Investigating the evolution of replication timing and monoallelic expression in mammals and birds

A thesis submitted for the degree of Doctor of Philosophy

December 2014

Megan Lynne Wright, BSc. (Hons), B.Media

School of Molecular and Biomedical Science, Discipline of Genetics

The University of Adelaide

Table of Contents

Index of Figures and Tables

CHAPTER 1: Introduction

BOOK CHAPTER: Replication Timing: Evolution, Nuclear Organization and Relevance for Human Disease

CHAPTER 2: Asynchronous replication of loci residing on platypus X- and chicken Z- chromosomes

PAPER: Replication asynchrony and differential condensation of X chromosomes in female platypus (Ornithorhynchus anatinus)

SUBMITTED MANUSCRIPT: Asynchronous replication timing of loci on the Z-chromosome in male chicken

CHAPTER 3: Conservation of replication timing in amniotes

MANUSCRIPT: Conservation of replication asynchrony and S-phase entry predates evolution of genomic imprinting in mammals

MANUSCRIPT: Co-detection of bromodeoxyuridine and directly-labelled fluorescent *in situ* hybridization signals in interphase cells of amniote species

CHAPTER 4: Monoallelic expression predates genomic imprinting

MANUSCRIPT: Monoallelic expression predates imprinting: platypus imprinted orthologs display an amount of monoallelic expression

CHAPTER 5: Conclusion

Declaration

For a thesis that contains publications. I certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

22.12.2014

Megan Wright **Date**

Acknowledgements

Foremost, I would like to thank my supervisor, Frank Grützner, for his loving support and guidance. I am extremely grateful to all members of the Grützner Lab, past and present, for their advice and encouragement, and members of the Discipline of Genetics, who have supported me in various ways over the years, with special mention to Michelle Coulson. And to Velta, who is a wonderful mentor, and a magnificent person.

To all my friends, in particular Wendy, Tassy, and Shuly, who have all helped me through these years with humour and love. To my partner, Jordan, whose 'super-logic' over the years has guided me through the hard times, and whose love and friendship has made me strong. And to my family, with special mention to Mum, Dad, Shannon, and Aunty Lee, I offer many thanks for their love, support, and encouragement, and editorial help. I could not have achieved as much as I have without you.

Abbreviations

°C	Degree celsius
µg	Microgram
µl	Microlitre
µm	Micrometre
3C	Chromosome conformation capture
4C	Chromosome conformation capture-on-chip
5C	Chromosome conformation capture carbon copy
ACT	Associated chromosome trap
AER	Allele expression ratio
BrdU	Bromodeoxyuridine
cDNA-seq	cDNA library sequencing
ChIA-PET	Chromatin Interaction Analysis by Paired-End Tag sequencing
CT	Chromosome territory
CTCF	CCCTC-binding factor
DD	Double-double dot DNA FISH signal (replicated locus)
DMR	Differentially methylated region
EIO	Eutherian imprinted ortholog
EIPO	Eutherian imprinted platypus ortholog
FISH	Fluorescent *in situ* hybridisation
G1	Gap 1
Hi-C	Genome-wide 3C experiment
Hox	Homeobox
ICR	Imprinting control regions
LCR	Locus control region
LINE	Interspersed nuclear element
MB	Mega base
MCM	Minichromosome maintenance
MHM	Male hypermethylated
ncRNA	Non-coding RNA
ND	No data avaliable
NK	Natural Killer-cell

OR	Origin of replication
ORC	Origin recognition complex
ORc	Olfactory receptor
Pre-RC	Pre-replicative complex
RNA-seq	RNA library sequencing
Rsx	RNA-on-the-silent X
RT-PCR	Reverse-transcriptase polymerase chain reaction
S-phase	DNA synthesis phase of the cell cycle
SD	Single-double dot DNA FISH signal (replicating locus)
SINE	Short interspersed nuclear element
SNP	Single nucleotide polymorphism
SS	Single-single dot DNA FISH signal (unreplicated locus)
TE	Transposable element
XIC	X-chromosome inactivation centre
Xist	X inactive specific transcript

Nomenclature

Throughout this thesis, various forms of conventional notations are observed which pertain to species-specific nomenclature, particularly mouse, human, platypus, and chicken.

Abstract

Monoallelic expression and replication timing are closely linked fundamental aspects of genome biology, yet their evolutionary trajectory has not been investigated in much detail. The monoallelic expression status of imprinted genes observed in therian species has previously not been found in the earlier-diverged monotreme mammals, or in birds, when measured using molecular techniques. Furthermore, the observation that eutherian imprinted and X-borne genes asynchronously replicate was traditionally thought to be linked to the dissimilar epigenetic states that existed at each allele controlling monoallelic expression. In this study, we use a combination of cytogenetic and molecular techniques to assess the replication status of sex chromosome genes in the platypus and chicken, as well as the replication status and expression pattern of platypus imprinted orthologs.

We find that asynchronous replication does occur at specific sex chromosome loci in platypus and chicken, although in chicken the amount of asynchronous replication changes over development. Furthermore, differential chromatin compaction is observed in platypus sex chromosomes, a characteristic observed in therian X-inactivation, suggesting that both asynchronous replication and chromatin compaction are features characteristic of amniote sex chromosomes. Asynchronous replication and monoallelic expression is observed at platypus imprinted orthologs, indicating that a ‘pre-imprinted’ status is observed at these genes in non-therian amniote species. These results show that monoallelic expression predates imprinting at these loci, suggesting that the partial monoallelic expression observed in monotreme mammals has evolved in therian mammals to become parentally-inherited imprinted expression.

CHAPTER 1: Introduction

This chapter consists of one published book chapter.

Chapter overview

This chapter consists of a literature review book chapter for the online open access publisher, *InTech*. The review discusses the broader aspects of eukaryotic DNA replication and how it is measured. It covers areas of research such as replication initiation, temporal programmes of replication initiation, asynchronous replication, replication timing, and long-range chromatin interactions. The affect of the aforementioned genomic processes are discussed in light of monoallelically expressing genes, including X-inactivated, imprinted, and random monoallelically expressed genes. The more technical aspects discussed in the review include past and present experiments used to measure replication asynchrony and replication timing. The review also outlines current gaps in knowledge in terms of replication asynchrony data, which is particularly pertinent for the asynchronous replication research discussed in chapters 2 and 3 of this thesis.

BOOK CHAPTER:

Replication Timing: Evolution, Nuclear Organization and Relevance for Human Disease

Statement of Authorship

Title of Book Chapter	**Replication Timing: Evolution, Nuclear Organization and Relevance for Human Disease**
Publication Status	Published
Publication Details	Wright, M.L. and Grutzner, F. 2011. Replication Timing: Evolution, Nuclear Organization and Relevance for Human Disease. In *DNA Replication - Current Advances* (ed. H. Seligmann), pp. 437-466. InTech.

Open access to electronic copies are available from:

http://www.intechopen.com/source/pdfs/16947/InTech-Replication_timing_evolution_nuclear_organization_and_relevance_for_human_disease.pdf

Author Contributions

By signing the Statement of Authorship, each author certifies that their stated contribution to the publication is accurate and that permission is granted for the publication to be included in the candidate's thesis.

Name of Principal Author (Candidate)	Megan Wright
Contribution to the Book Chapter	Researched all aspects in the chapter, and wrote the review. Created all figures and tables. Carried out all FISH experiments and obtained all fluorescent microscope images for the 'Cytogenetic FISH dot assay' (Figure 2).
Signature	19.12.2014

Name of Co-Author	Frank Grützner
Contribution to the Book Chapter	Supervised all research, gave topic suggestions, and aided in the chapter construction/outlining. Also proof-read drafts, was the primary correspondent during the review process, and made corrections.
Signature	19.12.2014

20

Replication Timing: Evolution, Nuclear Organization and Relevance for Human Disease

Megan Lynne Wright and Frank Grützner
The Robinson Institute, School of Molecular & Biomedical Sciences,
University of Adelaide, Adelaide
Australia

1. Introduction

DNA replication in eukaryotes is multifaceted, dynamic and highly organised. In contrast to bacterial cells, which replicate from single origins of replication, complex eukaryote genomes replicate from thousands of origins of replication. Although we know that the timing of replication depends on the chromatin environment, the function and evolution of mechanisms controlling replication timing are unclear. Many studies in species ranging from yeast to humans have demonstrated how replication timing depends on proximity to certain sequences such as telomeres and centromeres (Ferguson and Fangman, 1992; Friedman et al., 1996; Heun et al., 2001), chromatin status (euchromatin and heterochromatin) and is linked to gene function and expression (housekeeping genes versus tissue specific genes and monoallelically expressed genes) (Hiratani and Gilbert, 2009; Hiratani et al., 2009). Replication timing has been linked to fundamental epigenetic regulatory mechanisms including genomic imprinting (Kitsberg et al., 1993; Knoll et al., 1994), X chromosome inactivation (Gilbert, 2002; Takagi et al., 1982; Wutz and Jaenisch, 2000), interchromosomal interactions (Ryba et al. 2010) and is increasingly recognised to be important in human disease (DePamphilis, 2006).
This chapter integrates established knowledge with recent scientific breakthroughs, using genome-wide approaches linking different aspects of epigenetic control with replication timing, to provide a state-of-the-art overview and perspective for future work in this area of research. Despite detailed knowledge on replication timing in a select number of model organisms (e.g. yeast, drosophila, mouse) we are only beginning to understand how replication timing evolved in relation to other epigenetic mechanisms (e.g. genomic imprinting, X inactivation, and long-range chromatin interaction). The evolution of these epigenetic mechanisms will be presented together with novel ideas about how cytological and genome-wide approaches and methodologies can be combined to provide a comprehensive picture of spatial and temporal organization, the evolution of replication timing in eukaryotic genomes, and their relevance in human disease.

2. Background

2.1 Replication initiation

The complete and accurate replication of DNA during the S-phase is of fundamental importance for all organisms. The mechanism of replication is highly conserved across

evolution, whereby a cell must gather the proteins to initiate replication at specific origins of replication (OR)s, unwind the DNA, move the replication fork bi-directionally away from the OR in such a manner as to allow the replication of the new daughter strand of DNA using the old parental DNA strand, and then cease replication. However whilst the replication process is highly conserved, different eukaryotes use different proteins and forms of control over replication (Gilbert, 2010).

Whilst general similarities exist in the type of machinery required to copy and create a new DNA strand across organisms, some areas of genome replication remain elusive. One such area in eukaryotes is replication initiation and timeline. Linear eukaryotic chromosomes replicate from many ORs which are spread out along their structure and are recognized by the origin recognition complex (ORC) (reviewed in Masai et al., 2010). These OR sites are where replication forks form and move bi-directionally away from the OR, replicating the DNA sequence as they move, then terminating when they meet another fork approaching from the opposite direction. The ORCs recognize almost all ORs, and will assemble at these regions in a highly conserved manner across eukaryotes. However, whilst ORCs bind specific sequence motifs in some eukaryotes, such as in budding yeast (Bell and Stillman, 1992), in other eukaryotes specificity is not well defined through sequence. Fission yeast and *Drosophila* have ORCs that recognize AT-rich sequences (Austin et al., 1999; Chuang and Kelly, 1999), rather than specific motifs. Moreover, human ORCs, which are chosen as initiators of replication, have also been shown to require AT-rich sequences as well as various other features, including matrix attachment region sequences, dinucleotide repeats and asymmetrical purine-pyrimidine sequences (Altman and Fanning, 2004; Debatisse et al., 2004; Paixao et al., 2004; Schaarschmidt et al., 2004; Wang et al., 2004). Other factors that may affect the initiation of replication at certain ORs also include DNA topology, transcription factors, and elements of the pre-replicative complex (pre-RC) (reviewed in Masai et al., 2010).

During late mitosis and G1, the chromatin-bound ORCs are loaded with minichromosome maintenance (MCM) complex, and thus become pre-RCs, with the ability to gather the required components to start replication. The pre-RCs assemble at most of the OR regions, however only a few of these complexes start replication in their region. The cell's choice to start replication at some ORs as opposed to others is unclear; whilst it is thought that the assembly of the pre-RCs at most ORs is used as backup in case the cell runs into trouble during replication, the choice as to whether a Pre-RC becomes an active replication initiator is not well understood (Doksani et al., 2009; Ibarra et al., 2008; Koren et al., 2010; Woodward et al., 2006).

There are, however, some known factors that may contribute to a pre-RC site becoming an active OR (reviewed in Masai et al., 2010). Firstly, the selection of replication initiation sites may be controlled by both the existence of a pre-RC and its assembly in combination with events that actually cause initiation. For example, the firing of an OR appears to affect the firing of adjacent ORs, as shown in the example of budding yeast, where active ORs suppress the initiation of replication at adjacent ORs (Brewer and Fangman, 1993). In this example, the suppression of adjacent potential ORs may be caused by the disruption of pre-RC complexes at these sites by the replication process initiated at the active OR (reviewed in Masai et al., 2010). Also, read-through transcription may affect the firing of downstream ORs (Haase et al., 1994; Saha et al., 2004). Furthermore chromatin structure, which refers to the chemical characteristics of the chromatin strand, may influence the initiation of replication by affecting the pre-RC assembly. There is evidence to show that histone

acetylases and deacetylases play roles in the assembly of pre-RCs by interacting with, or disturbing the loading of, pre-RC elements such as the MCM complex (Burke et al., 2001; Iizuka et al., 2006; Pappas et al., 2004; Pasero et al., 2002).
Finally, distal elements, such as locus control regions (LCRs) are known to affect initiation (Hayashida et al., 2006; Kalejta et al., 1998), with the initiation of replication at regions such as the human *β-globin* locus being controlled by a 5′ LCR (Aladjem et al., 1995).

2.2 Temporal programmes of ORs in eukaryotic chromosomes

Replication of eukaryotic genomes follows a defined temporal program, whereby the firing of ORs occurs in a predetermined but tissue specific manner. Hence this process is dynamic in terms of the selection of OR activation, as the cellular environment also plays a role in the temporal regulation of replication across the genome. Experiments have shown that a reduction in cellular thymidine caused a reduction in replication fork speed. This caused more intermediate ORs to be activated in order to compensate for the reduction in replication speed (Anglana et al., 2003; Taylor, 1977), and showed that cellular environments indeed affect the dynamics of OR firing. This shows that a cell is able to change its pre-determined temporal replication program if it undergoes replication stress, with the most relevant aspect of OR activation being the genomic context and how it impacts the replication program.
Factors that are involved in OR firing include chromatin loops, dormant and active pre-RC complexes and fork replication rate, and finally nuclear organisation. Firstly, there is some evidence to suggest that chromatin loops affect replication firing. Studies in *Xenopus* egg extracts transferred with erythrocyte nuclei showed that cells that entered into M-phase instantly after somatic transfer took longer to replicate than cells which were held in mitosis and allowed to undergo a single mitosis event. This was due to the influence of the single round of mitosis on the chromatin structure; the round of mitosis supported the formation of smaller chromatin loops which correlated with higher ORC protein recruitment and more efficient genome replication (Lemaitre et al., 2005). Another study showed that the ORs closer to regions of chromatin loop anchorage in G1 initiated replication in the following S-phase earlier than ORs located further away from anchorage regions, indicating that loop-formation was part of the control mechanism for OR firing (Courbet et al., 2008).
Fork replication rate also appears to have a role in the temporal organization of OR firing. Genomic integrity may be aided by the presence of dormant origins of replication, as MCMs are often present in much greater amounts than those needed at pre-RCs, and the reduced presence or loss of pre-RCs result in genomic instability, S-phase arrest, and cell death (Edwards et al., 2002; Hyrien et al., 2003; Lengronne and Schwob, 2002; Shreeram et al., 2002; Tanaka and Diffley, 2002). Dormant ORs have been shown to activate when forks are stalled, with one model hypothesizing that OR activation occurs stochastically, whereby the presence of a stalled fork increases the chances of adjacent dormant ORs being activated (Blow and Ge, 2009; Ibarra et al., 2008). Other models propose that the presence of a stalled fork changes the topology of the DNA strand and the chromatin structure within the region, thus causing nearby and usually dormant ORs to activate (Ibarra et al., 2008).
Finally, nuclear organisation has a role to play in a cell's replication program. Distinct chromosome territories exist as separate nuclear architecture compartments in interphase cells. Within these territories, a higher order of chromatin structure exists, where domains containing specific chromosomal arms and bands have been found to be located in the

nucleus in similar regions of certain cell types (Dietzel et al., 1998). It has also been proposed that these chromatin-rich chromosome territories (CTs) are separated by chromatin-poor areas called 'interchromatin compartments', which contain transcriptional and splicing machinery, as well as DNA replication and damage-repair machinery (reviewed in Aten and Kanaar, 2006; Cremer and Cremer, 2001; Misteli, 2001). However recent work showed extensive intermingling of CTs contradicting the existence of the interchomatin compartment (reviewed in Aten and Kanaar, 2006; Branco and Pombo, 2006; reviewed in Cremer and Cremer, 2010). Within separate chromosome territories there are many replication foci, whereby early and late replicating DNA can be found in spatially separate and distinct regions (Zink et al., 1999). Overall late replicating DNA (including the late replicating inactive X chromosome) is often located at the nuclear periphery or around the nucleolus organizing region (Sadoni et al., 1999).

3. Asynchronous replication

Asynchronous replication is another variation in the eukaryotic temporal replication repertoire. Asynchronous replication occurs when the ORs present in the same regions on two homologous chromosomes, initiate replication at different times. This results in one of the alleles replicating earlier than the allele on the other homologue. Notably, the alleles of asynchronously replicating genes are also observed to locate to separate discrete foci in a nucleus. This form of replication is a feature of monoallelically expressed genes, including genes that undergo allelic exclusion, imprinted genes, and genes from the X-chromosome in female somatic cells.

3.1 Approaches to measuring asynchronous replication and its effects on genome biology and disease

3.1.1 Chromosome banding

Chromosome banding techniques gave the first insights into the epigenetics behind replication, and more specifically, asynchronous replication. It is now well established that replication timing is not uniform across eukaryotic genomes, with select chromosomal regions showing early or late replication in the S-phase. This phenomenon has been observed in distinct banding regions along condensed metaphase chromosomes.

The discovery of early and late replication banding on metaphase chromosomes using the Bromodeoxyuridine (BrdU) incorporation technique, can be attributed to Latt (1973). Latt discovered that the differential incorporation of BrdU, a thymidine replacement, during the S-phase between early and late replicating regions of DNA, could be measured using 33258 Hoechst fluorescence. An efficiency reduction of the Hoechst dye fluorescence occurs when it is bound to the incorporated poly(dA-BrdU) compared to the poly(dA-dT). Incorporation of BrdU into either the late or early replicating DNA can be adjusted by culturing cells in BrdU for different time periods; specifically early replication stage BrdU incorporation was achieved by first culturing in BrdU with the addition of a terminal pulse of [3H]-dT, whilst late replication BrdU incorporation was achieved by culturing in medium containing thymidine to which BrdU was only added 6 hours before harvest. This allowed identification of 5-10 megabasepair regions on chromosomes replicating either early or late in the S-phase.

Latt's early research defined a fundamental relationship between chromosome organisation and replication timing; eukaryotic chromosomes do not undergo equivalent amounts of

replication both within a chromosome and across a karyotype, whereby a distinct non-equivalence of replication is represented by the presence of discrete bands for early and late replicating regions on a chromosome. Furthermore, the late-replicating inactive X chromosome in human females, which is noted to have a slightly more condensed karyotype, showed distinctly opposing fluorescence to the less-condensed active-X chromosome.

Higher-resolution replication banding has since been established in humans and numerous vertebrate species (Biederman and Lin, 1979; Costantini and Bernardi, 2008; Drets et al., 1978). Currently, there are three tiers of replication resolution: 1) low-resolution banding (e.g. De Latt's BrdU bands, and Giemsa and Quinacrine bands); 2) higher resolution banding (GC content in grouped isochore regions); and 3) individual isochores (Costantini and Bernardi, 2008). Isochores are regions of DNA, above 300 kb (on average around 0.9 Mb in size in the human genome), that have a similar GC content, and also have similar gene content (Costantini and Bernardi, 2008; Costantini et al., 2006; Costantini et al., 2007). Specifically, there are five groups of isochores, whereby lower GC content is classed with the isochore groups L1 and L2 (less than 40% GC-content, and few genes), intermediate groups are H1 and H2 (with around 47% and 52% GC-content, and intermediate amounts of genes), and finally the highest group is H3 (with above 52% GC-content, and high amounts of genes) (Bernardi, 1995). A replicon is a genomic region around 50-400 kb in size, that replicates from a single origin of replication. It has been shown that replicons that exist within a certain isochore region, all undergo similar replication timing, with clusters of early replicating replicons being found next to each other, and clusters of late-replicating replicons being grouped as well (Watanabe et al., 2002). Through the comparisons of the three tiers of resolution, it was found that groups of early and late replicating isochores corresponded to, and approached the same size of, high-resolution replication banding regions (4-7 Mb).

The results of the highest-replication isochore banding when compared to the other banding techniques has indicated that in mammalian chromosomes there are three nested structures important to replication (Figure 1). The first structure is that of the replicon (50-450 kb), whereby individual replicons undergo dynamic firing of their ORs. These replicons however usually exist in clusters of 10 or more, and every replicon in the cluster will usually undergo replication at the same time during the S-phase. The second is that of the isochore (> 300 kb) which is a region that exists as a combination of replicons all with similar early or late replication status and GC content, which can undergo early or late replication in the cell cycle. The third structure is that of the cytogenetic bands, which indicates large regions on a chromosome undergoing early or late replication, and corresponds well to groups of all-early or all-late replicating isochores (Costantini and Bernardi, 2008). This shows that the arrangement of mitotic chromosome structure is closely related to replication timing, from the chromosome banding level, all the way through to the level of organisation of the individual replicons. This pattern is maintained in interphase, where chromosome territories in the S-phase have clusters of early and late replicating foci, which correspond to the R- and G/C bands observed in mitotic chromosomes respectively (Sadoni et al., 1999).

