski M, et al. A mammalian model for Laron syndrome produced by targeted disruption of the mouse growth hormone receptor/binding protein gene (the Laron mouse). Proc Natl Acad Sci USA 1997;94(24):13215–20.
[11] Denley A, Wang CC, McNeil KA, Walenkamp MJ, van Duyvenvoorde H, Wit JM, et al. Structural and functional characteristics of the Val44Met insulin-like growth factor I missense mutation: correlation with effects on growth and development. Mol Endocrinol (Baltimore MD) 2005;19(3):711–21.
[12] Powell-Braxton L, Hollingshead P, Warburton C, Dowd M, Pitts-Meek S, Dalton D, et al. IGF-I is required for normal embryonic growth in mice. Genes Dev 1993;7(12B):2609–17.
[13] Baker J, Liu JP, Robertson EJ, Efstratiadis A. Role of insulin-like growth factors in embryonic and postnatal growth. Cell 1993;75(1):73–82.
[14] DeChiara TM, Robertson EJ, Efstratiadis A. Parental imprinting of the mouse insulin-like growth factor II gene. Cell 1991;64(4):849–59.
[15] LeRoith D, Roberts Jr CT. The insulin-like growth factor system and cancer. Cancer Lett 2003;195(2):127–37.
[16] Reik W, Constancia M, Dean W, Davies K, Bowden L, Murrell A, et al. Igf2 imprinting in development and disease. Int J Dev Biol 2000;44(1):145–50.
[17] Clemmons DR. Insulin-like growth factor binding proteins and their role in controlling IGF actions. Cytokine Growth Factor Rev 1997;8(1):45–62.
[18] Constancia M, Hemberger M, Hughes J, Dean W, Ferguson-Smith A, Fundele R, et al. Placental-specific IGF-II is a major modulator of placental and fetal growth. Nature 2002;417(6892):945–8.
[19] Ohlsson R, Hedborg F, Holmgren L, Walsh C, Ekstrom TJ. Overlapping patterns of IGF2 and H19 expression during human development: biallelic IGF2 expression correlates with a lack of H19 expression. Development (Cambridge England) 1994;120(2):361–8.
[20] de Pagter-Holthuizen P, Jansen M, van Schaik FM, van der Kammen R, Oosterwijk C, Van den Brande JL, et al. The human insulin-like growth factor II gene contains two development-specific promoters. FEBS Lett 1987;214(2):259–64.
[21] Rotwein P, Hall LJ. Evolution of insulin-like growth factor II: characterization of the mouse IGF-II gene and identification of two pseudo-exons. DNA Cell Biol 1990;9(10):725–35.
[22] Lewis A, Reik W. How imprinting centres work. Cytogenet Genome Res 2006;113(1–4):81–9.
[23] Thorvaldsen JL, Duran KL, Bartolomei MS. Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2. Genes Dev 1998;12(23):3693–702.
[24] Ohlsson R, Renkawitz R, Lobanenkov V. CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. Trends Genet 2001;17(9):520–7.
[25] Bell AC, Felsenfeld G. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature 2000;405(6785):482–5.
[26] Chao W, Huynh KD, Spencer RJ, Davidow LS, Lee JT. CTCF, a candidate trans-acting factor for X-inactivation choice.. Science (New York NY) 2002;295(5553):345–7.
[27] Boumil RM, Ogawa Y, Sun BK, Huynh KD, Lee JT. Differential methylation of Xite and CTCF sites in Tsix mirrors the pattern of X-inactivation choice in mice. Mol Cell Biol 2006;26(6):2109–17.
[28] Stavropoulos N, Rowntree RK, Lee JT. Identification of developmentally specific enhancers for Tsix in the regulation of X chromosome inactivation. Mol Cell Biol 2005;25(7):2757–69.
[29] Lyon MF. Gene action in the X-chromosome of the mouse (*Mus musculus* L.). Nature 1961;190:372–3.
[30] Thorvaldsen JL, Verona RI, Bartolomei MS. X-tra! X-tra! News from the mouse X chromosome. Dev Biol 2006;298(2):344–53.
[31] Lee JT, Davidow LS, Warshawsky D, Tsix. a gene antisense to Xist at the X-inactivation centre. Nat Genet 1999;21(4):400–4.
[32] Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RD, et al. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell 2007;128(6):1231–45.
[33] Do Kim J, Hinz AK, Ha Choo J, Stubbs L, Kim J. YY1 as a controlling factor for the Peg3 and Gnas imprinted domains. Genomics 2007;89(2):262–9.
[34] Donohoe ME, Zhang LF, Xu N, Shi Y, Lee JT. Identification of a Ctcf cofactor, Yy1, for the X chromosome binary switch. Mol Cell 2007;25(1):43–56.
[35] Ling JQ, Li T, Hu JF, Vu TH, Chen HL, Qiu XW, et al. CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. Science (New York NY) 2006;312(5771):269–72.
[36] Lupski JR, Stankiewicz P. Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. PLoS Genet 2005;1(6):e49.
[37] Medema RH, Burgering BM. The X Factor: skewing X Inactivation towards Cancer. Cell 2007;129(7):1253–4.
[38] Zuo T, Wang L, Morrison C, Chang X, Zhang H, Li W, et al. FOXP3 is an X-linked breast cancer suppressor gene and an important repressor of the HER-2/ErbB2 oncogene. Cell 2007;129(7):1275–86.
[39] Foulstone E, Prince S, Zaccheo O, Burns JL, Harper J, Jacobs C, et al. Insulin-like growth factor ligands, receptors, and binding proteins in cancer. J Pathol 2005;205(2):145–53.
[40] Rivera MN, Kim WJ, Wells J, Driscoll DR, Brannigan BW, Han M, et al. An X chromosome gene, WTX, is commonly inactivated in Wilms tumor. Science (New York NY) 2007;315(5812):642–5.
[41] Moore T, Constancia M, Zubair M, Bailleul B, Feil R, Sasaki H, et al. Multiple imprinted sense and antisense transcripts, differential methylation and tandem repeats in a putative imprinting control region upstream of mouse Igf2. Proc Natl Acad Sci USA 1997;94(23):12509–14.
[42] Okutsu T, Kuroiwa Y, Kagitani F, Kai M, Aisaka K, Tsutsumi O, et al. Expression and imprinting status of human PEG8/IGF2AS, a paternally expressed antisense transcript from the IGF2 locus, in Wilms' tumors. J Biochem (Tokyo) 2000;127(3):475–83.
[43] Vu TH, Chuyen NV, Li T, Hoffman AR. Loss of imprinting of IGF2 sense and antisense transcripts in Wilms' tumor. Cancer Res 2003;63(8):1900–5.
[44] Denley A, Wallace JC, Cosgrove LJ, Forbes BE. The insulin receptor isoform exon 11- (IR-A) in cancer and other diseases: a review. Horm Metab Res Horm Stoffwechselforschung 2003;35(11–12):778–85.
[45] Denley A, Bonython ER, Booker GW, Cosgrove LJ, Forbes BE, Ward CW, et al. Structural determinants for high-affinity binding of insulin-like growth factor II to insulin receptor (IR)-A, the exon 11 minus isoform of the IR. Mol Endocrinol (Baltimore MD) 2004;18(10):2502–12.
[46] White MF. Regulating insulin signaling and beta-cell function through IRS proteins. Can J Physiol Pharmacol 2006;84(7):725–37.
[47] Denley A, Cosgrove LJ, Booker GW, Wallace JC, Forbes BE. Molecular interactions of the IGF system. Cytokine Growth Factor Rev 2005;16(4–5):421–39.
[48] Scott CD, Firth SM. The role of the M6P/IGF-II receptor in cancer: tumor suppression or garbage disposal? Horm Metab Res Horm Stoffwechselforschung 2004;36(5):261–71.