Replication banding techniques have allowed early and late timing replication zones to be delineated along metaphase chromosomes, where areas of similarly replicating replicons are grouped making larger replicon clusters (Watanabe et al., 2002). However, the large genomic regions that bridge the transition of an early-replicating replicon cluster to a late-replicating replicon cluster appear to lack any ORs, and rely on the continuous movement of

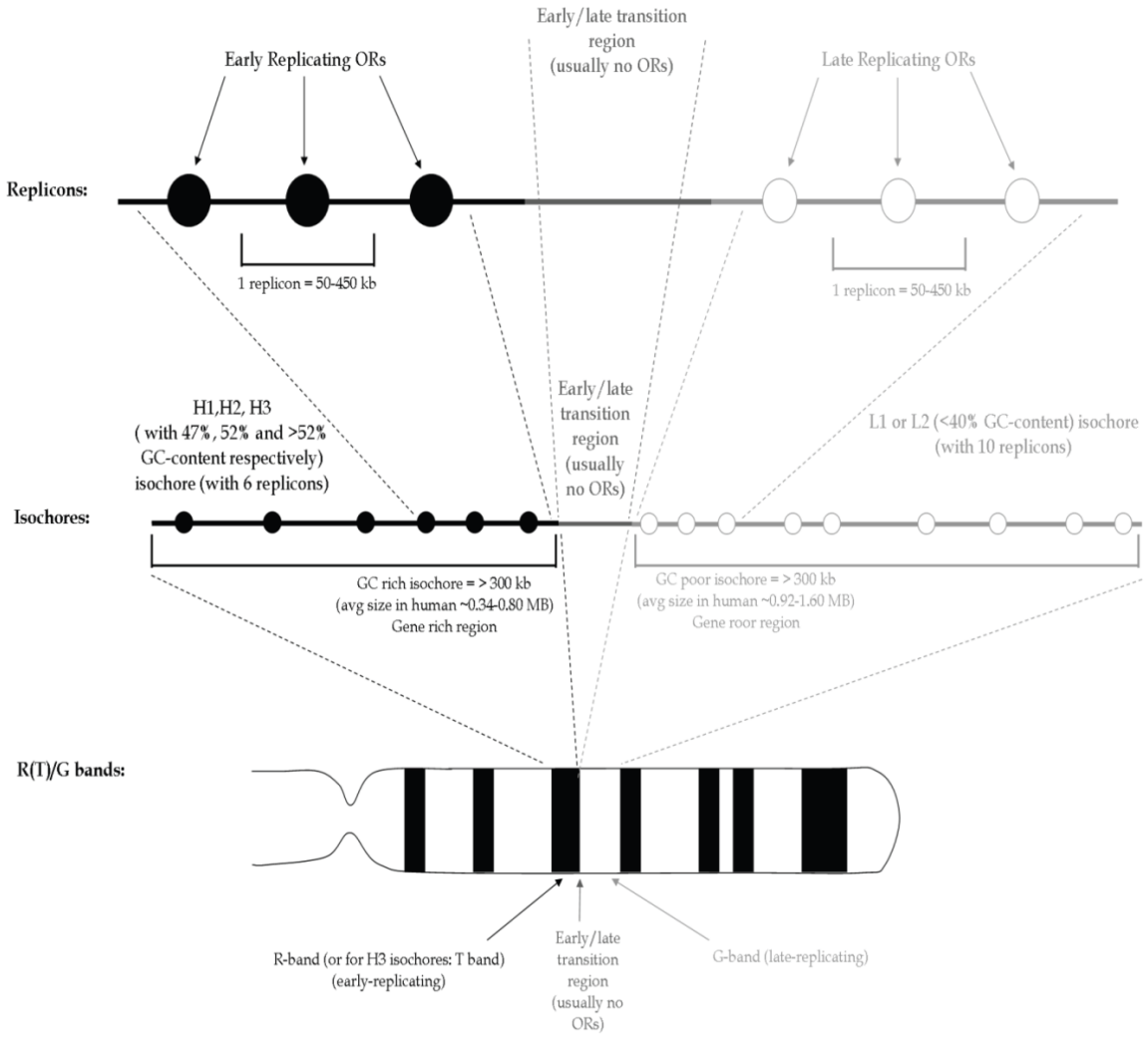


Fig. 1. The three nested structures of replication (see text for explanation).

forks from adjacent replicon-clusters/isochore regions for replication to occur in their region (reviewed in Farkash-Amar et al., 2008; Hiratani et al., 2008; Watanabe and Maekawa, 2010). This means that the fork from the earlier firing OR will have to move across the replication transition region, until it meets another fork from the late-replicating region. This will often pause replication in these early to late transition zones, which can cause genomic instability in the form of DNA breaks and rearrangements (Raghuraman et al., 2001; Rothstein et al., 2000). Furthermore, common genomic fragile sites frequently reside in early to late replication transition regions, and also lack backup ORs (Debatisse et al., 2006; Ge et al., 2007; Ibarra et al., 2008).

In addition to the increased genomic instability there is also an increase in the number of non-B-form DNA structures in replication transition regions (reviewed in Watanabe and Maekawa, 2010). Replication switch points (from early to late) are often associated with purine/pyrimidine rich areas, as these DNA regions can form structures called triplexes (H-DNA) that are known stop replication forks (Baran et al., 1991; Brinton et al., 1991; Ohno et al., 2000). The non-B-form structures however also have mutagenic properties, causing somatic recombination events (Kalish and Glazer, 2005; Knauert et al., 2006). It has thus been

proposed that these replication transition regions, which correspond to the regions between R/G bands, are subject to more genomic instability due to the increased presence of non-B-DNA structures in these genomic areas (Watanabe and Maekawa, 2010).
Replication timing is affected in regions of the human genome involved in disease. Generally it has been proposed that regions of the human genome that reside in areas where replication timing switches (early to late) would be unstable and more prone to DNA damage (reviewed in Watanabe and Maekawa, 2010). Notably, these regions of replication timing transition are also associated with many human diseases, including cancer (Watanabe et al., 2009; Watanabe et al., 2002; Watanabe et al., 2004). Regions or genes associated with other diseases, such as familial Alzheimer's, familial amyotrophic lateral sclerosis and phenylketonuria, are also found in these replication timing transition regions. Furthermore, there are over 70 human diseases associated with non-B DNA structures, including neurological and psychiatric diseases, and many genomic disorders, indicating that the increase of these structures in replication timing transition regions may be a first step in the mutational process associated with these diseases (reviewed in Watanabe and Maekawa, 2010).

3.1.2 Measuring asynchronous replication with the dot assay technique

Molecular cytogenetic techniques like Fluorescence *in situ* Hybridization (FISH) and an explosion of available genomic clones and whole chromosome probes has let to huge refinement of physical maps on metaphase and interphase chromosomes. This also enabled replication timing to be investigated on the single gene level. In these experiments, DNA probes designed to hybridise to a specific gene allowed the replication status to be observed in three states in a nucleus; two signals (single-single (SS) dot) represents an unreplicated status, whilst a three signal status (single-double (SD) dot) represents a locus undergoing

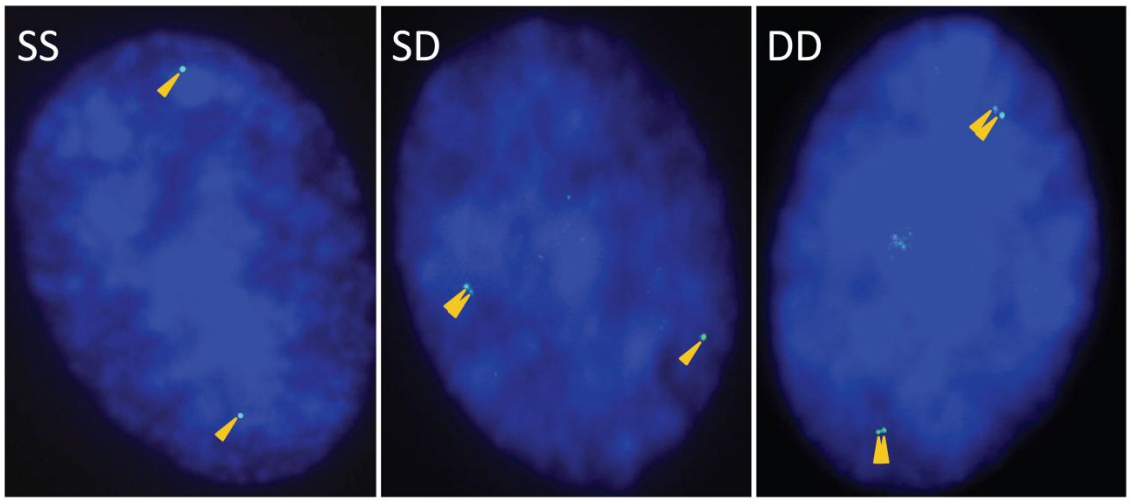


Fig. 2. Cytogenetic FISH dot assay
Mammalian interphase nuclei stained with DAPI (blue). Yellow arrows point to allele copies (green FISH signals) observed in each nucleus. The SS panel has two clearly defined green signals representing the two allele copies present in the nucleus, meaning the locus has not replicated. The SD panel shows three green signals, indicating that one allele has undergone replication, whilst the other allele is lagging behind and not yet replicated. The DD panel shows 4 green signals, indicating that both the alleles have replicated, and the locus has finished replicating.

replication, where one allele has replicated and the other is lagging, and finally a four signal status (double-double (DD) dot) represents a locus that is fully replicated (Selig et al., 1992). Asynchrony in this case is measured by the frequency of three-signal (SD dot) status observed in a cell line. However, the classification of asynchronous replication varies in the literature, with an asynchronously replicating state being assigned for loci with anywhere between 30-50% SD signal, and a non-asynchronously replicating locus generally having below 30% SD signal (Baumer et al., 2004; Wilson et al., 2007).

4. Replication timing in heteromorphic sex chromosomes

Replication banding and FISH dot assay techniques have not only shed light on how chromosome structure can affect replication, they have also allowed new insights into how replication timing of single genes has evolved. Changes in replication banding specific to one homolog in a karyotype have been used to identify early stage cytologically "homomorphic" sex chromosomes in various vertebrates (Nishida-Umehara et al., 1999; Schempp and Schmid, 1981). Heteromorphic sex chromosomes evolved from a pair of autosomes by a combination of suppression of recombination and accumulation of sexual antagonist genes (Ohno, 1967). The isolation of one of the sex chromosomes in one sex (Y chromosomes in mammals and some fish, the W chromosome in birds and many non-mammal vertebrates) has led to degeneration and massive gene loss. The evolution of heteromorphic sex chromosomes has been indicated to lead to a gene dosage difference between the sexes. In mammals this has resulted in the inactivation of one of the X chromosomes in female somatic cells.

X chromosome inactivation is a unique example where the status of chromatin can be changed from active to inactive (facultative heterochromatin) on a chromosome-wide level. In therian female mammals (marsupials and placental mammals), one of the X chromosomes in somatic cells is heterochromatic and late replicating (Holmquist, 1987; Lyon, 1961; Ohno et al., 1963; Schweizer et al., 1987; Takagi, 1974). This transcriptionally silenced and condensed X-chromosome is visible as a Barr body in somatic cells. In the third major group of mammals, the egg laying monotremes (platypuses and echidnas), it is less clear if X inactivation and late replication occurs. Earlier replication banding did not reveal obvious asynchronously replicating X chromosomes (Wrigley and Graves, 1988). More recently molecular cytological data suggests the platypus X-chromosomes display partial and gene specific forms of inactivation, but still undergo some level of asynchronous replication of X-specific genes (Deakin et al., 2008a; Ho et al., 2009). Furthermore, a wholesale shift in replication timing for the avian Z-chromosome, which shares extensive homology with the extraordinary ten sex chromosome system in monotremes, is not observed in male homogametic birds, indicating that this process is only present in therian mammals (Arnold et al., 2008; Grutzner et al., 2004; Rens et al., 2007; Veyrunes et al., 2008).

4.1 Chromatin marks behind X-inactivation

The X-inactivation process results in monoallelic expression of the vast majority of X-linked genes in humans and mice. Its process is dependent on critical elements which reside in the X-inactivation centre (XIC) on each X-chromosome, particularly the imprinted *Xist* and *Tsix* genes, and long-range chromatin elements (Boumil and Lee, 2001; Brockdorff et al., 1991; Brown et al., 1991; Clerc and Avner, 2003). The *Tsix* gene appears to regulate chromatin

structure at the *Xist* locus, causing its expression to be upregulated. This upregulation of Xist RNA corresponds to chromatin changes in the inactive X, most of which are associated with silencing (Heard, 2005). These *Xist*-induced marks on the inactive X include methylation of CpG dinucleotides in gene promoters, and histone modifications such as hypomethylation of H3K4 and hypoacetylation of H3K9 and H4, also monomethylation of H4K20 and trimethylation of H3K27, and finally H2AK119 ubiquitination (reviewed in Zakharova et al., 2009). Furthermore, the chromatin from the inactivated-X chromosome is enriched for the histone variant macroH2A1, and the final epigenetic mark is the late replication status of the inactive-X during the S-phase (reviewed in Zakharova et al., 2009). This inactivated state facilitates a change in the expression potential of the inactive X, and thus provides gene dosage compensation in female therian mammals (Hellman and Chess, 2007). It has also been observed that the active human X-chromosome is hypomethylated at gene-rich areas compared to the inactive X-chromosome, which displays hypermethylation (Hellman and Chess, 2007).

In placental mammals X inactivation of the maternal or paternal X chromosome is random, in marsupials and mouse extra-embryonic tissues only the paternal X is inactivated (reviewed in Lee, 2003). The epigenetic marks associated with marsupial X-inactivation include the loss or reduction of active histone marks on the inactive-X including H3K4 dimethylation, H4 acetylation, H3K9 acetylation marks (Koina et al., 2009; Wakefield et al., 1997). However, the absence of inactivating histone marks in marsupials, as observed on the inactive-X in placental mammals, may be due to the absence of a XIC region in marsupials (Duret et al., 2006; Hore et al., 2007; Koina et al., 2009). The evolution of the *Xist* non-coding RNA gene involves the pseudogenization of a protein-coding gene in the placental mammalian genome. As such, this gene is not present in marsupial and monotreme mammals, and cannot be found in the regions orthologous to the XIC in these mammalian clades. In marsupials and monotremes, the orthologous flanking genes to the placental mammal XIC region map to different ends of the X-chromosome and chromosome 6 respectively (Davidow et al., 2007; Deakin et al., 2008b; Duret et al., 2006; Hore et al., 2007; Shevchenko et al., 2007).

The FISH based dot assay was utilized to measure replication timing of genes from X-specific regions within the five platypus X-chromosomes. This did not reveal a clear cut replication asynchrony on X specific regions but one of the homologous pairs, namely the X3 chromosomes, showed significantly differential condensation, indicative of wholesale chromatin silencing (Ho et al., 2009). The other four sex chromosome pairs in platypus females, however, show no significant difference in condensation between homologs indicating that the X-inactivation process in monotremes may be region specific (Ho et al., 2009). In male homogametic birds (with ZZ sex chromosomes), studies have shown that whilst the entire chicken Z-chromosome replicates synchronously, the inactivation process appears to be partial and gene-specific, with dosage-compensation occurring stochastically, and in a stage and tissue-specific manner (Arnold et al., 2008; Deakin et al., 2008a; Ho et al., 2009; Kuroda et al., 2001; Kuroiwa et al., 2002; Mank and Ellegren, 2009). Moreover, there is evidence that dosage compensation in monotreme mammals operates in a similar manner as in birds, with platypus females showing stochastic transcriptional inhibition of genes from X-chromosomes (Deakin et al., 2008a). In this case, some X-genes were shown not to be dosage compensated, whilst monoallelic expression was observed at other X-chromosome loci (Deakin et al., 2008a).

5. Asynchronous replication in genes subject to genomic imprinting and allelic exclusion

Genomic imprinting refers to the parent of origin dependent monoallelic expression of an autosomal gene, engendered by the inheritance of parental-specific methylation at an allele. To date, imprinting mechanisms have only been found in therian mammals, which rely on extensive intrauterine foetal-maternal exchange during early development. The 'parental conflict hypothesis' proposed that imprinting is a way of parental genomes counteracting the effects of each other during foetal development, particularly in foetal-maternal placental nutrient exchange (Moore and Haig, 1991). Monotremes, unlike therian mammals, have a brief intrauterine foetal-maternal exchange and there is no competition of the parental genomes over maternal resources. In line with the 'parental conflict hypothesis' to date no imprinting has been discovered in this basal mammalian lineage, suggesting that imprinting evolved after their divergence from therian mammals (Renfree et al., 2009).

5.1 Imprinted genes

Imprinted genes are asynchronously replicated (Table 1), where the replication of one allele lags behind the other in the S-phase, even though the two alleles should be controlled by similarly situated ORs. Traditionally, imprinting involves DNA methylation at only one allele of a gene (i.e. the copy from just one parent is methylated) (Delcuve et al., 2009). In most cases the imprinted allele is methylated and transcriptionally silent. The active or silenced transcriptional state of an allele appears to go hand in hand with replication timing, whereby the expressed allele is early replicated, whilst the silenced allele undergoes late replication in the S-phase (reviewed in Zakharova et al., 2009).

Imprinting control regions (ICRs) are the elements which control the imprinting status of an allele (Bartolomei, 2009). The parentally inherited methylation status, which is established during gametogenesis, of an ICR dictates its control over an allele, meaning that maternal and paternal ICRs at a locus will interact differently with transcriptional control elements, due to their dissimilar methylation status (Bartolomei, 2009). Notably, maternally-imprinted ICRs are often found in the promoters for antisense transcripts, whilst paternally-imprinted ICRs usually reside in intergenic regions (reviewed in Edwards and Ferguson-Smith, 2007). Moreover, the formation of large imprinted gene clusters, where regions of maternally and paternally expressed genes are interspersed with non-imprinted genes, allows many imprinted genes to share regulatory elements, such as ICRs (reviewed in Bartolomei, 2009).

The asynchronously replicating status of imprinted loci has been linked to DNA methylation and other epigenetic marks associated with imprinted gene silencing (Dünzinger et al., 2005). However in birds, which have no fetal-maternal exchange and display no form of genomic imprinting, there are several conserved regions of mammalian imprinted gene orthologs that are asynchronously replicated (Dünzinger et al., 2005). These asynchronously replicating regions are found on chicken macrochromosomes which, compared to their microchromosome counterparts, are hypoacetylated, hypomethylated, late replicating, and display a lower recombination rate during meiosis (Consortium, 2004; Grutzner et al., 2001; McQueen et al., 1998; Schmid et al., 1989). This indicates that asynchronous replication predates imprinting, and that the common vertebrate ancestor of mammals and birds had genomic regions with a 'pre-imprinted' status, whereby asynchronous replication still occurred (Dünzinger et al., 2005). It will be interesting to see whether monotreme orthologs of imprinted genes also replicate asynchronously, as observed in birds (Dünzinger et al., 2005).

5.2 Allelic exclusion genes

Allelic exclusion is a process whereby the future expression from one allele of a locus is chosen in a cell, resulting in monoallelic expression at the locus. Allelic exclusion is a feature of many multigene families, with olfactory gene clusters and immunoglobulin gene clusters being two classic groups of genes utilizing this form of epigenetic control. However there are also other groups of genes which utilize allelic exclusion, including interleukins and the p120 catenin (Gimelbrant et al., 2005; Hollander et al., 1998). Many epigenetic elements control the cell's choice over which allele will be active, including cis and trans-acting DNA sequences, long-range interactions, and chromatin modification (reviewed in Zakharova et al., 2009).

5.2.1 Olfactory genes

Whilst some olfactory receptor (ORc) genes are dispersed in the mammalian genome, many exist in clusters (Kambere and Lane, 2007). The largest cluster in mouse consists of 244 ORc genes, whilst in human the largest cluster contains 116 genes (Godfrey et al., 2004; Malnic et al., 2004). Both species have individual ORc genes and ORc clusters spread across many different chromosomes, with a few chromosomes containing large clusters of ORcs (Glusman et al., 2001; Kambere and Lane, 2007). However, even though the eutherian genome contains around 1000 ORc genes, only a single ORc gene will be expressed in a single olfactory neuron, meaning that that neuron will only express one type of odorant receptor (Malnic et al., 1999). Furthermore in a process known as allelic inactivation, the locus that is being expressed undergoes differential epigenetic processes at each allele that cause one allele to be inactivated, and thus monoallelic expression of the gene (Chess et al., 1994).

Chromosome conformation capture (3C) assays have given an insight into the mechanisms surrounding the selection of a single ORc gene (Lomvardas et al., 2006; Serizawa et al., 2003). The recently developed 3C technique has become invaluable to studies on nuclear architecture, as it is able to detect and quantify long-range DNA interactions *in vivo*, at high resolution, between sequences in close nuclear proximity. The technique relies on the cross-linking of proteins using formaldehyde in intact nuclei or cells (Dekker et al., 2002). The result is that proteins are cross-linked to other proteins and to adjacent chromatin (Orlando et al., 1997). DNA regions that are actually touching at the time of fixation will be held together via the cross-linking of their DNA bound proteins. The cross-linked genomic DNA is then digested with DNA restriction enzymes and the resulting DNA segments are then ligated. Finally, PCR across these ligation sites detects long-range interacting regions at the DNA sequence level (Dekker et al., 2002).

The 3C experiments on olfactory neurons indicated that ORcs undergo an interaction with a long-range interacting region called the "H element", located within the mouse ORc gene cluster *MOR28*, and perhaps do so in a competitive manner in order to become the activated ORc gene (Fuss et al., 2007; Lomvardas et al., 2006; Serizawa et al., 2003) so that only one gene will be chosen and actively expressed (Lomvardas et al., 2006; Serizawa et al., 2003). However another study showed that deletion of the H element only affected proximal genes within its *MOR28* cluster, with no effect on genes outside this cluster, indicating that it cannot be the only factor involved in terms of activating ORc genes in long-range *cis* and *trans* conformations (Fuss et al., 2007).

ORc genes are observed to undergo asynchronous replication (Table 1), with different clusters and individual ORc genes on the same chromosome undergoing replication at the same time in the S-phase, and the establishment of this form of replication occurring in early embryogenesis (Chess et al., 1994; Mostoslavsky et al., 2001; Singh et al., 2003). The asynchronous replication of ORc loci is believed to be controlled in part by the Polycomb group methyltransferase *Eed*, as ORc genes lose their asynchronously replicating status in its absence (Alexander et al., 2007). This could explain how ORc genes located on the same chromosome are observed to undergo asynchronous replication, with *Eed* being a requirement for asynchronous replication, regardless of position on a chromosome (Alexander et al., 2007; Singh et al., 2003).

5.2.2 Immunoglobulin gene loci

It has been suggested that asynchronous replication plays an important role in the selection of which parental allele will undergo V(D)J rearrangement. The allelic exclusion process in mouse occurs for the genes which do not undergo intrachromosomal recombination, and thus are silenced. The rearrangement process of the immunoglobulin genes in mouse requires crosstalk between two loci from two different chromosomes, namely the IgH locus (containing V, D and J gene segments), and Igκ locus (containing V and J segments). The *de novo* methylation of all the VDJ alleles occurs at the implantation stage, and this is also when asynchronous replication is established (Table 1) (Mostoslavsky et al., 2001). However, the selection of one allele at each locus to undergo early replication puts this allele down a demethylation and chromatin opening pathway, allowing it to be rearranged and to become a functional gene (Goldmit et al., 2002). The other late replicating allele however, remains methylated and cannot be rearranged, and is therefore functionally silenced (Goldmit et al., 2002). The two alleles also have different histone marks with the inactive allele binding the heterochromatin specific protein HP1, and the active allele displaying active histone marks such as di- or trimethylated H3K4, and H3 and H4 acetylation (reviewed in Zakharova et al., 2009).

Asynchronous replication and monoallelic expression are hallmarks of genes which undergo imprinting, X-inactivation, and allelic exclusion. Whilst each might come with its own epigenetic makeup, there are also similarities in the types of epigenetic marks observed to differentiate the active allele (with active histone marks) from the inactive allele (with silencing histone marks). Furthermore, the very fact that asynchronous replication occurs together with different forms of epigenetic monoallelic expression suggests that asynchronous replication may have evolved as a mechanism to control the expression of underlying genes, helping to establish the correct epigenetic marks for monoallelic expression.

6. The CTCF protein and the interactome

The CCCTC-binding factor (CTCF) is a renowned genome organiser, and has roles in regulating long-range chromatin interactions (both intrachromsomal and interchromosomal), but also has roles in other processes such as transcriptional insulation, activation/repression, imprinting control, and X-inactivation (Ling et al., 2006; Murrell et al., 2004; Phillips and Corces, 2009). It is also implicated to have roles in sister chromatid cohesion during DNA replication, as CTCF has been shown to interact with the STAG1

		Placentals			Monotremes (Platypus)			Birds (Chicken)		
	Gene/ Region	% SD	N	Reference	% SD	N	Reference	% SD	N	Reference
Sex chromo-some specific regions	*Xist*	39% Mus	138	(Gribnau et al., 2005)	NA					
	Mecp2	33% Mus	108	(Gribnau et al., 2005)						
	Smcx	38% Mus	157	(Gribnau et al., 2005)						
	OGN	NA			22%	587	(Ho et al., 2009)			
	APC				29%	420	(Ho et al., 2009)			
Imprin-ted genes (in euthe-rians)	*Igf2*	23% Mus	>100	(Kitsberg et al., 1993)	NA			25%*	258	(Dünzinger et al., 2005)
	Igf2R	35% Mus	>100	(Kitsberg et al., 1993)				22%	279	(Dünzinger et al., 2005)
	Mest/Copg2	25% HSA	>200	(Bentley et al., 2003)				18%	299	(Dünzinger et al., 2005)
Allelic exclu-sion	*TCRβ*	46% Mus	100-300	(Mostoslavsky et al., 2001)						
	B-cell receptor (κ)	48% Mus	100-300	(Mostoslav-sky et al., 2001)						
	IL-2	68% Mus	100	(Hollander et al., 1998)						
	Olfactory receptor	31% Mus	> 99	(Simon et al., 1999)						

Table 1. A subset of the asynchronous replication data that exists for genes/genomic regions which are sex chromosome specific, imprinted (in eutherians), or undergo allelic exclusion.

subunit of cohesin and localize cohesin to specific CTCF binding sites on chromosome arms (Parelho et al., 2008; Rubio et al., 2008). Important in this context is that the CTCF protein has been shown to mediate asynchronous replication and imprinting control for the *Igf2-H19* cluster (Bergstrom et al., 2007).