[49] Petley T, Graff K, Jiang W, Yang H, Florini J. Variation among cell types in the signaling pathways by which IGF-I stimulates specific cellular responses. Horm Metab Res Horm Stoffwechselforschung 1999;31(2–3):70–6.
[50] Baserga R, Peruzzi F, Reiss K. The IGF-1 receptor in cancer biology. Int J Cancer 2003;107(6):873–7.
[51] Hennessy BT, Smith DL, Ram PT, Lu Y, Mills GB. Exploiting the PI3K/AKT pathway for cancer drug discovery. Nat Rev Drug Discov 2005;4(12):988–1004.
[52] Wilson EM, Hsieh MM, Rotwein P. Autocrine growth factor signaling by insulin-like growth factor-II mediates MyoD-stimulated myocyte maturation. J Biol Chem 2003;278(42):41109–13.
[53] Wilson EM, Rotwein P. Control of MyoD function during initiation of muscle differentiation by an autocrine signaling pathway activated by insulin-like growth factor-II. J Biol Chem 2006;281(40):29962–71.
[54] Minuto F, Palermo C, Arvigo M, Barreca AM. The IGF system and bone. J Endocrinol Invest 2005;28(Suppl. 8):8–10.
[55] Wang ZQ, Fung MR, Barlow DP, Wagner EF. Regulation of embryonic growth and lysosomal targeting by the imprinted Igf2/Mpr gene. Nature 1994;372(6505):464–7.
[56] Leighton PA, Ingram RS, Eggenschwiler J, Efstratiadis A, Tilghman SM. Disruption of imprinting caused by deletion of the H19 gene region in mice. Nature 1995;375(6526):34–9.
[57] Sun FL, Dean WL, Kelsey G, Allen ND, Reik W. Transactivation of Igf2 in a mouse model of Beckwith-Wiedemann syndrome. Nature 1997;389(6653):809–15.
[58] Morison IM, Reeve AE. Insulin-like growth factor 2 and overgrowth: molecular biology and clinical implications. Mol Med Today 1998;4(3):110–5.
[59] Sparago A, Cerrato F, Vernucci M, Ferrero GB, Silengo MC, Riccio A. Microdeletions in the human H19 DMR result in loss of IGF2 imprinting and Beckwith-Wiedemann syndrome. Nat Genet 2004;36(9):958–60.
[60] Prawitt D, Enklaar T, Gartner-Rupprecht B, Spangenberg C, Oswald M, Lausch E, et al. Microdeletion of target sites for insulator protein CTCF in a chromosome 11p15 imprinting center in Beckwith-Wiedemann syndrome and Wilms' tumor. Proc Natl Acad Sci USA 2005;102(11):4085–90.
[61] Feinberg AP, Ohlsson R, Henikoff S. The epigenetic progenitor origin of human cancer. Nat Rev 2006;7(1):21–33.
[62] Bae SK, Bae MH, Ahn MY, Son MJ, Lee YM, Bae MK, et al. Egr-1 mediates transcriptional activation of IGF-II gene in response to hypoxia. Cancer Res 1999;59(23):5989–94.
[63] Ravenel JD, Broman KW, Perlman EJ, Niemitz EL, Jayawardena TM, Bell DW, et al. Loss of imprinting of insulin-like growth factor-II (IGF2) gene in distinguishing specific biologic subtypes of Wilms tumor. J Natl Cancer Inst 2001;93(22):1698–703.
[64] McCann AH, Miller N, O'Meara A, Pedersen I, Keogh K, Gorey T, et al. Biallelic expression of the IGF2 gene in human breast disease. Hum Mol Genet 1996;5(8):1123–7.
[65] Kaneda A, Feinberg AP. Loss of imprinting of IGF2: a common epigenetic modifier of intestinal tumor risk. Cancer Res 2005;65(24):11236–40.
[66] Docquier F, Farrar D, D'Arcy V, Chernukhin I, Robinson AF, Loukinov D, et al. Heightened expression of CTCF in breast cancer cells is associated with resistance to apoptosis. Cancer Res 2005;65(12):5112–22.
[67] Hong JA, Kang Y, Abdullaev Z, Flanagan PT, Pack SD, Fischette MR, et al. Reciprocal binding of CTCF and BORIS to the NY-ESO-1 promoter coincides with derepression of this cancer-testis gene in lung cancer cells. Cancer Res 2005;65(17):7763–74.
[68] Vatolin S, Abdullaev Z, Pack SD, Flanagan PT, Custer M, Loukinov DI, et al. Conditional expression of the CTCF-paralogous transcriptional factor BORIS in normal cells results in demethylation and derepression of MAGE-A1 and reactivation of other cancer-testis genes. Cancer Res 2005;65(17):7751–62.
[69] Ulaner GA, Vu TH, Li T, Hu JF, Yao XM, Yang Y, et al. Loss of imprinting of IGF2 and H19 in osteosarcoma is accompanied by reciprocal methylation changes of a CTCF-binding site. Hum Mol Genet 2003;12(5):535–49.
[70] Jelinic P, Shaw P. Loss of imprinting and cancer. J Pathol 2007;211(3):261–8.
[71] Ruiz i Altaba A, Sanchez P, Dahmane N. Gli and hedgehog in cancer: tumours, embryos and stem cells. Nat Rev Cancer 2002;2(5):361–72.
[72] Ingram WJ, Wicking CA, Grimmond SM, Forrest AR, Wainwright BJ. Novel genes regulated by Sonic Hedgehog in pluripotent mesenchymal cells. Oncogene 2002;21(53):8196–205.
[73] Hahn H, Wojnowski L, Specht K, Kappler R, Calzada-Wack J, Potter D, et al. Patched target Igf2 is indispensable for the formation of medulloblastoma and rhabdomyosarcoma. J Biol Chem 2000;275(37):28341–4.
[74] Yoon JW, Kita Y, Frank DJ, Majewski RR, Konicek BA, Nobrega MA, et al. Gene expression profiling leads to identification of GLI1-binding elements in target genes and a role for multiple downstream pathways in GLI1-induced cell transformation. J Biol Chem 2002;277(7):5548–55.
[75] Firth SM, Baxter RC. Cellular actions of the insulin-like growth factor binding proteins. Endocr Rev 2002;23(6):824–54.
[76] Bates P, Fisher R, Ward A, Richardson L, Hill DJ, Graham CF. Mammary cancer in transgenic mice expressing insulin-like growth factor II (IGF-II). Br J Cancer 1995;72(5):1189–93.
[77] Rao G, Pedone CA, Valle LD, Reiss K, Holland EC, Fults DW. Sonic hedgehog and insulin-like growth factor signaling synergize to induce medulloblastoma formation from nestin-expressing neural progenitors in mice. Oncogene 2004;23(36):6156–62.
[78] Christofori G, Naik P, Hanahan D. A second signal supplied by insulin-like growth factor II in oncogene-induced tumorigenesis. Nature 1994;369(6479):414–8.
[79] Sivridis E, Giatromanolaki A, Koukourakis MI. The vascular network of tumours—what is it not for? J Pathol 2003;201(2):173–80.
[80] Semenza GL. Regulation of mammalian O_2 homeostasis by hypoxia-inducible factor 1. Annu Rev Cell Dev Biol 1999;15:551–78.
[81] Dvorak HF. Vascular permeability factor/vascular endothelial growth factor: a critical cytokine in tumor angiogenesis and a potential target for diagnosis and therapy. J Clin Oncol 2002;20(21):4368–80.
[82] Shima DT, Deutsch U, D'Amore PA. Hypoxic induction of vascular endothelial growth factor (VEGF) in human epithelial cells is mediated by increases in mRNA stability. FEBS Lett 1995;370(3):203–8.
[83] Loureiro RM, D'Amore PA. Transcriptional regulation of vascular endothelial growth factor in cancer. Cytokine Growth Factor Rev 2005;16(1):77–89.
[84] Kim KW, Bae SK, Lee OH, Bae MH, Lee MJ, Park BC. Insulin-like growth factor II induced by hypoxia may contribute to angiogenesis of human hepatocellular carcinoma. Cancer Res 1998;58(2):348–51.
[85] Kwon YW, Kwon KS, Moon HE, Park JA, Choi KS, Kim YS, et al. Insulin-like growth factor-II regulates the expression of vascular endothelial growth factor by the human keratinocyte cell line HaCaT. J Invest Dermatol 2004;123(1):152–8.
[86] Kim HJ, Kim TY. Regulation of vascular endothelial growth factor expression by insulin-like growth factor-II in human keratinocytes, differential involvement of mitogen-activated protein kinases and feedback inhibition of protein kinase C. Br J Dermatol 2005;152(3):418–25.
[87] Feldser D, Agani F, Iyer NV, Pak B, Ferreira G, Semenza GL. Reciprocal positive regulation of hypoxia-inducible factor 1alpha and insulin-like growth factor 2. Cancer Res 1999;59(16):3915–8.
[88] Chen J, Smith LE. Retinopathy of prematurity. Angiogenesis 2007;10(2):133–40.
[89] Guidry C, Feist R, Morris R, Hardwick CW. Changes in IGF activities in human diabetic vitreous. Diabetes 2004;53(9):2428–35.
[90] Lofqvist C, Chen J, Connor KM, Smith AC, Aderman CM, Liu N, et al. From the Cover: IGFBP3 suppresses retinopathy through