6.1 The evolution of CTCF

The CTCF protein is highly conserved across higher eukaryotes, and the active site shows close to 100% homology between mouse, human and chicken suggesting that the protein has a highly conserved role (Ohlsson et al., 2001). A *CTCF* gene duplication event is believed to have occurred in the amniote ancestor preceding the divergence of reptiles and birds, as they both have functional CTCF, but not its gene paralogue, *BORIS* (brother of regulator of imprinted sites) (Hore et al., 2008). BORIS has similar DNA binding capabilities to CTCF, but shows antagonistc epigenetic regulation to *CTCF*, as well as gonad-specific expression in placental and marsupial mammals (Hore et al., 2008). Conversely, BORIS appears to be widely expressed in monotremes and reptiles, indicating that the gene underwent a

functional change after the divergence of therian mammals, which is interesting because as yet, there is no evidence that CTCF binding sites exist in the genomes of earlier diverged monotreme mammals (Hore et al., 2008; Weidman et al., 2004). However, CTCF sites have been observed in the chicken genome which is an earlier-split vertebrate than the monotreme clade, and tied with the evidence that CTCF binding occurs in therian genomes (Baniahmad et al., 1990; Lobanenkov et al., 1990), it is likely that CTCF sites exist in the montreme genome.

6.2 CTCF and genome organization

It is hypothesised that although chromatin fibres are subjected to random contacts, and thus will always inhabit slightly different positions in the nucleus, the characteristics of the interacting regions on chromosomes allow interactions to occur (de Laat and Grosveld, 2007). Furthermore, it has been argued that genomic regions preferentially interact with other genomic regions that have similar characteristics to their own, such as regions that share CTCF binding (de Laat and Grosveld, 2007). It has been hypothesised that regions of a chromosome which undergo similar replication timing, like asynchronously replicating genes, may be pulled into similar replication domains (Ryba et al., 2010; Singh et al., 2003). Within the mammalian cell nucleus, chromatin from separate chromosomes is organised into the aforementioned chromosome territories. Within these CTs, a higher order of chromatin structure exists, where domains containing specific chromosomal arms and bands have been found to be located in the nucleus in similar regions of certain cell types (Dietzel et al., 1998). Genes are readily transcribed when they reside on the periphery of chromosome territories, and can even loop out of the territories. Furthermore, genes that are late-replicating and inactivated are often seen to reside on the outer regions of chromosome territories near the nuclear periphery. Looping of the chromatin fibres allows genes to easily interact with the transcriptional machinery residing in the interchromatin compartments (Cremer and Cremer, 2001; Osborne et al., 2004). Imprinted and allelic exclusion genes often 'loop out' and undergo long-range interactions for regulatory purposes (Ling and Hoffman, 2007; Lomvardas et al., 2006).

A good example of CTCF controlling some of the discussed epigenetic, replication, and transcriptional mechanisms occurs at the imprinted *Igf2/H19* domain. The ICR for this imprinted cluster lies between these two genes, in the 5' flanking sequence of *H19*, and the maternal allele interacts with CTCF (Kurukuti et al., 2006). CTCF regulates and insulates imprinted gene transcription for the *Igf2/H19* region by controlling the intrachromosomal interactions of the maternal and paternal alleles (Murrell et al., 2004). When endogenous CTCF is knocked-down in mice, loss of *Igf2* imprinting is observed, whilst deletion of the ICR leads to biallelic expression of *H19* (Ling et al., 2006). In mouse, the paternal chromosome forms a DNA loop between the differentially methylated region (DMR) 2, present in the *Igf2* gene, and the methylated ICR, aided by putative binding factors (Murrell et al., 2004). When the paternal *Igf2* allele promoter comes into close proximity with the *H19* enhancer elements, *Igf2* transcription occurs (Murrell et al., 2004). The DMR1 on the maternal chromosome interacts with the unmethylated ICR, which causes the maternal *Igf2* allele to be sequestered into a transcriptional silencing loop. This causes the maternal *H19* allele to become proximal to its enhancers, allowing it to be expressed (Murrell et al., 2004).

Conversely, CTCF also facilitates an interchromosomal interaction in mouse between the *Igf2/H19* domain, and the *Wsb1/Nf1* region on a different chromosome (Ling et al., 2006).

Specifically, the ICR in the imprinted *Igf2/H19* domain, which contains CTCF binding sites, has been found to interact with another region with CTCF binding sites between the *Wsb1* (WD repeat and SOCS box-containing 1) and *Nf1* (Neurofibromin 1) genes (Ling et al., 2006). Whilst the *Wsb1* and *Nf1* do not appear to be imprinted, as their expression is biallelic, only the paternal copy of the *Wsb1/Nf1* region interacts with CTCF (Krueger and Osborne, 2006; Ling et al., 2006). As explained before, CTCF only binds the maternal copy of the ICR region (flanked by *Igf2* and *H19*). It is consequently hypothesized that the long-range interaction observed between the ICR and *Wsb1/Nf1* region occurs between the maternal and paternal copies respectively, and is mediated by the genome-organizing protein CTCF (Ling et al., 2006).

6.3 Replication timing and CTCF

The specific binding of CTCF at the maternal ICR in the mouse *Igf2/H19* domain has been shown to mediate asynchronous replication in this imprinted region (Bergstrom et al., 2007). The inheritance of a mutated maternal ICR, which lacks CTCF binding, caused the usually late replicating maternal *Igf2/H19* domain to become early replicating (Bergstrom et al., 2007) showing that CTCF binding is required for asynchronous replication of these loci. The mechanism by which CTCF might regulate asynchronous replication at this domain, however, is still unclear. In addition to replication CTCF is involved in other epigenetic effects, including long-range interactions (both intrachromosomal and interchromosomal), insulator activity and transcriptional activation (Kurukuti et al., 2006; Ohlsson et al., 2001; Zhao et al., 2006). Notably, it has been shown that regions which undergo greater amounts of long-range chromatin interaction are subject to late replication timing (Ryba et al., 2010).

Another example of the close relationship between replication, CTCF, and methylation occurs at the differentially methylated silencer region controlling the expression of the *AWT1/ WT1-AS* genes (Hancock et al., 2007). The CTCF protein can only bind the late-replicating unmethylated paternal silencer region within the *AWT1/WT1-AS* cluster, allowing expression of the paternal alleles. The homologous early-replicating maternal region however, has a methylated silencer which does not facilitate CTCF binding and so the maternal *AWT1/WT1-AS* alleles are not expressed (Hancock et al., 2007). It is interesting to speculate as to whether CTCF also controls the asynchronous replication observed at the *WT1* locus in human, and perhaps even in birds (Bickmore and Carothers, 1995; Dünzinger et al., 2005). It is also interesting to note that in both cases the late-replicating allele at these imprinted loci, namely in the maternal *Igf2/H19* allele and the paternal *AWT1/WT1-AS* allele, is the allele which binds CTCF (Bergstrom et al., 2007; Bickmore and Carothers, 1995; Hancock et al., 2007). Whilst CTCF is observed to mediate asynchronous replication and imprinting at the *Igf2/H19* domain in eutherian mammals, the fact that the imprinted orthologs of *Igf2/H19* and *AWT1/WT1-AS* still asynchronously replicate could suggest that CTCF binding in these regions evolved before establishment of genomic imprinting.

6.4 The role of CTCF in replication timing changes in cancer

CTCF may also play a role in the progression of cancer and has many of the characteristics of a tumour suppressor gene; in the human genome it maps to a small region, 16q22.1, which characteristically undergoes loss of heterozygosity in many solid tumours (reviewed in Filippova et al., 1998). Furthermore, changes in DNA consensus sites and DNA methylation patterns in cancers are known to cause loss of CTCF binding, which could result in the loss of functional control of these regions (Filippova et al., 2002; Ohlsson et al., 2001). The regions required for zinc-finger formation, and their corresponding DNA binding

sites are often mutated in tumours, changing the CTCF binding-landscape of a genome (Filippova et al., 2002). Specifically, the presence of these mutations in tumours was observed to abolish CTCF's association with the *Igf2/H19* growth regulating genes, whilst not changing its association with non-growth regulating genes (Filippova et al., 2002; Ohlsson et al., 2001). The loss of CTCF association with the *Igf2/H19* region in tumours could be associated with a shift in replication asynchrony. As mentioned in the previous section, when CTCF binding is abolished in the maternal *Igf2/H19* region it results in the loss of asynchronous replication at the locus (Bergstrom et al., 2007). Furthermore, omission of CTCF binding to the maternal *Igf2/H19* ICR has also been observed to abrogate inter-chromosomal interactions for this region (Ling et al., 2006). These results all indicate that the loss of CTCF binding for specific genomic regions in tumours has downstream epigenetic effects, such as loss of replication asynchrony and chromatin interaction, for the genes usually involved in CTCF-interaction.

7. Evolution of replication timing and epigenetic control

7.1 The evolution of replication timing

At the genome level, recent work shows that asynchronous replication pre-dates the establishment of monoallelic expression and genomic imprinting (Zechner et al. 2006, Wright et al. in preparation). The bird genome, which lacks genomic imprinting, contains conserved regions of mammalian imprinted gene orthologs that are asynchronously replicated (Dünzinger et al., 2005). This indicates that asynchronous replication most likely predates imprinting, and that the common vertebrate ancestor of mammals and birds had genomic regions with a 'pre-imprinted' status which still underwent asynchronous replication without any form of traditional imprinting (Dünzinger et al., 2005). It is interesting to note that a recent genome-wide study has indicated that regions with conserved synteny also have conserved replication profiles among human and mouse (e.g. Ryba et al., 2010). Imprinted clusters are renowned for having conserved synteny, and it has been suggested that the selection of highly conserved arrays of imprinted gene orthologs occured during vertebrate evolution, however why these regions were selected for syntenic conservation has been difficult to explain (Dünzinger et al., 2005).

At the replicon level, there has been a model proposing that spatiotemporal properties of mammalian ORs contribute to a combination of pre-determined and stochastic DNA replication (Takahashi, 1987). This mechanism is echoed in budding yeast, which also shows OR activation in a combined chronological and stochastic manner (Barberis et al., 2010; Spiesser et al., 2009). This model, combined with the finding that conserved syntenic regions in human and mouse have very similar replication profiles, indicates that there is a conservation of the temporal programme controlling replicon firing. Furthermore there appears to be a highly conserved order in which amniote imprinted genes or imprinted gene orthologs replicate; with individual imprinted genes following similar temporal patterns when entering replication in birds, monotremes, and eutherians (Wright et al. in preparation). This indicates that in closer related clades of eukaryotes, this temporal replication program may be highly conserved.

7.2 The chromatin interactome and replication profiling

Developing molecular technologies are allowing greater insights into the many interactions occurring in a genome, but also showing how spatial organisation can affect other processes

in a genome, such as replication timing. Extensions of the previously discussed 3C molecular interaction technology include Associative Chromosome Trap (ACT), Circular Chromosome Conformation Capture or Chromosome Conformation Capture-on-Chip (4C), and Carbon-Copy Chromosome Conformation Capture (5C), all of which can measure more than a single to single region interaction (Dekker et al., 2002; Dostie et al., 2006; Ling et al., 2006; Simonis et al., 2006; Zhao et al., 2006). In addition to these technologies, new techniques are allowing interactions to be measured across entire genomes, resulting in the mapping of an "interactome", whereby all the long-range interactions occurring in a genome are measured (Fullwood et al., 2009; Lieberman-Aiden et al., 2009). Specifically, there are two techniques that have been developed to do this, Chromatin Interaction Analysis by Paired-End Tag sequencing (ChIA-PET) and Hi-C (which measures the three-dimensional architecture of a genome by coupling proximity-based ligation with parallel sequencing) (Fullwood et al., 2009; Lieberman-Aiden et al., 2009). These experiments, in conjunction with replication-timing profiling by microarrays, have indicated that the interactome of a genome is very closely aligned with replication timing (Ryba et al., 2010).
The chromatin "interactome" is now understood to play a critical part in genome organisation; allowing complex regulatory networks of interactions to occur, each of which with functional significance, all of which highly dynamic and organised within a nucleus by proteins such as CTCF and the Estrogen-receptor alpha (Botta et al., 2010; Fullwood et al., 2009). These interactions also appear to be conserved in similar cell types across mammalian evolution, suggesting that perhaps these long-range interactions are part of an evolutionary conserved mechanism of spatial organisation (Ryba et al., 2010). Furthermore, initiation of replication appears to be an evolutionarily conserved process across eukaryotic evolution, and the overlay of entire genome replication timing profiles with interactome maps have shown that late-replicating regions are often undergoing greater amounts of long-range interaction (Ryba et al., 2010). These findings, in conjunction with asynchronous replication data, could indicate that long-range interactions which occur in abundance at imprinted and monoallelically expressed loci, are affecting asynchronous replication. Specifically, there is data supporting the argument that the allele undergoing long-range interaction could also be the allele which undergoes late-replication. Firstly, it has been observed that asynchronously replicated alleles often localize to spatially distinct regions in a nucleus (Gribnau et al., 2003; Sadoni et al., 1999). Secondly, as mentioned previously, the late-replicating maternal *Igf2/H19* allele and the paternal *AWT1/WT1-AS* allele, are also the alleles which bind CTCF, in an imprinting dependent manner. It could be that the binding of proteins which mediate long-range chromatin interaction at these alleles is facilitating greater amounts of interaction, which is reflected in their late replicating status, and also in the asynchronous replication of these genes (Bergstrom et al., 2007; Bickmore and Carothers, 1995; Hancock et al., 2007).

7.3 Measuring replication to combat cancer

It has been proposed that measuring changes in replication profiles may be a way of detecting abnormalities associated with cancer, not observed through usual techniques (reviewed in Watanabe and Maekawa, 2010). Epigenetic reprogramming in diseased cells is often observed to occur with changes in replication timing patterns, with changes in replication being observed with chromosomal rearrangements in cancer cell lines (D'Antoni et al., 2004; Gondor and Ohlsson, 2009; State et al., 2003). Better detection of prostate cancer

may come in the form of measuring replication timing changes observed in peripheral blood lymphocytes undergoing aneuploidy (Dotan et al., 2004). In terms of protein detection of cancer, measuring the function of the tumour suppressor gene p53, may be a good determinant in the progression of cancer. P53 is the most commonly mutated gene in human cancers, and is a G1/S-phase and S-phase checkpoint regulator during DNA replication. Loss of its function is observed to affect the replication timing of human colon carcinoma cells (Watanabe et al., 2007).

Changes in replication timing may also be affected by altered function of CTCF in cancer. As mentioned previously, it has been observed that mutation of CTCF binding sites near growth factor genes, such as in the *Igf2/H19* region, occurs in many tumours (Filippova et al., 2002). These mutations may cause a loss of CTCF binding in the region, which has been observed to abolish asynchronous replication of the *Igf2* locus, and changes the replication timing of the gene (Bergstrom et al., 2007). However the mutation of CTCF binding sites would also change the interactome profile of a cell. Loss of CTCF-binding through mutation around genes like *Igf2* and *H19* would result in them no longer undergoing their "normal" chromatin interactions, perhaps causing different spatial organization of these loci in the nucleus of a cancerous cell.

7.4 The chromatin interactome: controlling eukaryotic replication timing

To date there is a lack of data that could provide insight about the evolution of an interactome. It has been observed that many long-range interacting regions share many of the same (but not necessarily all) epigenetic characteristics, such as asynchronous replication, monoallelic expression, differentially methylated regions and histone modifications and variants, imprinting, and CTCF binding. It is currently unknown how these epigenetic events evolved and investigating those epigenetic features in a range of vertebrate genomes could tease apart the sequence of events that has led to a complex network of epigenetic regulation.

Chromatin interactions may have evolved in many genomic control processes, but it is the binding of master genome regulators, like CTCF, which dictate where these interactions can occur. The CTCF protein is highly conserved among amniotes, conserved in vertebrates, and exists in *Drosophila* and subsets of nematodes (Heger et al., 2009; Ohlsson et al., 2001). Furthermore, there is evidence to suggest that CTCF binding and function are conserved in humans, mouse, and chicken, in genes such as β-globin, whereby CTCF binding at this locus allows cell-type specific intrachromosomal interactions to occur (Bell et al., 1999; Yusufzai et al., 2004). CTCF binding and chromatin interaction in this region suggest that CTCF spatial control of chromatin, at least in this region, was present in the common ancestor of amniotes. The evolutionary conservation of replication timing and the strikingly similar genomic interactome in similar cell types among human and mouse suggests that replication timing is intrinsically tied to long-range interaction. Moreover, there is evidence to suggest that replication timing relies on the presence of long-range interactions at specific loci, with the knockdown of long-range mediator proteins causing interactions to be abolished, and also causing replication asynchrony to cease (Bergstrom et al., 2007; Fullwood et al., 2009; Ling et al., 2006). The loss of replication asynchrony in this case could be due to ectopic spatial organisation of the alleles, whereby the loss of the interaction mediator protein causes the allele of a locus to reside in an atypical subnuclear domain. This irregular replication domain would not have the correct molecular and chemical

characteristics to allow the ORs of the spatially ectopic allele to fire in the normal temporal order. This could cause the erroneous firing of ORs in such a way as to abolish replication asynchrony at the locus.

8. Conclusion

Replication timing of DNA at S-phase is tightly regulated and affects gene activity, nuclear organisation, as well as other aspects of genome biology. Differences in replication timing have been used to identify individual chromosomes and differentiated sex chromosomes for several decades. Since then, an increasing number of proteins have been identified as important for regulating replication timing and genome-wide approaches are now used to study replication timing. A fascinating variation of the replication-timing theme is asynchronous replication, which appears to be closely aligned with other epigenetic mechanisms involved in long-range interaction, genomic imprinting and X chromosome inactivation. Whilst previous research has stipulated that asynchronous replication and long range interactions have evolved as a result of epigenetic control of (eg. monoallelic expression), there is emerging evidence that both predate the presence of other epigenetic processes. We suggest that the interactome has played a role in the evolution of spatial nuclear organisation. In addition, mutations in sequences important for long-range interaction and replication timing, and also changes in the replication timing program itself, are important factors influencing a diverse array of human diseases, including cancer. The study of replication timing in different organisms and in human disease will reveal the full extent to which replication timing contributes to the epigenetic landscape in normal and abnormal cells.

9. Acknowledgements

Megan L. Wright is funded by an Australian Postgraduate Award. Frank Grutzner is an ARC Senior Research Fellow.

10. References

Aladjem, M. I., Groudine, M., Brody, L. L., Dieken, E. S., Fournier, R. E., Wahl, G. M. and Epner, E. M. (1995). Participation of the human beta-globin locus control region in initiation of DNA replication. *Science* 270, 815-9.

Alexander, M. K., Mlynarczyk-Evans, S., Royce-Tolland, M., Plocik, A., Kalantry, S., Magnuson, T. and Panning, B. (2007). Differences between homologous alleles of olfactory receptor genes require the Polycomb Group protein Eed. *J Cell Biol* 179, 269-76.

Altman, A. L. and Fanning, E. (2004). Defined sequence modules and an architectural element cooperate to promote initiation at an ectopic mammalian chromosomal replication origin. *Mol Cell Biol* 24, 4138-50.

Anglana, M., Apiou, F., Bensimon, A. and Debatisse, M. (2003). Dynamics of DNA replication in mammalian somatic cells: nucleotide pool modulates origin choice and interorigin spacing. *Cell* 114, 385-94.

Arnold, A. P., Itoh, Y. and Melamed, E. (2008). A bird's-eye view of sex chromosome dosage compensation. *Annu Rev Genomics Hum Genet* 9, 109-27.

Aten, J. A. and Kanaar, R. (2006). Chromosomal organization: mingling with the neighbors. *PLoS Biol* 4, e155.

Austin, R. J., Orr-Weaver, T. L. and Bell, S. P. (1999). Drosophila ORC specifically binds to ACE3, an origin of DNA replication control element. *Genes Dev* 13, 2639-49.

Baniahmad, A., Steiner, C., Kohne, A. C. and Renkawitz, R. (1990). Modular structure of a chicken lysozyme silencer: involvement of an unusual thyroid hormone receptor binding site. *Cell* 61, 505-14.

Baran, N., Lapidot, A. and Manor, H. (1991). Formation of DNA triplexes accounts for arrests of DNA synthesis at d(TC)n and d(GA)n tracts. *Proceedings of the National Academy of Sciences* 88, 507-511.

Barberis, M., Spiesser, T. W. and Klipp, E. (2010). Replication origins and timing of temporal replication in budding yeast: how to solve the conundrum? *Curr Genomics* 11, 199-211.

Bartolomei, M. S. (2009). Genomic imprinting: employing and avoiding epigenetic processes. *Genes Dev* 23, 2124-33.

Baumer, A., Riegel, M. and Schinzel, A. (2004). Non-random asynchronous replication at 22q11.2 favours unequal meiotic crossovers leading to the human 22q11.2 deletion. *Journal of Medical Genetics* 41, 413-420.

Bell, A. C., West, A. G. and Felsenfeld, G. (1999). The Protein CTCF Is Required for the Enhancer Blocking Activity of Vertebrate Insulators. *Cell* 98, 387-396.

Bell, S. P. and Stillman, B. (1992). ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* 357, 128-34.

Bentley, L., Nakabayashi, K., Monk, D., Beechey, C., Peters, J., Birjandi, Z., Khayat, F. E., Patel, M., Preece, M. A., Stanier, P. et al. (2003). The imprinted region on human chromosome 7q32 extends to the carboxypeptidase A gene cluster: an imprinted candidate for Silver-Russell syndrome. *J Med Genet* 40, 249-56.

Bergstrom, R., Whitehead, J., Kurukuti, S. and Ohlsson, R. (2007). CTCF regulates asynchronous replication of the imprinted H19/Igf2 domain. *Cell Cycle* 6, 450-4.

Bernardi, G. (1995). The human genome: organization and evolutionary history. *Annu Rev Genet* 29, 445-76.

Bickmore, W. A. and Carothers, A. D. (1995). Factors affecting the timing and imprinting of replication on a mammalian chromosome. *J Cell Sci* 108 (Pt 8), 2801-9.

Biederman, B. and Lin, C. C. (1979). R-banding methods for chromosome identification using acridine orange following heat denaturation (RFA) and bromodeoxyuridine labeling (RBA). *Methods in Cell Science* 5, 1189-1192.

Blow, J. J. and Ge, X. Q. (2009). A model for DNA replication showing how dormant origins safeguard against replication fork failure. *EMBO Rep* 10, 406-12.

Botta, M., Haider, S., Leung, I. X. Y., Lio, P. and Mozziconacci, J. (2010). Intra- and inter-chromosomal interactions correlate with CTCF binding genome wide. *Mol Syst Biol* 6.