suppression of oxygen-induced vessel loss and promotion of vascular regrowth. Proc Natl Acad Sci USA 2007;104(25):10589–94.
[91] Chen JN, Haffter P, Odenthal J, Vogelsang E, Brand M, van Eeden FJ, et al. Mutations affecting the cardiovascular system and other internal organs in zebrafish. Development (Cambridge England) 1996;123: 293–302.
[92] Brown LA, Rodaway AR, Schilling TF, Jowett T, Ingham PW, Patient RK, et al. Insights into early vasculogenesis revealed by expression of the ETS-domain transcription factor Fli-1 in wild-type and mutant zebrafish embryos. Mech Dev 2000;90(2):237–52.
[93] Lawson ND, Vogel AM, Weinstein BM. Sonic hedgehog and vascular endothelial growth factor act upstream of the Notch pathway during arterial endothelial differentiation. Dev Cell 2002;3(1): 127–36.
[94] Dyer MA, Farrington SM, Mohn D, Munday JR, Baron MH. Indian hedgehog activates hematopoiesis and vasculogenesis and can respecify prospective neurectodermal cell fate in the mouse embryo. Development (Cambridge England) 2001;128(10):1717–30.
[95] Byrd N, Becker S, Maye P, Narasimhaiah R, St-Jacques B, Zhang X, et al. Hedgehog is required for murine yolk sac angiogenesis. Development (Cambridge England) 2002;129(2):361–72.
[96] Pepicelli CV, Lewis PM, McMahon AP. Sonic hedgehog regulates branching morphogenesis in the mammalian lung. Curr Biol 1998;8(19):1083–6.
[97] Rowitch DH, St.-Jacques B, Lee SM, Flax JD, Snyder EY, McMahon AP. Sonic hedgehog regulates proliferation and inhibits differentiation of CNS precursor cells. J Neurosci 1999;19(October (20)): 8954–65.
[98] Pola R, Ling LE, Silver M, Corbley MJ, Kearney M, Blake Pepinsky R, et al. The morphogen Sonic hedgehog is an indirect angiogenic agent upregulating two families of angiogenic growth factors. Nat Med 2001;7(6):706–11.
[99] Ruiz i Altaba A. Gli proteins encode context-dependent positive and negative functions: implications for development and disease. Development (Cambridge England) 1999;126(14):3205–16.
[100] Sasaki H, Nishizaki Y, Hui C, Nakafuku M, Kondoh H. Regulation of Gli2 and Gli3 activities by an amino-terminal repression domain: implication of Gli2 and Gli3 as primary mediators of Shh signaling. Development (Cambridge England) 1999;126(17):3915–24.



Wendy Chao received her PhD in genetics from Harvard Medical School in 2007. She is an editor at Natural Standard Research Collaboration in Cambridge, MA, which provides evidence-based information about integrative therapies to Harvard Medical School, National Institutes of Health, Susan G. Komen Foundation, and others. Dr. Chao is a contributing editor for *The Scientist* magazine, and has worked with the Journal of Visualized Experiments and the Foundation for Art and Healing. She has appeared on ABC News, The Learning Channel, and CNN.

Patricia D'Amore received her PhD in biology from Boston University in 1977. She was a postdoctoral fellow at Johns Hopkins Medical School before moving to the Children's Hospital in Boston where she is currently a research associate in surgery. In 1998, she became professor of ophthalmology (Pathology) at Harvard Medical School and a senior scientist at the Schepens Eye Research Institute. She is the recipient of numerous awards including the Jules & Doris Stein Research to Prevent Blindness, Senior Scientific Investigator Award, Cogan Award, and the A. Clifford Barger Excellence in Mentoring Award. She is currently the Associate Director of Research and the Ankeny Scholar of Retinal Molecular Biology at Schepens. Dr. D'Amore's research focuses on understanding the mechanism of vascular growth and development. She is the author of 112 publications.

ments for the complex pattern of transcriptional regulation of the *py235* genes remain to be elucidated. Py235 proteins have previously been shown to be involved in red blood cell invasion. Because a subset of these proteins is expressed in the sporozoite and is the target of antibodies that inhibit hepatocyte invasion, these proteins may be important in the recognition and/or invasion of the mosquito salivary glands and the liver. Merozoites released from both the liver and the infected erythrocyte invade red blood cells, so the need to express a distinct set of *py235* genes in the infected hepatocyte is puzzling. This differential expression of *py235* in the hepatic schizont reinforces the idea that the obligatory passage of the parasite through the liver not only amplifies the number of parasites injected by the mosquito but also preadapts the parasite to invade red blood cells. The presence of distinct rhoptry proteins in the sporozoite and the liver-stage malaria parasite may form the basis of an efficient vaccination strategy to target these pre-erythrocytic–stage parasites, which are present in small numbers and are at their most vulnerable. Conserved regions of the rhoptry proteins that are the target of protective immune responses may also form the basis of a vaccine against both pre-erythrocytic– and erythrocytic-stage parasites.

References and Notes

1. R. R. Freeman, A. J. Trejdosiewicz, G. A. Cross, *Nature* **284**, 366 (1980).
2. A. A. Holder, R. R. Freeman, *Nature* **294**, 361 (1981).
3. S. A. Ogun, A. A. Holder, *Mol. Biochem. Parasitol.* **76**, 321 (1996).
4. P. R. Preiser, W. Jarra, T. Capiod, G. Snounou, *Nature* **398**, 618 (1999).
5. G. Snounou, W. Jarra, P. R. Preiser, *Parasitol. Today* **16**, 28 (2000).
6. A. A. Holder, R. R. Freeman, *Philos. Trans. R. Soc. London Ser. B* **307**, 171 (1984).
7. M. R. Galinski, C. C. Medina, P. Ingravallo, J. W. Barnwell, *Cell* **69**, 1213 (1992).
8. M. R. Galinski, J. W. Barnwell, *Parasitol. Today* **12**, 20 (1996).
9. C. A. Owen, K. A. Sinha, J. K. Keen, S. A. Ogun, A. A. Holder, *Mol. Biochem. Parasitol.* **99**, 183 (1999).
10. M. R. Galinski, M. Xu, J. W. Barnwell, *Mol. Biochem. Parasitol.* **108**, 257 (2000).
11. J. C. Rayner, M. R. Galinski, P. Ingravallo, J. W. Barnwell, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9648 (2000).
12. H. M. Taylor *et al.*, *Infect. Immun.* **69**, 3635 (2001).
13. T. Triglia *et al.*, *Infect. Immun.* **69**, 1084 (2001).
14. M. Marussig *et al.*, *Int. Immunol.* **9**, 1817 (1997).
15. P. R. Preiser, W. Jarra, *Exp. Parasitol.* **89**, 50 (1998).
16. P. R. Preiser *et al.*, data not shown.
17. S. Khan, W. Jarra, H. Bayele, P. R. Preiser, *Mol. Biochem. Parasitol.* **114**, 197 (2001).
18. D. L. Narum, J. L. Green, S. A. Ogun, A. A. Holder, *Mol. Biochem. Parasitol.* **112**, 193 (2001).
19. J. Thompson *et al.*, *Mol. Microbiol.* **31**, 253 (1999).
20. Q. Chen *et al.*, *Nature* **394**, 392 (1998).
21. A. Scherf *et al.*, *EMBO J.* **17**, 5418 (1998).
22. R. E. Hayward, B. Tiwari, K. P. Piper, D. I. Baruch, K. P. Day, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4563 (1999).
23. J. D. Barry, R. McCulloch, *Adv. Parasitol.* **49**, 1 (2001).
24. P. Borst, S. Ulbert, *Mol. Biochem. Parasitol.* **114**, 17 (2001).
25. J. K. Keen, A. A. Holder, J. H. L. Playfair, M. J. Lockyer, A. P. Lewis, *Mol. Biochem. Parasitol.* **42**, 241 (1990).
26. S. A. Ogun, A. A. Holder, *Exp. Parasitol.* **79**, 270 (1994).
27. L. Rénia *et al.*, *Eur. J. Immunol.* **20**, 1445 (1990).
28. K. A. Sinha, J. K. Keen, S. A. Ogun, A. A. Holder, *Mol. Biochem. Parasitol.* **76**, 329 (1996).
29. Supplementary data are available on *Science* Online at www.sciencemag.org/cgi/content/full/295/5553/342/DC1.
30. J. Keen, K. A. Sinha, K. N. Brown, A. A. Holder, *Mol. Biochem. Parasitol.* **65**, 171 (1994).
31. L. Rénia, unpublished observations.
32. I. Landau, P. Gautret, in *Malaria: Parasite Biology, Pathogenesis, and Protection*, I. W. Sherman, Ed. (American Society for Microbiology, Washington, DC, 1998), pp. 401–417.
33. G. Snounou, S. Viriyakosol, W. Jarra, S. Thaithong, K. N. Brown, *Mol. Biochem. Parasitol.* **58**, 283 (1993).
34. P. R. Preiser *et al.*, *EMBO J.* **15**, 684 (1996).