Boumil, R. M. and Lee, J. T. (2001). Forty years of decoding the silence in X-chromosome inactivation. *Hum Mol Genet* 10, 2225-32.

Branco, M. R. and Pombo, A. (2006). Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS Biol* 4, e138.

Brewer, B. J. and Fangman, W. L. (1993). Initiation at closely spaced replication origins in a yeast chromosome. *Science* 262, 1728-31.

Brinton, B. T., Caddle, M. S. and Heintz, N. H. (1991). Position and orientation-dependent effects of a eukaryotic Z-triplex DNA motif on episomal DNA replication in COS-7 cells. *Journal of Biological Chemistry* 266, 5153-5161.

Brockdorff, N., Ashworth, A., Kay, G. F., Cooper, P., Smith, S., McCabe, V. M., Norris, D. P., Penny, G. D., Patel, D. and Rastan, S. (1991). Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* 351, 329-31.

Brown, C. J., Ballabio, A., Rupert, J. L., Lafreniere, R. G., Grompe, M., Tonlorenzi, R. and Willard, H. F. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* 349, 38-44.

Burke, T. W., Cook, J. G., Asano, M. and Nevins, J. R. (2001). Replication factors MCM2 and ORC1 interact with the histone acetyltransferase HBO1. *J Biol Chem* 276, 15397-408.

Chess, A., Simon, I., Cedar, H. and Axel, R. (1994). Allelic inactivation regulates olfactory receptor gene expression. *Cell* 78, 823-34.

Chuang, R. Y. and Kelly, T. J. (1999). The fission yeast homologue of Orc4p binds to replication origin DNA via multiple AT-hooks. *Proc Natl Acad Sci U S A* 96, 2656-61.

Clerc, P. and Avner, P. (2003). Multiple elements within the Xic regulate random X inactivation in mice. *Semin Cell Dev Biol* 14, 85-92.

Consortium, I. C. G. S. (2004). Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432, 695-716.

Costantini, M. and Bernardi, G. (2008). Replication timing, chromosomal bands, and isochores. *Proc Natl Acad Sci U S A* 105, 3433-7.

Costantini, M., Clay, O., Auletta, F. and Bernardi, G. (2006). An isochore map of human chromosomes. *Genome Res* 16, 536-41.

Costantini, M., Clay, O., Federico, C., Saccone, S., Auletta, F. and Bernardi, G. (2007). Human chromosomal bands: nested structure, high-definition map and molecular basis. *Chromosoma* 116, 29-40.

Courbet, S., Gay, S., Arnoult, N., Wronka, G., Anglana, M., Brison, O. and Debatisse, M. (2008). Replication fork movement sets chromatin loop size and origin choice in mammalian cells. *Nature* 455, 557-60.

Cremer, T. and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* 2, 292-301.

Cremer, T. and Cremer, M. (2010). Chromosome territories. *Cold Spring Harb Perspect Biol* 2, a003889.

D'Antoni, S., Mattina, T., Di Mare, P., Federico, C., Motta, S. and Saccone, S. (2004). Altered replication timing of the HIRA/Tuple1 locus in the DiGeorge and Velocardiofacial syndromes. *Gene* 333, 111-119.

Davidow, L. S., Breen, M., Duke, S. E., Samollow, P. B., McCarrey, J. R. and Lee, J. T. (2007). The search for a marsupial XIC reveals a break with vertebrate synteny. *Chromosome Res* 15, 137-46.

de Laat, W. and Grosveld, F. (2007). Inter-chromosomal gene regulation in the mammalian cell nucleus. *Curr Opin Genet Dev* 17, 456-64.

Deakin, J. E., Hore, T. A., Koina, E. and Marshall Graves, J. A. (2008a). The status of dosage compensation in the multiple X chromosomes of the platypus. *PLoS Genet* 4, e1000140.

Deakin, J. E., Koina, E., Waters, P. D., Doherty, R., Patel, V. S., Delbridge, M. L., Dobson, B., Fong, J., Hu, Y., van den Hurk, C. et al. (2008b). Physical map of two tammar wallaby chromosomes: a strategy for mapping in non-model mammals. *Chromosome Res* 16, 1159-75.

Debatisse, M., El Achkar, E. and Dutrillaux, B. (2006). Common fragile sites nested at the interfaces of early and late-replicating chromosome bands: cis acting components of the G2/M checkpoint? *Cell Cycle* 5, 578-81.

Debatisse, M., Toledo, F. and Anglana, M. (2004). Replication initiation in mammalian cells: changing preferences. *Cell Cycle* 3, 19-21.

Dekker, J., Rippe, K., Dekker, M. and Kleckner, N. (2002). Capturing chromosome conformation. *Science* 295, 1306-11.

Delcuve, G. P., Rastegar, M. and Davie, J. R. (2009). Epigenetic control. *Journal of Cellular Physiology* 219, 243-250.

DePamphilis, M. L. (2006). DNA replication and human disease. New York: Cold Spring Harbor Laboratory Press.

Dietzel, S., Jauch, A., Kienle, D., Qu, G., Holtgreve-Grez, H., Eils, R., Munkel, C., Bittner, M., Meltzer, P. S., Trent, J. M. et al. (1998). Separate and variably shaped chromosome arm domains are disclosed by chromosome arm painting in human cell nuclei. *Chromosome Res* 6, 25-33.

Doksani, Y., Bermejo, R., Fiorani, S., Haber, J. E. and Foiani, M. (2009). Replicon dynamics, dormant origin firing, and terminal fork integrity after double-strand break formation. *Cell* 137, 247-58.

Dostie, J. e., Richmond, T. A., Arnaout, R. A., Selzer, R. R., Lee, W. L., Honan, T. A., Rubio, E. D., Krumm, A., Lamb, J., Nusbaum, C. et al. (2006). Chromosome Conformation Capture Carbon Copy (5C): A massively parallel solution for mapping interactions between genomic elements. *Genome Research* 16, 1299-1309.

Dotan, Z. A., Dotan, A., Ramon, J. and Avivi, L. (2004). Altered mode of allelic replication accompanied by aneuploidy in peripheral blood lymphocytes of prostate cancer patients, vol. 111 (ed., pp. 60-66: Wiley Subscription Services, Inc., A Wiley Company.

Drets, M. E., Folle, G. A. and Comings, D. E. (1978). Mechanisms of chromosome banding. X. Chromosome and nuclear changes induced by photo-oxidation and their relation to R-banding with anti-C antibodies. *Chromosoma* 69, 101-12.

Dünzinger, U., Nanda, I., Schmid, M., Haaf, T. and Zechner, U. (2005). Chicken orthologues of mammalian imprinted genes are clustered on macrochromosomes and replicate asynchronously. *Trends in Genetics* 21, 488-492.

Duret, L., Chureau, C., Samain, S., Weissenbach, J. and Avner, P. (2006). The Xist RNA gene evolved in eutherians by pseudogenization of a protein-coding gene. *Science* 312, 1653-5.

Edwards, C. A. and Ferguson-Smith, A. C. (2007). Mechanisms regulating imprinted genes in clusters. *Current Opinion in Cell Biology* 19, 281-289.

Edwards, M. C., Tutter, A. V., Cvetic, C., Gilbert, C. H., Prokhorova, T. A. and Walter, J. C. (2002). MCM2-7 complexes bind chromatin in a distributed pattern surrounding the origin recognition complex in Xenopus egg extracts. *J Biol Chem* 277, 33049-57.

Farkash-Amar, S., Lipson, D., Polten, A., Goren, A., Helmstetter, C., Yakhini, Z. and Simon, I. (2008). Global organization of replication time zones of the mouse genome. *Genome Res* 18, 1562-70.

Ferguson, B. M. and Fangman, W. L. (1992). A position effect on the time of replication origin activation in yeast. *Cell* 68, 333-9.

Filippova, G. N., Lindblom, A., Meincke, L. J., Klenova, E. M., Neiman, P. E., Collins, S. J., Doggett, N. A. and Lobanenkov, V. V. (1998). A widely expressed transcription factor with multiple DNA sequence specificity, CTCF, is localized at chromosome segment 16q22.1 within one of the smallest regions of overlap for common deletions in breast and prostate cancers. *Genes Chromosomes Cancer* 22, 26-36.

Filippova, G. N., Qi, C. F., Ulmer, J. E., Moore, J. M., Ward, M. D., Hu, Y. J., Loukinov, D. I., Pugacheva, E. M., Klenova, E. M., Grundy, P. E. et al. (2002). Tumor-associated zinc finger mutations in the CTCF transcription factor selectively alter tts DNA-binding specificity. *Cancer Res* 62, 48-52.

Friedman, K. L., Diller, J. D., Ferguson, B. M., Nyland, S. V., Brewer, B. J. and Fangman, W. L. (1996). Multiple determinants controlling activation of yeast replication origins late in S phase. *Genes Dev* 10, 1595-607.

Fullwood, M. J., Liu, M. H., Pan, Y. F., Liu, J., Xu, H., Mohamed, Y. B., Orlov, Y. L., Velkov, S., Ho, A., Mei, P. H. et al. (2009). An oestrogen-receptor-[agr]-bound human chromatin interactome. *Nature* 462, 58-64.

Fuss, S. H., Omura, M. and Mombaerts, P. (2007). Local and cis effects of the H element on expression of odorant receptor genes in mouse. *Cell* 130, 373-84.

Ge, X. Q., Jackson, D. A. and Blow, J. J. (2007). Dormant origins licensed by excess Mcm2-7 are required for human cells to survive replicative stress. *Genes Dev* 21, 3331-41.

Gilbert, D. M. (2002). Replication timing and transcriptional control: beyond cause and effect. *Curr Opin Cell Biol* 14, 377-83.

Gilbert, D. M. (2010). Evaluating genome-scale approaches to eukaryotic DNA replication. *Nat Rev Genet* 11, 673-84.

Gimelbrant, A. A., Ensminger, A. W., Qi, P., Zucker, J. and Chess, A. (2005). Monoallelic expression and asynchronous replication of p120 catenin in mouse and human cells. *J Biol Chem* 280, 1354-9.

Glusman, G., Yanai, I., Rubin, I. and Lancet, D. (2001). The complete human olfactory subgenome. *Genome Res* 11, 685-702.

Godfrey, P. A., Malnic, B. and Buck, L. B. (2004). The mouse olfactory receptor gene family. *Proc Natl Acad Sci U S A* 101, 2156-61.

Goldmit, M., Schlissel, M., Cedar, H. and Bergman, Y. (2002). Differential accessibility at the kappa chain locus plays a role in allelic exclusion. *EMBO J* 21, 5255-61.

Gondor, A. and Ohlsson, R. (2009). Replication timing and epigenetic reprogramming of gene expression: a two-way relationship? *Nat Rev Genet* 10, 269-276.

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

Gribnau, J., Luikenhuis, S., Hochedlinger, K., Monkhorst, K. and Jaenisch, R. (2005). X chromosome choice occurs independently of asynchronous replication timing. *J Cell Biol* 168, 365-73.

Grutzner, F., Rens, W., Tsend-Ayush, E., El-Mogharbel, N., O'Brien, P. C., Jones, R. C., Ferguson-Smith, M. A. and Marshall Graves, J. A. (2004). In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes. *Nature* 432, 913-7.

Grutzner, F., Zend-Ajusch, E., Stout, K., Munsche, S., Niveleau, A., Nanda, I., Schmid, M. and Haaf, T. (2001). Chicken microchromosomes are hypermethylated and can be identified by specific painting probes. *Cytogenet Cell Genet* 93, 265-9.

Haase, S. B., Heinzel, S. S. and Calos, M. P. (1994). Transcription inhibits the replication of autonomously replicating plasmids in human cells. *Mol Cell Biol* 14, 2516-24.

Hancock, A. L., Brown, K. W., Moorwood, K., Moon, H., Holmgren, C., Mardikar, S. H., Dallosso, A. R., Klenova, E., Loukinov, D., Ohlsson, R. et al. (2007). A CTCF-binding silencer regulates the imprinted genes AWT1 and WT1-AS and exhibits sequential epigenetic defects during Wilms' tumourigenesis. *Hum Mol Genet* 16, 343-54.

Hayashida, T., Oda, M., Ohsawa, K., Yamaguchi, A., Hosozawa, T., Locksley, R. M., Giacca, M., Masai, H. and Miyatake, S. (2006). Replication initiation from a novel origin identified in the Th2 cytokine cluster locus requires a distant conserved noncoding sequence. *J Immunol* 176, 5446-54.

Heard, E. (2005). Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome. *Curr Opin Genet Dev* 15, 482-9.

Heger, P., Marin, B. and Schierenberg, E. (2009). Loss of the insulator protein CTCF during nematode evolution. *BMC Molecular Biology* 10, 84.

Hellman, A. and Chess, A. (2007). Gene body-specific methylation on the active X chromosome. *Science* 315, 1141-3.

Heun, P., Laroche, T., Raghuraman, M. K. and Gasser, S. M. (2001). The positioning and dynamics of origins of replication in the budding yeast nucleus. *J Cell Biol* 152, 385-400.

Hiratani, I. and Gilbert, D. M. (2009). Replication timing as an epigenetic mark. *Epigenetics* 4, 93-7.

Hiratani, I., Ryba, T., Itoh, M., Yokochi, T., Schwaiger, M., Chang, C. W., Lyou, Y., Townes, T. M., Schubeler, D. and Gilbert, D. M. (2008). Global reorganization of replication domains during embryonic stem cell differentiation. *PLoS Biol* 6, e245.

Hiratani, I., Takebayashi, S., Lu, J. and Gilbert, D. M. (2009). Replication timing and transcriptional control: beyond cause and effect--part II. *Curr Opin Genet Dev* 19, 142-9.

Ho, K. K. K., Deakin, J. E., Wright, M. L., Graves, J. A. M. and GrÃ¼tzner, F. (2009). Replication asynchrony and differential condensation of X chromosomes in female platypus (Ornithorhynchus anatinus). *Reproduction, Fertility and Development* 21, 952-963.

Hollander, G. A., Zuklys, S., Morel, C., Mizoguchi, E., Mobisson, K., Simpson, S., Terhorst, C., Wishart, W., Golan, D. E., Bhan, A. K. et al. (1998). Monoallelic expression of the interleukin-2 locus. *Science* 279, 2118-21.

Holmquist, G. P. (1987). Role of replication time in the control of tissue-specific gene expression. *Am J Hum Genet* 40, 151-73.

Hore, T. A., Deakin, J. E. and Marshall Graves, J. A. (2008). The evolution of epigenetic regulators CTCF and BORIS/CTCFL in amniotes. *PLoS Genet* 4, e1000169.

Hore, T. A., Koina, E., Wakefield, M. J. and Marshall Graves, J. A. (2007). The region homologous to the X-chromosome inactivation centre has been disrupted in marsupial and monotreme mammals. *Chromosome Res* 15, 147-61.

Hyrien, O., Marheineke, K. and Goldar, A. (2003). Paradoxes of eukaryotic DNA replication: MCM proteins and the random completion problem. *Bioessays* 25, 116-25.

Ibarra, A., Schwob, E. and Mendez, J. (2008). Excess MCM proteins protect human cells from replicative stress by licensing backup origins of replication. *Proc Natl Acad Sci U S A* 105, 8956-61.

Iizuka, M., Matsui, T., Takisawa, H. and Smith, M. M. (2006). Regulation of replication licensing by acetyltransferase Hbo1. *Mol Cell Biol* 26, 1098-108.

Kalejta, R. F., Li, X., Mesner, L. D., Dijkwel, P. A., Lin, H. B. and Hamlin, J. L. (1998). Distal sequences, but not ori-beta/OBR-1, are essential for initiation of DNA replication in the Chinese hamster DHFR origin. *Mol Cell* 2, 797-806.

Kalish, J. M. and Glazer, P. M. (2005). Targeted genome modification via triple helix formation. *Ann N Y Acad Sci* 1058, 151-61.

Kambere, M. B. and Lane, R. P. (2007). Co-regulation of a large and rapidly evolving repertoire of odorant receptor genes. *BMC Neurosci* 8 Suppl 3, S2.

Kitsberg, D., Selig, S., Brandeis, M., Simon, I., Keshet, I., Driscoll, D. J., Nicholls, R. D. and Cedar, H. (1993). Allele-specific replication timing of imprinted gene regions. *Nature* 364, 459-63.

Knauert, M. P., Kalish, J. M., Hegan, D. C. and Glazer, P. M. (2006). Triplex-stimulated intermolecular recombination at a single-copy genomic target. *Mol Ther* 14, 392-400.

Knoll, J. H., Cheng, S. D. and Lalande, M. (1994). Allele specificity of DNA replication timing in the Angelman/Prader-Willi syndrome imprinted chromosomal region. *Nat Genet* 6, 41-6.

Koina, E., Chaumeil, J., Greaves, I. K., Tremethick, D. J. and Graves, J. A. (2009). Specific patterns of histone marks accompany X chromosome inactivation in a marsupial. *Chromosome Res* 17, 115-26.

Koren, A., Soifer, I. and Barkai, N. (2010). MRC1-dependent scaling of the budding yeast DNA replication timing program. *Genome Res* 20, 781-90.

Krueger, C. and Osborne, C. S. (2006). Raising the curtains on interchromosomal interactions. *Trends in Genetics* 22, 637-639.

Kuroda, Y., Arai, N., Arita, M., Teranishi, M., Hori, T., Harata, M. and Mizuno, S. (2001). Absence of Z-chromosome inactivation for five genes in male chickens. *Chromosome Res* 9, 457-68.

Kuroiwa, A., Yokomine, T., Sasaki, H., Tsudzuki, M., Tanaka, K., Namikawa, T. and Matsuda, Y. (2002). Biallelic expression of Z-linked genes in male chickens. *Cytogenet Genome Res* 99, 310-4.

Kurukuti, S., Tiwari, V. K., Tavoosidana, G., Pugacheva, E., Murrell, A., Zhao, Z., Lobanenkov, V., Reik, W. and Ohlsson, R. (2006). CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proc Natl Acad Sci U S A* 103, 10684-9.

Lee, J. T. (2003). Molecular links between X-inactivation and autosomal imprinting: X-inactivation as a driving force for the evolution of imprinting? *Curr Biol* 13, R242-54.

Lemaitre, J. M., Danis, E., Pasero, P., Vassetzky, Y. and Mechali, M. (2005). Mitotic remodeling of the replicon and chromosome structure. *Cell* 123, 787-801.

Lengronne, A. and Schwob, E. (2002). The yeast CDK inhibitor Sic1 prevents genomic instability by promoting replication origin licensing in late G(1). *Mol Cell* 9, 1067-78.

Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O. et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289-93.

Ling, J. Q. and Hoffman, A. R. (2007). Epigenetics of Long-Range Chromatin Interactions. *Pediatric Research* 61, 11R-16R 10.1203/pdr.0b013e31804575db.

Ling, J. Q., Li, T., Hu, J. F., Vu, T. H., Chen, H. L., Qiu, X. W., Cherry, A. M. and Hoffman, A. R. (2006). CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. *Science* 312, 269-72.

Lobanenkov, V. V., Nicolas, R. H., Adler, V. V., Paterson, H., Klenova, E. M., Polotskaja, A. V. and Goodwin, G. H. (1990). A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. *Oncogene* 5, 1743-53.

Lomvardas, S., Barnea, G., Pisapia, D. J., Mendelsohn, M., Kirkland, J. and Axel, R. (2006). Interchromosomal interactions and olfactory receptor choice. *Cell* 126, 403-13.

Lyon, M. F. (1961). Gene action in the X-chromosome of the mouse (Mus musculus L.). *Nature* 190, 372-3.

Malnic, B., Godfrey, P. A. and Buck, L. B. (2004). The human olfactory receptor gene family. *Proc Natl Acad Sci U S A* 101, 2584-9.

Malnic, B., Hirono, J., Sato, T. and Buck, L. B. (1999). Combinatorial receptor codes for odors. *Cell* 96, 713-23.

Mank, J. E. and Ellegren, H. (2009). All dosage compensation is local: gene-by-gene regulation of sex-biased expression on the chicken Z chromosome. *Heredity* 102, 312-320.

Masai, H., Matsumoto, S., You, Z., Yoshizawa-Sugata, N. and Oda, M. (2010). Eukaryotic chromosome DNA replication: where, when, and how? *Annu Rev Biochem* 79, 89-130.

McQueen, H. A., Siriaco, G. and Bird, A. P. (1998). Chicken microchromosomes are hyperacetylated, early replicating, and gene rich. *Genome Res* 8, 621-30.

Misteli, T. (2001). Protein dynamics: implications for nuclear architecture and gene expression. *Science* 291, 843-7.

Moore, T. and Haig, D. (1991). Genomic imprinting in mammalian development: a parental tug-of-war. *Trends in Genetics* 7, 45-49.

Mostoslavsky, R., Singh, N., Tenzen, T., Goldmit, M., Gabay, C., Elizur, S., Qi, P., Reubinoff, B. E., Chess, A., Cedar, H. et al. (2001). Asynchronous replication and allelic exclusion in the immune system. *Nature* 414, 221-5.

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

Nishida-Umehara, C., Fujiwara, A., Ogawa, A., Mizuno, S., Abe, S. and Yoshida, M. C. (1999). Differentiation of Z and W Chromosomes Revealed by Replication Banding

and FISH Mapping of Sex-chromosome-linked DNA Markers in the Cassowary (Aves, Ratitae). *Chromosome Research* 7, 635-640.

Ohlsson, R., Renkawitz, R. and Lobanenkov, V. (2001). CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet* 17, 520-7.

Ohno, M., Tenzen, T., Watanabe, Y., Yamagata, T., Kanaya, S. and Ikemura, T. (2000). Non-B DNA structures spatially and sequence-specifically associated with individual centromeres in the human interphase nuleus. In *Chromosomes Today*, vol. 13 (ed. E. Olmo and C. A. Redi), pp. 57-69. Berlin: Birkhauser.

Ohno, S. (1967). Sex Chromosomes and sex-linked genes. Berlin: Springer-Verlag

Ohno, S., Jainchill, J. and Stenius, C. (1963). The Creeping Vole (Microtus Oregoni) as a Gonosomic Mosaic. I. The Oy/Xy Constitution of the Male. *Cytogenetics* 14, 232-9.

Orlando, V., Strutt, H. and Paro, R. (1997). Analysis of chromatin structure by in vivo formaldehyde cross-linking. *Methods* 11, 205-14.

Osborne, C. S., Chakalova, L., Brown, K. E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J. A., Lopes, S., Reik, W. et al. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat Genet* 36, 1065-71.

Paixao, S., Colaluca, I. N., Cubells, M., Peverali, F. A., Destro, A., Giadrossi, S., Giacca, M., Falaschi, A., Riva, S. and Biamonti, G. (2004). Modular structure of the human lamin B2 replicator. *Mol Cell Biol* 24, 2958-67.

Pappas, D. L., Jr., Frisch, R. and Weinreich, M. (2004). The NAD(+)-dependent Sir2p histone deacetylase is a negative regulator of chromosomal DNA replication. *Genes Dev* 18, 769-81.

Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H. C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T. et al. (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* 132, 422-33.

Pasero, P., Bensimon, A. and Schwob, E. (2002). Single-molecule analysis reveals clustering and epigenetic regulation of replication origins at the yeast rDNA locus. *Genes Dev* 16, 2479-84.

Phillips, J. E. and Corces, V. G. (2009). CTCF: master weaver of the genome. *Cell* 137, 1194-211.

Raghuraman, M. K., Winzeler, E. A., Collingwood, D., Hunt, S., Wodicka, L., Conway, A., Lockhart, D. J., Davis, R. W., Brewer, B. J. and Fangman, W. L. (2001). Replication dynamics of the yeast genome. *Science* 294, 115-21.

Renfree, M. B., Papenfuss, A. T., Shaw, G. and Pask, A. J. (2009). Eggs, embryos and the evolution of imprinting: insights from the platypus genome. *Reproduction, Fertility and Development* 21, 935-942.