35. We thank R. J. M. Wilson, A. Scherf, P. Druilhe, and A. C. Grüner for a critical review of the manuscript and we thank A. M. Vigario for help. F.T.M.C. is supported by a Brazilian fellowship from the CAPES Foundation. Partly funded by the Royal Society's Joint Project under European Science Exchange Programme, the British Council's Alliance: Franco-British Partnership Programme, and European Union contract IC18 CT98 0369.

31 July 2001; accepted 25 October 2001

CTCF, a Candidate *Trans*-Acting Factor for X-Inactivation Choice

Wendy Chao, Khanh D. Huynh, Rebecca J. Spencer, Lance S. Davidow, Jeannie T. Lee*

In mammals, X-inactivation silences one of two female X chromosomes. Silencing depends on the noncoding gene, *Xist* (inactive X-specific transcript), and is blocked by the antisense gene, *Tsix*. Deleting the choice/imprinting center in *Tsix* affects X-chromosome selection. Here, we identify the insulator and transcription factor, CTCF, as a candidate *trans*-acting factor for X-chromosome selection. The choice/imprinting center contains tandem CTCF binding sites that function in an enhancer-blocking assay. In vitro binding is reduced by CpG methylation and abolished by including non-CpG methylation. We postulate that *Tsix* and CTCF together establish a regulatable epigenetic switch for X-inactivation.

Dosage compensation ensures equal expression of X-linked genes in XX females and XY males. In mammals, this process results in inactivation of one female X chromosome (XCI) (*1*) in a random or imprinted manner. In the random form (eutherian), a zygotic counting mechanism initiates dosage compensation and enables a choice mechanism to randomly designate one active (Xa) and one inactive (Xi) X [reviewed in (*2*)]. In the imprinted form, zygotic counting and choice are superseded by parental imprints that direct exclusive paternal X-silencing (*3, 4*). Imprinted XCI is found in ancestral marsupials (*3*) but vestiges remain in the extraembryonic tissues of eutherians such as mice (*4*).

An epigenetic mark for random and imprinted XCI has long been postulated (*2*). The marks are placed at the X-inactivation center (*Xic*) (*5*), which includes the *cis*-acting noncoding gene, *Xist* (*6, 7*), and its antisense counterpart, *Tsix* (*8*). *Xist* RNA accumulation along the Xi initiates the silencing step (*9, 10*), whereas *Tsix* represses silencing by blocking *Xist* RNA accumulation (*11, 12*). A *cis*-acting center for choice and imprinting lies at the 5′ end of *Tsix*, as its deletion abolishes random choice in epiblast-derived cells to favor inactivation of the mutated X (*11, 13*) and disrupts maternal *Xist* imprinting in extraembryonic tissues (*14, 15*). Thus, while imprinted XCI is parentally directed and random XCI is zygotically controlled, both work through *Tsix* to regulate *Xist*.

To date, only X-linked *cis*-elements have been identified as XCI regulators. Yet, virtually all models invoke *trans*-acting factors which interact with the X-linked sites. In one model for imprinted XCI, a maternal-specific *trans*-factor confers resistance to XCI (*16*). In models for random XCI, an autosomally expressed "blocking factor" protects a single X from silencing (*2*). We have proposed that *Tsix* is the *cis*-target of both *trans*-factors (*11, 14*).

To isolate candidate *trans*-factors, we now used computational analysis (Fig. 1) to identify mouse-to-human conserved elements within the 2- to 4-kilobase (kb) sequence implicated in choice and imprinting (*11, 13–15*), a region including *DXPas34* (*17*). We found that the region is composed almost entirely of 60– to 70–base pair (bp) repeats with striking resemblance to known binding sites for CTCF, a transcription factor with a 60-bp footprint and 11 zinc fingers that work in various combinations to generate a wide range of DNA-binding activities (*18*). CTCF functions as a boundary element at the globin locus (*19*), regulates enhancer access to the *H19-Igf2* imprinted genes (*20–23*), and associates with CTG/CAG repeats

Howard Hughes Medical Institute, Department of Molecular Biology, Massachusetts General Hospital, Department of Genetics, Harvard Medical School, Boston, MA 02114, USA.

*To whom correspondence should be addressed. E-mail: lee@frodo.mgh.harvard.edu

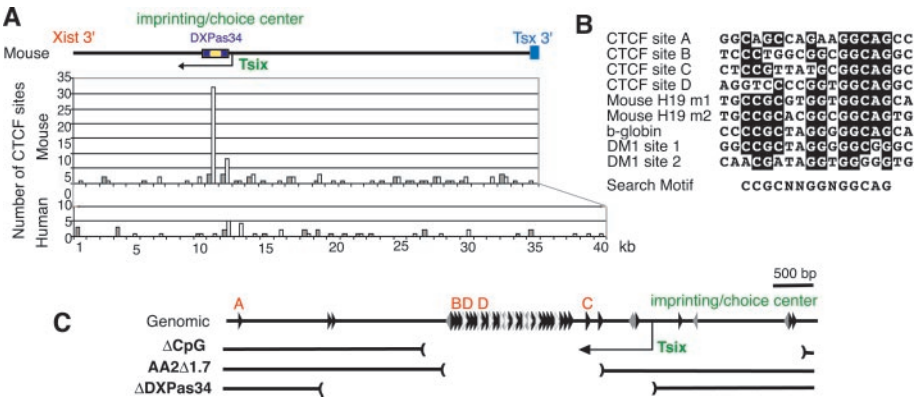


Fig. 1. Tandem CTCF-like binding sites in the *Tsix* imprinting/choice center. (**A**) Histogram of conserved human and mouse sites with 0 to 3 mismatches to the CTCF consensus (*20, 21*). Open and shaded bars represent two orientations. (**B**) Alignment of mouse *Tsix*, *H19*, *DM1*, and chicken β-globin sites. Shading indicates identity with the consensus. (**C**) Clustering of CTCF motifs. ΔCpG (*11, 14*), ΔDXPas34 (*13*), and AA2Δ1.7 (*15*). Filled triangles, sites with 0 to 3 mismatches. Open triangles, sites in the center with >3 mismatches. Forward sites, gray; reverse sites, black. Tested CTCF sites are indicated by red letters.