Rens, W., O'Brien, P. C., Grutzner, F., Clarke, O., Graphodatskaya, D., Tsend-Ayush, E., Trifonov, V. A., Skelton, H., Wallis, M. C., Johnston, S. et al. (2007). The multiple sex chromosomes of platypus and echidna are not completely identical and several share homology with the avian Z. *Genome Biol* 8, R243.

Rothstein, R., Michel, B. and Gangloff, S. (2000). Replication fork pausing and recombination or "gimme a break". *Genes Dev* 14, 1-10.

Rubio, E. D., Reiss, D. J., Welcsh, P. L., Disteche, C. M., Filippova, G. N., Baliga, N. S., Aebersold, R., Ranish, J. A. and Krumm, A. (2008). CTCF physically links cohesin to chromatin. *Proc Natl Acad Sci U S A* 105, 8309-14.

Ryba, T., Hiratani, I., Lu, J., Itoh, M., Kulik, M., Zhang, J., Schulz, T. C., Robins, A. J., Dalton, S. and Gilbert, D. M. (2010). Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Res* 20, 761-70.

Sadoni, N., Langer, S., Fauth, C., Bernardi, G., Cremer, T., Turner, B. M. and Zink, D. (1999). Nuclear organization of mammalian genomes. Polar chromosome territories build up functionally distinct higher order compartments. *J Cell Biol* 146, 1211-26.

Saha, S., Shan, Y., Mesner, L. D. and Hamlin, J. L. (2004). The promoter of the Chinese hamster ovary dihydrofolate reductase gene regulates the activity of the local origin and helps define its boundaries. *Genes Dev* 18, 397-410.

Schaarschmidt, D., Baltin, J., Stehle, I. M., Lipps, H. J. and Knippers, R. (2004). An episomal mammalian replicon: sequence-independent binding of the origin recognition complex. *EMBO J* 23, 191-201.

Schempp, W. and Schmid, M. (1981). Chromosome banding in Amphibia. VI. BrdU-replication patterns in anura and demonstration of XX/XY sex chromosomes in Rana esculenta. *Chromosoma* 83, 697-710.

Schmid, M., Enderle, E., Schindler, D. and Schempp, W. (1989). Chromosome banding and DNA replication patterns in bird karyotypes. *Cytogenet Cell Genet* 52, 139-46.

Schweizer, D., Loidl, J. and Hamilton, B. (1987). Heterochromatin and the phenomenon of chromosome banding. *Results Probl Cell Differ* 14, 235-54.

Selig, S., Okumura, K., Ward, D. C. and Cedar, H. (1992). Delineation of DNA replication time zones by fluorescence in situ hybridization. *EMBO J* 11, 1217-25.

Serizawa, S., Miyamichi, K., Nakatani, H., Suzuki, M., Saito, M., Yoshihara, Y. and Sakano, H. (2003). Negative feedback regulation ensures the one receptor-one olfactory neuron rule in mouse. *Science* 302, 2088-94.

Shevchenko, A. I., Zakharova, I. S., Elisaphenko, E. A., Kolesnikov, N. N., Whitehead, S., Bird, C., Ross, M., Weidman, J. R., Jirtle, R. L., Karamysheva, T. V. et al. (2007). Genes flanking Xist in mouse and human are separated on the X chromosome in American marsupials. *Chromosome Res* 15, 127-36.

Shreeram, S., Sparks, A., Lane, D. P. and Blow, J. J. (2002). Cell type-specific responses of human cells to inhibition of replication licensing. *Oncogene* 21, 6624-32.

Simon, I., Tenzen, T., Reubinoff, B. E., Hillman, D., McCarrey, J. R. and Cedar, H. (1999). Asynchronous replication of imprinted genes is established in the gametes and maintained during development. *Nature* 401, 929-32.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B. and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat Genet* 38, 1348-54.

Singh, N., Ebrahimi, F. A., Gimelbrant, A. A., Ensminger, A. W., Tackett, M. R., Qi, P., Gribnau, J. and Chess, A. (2003). Coordination of the random asynchronous replication of autosomal loci. *Nat Genet* 33, 339-41.

Spiesser, T., Klipp, E. and Barberis, M. (2009). A model for the spatiotemporal organization of DNA replication in <i>Saccharomyces cerevisiae</i&gt. *Molecular Genetics and Genomics* 282, 25-35.

State, M. W., Greally, J. M., Cuker, A., Bowers, P. N., Henegariu, O., Morgan, T. M., Gunel, M., DiLuna, M., King, R. A., Nelson, C. et al. (2003). Epigenetic abnormalities

associated with a chromosome 18(q21-q22) inversion and a Gilles de la Tourette syndrome phenotype. *Proc Natl Acad Sci U S A* 100, 4684-9.

Takagi, N. (1974). Differentiation of X chromosomes in early female mouse embryos. *Exp Cell Res* 86, 127-35.

Takagi, N., Sugawara, O. and Sasaki, M. (1982). Regional and temporal changes in the pattern of X-chromosome replication during the early post-implantation development of the female mouse. *Chromosoma* 85, 275-86.

Takahashi, M. (1987). A model for the spatio-temporal organization of DNA replication in mammalian cells. *J Theor Biol* 129, 91-115.

Tanaka, S. and Diffley, J. F. (2002). Deregulated G1-cyclin expression induces genomic instability by preventing efficient pre-RC formation. *Genes Dev* 16, 2639-49.

Taylor, J. H. (1977). Increase in DNA replication sites in cells held at the beginning of S phase. *Chromosoma* 62, 291-300.

Veyrunes, F., Waters, P. D., Miethke, P., Rens, W., McMillan, D., Alsop, A. E., Grutzner, F., Deakin, J. E., Whittington, C. M., Schatzkamer, K. et al. (2008). Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. *Genome Res* 18, 965-73.

Wakefield, M. J., Keohane, A. M., Turner, B. M. and Graves, J. A. (1997). Histone underacetylation is an ancient component of mammalian X chromosome inactivation. *Proc Natl Acad Sci U S A* 94, 9665-8.

Wang, L., Lin, C. M., Brooks, S., Cimbora, D., Groudine, M. and Aladjem, M. I. (2004). The human beta-globin replication initiation region consists of two modular independent replicators. *Mol Cell Biol* 24, 3373-86.

Watanabe, Y., Abe, T., Ikemura, T. and Maekawa, M. (2009). Relationships between replication timing and GC content of cancer-related genes on human chromosomes 11q and 21q. *Gene* 433, 26-31.

Watanabe, Y., Fujiyama, A., Ichiba, Y., Hattori, M., Yada, T., Sakaki, Y. and Ikemura, T. (2002). Chromosome-wide assessment of replication timing for human chromosomes 11q and 21q: disease-related genes in timing-switch regions. *Hum Mol Genet* 11, 13-21.

Watanabe, Y., Ikemura, T. and Sugimura, H. (2004). Amplicons on human chromosome 11q are located in the early/late-switch regions of replication timing. *Genomics* 84, 796-805.

Watanabe, Y. and Maekawa, M. (2010). Spatiotemporal regulation of DNA replication in the human genome and its association with genomic instability and disease. *Curr Med Chem* 17, 222-33.

Watanabe, Y., Shibata, K., Sugimura, H. and Maekawa, M. (2007). p53-Dependent change in replication timing of the human genome. *Biochemical and Biophysical Research Communications* 364, 289-293.

Weidman, J. R., Murphy, S. K., Nolan, C. M., Dietrich, F. S. and Jirtle, R. L. (2004). Phylogenetic footprint analysis of IGF2 in extant mammals. *Genome Res* 14, 1726-32.

Wilson, N. D., Ross, L. J., Close, J., Mott, R., Crow, T. J. and Volpi, E. V. (2007). Replication profile of PCDH11X and PCDH11Y, a gene pair located in the non-pseudoautosomal homologous region Xq21.3/Yp11.2. *Chromosome Res* 15, 485-98.

Woodward, A. M., Gohler, T., Luciani, M. G., Oehlmann, M., Ge, X., Gartner, A., Jackson, D. A. and Blow, J. J. (2006). Excess Mcm2-7 license dormant origins of replication that can be used under conditions of replicative stress. *J Cell Biol* 173, 673-83.

Wrigley, J. M. and Graves, J. A. (1988). Sex chromosome homology and incomplete, tissue-specific X-inactivation suggest that monotremes represent an intermediate stage of mammalian sex chromosome evolution. *J Hered* 79, 115-8.

Wutz, A. and Jaenisch, R. (2000). A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol Cell* 5, 695-705.

Yusufzai, T. M., Tagami, H., Nakatani, Y. and Felsenfeld, G. (2004). CTCF Tethers an Insulator to Subnuclear Sites, Suggesting Shared Insulator Mechanisms across Species. *Molecular Cell* 13, 291-298.

Zakharova, I. S., Shevchenko, A. I. and Zakian, S. M. (2009). Monoallelic gene expression in mammals. *Chromosoma* 118, 279-90.

Zhao, Z., Tavoosidana, G., Sjolinder, M., Gondor, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K. S., Singh, U. et al. (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat Genet* 38, 1341-7.

Zink, D., Bornfleth, H., Visser, A., Cremer, C. and Cremer, T. (1999). Organization of early and late replicating DNA in human chromosome territories. *Exp Cell Res* 247, 176-88.

CHAPTER 2: Asynchronous replication of loci residing on platypus X- and chicken Z-chromosomes

This chapter consists of one published paper and one publication style manuscript.

Chapter overview

Asynchronous replication is a hallmark of therian female x-inactivation, where the inactive-X replicates late compared to the active-X, but this characteristic is not observed in earlier-diverged amniotes, such as male homogametic birds (ZZ). This chapter consists of a published research paper and a submitted (and revised after peer-reviewed comments) manuscript, which both examine the presence of locus-specific asynchronous replication as an indicator of dosage-compensatory processes on the sex chromosomes in platypus and birds. Platypus dosage compensation has been of interest since the discovery that dosage-compensation of Z-genes in birds occurs in a region-specific manner. Dosage compensation requires two mechanisms in order to achieve compensated expression from the homogametic sex chromosomes; the first mechanism involves the upregulation of expression from the active-X, and the second mechanism involves the epigenetic silencing and chromosome condensation of the silent-X. Furthermore, asynchronous replication is observed for the X-chromosomes in therian females, leading to the hypothesis that the observed dissimilar epigenetic states at the two X-chromosomes also result in this characteristic. The research paper, published in the journal *Reproduction, Fertility and Development*, specifically looks at asynchronous replication and differential condensation of X-homologs in the female platypus, and finds that replication asynchrony occurs in a locus-specific manner on sex-specific regions, whilst differential condensation is only observed at the X3 homologs. The manuscript,

submitted to the journal *Genesis*, looks at locus-specific replication asynchrony on the Z-chromosome in skin cells derived from pre-hatchling, and 6 week-old, homogametic male chickens. The findings show that replication asynchrony does occur at specific loci on the male chicken Z-chromosome, however, the amount of asynchrony declines later in development in these skin-derived cell lines. Furthermore, the Z-linked *Dmrt1* gene, which undergoes a process akin to therian dosage-compensation in female chicken, does not undergo asynchronous replication in male chicken at any tested age, indicating that a replication asynchrony status for loci on the Z-chromosome may only be associated with genes subject to dosage compensation in male chickens.

Together, this research shows that locus-specific replication asynchrony occurs on the sex chromosomes of platypus and chicken, indicating that this characteristic is not specific to therian sex chromosomes, and suggesting that there may also be dosage compensatory mechanisms occurring at the sex chromosomes in these species.

PAPER:

Replication asynchrony and differential condensation of X chromosomes in female platypus (Ornithorhynchus anatinus)

Statement of Authorship

Title of paper	**Replication asynchrony and differential condensation of X chromosomes in female platypus (Ornithorhynchus anatinus)**
Publication status	Published
Publication details	Ho, K.K.K., Deakin, J.E., Wright, M.L., Graves, J.A.M., and Grutzner, F. 2009. Replication asynchrony and differential condensation of X chromosomes in female platypus (Ornithorhynchus anatinus). *Reproduction, Fertility and Development* **21:** 952-963.

Access to electronic copies are available from:

http://www.publish.csiro.au/paper/RD09099

Author Contributions

By signing the Statement of Authorship, each author certifies that their stated contribution to the publication is accurate and that permission is granted for the publication to be included in the candidate's thesis.

Name of Principal Author	Kristen Ho
Contribution to the Paper	Completed all groundwork and the first-round of experiments, including FISH experiments, dot-assay counting (for asynchronous replication), and metaphase chromosome measurements (for chromatin compaction). Deduced necessary statistical formulae for analysis and carried out statistics on primary results. Wrote the manuscript and compiled figures and tables.
Signature	27.09.2014

Name of Co-Author	Janine Deakin
Contribution to the Paper	Intellectual input, aided with data analysis, and helped writing the manuscript.
Signature	24.11.2014

Name of Co-Author (Candidate)	Megan Wright
Contribution to the Paper	Completed repeat FISH experiments, dot-assay counting (for asynchronous replication), and metaphase chromosome measurements (for chromatin compaction) in order to increase the size of the existing datasets (see Table 2 and Figure 3). Also carried out additional FISH experiments and measurements, after a peer-review request, on human metaphase chromosomes (Table 3). Finally, merged all extra dot-assay data, and chromosome measurement (both platypus and human) data, into the final data sets, and carried out a statistical analysis (deduced originally by Kristen) on the final merged datasets (these are the statistics seen in the paper in Table 2 and Table 3).
Signature	19.12.2014

Name of Co-Author	Jenny Graves
Contribution to the Paper	Supervised research, commented on drafts of manuscript.
Signature	26.11.2014

Name of Co-Author	Frank Grützner
Contribution to the Paper	Supervised the asynchronous replication and metaphase chromatin compaction research, and the writing of the manuscript. Was the primary correspondent during the review process. Completed some dot-assay counting for the asynchronous replication results.
Signature	19.12.2014

Ho, K.K.K., Deakin, J.E., Wright, M.L., Graves, J.A.M., and Grutzner, F. (2009).
Replication asynchrony and differential condensation of X chromosomes in female platypus (Ornithorhynchus anatinus).
Reproduction, Fertility and Development, 21(8), 952-963.

SUBMITTED MANUSCRIPT:

Asynchronous replication timing of loci on the Z-chromosome in male chicken

Statement of Authorship

Title of paper	**Asynchronous replication timing of loci on the Z-chromosome in male chicken**
Publication status	Submitted to the journal *Genesis* in November 2014 (this is the revised version for resubmission)
Publication details	Wright, M.L. and Grutzner, F. Asynchronous replication of genes on the Z-chromosome in male chicken. *Unpublished Manuscript.*

Author Contributions

By signing the Statement of Authorship, each author certifies that their stated contribution to the publication is accurate and that permission is granted for the publication to be included in the candidate's thesis.

Name of Principal Author (Candidate)	Megan Wright
Contribution to the Paper	Completed all experiments (Sourcing and growing new BACs, Tissue culture, PCR-testing of new BACs, FISH dot-assay), data-collection (counting and taking microscope images), and statistical analysis. Wrote the manuscript, and compiled all figures and tables.
Signature	19.12.2014

Name of Co-Author	Frank Grützner
Contribution to the Paper	Supervised all research and the manuscript construction/outlining.
Signature	19.12.2014

Manuscript Title: Asynchronous replication timing of loci on the Z-chromosome in male chicken

Manuscript in revision for the journal *Genesis* (Letters) – please note that the following manuscript is the revised version for resubmission.

Author affiliations and contact details:

Megan Lynne Wright[1]* and Frank Grutzner[1]

*all correspondence to this author.

[1]The Robinson Research Institute, School of Molecular & Biomedical Sciences, The University of Adelaide, Australia

Corresponding author contact information:

megan.wright@adelaide.edu.au

Phone: + 61 430 402 921
Fax: + 61 8 8313 4362

The Robinson Research Institute
School of Molecular & Biomedical Sciences
The University of Adelaide
Adelaide
S.A. 5005
AUSTRALIA

Running Head: Asynchronous replication male chicken

Keywords: DNA, asynchronous, replication, chicken

Grant Support: Frank Grützner is an Australian Research Council (ARC) research fellow

Megan Wright's PhD is funded through an Australian Postgraduate Award (bestowed by the Australian Government)

Abstract

A hallmark of monoallelic gene expression is replication asynchrony during mitotic synthase phase. In the case of the therian homogametic female X-chromosome, the observed replication asynchrony reflects a chromosome-wide wholesale shift to late replication of the transcriptionally inactive X-chromosome in somatic cells of the homogametic sex. In birds, males are the homogametic sex (ZZ) and replication banding studies have shown that neither Z-chromosome undergoes an observable shift to late replication timing in male somatic cells. Here, we investigate if asynchronous replication occurs at six specific loci on the Z-chromosome of male chicken, and find that five of them undergo asynchronous replication. Furthermore, we observe changes in the amount of replication asynchrony present in cell lines derived from pre-hatchling and adolescent (6-week old) chickens. These results suggest that gene specific regions on the chicken Z-chromosome replicate asynchronously, however, this replication status may change during development.

Introduction

Replication asynchrony is the process whereby one allele at a locus initiates DNA replication earlier than the other allele during S-phase. In homogametic therian female somatic cells, X-chromosome homologs replicate at different time points during S-phase; the actively transcribed X-chromosome replicates during early S-phase, whilst the epigenetically silenced X-chromosome replicates late in S-phase (Graves, 1967; Latt, 1973; Robinson *et al.*, 1994; Sharman, 1971; Toder *et al.*, 1997). Late replication of the silent X-chromosome in therian mammals (marsupials and placental mammals) is believed to be a reflection of its inactive state. During embryonic development, when X inactivation is established, a loss in cell potency is

accompanied with a gain in the amount of late-replicating DNA. This observation acutely follows the expectation that late-replicating DNA accompanies the transcriptional inactivation of non-required genes (Holmquist, 1987).

In therians, evidence suggests that most X-borne genes asynchronously replicate (Boggs and Chinault, 1994; Dutta *et al.*, 2009; Gribnau *et al.*, 2005). However, there are a few genes on the eutherian X-chromosome that do escape the X-inactivation process and, for the most part, replicate synchronously (Boggs and Chinault, 1994). Interestingly, the replication asynchrony pattern that exists in a cell before the X-inactivation process occurs has no bearing on the X-inactivation choice, suggesting that asynchronous replication plays no part in the choice of which X chromosome undergoes inactivation (Gribnau *et al.*, 2005).

Bird sex chromosomes share extensive homology with monotreme sex chromosomes (Grutzner *et al.*, 2004; Veyrunes *et al.*, 2008). In monotremes, early cytological work and transcriptome analysis suggested that chromosome wide X-inactivation did not exist (Wrigley and Graves, 1988). However, more recent studies have shown that locus-specific asynchronous replication is observed to occur along the X-chromosomes of the platypus (Ho *et al.*, 2009), as well as locus-specific monoallelic expression (Deakin et al. 2008). Similarly, in birds replication banding indicates no difference in replication timing between the Z-chromosomes in males, and transcriptome work has also shown that a chromosome-wide inactivation mechanism does not exist in birds (Julien *et al.*, 2012; Schmid *et al.*, 1989). However, a skewed gene dosage ratio does exist for many genes on the chicken Z, with certain regions showing Male:Female expression ratios below 1.2-1.5, as opposed to a completely non-dosage compensated ratio of 2 (Arnold *et al.*, 2008; Julien *et al.*,

2012; Mank and Ellegren, 2009; Melamed and Arnold, 2007). A region of particular interest, on the short arm of the chicken Z-chromosome, is the male hypermethylated (MHM) region which expresses a non-coding RNA that colocalises with the *doublesex and Mab-3-related transcription factor 1* (*DMRT1*) gene in the nuclei of female chickens (Teranishi *et al.*, 2001). *DMRT1* expression levels in male chicken embryos are higher than those observed in female embryos, with expression of this gene only observed in the testis of post-hatchlings, suggesting a potential role of *DMRT1* in the determination of sex in galliformes that is reliant on dosage (Raymond *et al.*, 1999; Teranishi *et al.*, 2001). The existence of dosage-compensatory mechanisms within the MHM region on the Z-chromosome of galliformes, akin to those observed in therians, suggests that other epigenetic mechanisms surrounding inactivation, such as asynchronous replication, may also be present.

To date, no study has looked at single-locus replication on the Z-chromosome of homogametic male birds. In this study we investigate the asynchronous replication status of six individual Z-borne loci, including the *DMRT1* region, in both adolescent and pre-hatchling skin-derived male chicken fibroblast cell lines.

Results and discussion

We utilised the well established FISH dot-assay approach to measure asynchronous replication in skin-derived primary cell cultures from pre-hatchling and 6-week old male chickens (**Figure 1**, and see **Table 1** for gene and BAC information). The results from the assay demonstrated that five of the Z-chromosome loci (located at positions 19.6Mb, 23.5Mb, 58.5MB, 60.5Mb and 64.2Mb) in the pre-hatchling cell line underwent asynchronous replication in a locus-specific manner, whilst the locus

located at 26.7 Mb did not (**Figure 2**). These five Z-borne loci showed between 38-46% SD signal, which is within the previously observed 25-50% SD signal range expected for asynchronously replicating loci in eutherians (Baumer *et al.*, 2004; Dutta *et al.*, 2009; Wilson *et al.*, 2007). This was significantly higher than the autosomal control locus (Chr.15q at 3.3Mb) which showed around 28.5% SD signal, and thus is not asynchronously replicating in accordance with previous studies in eutherians which classify below 25-30% SD signal as synchronous replication (Baumer *et al.*, 2004; Dutta *et al.*, 2009; Wilson *et al.*, 2007). It has been noted that the overlap between the synchronously and asynchronously replicating loci ranges (between 25-30%) from previous studies most likely represents differences between how asynchrony has been measured between research groups. However, it is generally accepted that below 30% SD signal is concurrent with synchronous replication (Wilson et al., 2007).

Interestingly, in the 6-week old chicken cell line, we found that none of the six Z-loci had significantly higher SD values than the autosomal control locus, indicating that all of the Z-loci tested did not undergo asynchronous replication in adolescent chickens. This could reflect a developmentally driven change in the replication asynchrony pattern on the male chicken Z-chromosome.

We observed around 33% SD signal at the Zp 26.7Mb region in the nuclei of pre-hatchling derived cells, and only 19% at the same locus in 6-week old chicks, with neither SD value being significantly different to the autosomal control, indicating synchronous replication (**Figure 2**). The chicken Zp 26.7Mb region includes two genes in the doublesex gene cluster, *DMRT1 and DMRT3,* and the *KANK1* gene, which has roles in actin polymerization (**Table 1**). Notably, the *DMRT1* gene is

involved in the sexual development of chickens, with the gene undergoing RNA-mediated inactivation in females and no inactivation in males (Raymond *et al.*, 1999). The fact that this region is not asynchronously replicated, like the other Z-regions measured, could indicate that it is escaping dosage compensatory processes, a situation that would be akin to the genes which escape X-inactivation in therian mammals (Boggs and Chinault, 1994). Notably, female platypus X-chromosome genes are known to dosage-compensate using a combination of mono and biallelic expression (Deakin *et al.*, 2008), raising the possibility that some Z-genes in male birds may also show this pattern. If this were the case in chicken, the highly-expressed *DMRT1* gene in male birds could be undergoing synchronous replication due to a greater amount of biallelic expression, whilst other male Z-genes, subject to greater dosage-compensatory effects, asynchronously replicate.

We observed a reduction in the amount of asynchronous replication across development when we looked in cultured cells derived from pre-hatchling and adolescent chicks (**Table 2**). Four loci were observed to undergo significantly less asynchronous replication in the adolescent compared to the pre-hatchling chick cell lines (these regions were located at 19.6Mb, 23.5Mb, 60.5Mb and 64.2Mb) (**Table 2**). For one Z-borne locus, located at 23.5Mb, the amount of asynchronous replication observed in the cells cultured from the pre-hatchling chick cell line was not significantly different from the amounts observed in cells cultured from the adolescent chick.

Interestingly, three of the four Z-borne loci that showed a significant reduction in asynchronous replication across development were located on the long arm, which has previously been shown to have fewer dosage compensated genes compared to

the short arm in chicken embryos (Melamed and Arnold, 2007). Future work will need to address whether the general reduction in replication timing observed for the chicken Z-chromosome could reflect changes to this chromosome during development.

Conclusion

There is now considerable evidence to show the lack of a chromosome-wide dosage-compensation-by-inactivation mechanism on chicken sex chromosomes. However, whether replication asynchrony, one of the hallmarks of sex chromosome inactivation, exists on the chicken Z-chromosome, has not been addressed. We looked at six loci scattered across the chicken Z-chromosome in two primary cell lines derived from pre-hatchling and adolescent male chickens. Surprisingly we find evidence of asynchronous replication at five out of six Z-borne loci tested in pre-hatchling chicken cells. We also detected a reduction in the amount of replication asynchrony observed at all Z-borne loci in the adolescent chicken cells, possibly reflecting a chromosome-wide reduction in the amount of asynchronous replication for this chromosome during chicken development. We speculate that genes like *DMRT1*, which are highly expressed on the Z-chromosome of male chickens, may undergo no asynchronous replication due to region-specific alterations that cause differences in expression. These findings are in line with previous studies showing that regional specific changes in gene activity occur to achieve dosage compensation in birds, and raise the possibility of locus-specific changes occurring on the chicken Z-chromosome during development.

Methods

Cell lines and interphase preparations

Skin-derived primary chicken fibroblasts, derived from pre-hatchling (2-3 days pre-hatch) and 6-week old adolescent chicks, were cultured in Advanced DMEM medium (12491-023, Life Technologies), and supplemented with 10% Fetal Calf Serum and 1% L-glutamine. Interphase nuclei fixed preparations were prepared using a standard protocol (Grutzner *et al.*, 2004); in brief, cultured cells were treated with hypotonic solution (0.075 M KCl) and incubated at 37^{o}C for 20 minutes, before being fixed in methanol/acetic acid (3:1), and dropped onto freshly cleaned slides (slides cleaned with 1% HCl/Methanol), and dried under humid conditions (around 55-60% humidity).

BAC clones, culturing, DNA extraction, and PCR-testing

The set of six chicken Z-borne loci tested, their corresponding Bacterial Artificial Chromosomes (BACs), and specific gene information, are all outlined in **Table 1**. The Z-linked BAC clones were selected due to their distribution on the Z chromosome, with both p and q arms represented. The control BAC, present on the q arm of chicken chromosome 15, was used as an autosomal control, and to our knowledge, contains no genes implicated to asynchronously replicate in chicken or any mammalian species. BAC clones were purchased from Children's Hospital Oakland Research Institute (CHORI, Oakland, CA, USA), and grown following the company's instructions. BAC clone DNA was extracted using The Wizard Minipreps DNA Purification System (Promega, Madison, WI, USA). Chicken BAC clone identification and mapping are described in previous literature for all BACs except CH261-125B9 (contains *DMRT1*) and CH261-3H21 (contains *RIMBP2*) (Tsend-Ayush *et al.*, 2009). Presence of the *DMRT1* and *RIMBP2* genes on the corresponding BACs was tested

via PCR, using the following primers represented in 5'-3' orientation; *DMRT1* primers (Tm oC 54) F: GAGCCAGTTGTCAAGAAGAG, R: ACTGTTCTACCTATCCTCTGC, and *RIMBP2* primers (Tm 56^{o}C) F: ATGCCGGACACAAAGTGAAC, R: TGCCTAAAGAAGTGGCTATTCC. Cycling parameters were: 95^{o}C for 3 min, then 35 cycles of 95^{o}C for 30 s, aforementioned primer annealing temperature for 30 s; 72^{o}C for 2 min; then a final cycle of 72^{o}C for 7 min, before being stored at 4^{o}C indefinitely.

Fluorescent labelling of BAC DNA

For a 50 µl fluorescent DNA labelling reaction: 5 µl BAC DNA (1µg total), 10 µl of 1 µg/µl random 9-mer primer (Geneworks), and 22 µl of water were incubated for 15 minutes at 95^{o}C, put on ice for 5 minutes, before adding 9.2 µl of 5× dNTP supplemented NEB buffer 2 (supplemented with a 200 µM dNTP solution containing: 100 µM dACG-TPs and 40 µM dTTPs), 0.7 µl SpectrumGreen/Orange fluorescent dUTPs (Abbott Molecular), and 5 units (1 µl) Exo-Klenow polymerase (M0212, New England Biolabs). The reaction was then incubated at 37^{o}C overnight.

FISH

25 µl of labelling mix, salmon sperm (50 µg), and 10 µg sonicated genomic chicken DNA (300-600 bp sized fragments) were precipitated in ethanol for 20 minutes at −80^{o}C, centrifuged at 14,000 rpm for 30 min at 4^{o}C, and the pellet dried for 8 min at 37^{o}C, before being dissolved in 5µl of deionised formamide (F9037, Sigma) for 15 min at 37^{o}C, and 5 µl of hybridization mix (10% dextran sulphate/2×SSC) at 37^{o}C for 15 min. The probe was then denatured at 80^{o}C for 10 min, and pre-annealed at 37^{o}C for 40 min.

Freshly-dropped slides were washed in 2xSSC for 5 min, then treated with 100 µg/ml RNase/2xSSC at 37^{o}C for 30 min, washed three times in 2xSSC for 5 min and treated with 0.01% pepsin/10 mM HCl at 37^{o}C for 10 min, washed twice in 1xPBS for 5 min, and equilibrated in 50 mM MgCl2/1 x PBS, before fixation in 1.3% Formaldehyde/1xPBS/50 mM MgCl2. After ethanol dehydration slides were denatured for 2.5 min in 70% (v/v) formamide/2×SSC at 70^{o}C, then dehydrated in another ethanol air-dried and hybridised with 10 µl labelled BAC/formamide/hybridization mix over night at at 37^{o}C. The slides were washed 3 x 5 min in 50% Formamide/2xSSC at 42^{o}C, in 2xSSC at 42^{o}C for 5 min, 0.1xSSC at 60^{o}C for 5 min, and again in 2xSSC at 42^{o}C for 5 min. Slides were then counterstained with 1 µg/ml DAPI/2×SSC for 1 min, rinsed twice in MQ water and mounted with Vectashield (Vector Laboratories).

Microscopy and measuring replication asynchrony

Slides were analysed on a Zeiss AxioImagerZ.1 epifluorescence microscope equipped with a CCD camera, and images were acquired using Zeiss Axiovision software. To measure asynchronous replication we used the well established cytogenetic dot-assay approach, which relies on fluorescently-labelled BAC probe hybridisation to indicate the allele-copy number of a specific locus in a nucleus (see **Figure 1**). Specifically, the amount of fluorescent signals observed indicate the replication status of a specific locus, with a single-single (SS) dot signal representing an unreplicated state, a double-double (DD) dot signal representing a fully replicated state, and a single-double (SD) dot signal representing a locus currently undergoing replication. Two technical replicates of 100 nuclei were counted for each BAC using the dot-assay approach.

Acknowledgements

The authors would like to acknowledge Diana Demiyah Mohd Hamdan for her help with setting up the primary fibroblast cultures.

Funding

Frank Grützner is an Australian Research Council (ARC) research fellow. Megan Wright's PhD is funded through an Australian Postgraduate Award (bestowed by the Australian Government).

Conflict of Interest

The authors declare that they have no conflict of interest.

References

Arnold AP, Itoh Y, Melamed E. 2008. A bird's-eye view of sex chromosome dosage compensation. Annu Rev Genomics Hum Genet 9: 109-127.

Baumer A, Riegel M, Schinzel A. 2004. Non-random asynchronous replication at 22q11.2 favours unequal meiotic crossovers leading to the human 22q11.2 deletion. Journal of Medical Genetics 41: 413-420.

Boggs BA, Chinault AC. 1994. Analysis of replication timing properties of human X-chromosomal loci by fluorescence in situ hybridization. Proc Natl Acad Sci U S A 91: 6083-6087.

Deakin JE, Hore TA, Koina E, Marshall Graves JA. 2008. The status of dosage compensation in the multiple X chromosomes of the platypus. PLoS Genet 4: e1000140.

Dutta D, Ensminger AW, Zucker JP, Chess A. 2009. Asynchronous replication and autosome-pair non-equivalence in human embryonic stem cells. PLoS One 4: e4970.

Graves JA. 1967. DNA synthesis in chromosomes of cultured leucocytes from two marsupial species. Exp Cell Res 46: 37-57.

Gribnau J, Luikenhuis S, Hochedlinger K, Monkhorst K, Jaenisch R. 2005. X chromosome choice occurs independently of asynchronous replication timing. J Cell Biol 168: 365-373.

Grutzner F, Rens W, Tsend-Ayush E, El-Mogharbel N, O'Brien PC, Jones RC, Ferguson-Smith MA, Marshall Graves JA. 2004. In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes. Nature 432: 913-917.

Ho KKK, Deakin JE, Wright ML, Graves JAM, Grutzner F. 2009. Replication asynchrony and differential condensation of X chromosomes in female platypus (Ornithorhynchus anatinus). Reproduction, Fertility and Development 21: 952-963.

Holmquist GP. 1987. Role of replication time in the control of tissue-specific gene expression. Am J Hum Genet 40: 151-173.

Julien P, Brawand D, Soumillon M, Necsulea A, Liechti A, Schutz F, Daish T, Grutzner F, Kaessmann H. 2012. Mechanisms and evolutionary patterns of mammalian and avian dosage compensation. PLoS Biol 10: e1001328.

Latt SA. 1973. Microfluorometric detection of deoxyribonucleic acid replication in human metaphase chromosomes. Proc Natl Acad Sci U S A 70: 3395-3399.

Mank JE, Ellegren H. 2009. All dosage compensation is local: gene-by-gene regulation of sex-biased expression on the chicken Z chromosome. Heredity 102: 312-320.

Melamed E, Arnold AP. 2007. Regional differences in dosage compensation on the chicken Z chromosome. Genome Biol 8: R202.

Raymond CS, Kettlewell JR, Hirsch B, Bardwell VJ, Zarkower D. 1999. Expression of Dmrt1 in the genital ridge of mouse and chicken embryos suggests a role in vertebrate sexual development. Dev Biol 215: 208-220.

Robinson ES, Samollow PB, VandeBerg JL, Johnston PG. 1994. X-chromosome replication patterns in adult, newborn and prenatal opossums. Reprod Fertil Dev 6: 533-540.

Schmid M, Enderle E, Schindler D, Schempp W. 1989. Chromosome banding and DNA replication patterns in bird karyotypes. Cytogenet Cell Genet 52: 139-146.

Sharman GB. 1971. Late DNA replication in the paternally derived X chromosome of female kangaroos. Nature 230: 231-232.

Teranishi M, Shimada Y, Hori T, Nakabayashi O, Kikuchi T, Macleod T, Pym R, Sheldon B, Solovei I, Macgregor H, Mizuno S. 2001. Transcripts of the MHM region on the chicken Z chromosome accumulate as non-coding RNA in the nucleus of female cells adjacent to the DMRT1 locus. Chromosome Res 9: 147-165.

Toder R, Wienberg J, Voullaire L, O'Brien PC, Maccarone P, Graves JA. 1997. Shared DNA sequences between the X and Y chromosomes in the tammar wallaby - evidence for independent additions to eutherian and marsupial sex chromosomes. Chromosoma 106: 94-98.

Tsend-Ayush E, Dodge N, Mohr J, Casey A, Himmelbauer H, Kremitzki CL, Schatzkamer K, Graves T, Warren WC, Grutzner F. 2009. Higher-order genome organization in platypus and chicken sperm and repositioning of sex chromosomes during mammalian evolution. Chromosoma 118: 53-69.

Veyrunes F, Waters PD, Miethke P, Rens W, McMillan D, Alsop AE, Grutzner F, Deakin JE, Whittington CM, Schatzkamer K, Kremitzki CL, Graves T, Ferguson-Smith MA, Warren W, Marshall Graves JA. 2008. Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. Genome Res 18: 965-973.

Wilson ND, Ross LJ, Close J, Mott R, Crow TJ, Volpi EV. 2007. Replication profile of PCDH11X and PCDH11Y, a gene pair located in the non-pseudoautosomal homologous region Xq21.3/Yp11.2. Chromosome Res 15: 485-498.

Wrigley JM, Graves JA. 1988. Karyotypic conservation in the mammalian order monotremata (subclass Prototheria). Chromosoma 96: 231-247.

Figure Legends

Figure 1: Replication states of chicken loci as measured by FISH

A single-single dot pattern, or ‘SS’ signal, indicates that the locus has not yet replicated, whilst a double-double dot pattern, or ‘DD’ signal, indicates that the locus has fully replicated. A single-double dot pattern, or ‘SD’ signal, indicates that one allele has replicated (the double-dot), whilst the other allele is lagging and has not replicated yet (the single-dot).

Figure 2: Replication of chicken Z and autosomal genes in fibroblasts derived from adolescent and pre-hatchling chicken cells

The average SS, SD and DD counts, measured in embryonic and 6-week-old male chicken skin fibroblasts, of Z-chromosome loci and an autosomal control locus. The SD counts that are significantly higher in a t-test ($p<0.05$) than the autosomal control locus are indicated by asterisk in each group.

Tables

Table 1: BAC clones used in this study

Position on chromosome (Mb)	BAC covering region	Genes on BACs	Role of protein	Ensembl gene ID
Chr. Zp, 19.6	CH261-78C3	*SREK1IP1*	Possible splicing regulator involved in the control of cellular survival	ENSGALG00000014745
		CWC27	Spliceosome-Associated Protein Homolog (accelerates the folding of proteins)	ENSGALG00000014747
		ADAMTS6	Disintegrin-like and metalloprotease domain (reprolysin type) with thrombospondin type 1 motif 6, expressed in the periimplantation period in mouse embryo and adult tissues,extracellular matrix protein	ENSGALG00000014751
Chr. Zp, 23.5	CH261-89F20	*COL4A3BP*	ceramide transfer to endoplasmic reticulum	ENSGALG00000014952
		ANKRD31	ankyrin repeat domain (protein-protein interaction)	ENSGALG00000014945
		GCNT4	acetylglucosaminyltransferase (biosynthesis of mucin type O-glycans)	ENSGALG00000014944
Chr. Zp, 26.7	CH261-125B9	*KANK1*	Kank family of proteins, containing multiple ankyrin repeat domains. This family member functions in cytoskeleton formation by regulating actin polymerization	ENSGALG00000010158
		DMRT1	Transcription factor that is a key sex-determining factor in chickens (contains zinc finger-like DNA-binding region)	ENSGALG00000010160
		DMRT3	Transcription factor that is a key regulator of male development in flies and nematode (contains zinc finger-like DNA-binding region)	ENSGALG00000010161
Chr. Zq, 58.5		*GPR98*	The protein binds calcium and is expressed in the central nervous system	ENSGALG00000014657
Chr. Zq, 60.5		-	-	-
Chr. Zq, 64.2		*PALM2*	Implicated in plasma membrane dynamics	ENSGALG00000015655
Chr. 15q, 3.3		*RIMBP2*	Plays a role in the synaptic transmission as bifunctional linker that interacts simultaneously with RIMS1, RIMS2, CACNA1D and CACNA1B (By similarity)	ENSGALG00000002579

Table 2: %SD data for the genes tested in fibroblasts derived from pre-hatchling and adolescent chicken cells

Position on chromosome (Mb)	% SD	
	Pre-hatchling	Adolescent (6 weeks old)
Chr. Zp, 19.6	41.5 ± 2.1	34.5 ± 0.7*
Chr. Zp, 23.5	43.5 ± 0.7	35.5 ± 3.5
Chr. Zp, 26.7	33 ± 1.4	19 ± 5.7
Chr. Zq, 58.5	40.5 ± 2.1	25.5 ± 2.1*
Chr. Zq, 60.5	45.5 ± 0.7	21 ± 0*
Chr. Zq, 64.2	43 ± 1.4	32 ± 1.4*
Chr. 15q, 3.3	28.5 ± 2.1	27 ± 4.2

* A significant difference exists between the SD data of the embryo and 6 week old chick

Illustrations

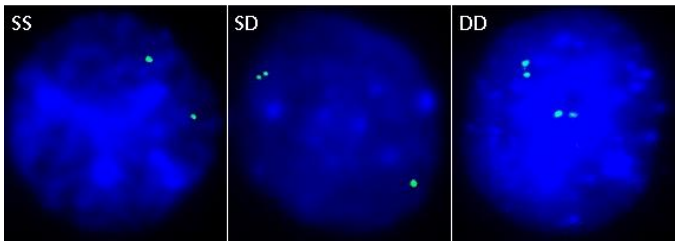


Figure 1: Replication states of chicken loci as measured by FISH

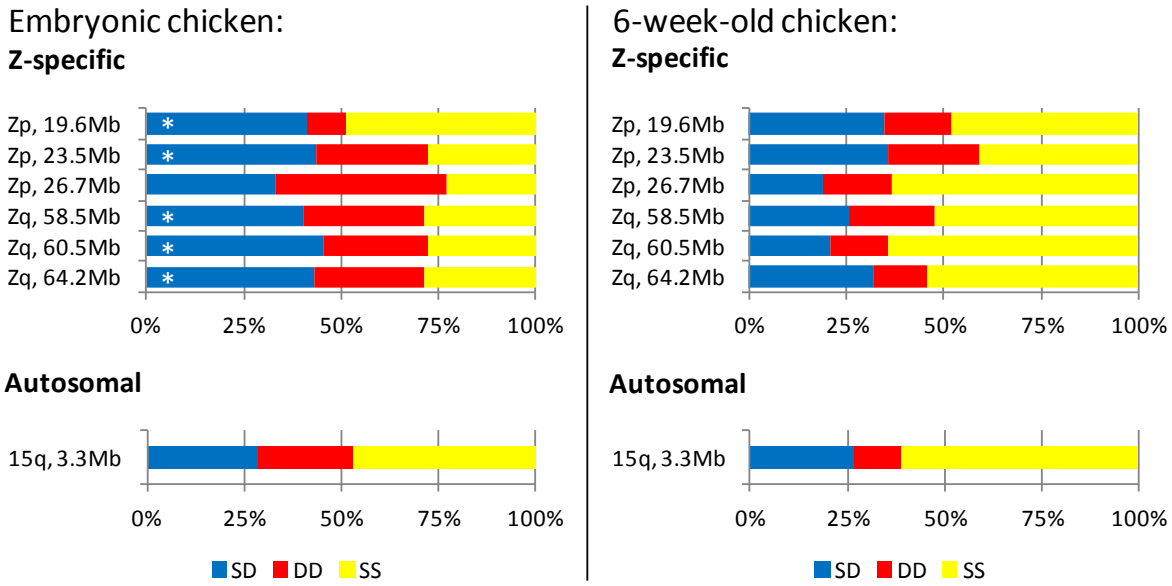


Figure 2: Replication of Chicken Z and autosomal genes in fibroblasts derived from adolescent and pre-hatchling chicken cells

CHAPTER 3: Conservation of replication timing in amniotes

This chapter consists of two publication style manuscripts (unpublished).

Chapter overview

This chapter consists of two manuscripts (unpublished), which encompass research on locus-specific asynchronous replication of imprinted orthologs in platypus, and a BrdU staining protocol combined with directly-labelled FISH in different species.

The first manuscript, formatted for submission to the journal *Chromosoma*, outlines results showing that asynchronous replication does occur at imprinted orthologs in the platypus, and that in similar cell types, different species will show similar levels of asynchronous replication for imprinted loci and their orthologs. Furthermore, using a consecutive FISH on the same set of nuclei, it was shown for the first time that these genes enter S-phase in relatively the same order in amniote species. This conserved temporal replication programme may indicate allele specific differences that may affect expression in different species.

The second manuscript, formatted for submission to the journal *Molecular Cytogenetics* (Methodologies), is a technical paper outlining a new way of performing cytogenetic BrdU staining coupled with directly-labelled DNA FISH in interphase nuclei. The experimental procedure is an improvement on previous methods that require harsh acidic denaturation treatments and secondary-antibodies for FISH signal detection, and allows BrdU incorporation to be measured when analysing DNA replication at the gene specific level.

.

MANUSCRIPT:

Conservation of replication asynchrony and S-phase entry predates evolution of genomic imprinting in mammals

Statement of Authorship

Title of paper	**Conservation of replication asynchrony and S-phase entry predates evolution of genomic imprinting in mammals**
Publication status	Unpublished
Publication details	Wright, M.L., Zechner, U., Haaf, T. and Grutzner, F. Conservation of replication asynchrony and S-phase entry predates evolution of genomic imprinting in mammals. *Unpublished manuscript.* For submission in 2015 to *Chromosoma*

Author Contributions

By signing the Statement of Authorship, each author certifies that their stated contribution to the publication is accurate and that permission is granted for the publication to be included in the candidate's thesis.

Name of Principal Author (Candidate)	Megan Wright
Contribution to the Paper	Performed most of the experiments (except for platypus imprinted ortholog BAC identification and isolation) including sourcing and growing of BACs, tissue culture, PCR-testing of BACs, FISH dot-assay and consecutive FISH, data-collection (counting and taking microscope images), and statistical analysis. Wrote the manuscript, and compiled all figures and tables.
Signature	19.12.2014

Name of Co-Author	Ulrich Zechner
Contribution to the Paper	Completed PCR and BAC screening for imprinted orthologs in platypus, aided in manuscript construction/outlining.
Signature	24.11.2014

Name of Co-Author	Thomas Haaf
Contribution to the Paper	Supervised all research and aided in the manuscript construction/outlining.
Signature	10.12.2014

Name of Co-Author	Frank Grützner
Contribution to the Paper	Supervised all research and aided in the manuscript construction/outlining.
Signature	19.12.2014

Manuscript Title: Conservation of replication asynchrony and S-phase entry predates evolution of genomic imprinting in mammals

Manuscript prepared for submission to the journal *Chromosoma* (Research article)

Authors, affiliations and contact details

Megan Lynne Wright[1], Ulrich Zechner[2], Thomas Haaf[3], and Frank Grutzner[1]*

*all correspondence to this author.

[1]The Robinson Research Institute, School of Molecular & Biomedical Sciences, The University of Adelaide, Australia
[2]Institute of Human Genetics, University Medical Centre of the Johannes Gutenberg University Mainz, Mainz, Germany.
[3]Institute of Human Genetics, University of Würzburg, Germany

Corresponding author contact information:

frank.grutzner@adelaide.edu.au

Phone: + 61 8 8331 4812
Fax: + 61 8 8313 4362

The Robinson Research Institute
School of Molecular & Biomedical Sciences
The University of Adelaide
Adelaide
S.A. 5005
AUSTRALIA

Running Title: Platypus asynchronous replication

Keywords: asynchronous replication, replication timing, monotreme, amniote

Abstract

Replication timing during S-phase is a highly regulated process that appears to be conserved within clades of eukaryotes; the conservation of temporal DNA replication timing programs is observed among species of budding yeast, as well as among orthologous genomic regions in human and mouse. Asynchronous replication occurs when one allele initiates replication during S-phase earlier than the other allele, and is a hallmark of all monoallelically expressed genes, including imprinted genes. The asynchronous replication of imprinted genes is observed in therian mammals, however, it has never been measured in orthologous loci of egg-laying monotreme mammals, which represent the most basal mammalian lineage and lack genomic imprinting. In order to investigate the evolutionary trajectory of replication asynchrony and timing we analysed four genes, *Igf2*, *Igf2r*, *Mest* and *Wt1*, that are imprinted in eutherian mammals, and four genes, *Wsb1*, *Sox1*, *RIMBP2*, and *B3galt1*, that do not undergo therian imprinting in chicken, platypus and mouse. For genes imprinted in eutherian mammals we find that replication asynchrony and S-phase entry is conserved independently of the evolution of genomic imprinting. We also report, for the first time, that the order in which these genes commence S-phase is conserved in amniotes. These results show the evolutionary conservation of replication asynchrony and order of S-phase entry, indicating that a preserved replication program exists among amniotes that preceded the evolution of genomic imprinting in therian mammals.