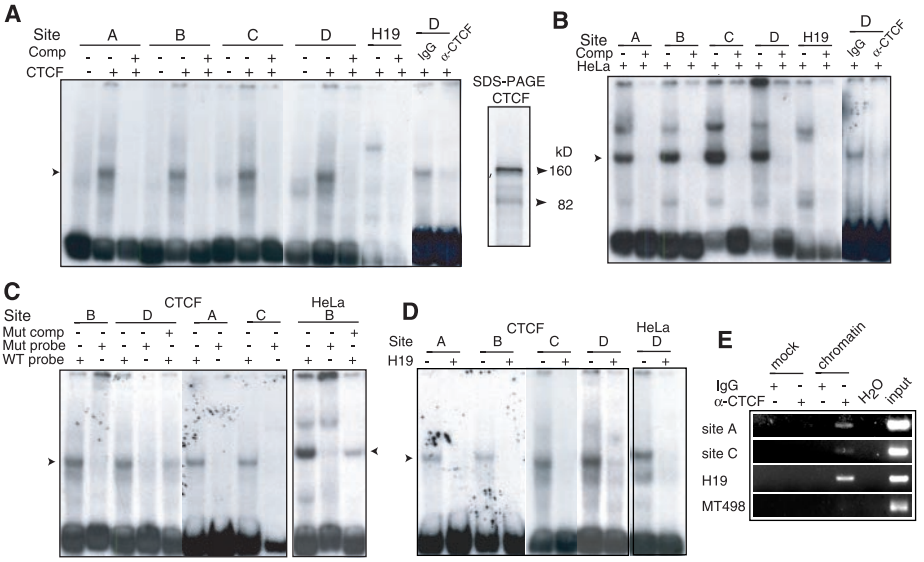
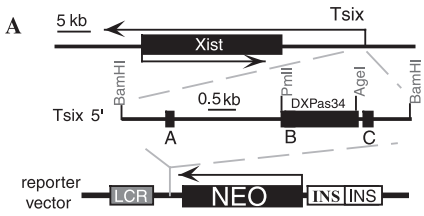


Fig. 2. *Tsix* elements bind CTCF in vitro and in vivo. (**A**) Gel-shift assay of P32-labeled *Tsix* oligos and CTCF protein. Reactions were carried out for 30 min at room temperature with 0.5 to 5.0 μl in vitro–synthesized CTCF protein (see SDS-PAGE) and 10 fmol double-stranded DNA probes in 20 mM HEPES (pH 7.5), 50 mM KCl, 5 mM $MgCl_2$, 1 mM dithiothreitol, 0.3 mg/ml BSA, 5% glycerol, 0.5% Triton X-100, and 1 μg poly-dI:dC before resolution in 5% acrylamide, 0.5× TBE gels at 4°C. Cold competitors here and below (comp) were added at 200× molar excess. Supershifts were carried out using normal IgG or COOH-terminal CTCF antibodies (*19*). Site A, 5′-TGGAGCCTAAACCTGTCTGTCTCTTTACCAGACGCAGGGCAGCCAGAAGGCAGCCATTCACAATCCAGGAAGACAGGAAGGG-3′; site B, GGGGTTGGTTATAAGGCAGGGATTTTAGCGATCTCCCCAGGTCCCTGGCGGCGGCAGGCATTTTAGTGATAGCCCAGGTCCCCG; site C, ATTTTGGCTCCAGGACCCAGCAGACATTTTAGTTATTCCTCCGTTATGCGGCAGGCATTTTAACTATCGGTTCGGGACTACGCAGG; site D, CAGATCCCCAGTGGCAGACATTTTAGTGATAGCCCAGGTCCCCGGTGGCAGGCATTTTAGTGATAGCCCAGGTCCCCGGTGGCA. H19, MS1 (*20*). Arrowhead, *Tsix* DNA-protein complex. (**B**) An activity in HeLa nuclear extract (1 to 2 μg/reaction) also binds *Tsix* sites. (**C**) Mutated CTCF sites show reduced binding. Mut, mutated; WT, wild type. MutA, 5′-TGGAGCCTAAACCTGTCTGTCTCTTTACCAGTAATAGAATTCATGTAATATATCCATTCACAATCCAGGAAGACAGGAAGGG-3′; MutB, GGGGTTGGTTATAAGGCAGGGATTTTAGCGATCTCCCCAGGTCTAATAGAATTCATGGCATTTTTAGTGATAGCCCAGGTCCCCG; MutC, ATTTTGGCTCCAGGACCCAGCAGACATTTTAGTTATTCCTTAATAGAATTCATGGCATTTTAACTATCGGTTCGGGACTACGCAGG; MutD, CAGATCCCCAGTGGCAGACATTTTAGTGATAGCCCAGTAATAGAATTCATGGCATTTTAGTGATAGCCCAGGTCCCCGGTGGCA. (**D**) Unlabeled *H19* sites compete against *Tsix* sites for CTCF. (**E**) CTCF binds *Tsix* in vivo (female fibroblasts) using ChIP analysis as described (*28*). Immunoprecipitations were performed overnight at 4°C with anti-CTCF antibodies (Upstate) or normal IgG. Primers pairs GTGTGTCATAGCTCAAGAGG, GGAGCCTAAACCTGTCTGTC (site A); AATGCTTGCCAGCTATGCGG, TAACCACCTGTAAGGGACAG (site C).