Introduction

Replication timing refers to the temporal program governing DNA replication, and corresponds to the initiation of replication during S-phase. The replication of a

eukaryotic genome is completed during S-phase, and initiates at thousands of discrete sites called origins of replication (ORs). Work in yeast, human, and mouse indicates that eukaryotic temporal programs exist, but are more likely affected by cis-regulatory processes and not necessarily intrinsic to a region, thus leaving room for stochastic changes and variation across different cell types (Audit et al. 2009; Farkash-Amar et al. 2008; Hansen et al. 2010; Heun et al. 2001; Hiratani et al. 2008; reviewed in Renard-Guillet et al. 2014). Ubiquitously expressed “housekeeping” genes (where transcription is required in all cell types for normal growth) are reported to replicate during early S-phase, whilst tissue-specific genes will only undergo early S-phase replication in a cell when there is a potential requirement of their expression. Thus, a switch in replication timing of a gene can indicate a change in its transcriptional status, with silenced expression inducing late S-phase replication, and activated expression inducing early S-phase replication (Goldman et al. 1984; Hatton et al. 1988; Taljanidisz et al. 1989). The conservation of replication timing among different budding yeast species and the maintenance of early or late replication in orthologous regions, among similar cell types in human and mouse, indicate the importance of replication timing for eukaryotic genome function (Muller and Nieduszynski 2012; Ryba et al. 2010).

In contrast to replication timing, the asynchronous replication status of a gene does not vary in different cell types within an organism, and instead is upheld in all cell types, even if a particular cell is not expressing the asynchronously replicating gene (Chess et al. 1994). Thus, while asynchronous replication is a property that is intrinsically tied to replication timing, they are not one and the same. When an allele initiates S-phase replication earlier than its homologous allele, it is known as asynchronous replication, a process characteristic of all monoallelically expressed

genes. In therian mammals, regions displaying genomic imprinting, X chromosome inactivation, and random autosomal monoallelically expression, all feature asynchronous replication (Bickmore and Carothers 1995; reviewed in Goldmit and Bergman 2004; Gribnau et al. 2003; reviewed in Wright and Grutzner 2011; reviewed in Zakharova et al. 2009).

Genomic imprinting is a specific form of monoallelic expression observed in therian mammals, and entails monoallelic expression in a parent-of-origin specific manner (reviewed in Savova et al. 2013). Generally, imprinting involves DNA methylation at only one allele of a gene (i.e. the copy from just one parent is methylated), and the methylated (imprinted allele) is transcriptionally silenced (Delcuve et al. 2009). Whilst the asynchronous replication of imprinted genes is also determined in a parent-of-origin specific manner during gametogenesis (Simon et al. 1999), the differential methylation pattern that guides the parentally-determined expression pattern at imprinted loci is not required to set up asynchronous replication at imprinted loci (Gribnau et al. 2003; Simon et al. 1999). Instead, DNA-binding proteins, such as the CTCF protein, are implicated to control replication timing at these genes (Bergstrom et al. 2007). The expressed allele of an imprinted locus is generally understood to replicate earlier during S-phase than the silenced allele, which generally replicates later (Greally et al. 1998; Gribnau et al. 2003; Zakharova et al. 2009), although there is evidence at some loci, such as *Igf2* that this may not always be the case (Kitsberg et al. 1993). Furthermore, imprinted genes are asynchronously replicated in every cell of an organism, even though many are not ubiquitously expressed across all tissue types, with the majority undergoing only tissue and developmental stage specific expression (Gimelbrant et al. 2005).

Of the few hundred predicted imprinted genes (DeVeale et al. 2012; Nikaido et al. 2003), many exist in imprinted clusters, where regions of maternally and paternally expressed genes are interspersed with non-imprinted genes, and share coordinated transcriptional control (Goncalves et al. 2012). This form of coordinated control is often carried out by a single differentially methylated imprinting control region (ICR), with many imprinted genes within the cluster requiring the ICR as an essential transcriptional regulatory element. The clustered organisation of orthologs of genes that are imprinted in eutherian mammals, which we will name eutherian imprinted gene orthologs (EIOs), is also conserved in the chicken and platypus where there is no evidence of genomic imprinting (Dünzinger et al. 2005; Pask et al. 2009). However, asynchronous replication is also maintained at these EIOs in chicken (Dünzinger et al. 2005).

Here, we investigate the conservation of replication asynchrony, and also, for the first time, determine S-phase replication timing in amniotes using a *Fluorescent in situ hybridisation* (FISH) approach to study replication at four eutherian imprinted genes and their orthologs, namely *Igf2*, *Igf2r*, *Mest* and *Wt1*. Comparisons across species show that replication asynchrony and replication timing profiles are conserved in amniote species independent of the evolution of genomic imprinting.

Results

Asynchronous replication of imprinted mouse genes and chicken EIOs

Previous studies have shown that eutherian imprinted genes and chicken EIOs replicate asynchronously (Baumer et al. 2004; Bentley et al. 2003; Dünzinger et al. 2005; Kitsberg et al. 1993; Wilson et al. 2007). We sought to determine the conservation of asynchronous replication at platypus EIOs, and also to analyse how

the frequency of replication asynchrony has changed during amniote evolution. Utilising the established FISH dot-assay approach to observe replication asynchrony of a single locus in individual nuclei (see **Table 1** for locus information and **Supplementary figure 1** for FISH dot-assay), we measured the single/double-dot signal (SD signal, 3 dots) which corresponds to the amount of asynchronous replication in a cell type. In eutherians, between 25-50% SD signal is representative of asynchronous replication, and below 25-30% SD signal is representative of synchronous replication (Baumer et al. 2004; Wilson et al. 2007). Concurrent with previous literature, we identified asynchronous replication at all the mouse imprinted loci measured, namely *Igf2*, *Igf2r*, *Mest*, and *Wt1*, with all four loci showing between 35-37% SD signal (**Table 2**, and see **Supplementary Table 1** for comparison of our results to results in previous studies in eutherian mammals) (Bentley et al. 2003; Bickmore and Carothers 1995; Kitsberg et al. 1993). We performed two-tailed T-tests, and found that all of the four mouse imprinted genes underwent significantly more frequently asynchronous replication compared to the control locus, *Wsb1* (**Table 2**). Furthermore, there was no significant difference observed between the amount of asynchronous replication between *Wsb1* and the other non-imprinted gene *B3galt1* (p-value = 0.64).

The chicken EIOs, *IGF2*, *IGF2R*, *MEST*, and *WT1,* all replicated asynchronously, showing between 36-46% SD signal (**Table 2** and **Supplementary Table 2,** for comparison of our results to results in a previous study in chicken) (Dünzinger et al. 2005)**.** All four chicken EIOs underwent significantly more frequently asynchronous replication compared to *WSB1* (T-test all p-values < 0.02). The two non-imprinted genes, *WSB1* and *RIMBP2*, showed no significant difference in their asynchronous replication status (**Table 2**).

Asynchronous replication occurs at three EIOs in the platypus

Conservation of asynchronous replication was observed for three of the four platypus EIOs, with the platypus *Igf2*, *Igf2r* and *Mest* all replicating asynchronously with 36-39% SD signal (**Table 2**). Platypus *Igf2*, *Igf2r* and *Mest* all showed significantly higher amounts of asynchronous replication compared to the control (**Table 2**, T-test all p-values < 0.05). Whilst the platypus EIO gene *Wt1* had 33% SD signal, this result was not significantly different to the amount of SD measured for the platypus *Wsb1* control gene (p-value = 0.11). The platypus *Wt1* gene therefore was the only example of an EOI that replicated synchronously. *Sox1* in platypus, did not have a significantly different SD value to the platypus *Wsb1* control gene, with both genes replicating synchronously in the platypus.

The level of replication asynchrony measured at mouse imprinted genes is similar to that observed at corresponding platypus and chicken EIOs

Next we measured if there was a significant difference in the level at which asynchronous replication was observed, given that some genes are imprinted in therian mammals but not in platypus and chicken. A pair wise comparison, using two-tailed T-tests, showed that mouse imprinted genes *Igf2*, *Igf2r* and *Mest*, and the *Wt1* gene (which is imprinted in humans but has not yet had its imprinted status tested in mouse), and their EIOs in platypus and chicken, have similar amounts of measured asynchronous replication (i.e. *Igf2* from mouse compared to *IGF2* in chicken; *Igf2* in mouse compared to *Igf2* in platypus; and *Igf2* in platypus compared to *IGF2* in chicken) (**Table 2**, all p-values: $0.54<p<0.96$). This result indicates a conservation in the level of replication asynchrony observed at these loci in similar cell types across amniote clades.

Replication timeline of S-phase entry is conserved at eutherian imprinted genes in mouse, and EIOs in platypus and chicken

We then investigated whether the conservation of replication asynchrony could be linked with the timing of S-phase entry of a gene, and performed consecutive FISH experiments on a common set of interphase nuclei (n=100) to measure the relative S-phase entry of these genes. We measured the relative replication time of the four imprinted genes and their corresponding EIOs, as well as the non-imprinted/non-EIO *Wsb1* gene, using the dot-assay approach (**Figure 1**). The resulting data gave an in-depth representation of how each locus replicated within a single nucleus in comparison to other loci measured within the same nucleus. A single-single signal (SS, 2 dot signal) corresponds to the unreplicated state of the locus, and thus can be measured to determine the average time it takes for the locus to enter S-phase, with a high amount of SS signal measured at a locus representing a late replicating status. We found that the non-imprinted/non-EIO control gene, *Wsb1* (*WSB1*), showed the most variation with a 21% SS variance across the species (**Table 3**). The other four imprinted genes, and their corresponding EIOs, entered S-phase in relatively the same order, starting with *Wt1* (*WT1*), then *Mest* (*MEST*) and *Igf2r* (*IGF2R*), with *Igf2* (*IGF2*) generally entering S-phase last (**Table 3**).

Conservation of S-phase entry replication timing across imprinted genes and EIOs

Finally, we tested whether the S-phase entry timing is conserved for mouse imprinted genes and EIOs in platypus and chicken, by performing a rank test on the S-phase entry data for each gene from each animal. We measured no significant difference between the order of S-phase entry for the four imprinted genes and EIOs (with SS values normalised against *Wsb1/WSB1* in each species) between the three species

(p–value = 0.037) (**Table 3**). To further determine whether the S-phase entry order was conserved among species, statistical T-tests were performed to determine whether the combined mouse, platypus, and chicken, SS values for a particular imprinted/corresponding EIO locus were significantly different to the other loci tested (**Figure 2**). Interestingly, pair wise comparisons between genes that appeared directly before or after each other in the replication timing pathway did not show any significant difference (all p-values > 0.07) (**Figure 2**). However, most of the loci which had at least one locus fall between them in the replication pathway did show a significant difference between genes, although this relationship did not hold true for *Wsb1* and *Mest* (**Figure 2**).

Discussion

We report a high level of conservation of asynchronous replication for a set of genes that are imprinted only in therian mammals in amniotes including the platypus (**Table 2**), with the exception of platypus *Wt1* (**Table 2**). Interestingly, whilst the *Wt1* gene is imprinted and undergoes around 25% asynchronous replication in humans (Bickmore and Carothers 1995), there are no reports of its imprinting status in mouse to date (Nishiwaki et al. 1997). However, here we observe that the asynchronous replication timing for the *Wt1* region is conserved in some amniotes.

Asynchronous replication: a characteristic of monoallelically expressing loci in amniotes?

In this study we report that chicken and platypus genes replicate asynchronously, suggesting that these loci may be monoallelically expressed in some cells within these organisms. It has previously been suggested that asynchronous replication may occur in imprinted orthologs in chicken due to transient monoallelic expression

during chicken embryogenesis. One study reported a monoallelically expressed state for *Igf2* in early chicken embryogenesis (Koski et al. 2000), however, other research has suggested that the locus is biallelically expressed during embryogenesis (Dünzinger et al. 2005; O'Neill et al. 2000; Yokomine et al. 2001). Notably, if these genes were found to undergo monoallelic expression in monotremes and birds it would be particularly pertinent to reassess current theories on the evolution of imprinting. It is interesting to speculate as to whether imprinting could be an adaptation of monoallelic expression, meaning that it has evolved at loci that were already undergoing random monoallelic expression. This could explain why earlier diverged amniotes which lack imprinting still have characteristics of monoallelic expression, namely asynchronous replication, at their EIOs. Hence, this would mean that in species such as the platypus and chicken, EIOs may be monoallelically expressed but not imprinted. Future work will need to investigate whether imprinted orthologs are undergoing transient monoallelic expression in certain cell types in birds and monotremes.

We report that imprinted genes and EIOs display similar amounts of asynchronous replication in similar cell types across three amniote species, indicating that the amount of replication asynchrony is also a conserved characteristic of these loci (**Table 2**). Notably, our replication asynchrony results for chicken EIOs had much higher SD values than in previous studies (Dünzinger et al. 2005). This could be due to the fact that the fibroblast cells in both studies were derived from different tissues, with evidence indicating that the same locus measured in different cell types will display distinct amounts of replication asynchrony in each cell type (Wilson et al. 2007). We also recently found that asynchronous replication of loci on the chicken Z

differed in cell lines derived from different developmental stages (Wright and Grutzner 2014 *manuscript in revision*).

Whilst no imprinting is observed in the platypus or chicken genome, there is evidence to suggest that imprinted genes in eutherians do not require differential allelic methylation in order to replicate asynchronously, suggesting the involvement of other epigenetic factors, such as DNA-binding proteins, in controlling replication asynchrony (Gribnau et al. 2003). Thus, the replication asynchrony at imprinted orthologs in platypus and chicken could still be controlled by differential epigenetic landscapes, such as differential histone modification, that could have promoted the evolution of imprinting at these loci in eutherian mammals.

Conservation of S-phase entry replication timing in amniotes

To our knowledge, we are the first to show that a conservation of S-phase entry timing exists among the eutherian imprinted genes *Igf2*, *Igf2r*, *Mest*, and *Wt1* in mouse, and their corresponding EIOs in platypus and chicken (**Table 3**, **Figure 2**). These imprinted genes and their corresponding EIOs also undergo similar amounts of replication asynchrony, indicating a greater conservation of mechanisms controlling replication timing at these loci than previously thought. It has previously been suggested that differences in allelic replication timing observed in eutherians are due to disparate chromatin landscapes at each allele (Bickmore and Carothers 1995). Thus, conserved chromatin landscapes could allow imprinted genes and their orthologs in chicken and monotreme amniotes to undergo similar replication entry into S-phase. Future work will need to assess the presence of differential chromatin marks, similar to those observed in therian imprinted genes, at platypus and chicken

EIOs, and whether the presence or absence of these marks change the replication timing of a locus.

While other studies have shown that larger genetic regions (10MB or more) appear to undergo similar replication timelines, in similar cell types among eutherians (Ryba et al. 2010), our research indicates that even at the level of individual loci a conservation of S-phase entry replication timing exists among imprinted genes and their orthologs across all amniotes (**Figure 2 and Table 2**). In light of previous research, which stipulates that expression is reflected in S-phase replication timing (Goldman et al. 1984), the presence of a conserved S-phase entry pattern suggests that the expression pattern of these genes may be more closely related among amniotes species than previously thought. We have uncovered a novel layer of regulated replication timing for imprinted genes and their corresponding EIOs, which is conserved among amniotes. This indicates that further replication timing program regulatory networks may exist for particular subsets of amniote genes. Whether these same mechanisms have set the scene for the evolution of monoallelic expression and genomic imprinting in these regions is currently unknown.

Methods

Cell lines and Interphase Preparations

Female mouse (C57), chicken, and platypus skin-derived fibroblasts were cultured using standard culture techniques, and passaged (and harvested) once they reached 70% confluence in order to promote exponential growth and increase the amount of cells in S-phase (Clausen 1987). Interphase nuclei were prepared following previously established protocols (Grutzner et al. 2004). Briefly, the cultured cells were treated with hypotonic solution (0.075 M KCl) and incubated at 37^{o}C for 20

minutes, before being fixed in methanol/acetic acid (3:1), dropped onto freshly cleaned slides (slides cleaned with 1% HCl/Methanol), and dried under humid conditions (around 55-60% humidity).

BAC DNA probes

Chicken and mouse BAC clones, corresponding to the genes in **Table 1**, were found using the BAC alignment in the UCSC Genome Browser. Clones were purchased from Children's Hospital Oakland Research Institute (CHORI, Oakland, CA, USA). Platypus imprinted ortholog BACs were identified and provided by Professor Thomas Haaf and Dr. Ulrich Zechner, from the Johannes Gutenberg University of Mainz, Germany. The non-imprinted platypus *Sox1* BAC was sourced from literature (Delbridge et al. 2006).

A BAC screen was performed to obtain a platypus *Wsb1* BAC using a previously established protocol (Dohm et al. 2007). Briefly, a PCR reaction was run on female platypus genomic DNA for template, with the platypus exon 4 *Wsb1* primers: Forward Primer 5' CCGGCTCCACCAAGAATGTC 3'; Reverse Primer 5' TTGTTTAACCCTGTTGCGAG 3'. The thermocycler conditions were the same as outlined in the next section, except that the annealing temperature for this primer set was 55^{o}C. The 223 bp PCR product (amplified from platypus exon 4) was gel-extracted (Wizard SV gel and PCR Clean-Up System), and 25 pmol of the extracted product was labelled using a random primed labelling kit (Roche Diagnostics, Mannheim, Germany), using 25 µCi of γ32PdATP (Amersham) following the manufacturer's instructions. The unincorporated γ32PdATP was removed using G-50 MicroColumns (Amersham), and the clean radioactively labelled oligonucleotide was then hybridised to the platypus BAC library filters (Oa_Bb BAC library with 11-fold

coverage of the Ornithorhynchus anatinus genomic sequence) in Church buffer (0.25 M Na2HPO4, 1 mM EDTA, 5% SDS) at 65^{o}C over night. The filters were then washed in 2×SSC/0.1% SDS at 65^{o}C for 20 minutes, and 0.5×SSC/0.1% SDS at 65^{o}C for 20 minutes, and exposed to Kodak XAR X-ray film (Sigma-Aldrich, Taufkirchen, Germany) for around 12-14 days at -80^{o}C, before being developed. Clones were purchased from Children's Hospital Oakland Research Institute (CHORI, Oakland, CA, USA), and the CUGI BAC/EST resource center (Clemson, SC, USA).

Confirming identity of BACs

Primers were designed to test for exonic gene regions on the BACs. Primer sequences and annealing temperatures are outlined in **Supplementary Table 1**. The presence of a gene on their respective BAC was confirmed using PCR. PCR products were amplified from mouse, platypus, and chicken DNA using the primers in **Supplementary Table 1**. Cycling parameters were: 95^{o}C for 3 min, then 35 cycles of 95^{o}C for 30 s; primer annealing temperature given in **Supplementary Table 1** for 30 s; 72^{o}C for 2 min (except for chicken and platypus *Igf2* primers, which were run for 7 and 12 minutes respectively); then a final cycle of 72^{o}C for 7 min, before being stored at 4^{o}C indefinitely.

BAC DNA labelling and precipitation

BAC clone DNA was extracted using The Wizard Minipreps DNA Purification System (Promega, Madison, WI, USA) following the manufacturer's protocol. 300-600 ng of BAC DNA was labelled using random 9-mer primers (Geneworks) in 5× supplemented dNTP buffer, Exo-Klenow polymerase (New England Biolabs, Ipswich, MA, USA) and green/orange fluorescent dUTPs (Abbott Molecular), the reaction was

then incubated at 37^{o}C overnight. Salmon sperm (50 µg), and sonicated genomic DNA (10 µg; 300-600 bp fragments) were then added to the labelled BACs, and precipitated for *20* minutes at −80^{o}C with and 5 vol. 100% ice-cold ethanol. The labelled probes were then pelleted and dissolved in formamide and hybridization mix (50% formamide/10% dextran sulphate/2×SSC) at 37^{o}C for 30 min, before being denatured at 80^{o}C for 10 min, and pre-annealed at 37^{o}C for 40 min.

Fluorescence in situ hybridisation

FISH was performed using a previously established protocol (Tsend-Ayush et al. 2009). Briefly, slides were washed in 2xSSC, then treated with 100 µgmL^{-1} RNase/2xSSC at 37^{o}C for 30 min, followed by a 0.01% pepsin/10 mM HCl treatment at 37^{o}C for 10 min. Slides were then equilibrated in 50 mM MgCl2, before being fixed in 1.3% formaldehyde/1xPBS/50 mM MgCl2. The slides were then dehydrated in a 70%, 90% and 100% ethanol series, and then denatured in 70% (v/v) formamide/2×SSC at 70^{o}C, before being dehydrated in another ethanol series. Slides were then hybridised with the labelled BAC in a moist chamber at 37^{o}C overnight. The following day, slides were washed three times in 50% formamide/2xSSC at 42^{o}C, washed in 2xSSC at 42^{o}C, 0.1xSSC at 60^{o}C, and again in 2xSSC at 42^{o}C, before being counterstained with 1 µg/ml DAPI/2×SSC for 1 min, dried and mounted with Vectashield (Vector Laboratories). Slides were viewed using a Zeiss AxioImagerZ.1 epifluorescence microscope equipped with a CCD camera, and images were taken using Zeiss Axiovision software. When slides underwent consecutive FISH experiments, coverslips were removed by washing slides in 2xSSC, running them through an ethanol (70%, 90%, 100%) series, then denaturing in 70% (v/v) formamide/2×SSC at 70^{o}C for 1 minute. Slides were then dehydrated in another ethanol series, allowed to air-dry completely, before being re-hybridised with

a new set of co-precipitated fluorescently labelled probes overnight at 37^{o}C, followed by subsequent washing and mounting steps as described above.

Microscopy and data collection

Slides were visualised with a Zeiss AxioImagerZ.1 epifluorescence microscope equipped with a CCD camera, and images acquired using Zeiss Axiovision software. For all BAC and gene information see **Table 1**. As an across-species control, we used *Wsb1*, which is not imprinted in mouse (Ling et al. 2006). We also used mouse *B3galt1* as a non-imprinted control, as well as platypus *Sox1* and chicken *RIMBP2* as non-EIO controls, as all three genes, to our knowledge, show no evidence of imprinting in therians. Measuring asynchronous replication: Four to six replicates of 50-61 nuclei were counted for each locus (total N counted is shown in **Table 2** for each locus). Measuring replication timing through consecutive FISH experiments: around 150 nuclei were counted in the first FISH experiment (to account for a third of data attrition across the experiment). After the three consecutive FISH experiments were completed (and images taken after each FISH) data was collated, and any nucleus, which showed no fluorescent signals or too much background to decipher a signal in any of the three FISH experiments, was discarded. Thus, only nuclei which showed clear signals in all three FISH experiments were included in the final dataset. T-tests were carried out to determine asynchronous replication for individual loci, and also to test for statistical differences in the amount of asynchronous replication at orthologous loci. A Friedman's rank test was performed to determine whether S-phase entry order was conserved among the mouse imprinted genes and EIOs in chicken and platypus.