B

	Mean number of colonies	S.D.	*P*-value
No insulator	100	0.00	--
2.3 kb λ DNA	96.9	4.04	0.07
Globin insulators x2	14.1	4.76	<0.0001
BamHI-BamHI (F)	90.2	32.4	0.45
BamHI-BamHI (R)	18.2	7.25	<0.0001
AgeI-PmlI (F)	86.4	3.53	<0.0001
AgeI-PmlI (R)	56.1	20.9	<0.0001

C

Mutated site B	100	0.00	–
Globin insulators x2	18.9	12.3	<0.0001
site A	64.7	26.8	<0.05
site B	65.9	18.8	<0.05
site C	62.5	20.5	<0.05

Fig. 3. The 5′ end of *Tsix* contains enhancer-blocking activity. (**A**) The enhancer-blocking assay (*26*) for *Tsix* sites in K562 cells. Sites A, B, and C are indicated by black boxes. Fragments in both forward (F) and reverse (R) orientations ("F," *Tsix* and *Neo* transcription in same direction) were inserted between the β-globin LCR and a neomycin-resistance reporter (*Neo*). Flanking globin insulators (Ins) protects against position effects (*26*). + control, globin insulators (pJC13-1) (*26*). (**B**) Results of enhancer-blocking assay. We transfected 1.5 pmol each of test plasmid and pTK-Hygromycin (transfection efficiency control). Neo-resistant colonies were counted 2 to 3 weeks after transfection and normalized to hygromycin-resistant colonies. Three to four experiments were averaged. *P*-values, unpaired one-tailed Student's *t* test in pairwise comparisons against the no-insulator control. (**C**) Enhancer-blocking activities for sites A, B, C, and mutated B. Constructs contained 1.5 kb of spacer to maintain equal distance. *P*-values, unpaired one-tailed Student's *t* test in pairwise comparisons against mutated B.

at *DM1* (*24*). Murine *Tsix* contains >40 CTCF motifs and the human sequence has >10 (Fig. 1A). Dotplot analysis indicated a contiguous head-to-tail arrangement of highly homologous *DXPas34* repeats (*25*). This clustering is rare, with only three other loci of comparable density (40 sites per 1629 bp) occurring in 40.4 Mb of available sequence (ScanACE, http://twod.med.harvard.edu). The clustering of nine human elements is not above genome average (test of 933 random 100-kb fragments; random sequence selection program, J. Aach). CTCF function, however, does not require a clustering of sites (*20–23*).

To determine if the sites could bind CTCF in vitro, we performed gel retardation analysis of representative sites A, B, C, and D (Fig. 1, B and C). Using in vitro–translated murine CTCF, we observed a protein-DNA complex at all sites that was eliminated by unlabeled self-competitor DNA (Fig. 2A). The complex migrated more rapidly than that formed by *H19*, possibly due to differential binding of CTCF

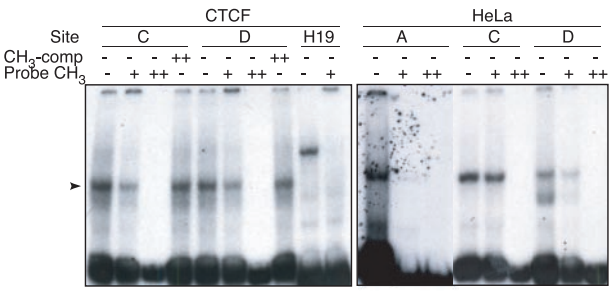


Fig. 4. CTCF binding is sensitive to DNA methylation in vitro. Gel-retardation analysis using *Tsix* probes which were unmethylated (–), methylated at CpGs only (+), or methylated at all C-nucleotides (++). Cold competitor (CH_3-comp) at 200× was methylated at all Cs. CpG methylation, achieved by SssI methylase and confirmed by insensitivity to HpaII or AciI digestion. Non-CpG methylation, achieved by direct synthesis. Arrow, *Tsix* DNA-protein complex.

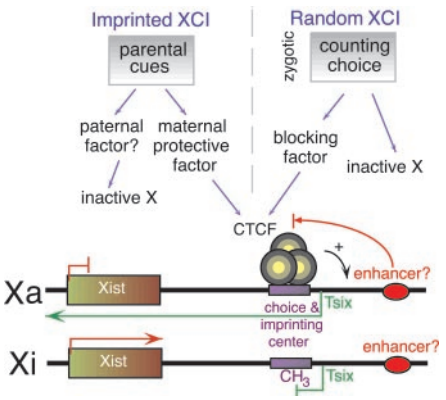


Fig. 5. Model of a regulatable epigenetic switch created by CTCF and *Tsix*.

isoforms (Fig. 2A; SDS-PAGE) or differential DNA bending induced by CTCF (*22*). Unprogrammed lysates did not shift the probe, indicating that the activity was specific to CTCF. HeLa extracts yielded two bands (Fig. 2B), one similar to that seen with in vitro–synthesized CTCF and one of lower intensity with a mobility similar to that for *H19* (this band was not always seen, e.g., Fig. 2D). Preincubation with polyclonal anti-CTCF antibodies blocked complex formation (Fig. 2, A and B). Mutating the 14-bp consensus (*20, 21*) within the 70-bp sites reduced binding (Fig. 2C) and unlabeled *H19* DNA effectively competed against *Tsix* for CTCF binding (Fig. 2D). Thus, CTCF specifically binds *Tsix* in vitro.

To test if CTCF binds *Tsix* in vivo, we carried out chromatin immunoprecipitation (ChIP) using anti-CTCF antibodies followed by *Tsix*-specific polymerase chain reaction in female mouse fibroblasts. Because the CTCF sites are tandemly repetitive, only sites A and C could be tested. Like the *H19* site [MS2 (*20*)], both sites were specifically coimmunoprecipitated with CTCF (Fig. 2E). In contrast, random loci on mouse chromosome 12 (MT498; www.jax.org) and in *Xist* (cDNA bp 13,177 to 13,428) did not coimmunoprecipitate (MT498 shown). Thus, CTCF complexes with *Tsix* DNA in vivo.

At some loci, CTCF sites act as chromatin insulators (*19–21*). In the established assay, insertion of these sites between the globin LCR and a neomycin (neo)–resistance reporter results in fewer neo-resistant K562 colonies (*26*). When a 4.3-kb Bam HI–Bam HI fragment containing all the *Tsix* sites was tested, we observed a dramatic reduction in colony number which was stronger in the R-orientation (Fig. 3, A and B). A 1.1-kb Pml–Age I fragment containing only sites B, D, and *DXPas34* also reduced colony number more strongly in the R-orientation (Student's t test, $P < 0.0001$; ANOVA, $P < 0.0001$). This modest orientation-dependent effect is consistent with published reports (*19–23*). The greater activity in the Bam HI–Bam HI fragment might be attributable to additional CTCF sites outside of *DXPas34* or to possible unmapped *Tsix* promoter activity in the Bam HI–Bam HI fragment that would be antisense to *Neo*. Individual sites A, B, and C each exhibited fewer colonies relative to mutated site B (Fig. 3C; t test, $P < 0.05$; ANOVA, $P < 0.05$). Thus, *Tsix* can block enhancer-promoter interaction and insulating activity correlates with CTCF binding in vitro.

Since CTCF responds to CpG methylation at some loci (*20–22*), we tested methylation-sensitivity at *Tsix* using gel retardation analysis. Unexpectedly, CTCF binding was only partially blocked by CpG methylation but was abolished when non-CpG methylation was included (Fig. 4). This contrasted with total inhibition at *H19* by CpG methylation alone. Relevant to this, *H19* sites contain three to four CpG's (*20, 21*), whereas many *Tsix* sites contain zero or one CpG in the consensus despite being strongly C-rich (Fig. 1B). These findings raised the possibility that non-CpG- together with CpG-methylation might regulate CTCF binding to *Tsix*. Notably, recent bisulfite sequencing has not uncovered differential CpG methylation in *DXPas34* (*27*). In light of our findings, the methylation status of non-CpG sites in the CTCF array will be critical in future work.

In summary, we have identified CTCF as a binding protein for the *cis*-acting choice/imprinting center in *Tsix*. We propose that CTCF and *Tsix* coordinately establish the epigenetic switch for *Xist* (Fig. 5). Because knocking out the CTCF array (choice/imprinting center) results in inactivation of the mutated X (*11, 13–15*), we favor a model in which binding of CTCF designates the future Xa. In this model, the zygotic blocking factor and the maternal protective factor work through CTCF to promote *Tsix* expression on the Xa. CTCF could directly stimulate *Tsix* transcription or do so by default through blocking *Xist*'s access to unidentified shared enhancers (*20–23*). *Tsix* transcription would in turn block *Xist* RNA accumulation (*12*). On the Xi, CTCF binding is excluded from *Tsix*, possibly by methylation (CH_3) of the CTCF array, thereby allowing the up-regulation of *Xist*. In the future, finer mutational analysis and the identification of differentially methylated regions will be required to test details of the model. Because CTCF is ubiquitous, developmental specificity must be achieved combinatorially with stage- and locus-specific factors. Identification of these protein-protein interactions will be instrumental in defining the long-postulated zygotic and maternal factors.

References and Notes

1. M. F. Lyon, *Nature* **190**, 372 (1961).
2. P. Avner, E. Heard, *Nature Rev. Genet.* **2**, 59 (2001).
3. G. B. Sharman, *Nature* **230**, 231 (1971).
4. N. Takagi, M. Sasaki, *Nature* **256**, 640 (1975).
5. C. J. Brown *et al.*, *Nature* **349**, 82 (1991).
6. C. J. Brown *et al.*, *Nature* **349**, 38 (1991).
7. N. Brockdorff *et al.*, *Cell* **71**, 515 (1992).
8. J. T. Lee, L. S. Davidow, D. Warshawsky, *Nature Genet.* **21**, 400 (1999).
9. G. D. Penny, G. F. Kay, S. A. Sheardown, S. Rastan, N. Brockdorff, *Nature* **379**, 131 (1996).
10. Y. Marahrens, B. Panning, J. Dausman, W. Strauss, R. Jaenisch, *Genes Dev.* **11**, 156 (1997).
11. J. T. Lee, N. Lu, *Cell* **99**, 47 (1999).
12. N. Stavropoulos, N. Lu, J. T. Lee, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10232 (2001).
13. E. Debrand, C. Chureau, D. Arnaud, P. Avner, E. Heard, *Mol. Cell. Biol.* **19**, 8513 (1999).
14. J. T. Lee, *Cell* **103**, 17 (2000).
15. T. Sado, Z. Wang, H. Sasaki, E. Li, *Development* **128**, 1275 (2001).
16. G. F. Kay, S. C. Barton, M. A. Surani, S. Rastan, *Cell* **77**, 639 (1994).
17. B. Courtier, E. Heard, P. Avner, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3531 (1995).
18. E. M. Klenova *et al.*, *Mol. Cell. Biol.* **13**, 7612 (1993).
19. A. C. Bell, A. G. West, G. Felsenfeld, *Cell* **98**, 387 (1999).
20. A. T. Hark *et al.*, *Nature* **405**, 486 (2000).
21. A. Bell, G. Felsenfeld, *Nature* **405**, 482 (2000).
22. C. Kanduri *et al.*, *Curr. Biol.* **10**, 853 (2000).
23. C. Kanduri *et al.*, *Curr. Biol.* **10**, 449 (2000).
24. G. N. Filippova *et al.*, *Nature Genet.* **28**, 335 (2001).
25. Web material is available on *Science* Online at www.sciencemag.org/cgi/content/full/1065982/DC1.
26. J. H. Chung, M. Whiteley, G. Felsenfeld, *Cell* **74**, 505 (1993).
27. M. Prissette, O. El-Maarri, D. Arnaud, J. Walter, P. Avner, *Hum. Mol. Genet.* **10**, 31 (2001).
28. K. E. Boyd, J. Wells, J. Gutman, S. M. Bartley, P. J. Farnham, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13887 (1998).

29. We thank G. Felsenfeld for pJC13-1 and anti-CTCF antibodies, J. Hughes and J. Aach for computational analysis, I. King for the HeLa extract, R. Kingston and M. Oettinger for advice on electrophoretic mobility shift assays, P. Farnham for a ChIP protocol, and J. E. Kirby for manuscript critique. This work was funded by the NIH, the Pew Scholars Program, and the Howard Hughes Medical Institute.

4 September 2001; accepted 6 November 2001
Published online 6 December 2001;
10.1126/science.1065982
Include this information when citing this paper.