Acknowledgements

The authors would like to acknowledge Mara Waldschmidt and Enkhjargal Tsend-Ayush, for their help during the BAC screening process, and confirming the presence of *Wt1* on the corresponding platypus BACs. They would also like to acknowledge Shu Ly Lim for her advice during the FISH and dot-assay counting procedures.

Funding

Frank Grützner is an Australian Research Council (ARC) research fellow. Megan Wright's PhD is funded through an Australian Postgraduate Award (bestowed by the Australian Government). Thomas Haaf and Ulrich Zechner were supported by research grants from the Deutsche Forschungsgemeinschaft (DFG).

Conflict of Interest

The authors declare that they have no conflict of interest.

Ethics

This article does not contain any studies with human participants or animals performed by any of the authors.

Tables

Table 1 Imprinted and non-imprinted genes and corresponding BACs used in FISH experiments

Gene status	Gene	Mouse BAC	Platypus BAC	Chicken BAC
Imprinted in eutherians	*Igf2*	RP23-51J21	Oa_Bb-349H20	CH261-122O12
	Igf2r	RP23-402J22	Oa_Bb-300J22	CH261-11C24
	Mest	RP24-211G11	Oa_Bb-491E23	CH261-56H17
	Wt1	RP23-357L23	Oa_Bb-274P3	CH261-98H7
Non-imprinted in eutherians	*Wsb1*	RP24-256H2	Oa_Bb-155A12	CH261-113K9
	B3galt1	RP23-346P15	-	-
	Sox1	-	Oa_Bb-117J7	-
	RIMBP2	-	-	CH261-3H21

Table 2 Replication asynchrony in platypus imprinted and non-imprinted orthologs, compared to mouse and chicken

Gene status	Gene†	Mouse % SD	N	Platypus % SD	N	Chicken % SD	N
Imprinted in eutherians	*Igf2 (IGF2)*	36%*	330	38%*	339	40%*	338
	Igf2r (IGF2R)	36%*	255	36%*	251	46%*	241
	Mest (MEST)	37%*	251	39%*	240	36%*	226
	Wt1 (WT1)	35%*	255	33%	251	45%*	241
Not imprinted in eutherians	*Wsb1 (WSB1)*	25%	263	26%	339	26%	257
	B3galt (B3GALT)	27%	299	NA		NA	
	Sox1 (SOX1)	NA		25%	328		
	Rimbp2 (RIMBP2)			NA		29%	269

†Chicken gene names are in brackets. *Indicates significant difference in replication asynchrony values ($p < 0.05$), as compared to the *Wsb1* (non-imprinted/non-imprinted ortholog) control gene.

Table 3 SS results from consecutive FISH experiment indicating conserved entry timing into S-phase

Gene	% SS			SS normalised to *Wsb1*			Order entering S-phase		
	M	P	C	Mouse	Platypus	Chicken	M	P	C
Igf2 (IGF2)	18%	20%	25%	0.642857	1.42857142	0.71428571	4	3	4
Igf2r (IGF2R)	8%	5%	12%	0.285714	0.35714285	0.34285714	2	1	2
Mest (MEST)	16%	16%	15%	0.571429	1.14285714	0.42857142	3	2	3
Wt1 (WT1)	6%	5%	7%	0.214286	0.35714285	0.2	1	1	1
Wsb1 (WSB1)	28%	14%	35%				Freidman's rank test p-value= 0.037		
Total nuclei counted	100	100	100						

Figures

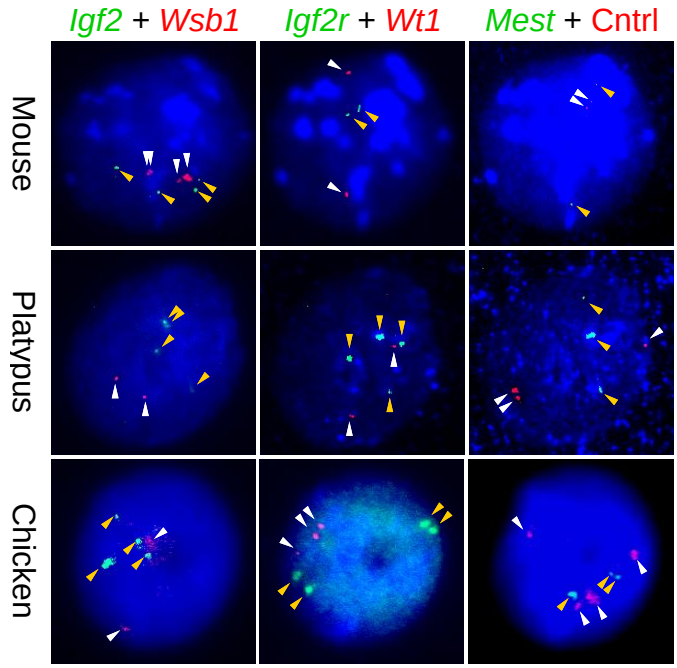


Fig1 Consecutive FISH experiment

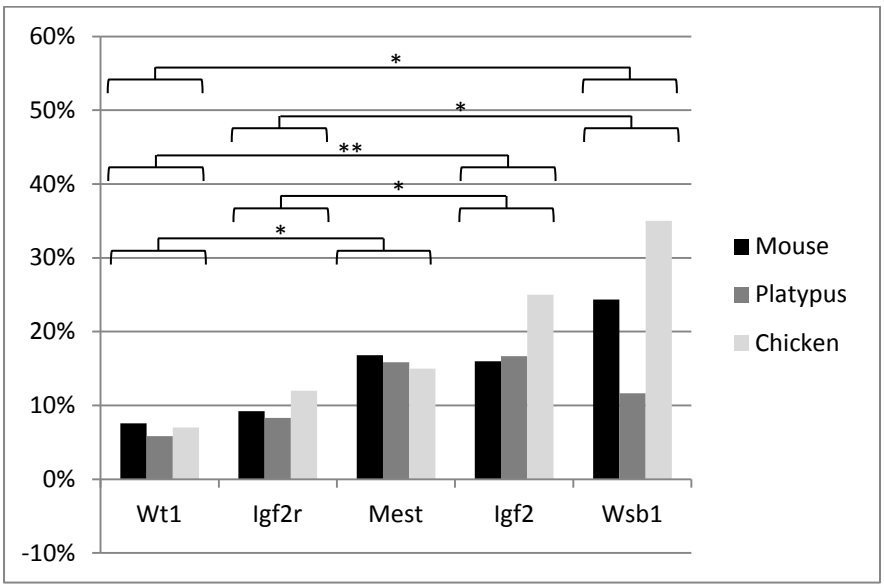


Fig2 Graph of the percentage of SS signal measured for each locus

Figure legends

Fig 1 Consecutive FISH experiment

Figure shows a single nucleus from mouse, platypus and chicken subjected to consecutive FISH experiments, with labelled BACs hybridising to the allele copies of the genes of interest. Yellow arrows indicate green signals, and white arrows indicate red signals. The control gene (Cntrl) is the non-imprinted control gene, and is as follows for each species: mouse *B3galt1*, platypus *Sox1*, and chicken *RIMBP2*.

Fig 2 Graph of the percentage of SS signal measured for each locus

The % SS is shown for *Wt1*, *Igf2r*, *Mest*, *Igf2* and *Wsb1*. Pairwise T-tests were performed between grouped species SS values for separate genes. Pairs of genes with a significant difference between their SS values are indicated on the graph. *P-value <0.05; **P-value<0.01.

References

Audit B, Zaghloul L, Vaillant C, Chevereau G, d'Aubenton-Carafa Y, Thermes C, Arneodo A (2009) Open chromatin encoded in DNA sequence is the signature of 'master' replication origins in human cells Nucleic Acids Res 37:6064-6075

Baumer A, Riegel M, Schinzel A (2004) Non-random asynchronous replication at 22q11.2 favours unequal meiotic crossovers leading to the human 22q11.2 deletion Journal of Medical Genetics 41:413-420

Bentley L et al. (2003) The imprinted region on human chromosome 7q32 extends to the carboxypeptidase A gene cluster: an imprinted candidate for Silver-Russell syndrome J Med Genet 40:249-256

Bergstrom R, Whitehead J, Kurukuti S, Ohlsson R (2007) CTCF regulates asynchronous replication of the imprinted H19/Igf2 domain Cell Cycle 6:450-454

Bickmore WA, Carothers AD (1995) Factors affecting the timing and imprinting of replication on a mammalian chromosome J Cell Sci 108 (Pt 8):2801-2809

Chess A, Simon I, Cedar H, Axel R (1994) Allelic inactivation regulates olfactory receptor gene expression Cell 78:823-834

Clausen OP (1987) DNA synthesis rate and unlabeled cells with S-phase DNA content in P388 leukemic cells in vivo studied by flow cytometry and cell sorting Cytometry 8:612-617

Delbridge ML, Wallis MC, Kirby PJ, Alsop AE, Grutzner F, Graves JA (2006) Assignment of SOX1 to platypus chromosome 20q by fluorescence in situ hybridization Cytogenet Genome Res 112:342L

Delcuve GP, Rastegar M, Davie JR (2009) Epigenetic control Journal of Cellular Physiology 219:243-250

DeVeale B, van der Kooy D, Babak T (2012) Critical Evaluation of Imprinted Gene Expression by RNA-Seq: A New Perspective PLoS Genet 8:e1002600

Dohm JC, Tsend-Ayush E, Reinhardt R, Grutzner F, Himmelbauer H (2007) Disruption and pseudoautosomal localization of the major histocompatibility complex in monotremes Genome Biol 8:R175

Dünzinger U, Nanda I, Schmid M, Haaf T, Zechner U (2005) Chicken orthologues of mammalian imprinted genes are clustered on macrochromosomes and replicate asynchronously Trends in Genetics 21:488-492

Farkash-Amar S, Lipson D, Polten A, Goren A, Helmstetter C, Yakhini Z, Simon I (2008) Global organization of replication time zones of the mouse genome Genome Res 18:1562-1570

Gimelbrant AA, Ensminger AW, Qi P, Zucker J, Chess A (2005) Monoallelic expression and asynchronous replication of p120 catenin in mouse and human cells J Biol Chem 280:1354-1359

Goldman MA, Holmquist GP, Gray MC, Caston LA, Nag A (1984) Replication timing of genes and middle repetitive sequences Science 224:686-692

Goldmit M, Bergman Y (2004) Monoallelic gene expression: a repertoire of recurrent themes Immunol Rev 200:197-214

Goncalves A et al. (2012) Extensive compensatory cis-trans regulation in the evolution of mouse gene expression Genome Res 22:2376-2384

Greally JM, Starr DJ, Hwang S, Song L, Jaarola M, Zemel S (1998) The mouse H19 locus mediates a transition between imprinted and non-imprinted DNA replication patterns Hum Mol Genet 7:91-95

Gribnau J, Hochedlinger K, Hata K, Li E, Jaenisch R (2003) Asynchronous replication timing of imprinted loci is independent of DNA methylation, but consistent with differential subnuclear localization Genes Dev 17:759-773

Grutzner F et al. (2004) In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes Nature 432:913-917

Hansen RS et al. (2010) Sequencing newly replicated DNA reveals widespread plasticity in human replication timing Proc Natl Acad Sci U S A 107:139-144

Hatton KS, Dhar V, Brown EH, Iqbal MA, Stuart S, Didamo VT, Schildkraut CL (1988) Replication program of active and inactive multigene families in mammalian cells Mol Cell Biol 8:2149-2158

Heun P, Laroche T, Raghuraman MK, Gasser SM (2001) The positioning and dynamics of origins of replication in the budding yeast nucleus J Cell Biol 152:385-400

Hiratani I et al. (2008) Global reorganization of replication domains during embryonic stem cell differentiation PLoS Biol 6:e245

Kitsberg D et al. (1993) Allele-specific replication timing of imprinted gene regions Nature 364:459-463

Koski LB, Sasaki E, Roberts RD, Gibson J, Etches RJ (2000) Monoalleleic transcription of the insulin-like growth factor-II gene (Igf2) in chick embryos Mol Reprod Dev 56:345-352

Ling JQ et al. (2006) CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1 Science 312:269-272

Muller CA, Nieduszynski CA (2012) Conservation of replication timing reveals global and local regulation of replication origin activity Genome Res 22:1953-1962

Nikaido I et al. (2003) Discovery of imprinted transcripts in the mouse transcriptome using large-scale expression profiling Genome Res 13:1402-1409

Nishiwaki K, Niikawa N, Ishikawa M (1997) Polymorphic and tissue-specific imprinting of the human wilms tumor gene,WT1 Jap J Human Genet 42:205-211

O'Neill MJ, Ingram RS, Vrana PB, Tilghman SM (2000) Allelic expression of IGF2 in marsupials and birds Dev Genes Evol 210:18-20

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

Renard-Guillet C, Kanoh Y, Shirahige K, Masai H (2014) Temporal and spatial regulation of eukaryotic DNA replication: from regulated initiation to genome-scale timing program Semin Cell Dev Biol 30:110-120

Ryba T et al. (2010) Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types Genome Res 20:761-770

Savova V, Vigneau S, Gimelbrant AA (2013) Autosomal monoallelic expression: genetics of epigenetic diversity? Curr Opin Genet Dev 23:642-648

Simon I, Tenzen T, Reubinoff BE, Hillman D, McCarrey JR, Cedar H (1999) Asynchronous replication of imprinted genes is established in the gametes and maintained during development Nature 401:929-932

Taljanidisz J, Popowski J, Sarkar N (1989) Temporal order of gene replication in Chinese hamster ovary cells Mol Cell Biol 9:2881-2889

Tsend-Ayush E et al. (2009) Higher-order genome organization in platypus and chicken sperm and repositioning of sex chromosomes during mammalian evolution Chromosoma 118:53-69

Wilson ND, Ross LJ, Close J, Mott R, Crow TJ, Volpi EV (2007) Replication profile of PCDH11X and PCDH11Y, a gene pair located in the non-pseudoautosomal homologous region Xq21.3/Yp11.2 Chromosome Res 15:485-498

Wright ML, Grutzner F (2011) Replication Timing: Evolution, Nuclear Organization and Relevance for Human Disease. In: Seligmann H (ed) DNA Replication - Current Advances. InTech, pp 437-466

Wright ML, Grutzner F (2014) Asynchronous replication timing of loci on the Z-chromosome in male chicken *manuscript in revision*

Yokomine T, Kuroiwa A, Tanaka K, Tsudzuki M, Matsuda Y, Sasaki H (2001) Sequence polymorphisms, allelic expression status and chromosome locations of the chicken IGF2 and MPR1 genes Cytogenet Cell Genet 93:109-113

Zakharova IS, Shevchenko AI, Zakian SM (2009) Monoallelic gene expression in mammals Chromosoma 118:279-290

Supplementary Material

Supplementary table 1: Our results vs literature results of replication asynchrony in eutherians of imprinted and non-imprinted genes

Gene status	Gene†	% SD^{Mus}	N	Literature % SD	Literature N	Asyn-chronous in literature?	Reference
Imprinted	*Igf2*	36%*	330	23% ^{Mus}	>100	Yes	(Kitsberg et al. 1993)
	Igf2r	36%*	255	35% ^{Mus}	>100	Yes	(Kitsberg et al. 1993)
	Mest (MEST)	37%*	251	25% ^{HSA}	>200	Yes	(Bentley et al. 2003)
	Wt1 (WT1)	35%*	255	25% ^{HSA}	>200	Yes	(Bickmore and Carothers 1995)
Non-imprinted	*Wsb1*	25%	263	NA			
	B3galt1	27%	299				
	p53	NA		12% ^{Mus}	>100	No	(Kitsberg et al. 1993)
	PfkL			11% ^{Mus}	>100	No	(Kitsberg et al. 1993)
	(D7S649) (marker on HSA chr.7)			9.7% ^{HSA}	>200	No	(Bentley et al. 2003)

†Human gene names are shown in brackets. *Indicates significant difference in replication asynchrony (probability <0.05), as compared to the *Wsb1* (non-imprinted) control gene.

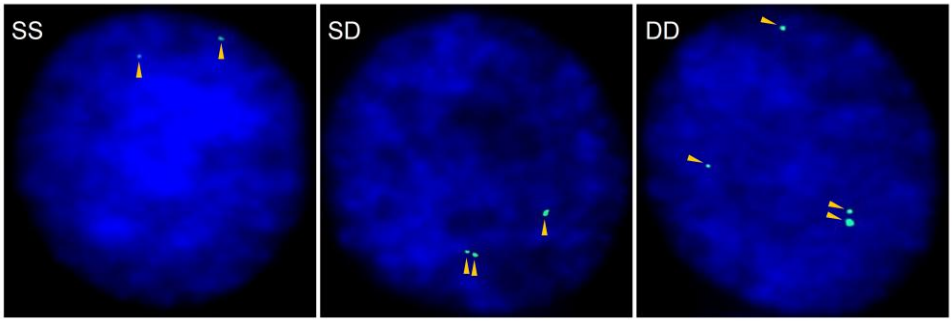
Supplementary table 2: Our results vs literature results of replication asynchrony in chicken around imprinted and non-imprinted orthologs

Gene status	Gene	% SD	N	Literature % SD	Literature N	Asynchronous in literature?	Reference
Imprinted orthologs	*IGF2*	40%*	338	25%*	258	Yes	(Dünzinger et al. 2005)
	IGF2R	45%*	241	22%	279	Yes	
	MEST	36%*	226	18%	299	Yes	
	WT1	46%*	241	NA			
Non-imprinted	*WSB1*	26%	257	NA			
	RIMBP2	29%	269				
	GAPDH	NA		10%	267	No	(Dünzinger et al. 2005)
	L23MRP			12%†	284	No	
	SLC6A8			4%	274	No	

*Indicates significant difference in replication asynchrony values (probability <0.05), as compared to the *Wsb1* (non-imprinted/non-imprinted ortholog) control gene. †More than 1 replicate in the literature, so only results from replicate with the highest N were included in this table.

Supplementary Table 3: Primers for testing genes on BACs, primer annealing temperatures, and amplified fragment size

Gene	Species	5' – 3' Primer Sequence	Primer anneal ^{o}C	PCR fragment size (bp)
Igf2	Mouse Platypus Chicken	F: GACACACTTCAGTTCGTCTGT R: AGCACTCCTCCACGATGCC	56	M: 1303 P: 12255 C: 6522
Igf2r	Mouse Platypus Chicken	F: AGAGATTATCTGGAAAGT R: CCACACACATTGAAA(T/A)AA(T/A)A	50	M: 458 P: 1673 C: 331
Wt1	Mouse Platypus Chicken	F: GAGAAACGTCCCTTCATGTG R: CTGTGCATCTGTAGGTGGG	55	M: 78 P: 77 C: 77
Mest	Mouse	F: GTAGACCATTCCTCAGCCTTCAG R: TCCTCTGAATACTTGCCACATTCC	57	181
Mest	Platypus	F: CTTGGCAGGATGAGAGAGTGG R: CCTGAAGACTTCCATGCCTGG	57	122
Mest	Chicken	F: CCACCACTACTCCATCTTTGAGC R: CCATAATCGTGGGACAGGAGG	57	102
Wsb1	Mouse	F: TGAAAGTATTGCGGGCACATCAG R: TTAAGACATTCAGGACACCGAGC	57	207
Wsb1	Platypus Chicken	F: GTGAATATAGAATGGCATCG R: TGGGGCAAAAGTTAAATCTC	50	P: 1339 C: 611
B3galt1	Mouse	F: GTTGTCAGTGGAAGTTTCCGTC R: CTGAAGCACATGCCCAGAAC	57	224
Sox1	Platypus	F: CACAACTCGGAGATCAGCAA R: GTTGGTCCAGCCGTTGAC	56	? (Delbridge et al. 2006)
Rimbp2	Chicken	F: ATGCCGGACACAAAGTGAAC R: TGCCTAAAGAAGTGGCTATTCC	56	176



Supplementary Fig 1 FISH dot-assay approach to measure replication timing

The fluorescent probes (green) hybridise to the allele copies of the *IGF2* locus in individual platypus nuclei, and indicate the replication status in one of three states; SS: two single dot signals show that locus is unreplicated; SD: a single-double signal pair show that the locus is replicating (this is counted in the dot assay as asynchronous replication, as one allele has replicated whilst the other has not yet done so); DD: a double-double dot signal showing that the locus has replicated.

MANUSCRIPT:

Co-detection of bromodeoxyuridine and directly-labelled fluorescent *in situ* hybridization signals in interphase cells of amniote species

Statement of Authorship

Title of paper	**Co-detection of bromodeoxyuridine and directly-labelled fluorescent *in situ* hybridization signals in interphase cells of amniote species**
Publication status	Unpublished
Publication details	Wright, M.L. and Grutzner, F. Co-detection of bromodeoxyuridine and directly-labelled fluorescent *in situ* hybridization signals in interphase cells of amniote species. *Unpublished manuscript*. For planned submission in 2015 to *Molecular Cytogenetics*.

Authors have copyright and right to re-use the article in this thesis.

Author Contributions

By signing the Statement of Authorship, each author certifies that their stated contribution to the publication is accurate and that permission is granted for the publication to be included in the candidate's thesis.

Name of Principal Author (Candidate)	Megan Wright
Contribution to the Paper	Completed all experiments (sourcing and growing BACs, tissue culture, FISH dot-assay, BrdU staining, and combined Brdu and directly-labelled FISH experiments), data-collection (counting and taking microscope images), and statistical analysis. Wrote the manuscript, and compiled all figures and tables.
Signature	19.12.2014

Name of Co-Author	Frank Grützner
Contribution to the Paper	Supervised all research and aided in the manuscript construction/outlining.
Signature	19.12.2014

Manuscript Title: Co-detection of bromodeoxyuridine and directly-labelled fluorescent *in situ* hybridization signals in interphase cells of amniote species

Manuscript prepared for submission to the journal *Molecular Cytogenetics* (Methodologies)

Authors and Affiliations

Megan Lynne Wright[1]* & Frank Grutzner[1]

*all correspondence to this author.

[1]The Robinson Research Institute, School of Molecular & Biomedical Sciences, The University of Adelaide, Australia

Corresponding author contact information

megan.wright@adelaide.edu.au

Phone: + 61 430 402 921

Fax: + 61 8 8313 4362

The Robinson Research Institute
School of Molecular & Biomedical Sciences
The University of Adelaide
Adelaide
S.A. 5005
AUSTRALIA

Abstract

Background: Bromodeoxyuridine (BrdU), a thymidine replacement, is incorporated during the Synthesis phase (S-phase) of the cell growth cycle, and has been used as a marker for cells in synthesis phase of mitosis. In combination with directly-labelled fluorescent *in situ* hybridisation (FISH), BrdU incorporation has been used to determine replication timing or replication asynchrony. Previously this powerful approach has been performed using indirectly-labelled FISH probes and immunostaining. Here, we present a more reliable method that allows directly-labelled FISH to be carried out in conjunction with BrdU immunostaining. We also investigate the feasibility of this approach in different amniote species including mouse, platypus and chicken.

Results: Using skin derived fibroblasts from mouse, platypus, and chicken, we perform BrdU-immnuostaining with an optimised blocking procedure on nuclei that reduces background signals. Directly-labelled FISH experiments are then performed successfully in combination with the optimised BrdU immunostaining procedure in each of the three species. Finally, we measure S-phase through BrdU immunostaining and correlate it with the FISH-measured replication status of the mouse *Insulin-like growth factor II* (*Igf2*) locus, as well as its orthologs in platypus (*Igf2*) and chicken (*IGF2*). We find that across all three species, an average of 22-23% nuclei show a combined positive BrdU immunostaining status with a FISH signal indicating a current DNA replication event.

Conclusion: We have developed an improved technique that couples BrdU immunostaining with a direct-labelled FISH experiment, negating the requirement for secondary-antibodies during FISH detection, and reducing experimental background and increasing sensitivity for cross species experiments. Furthermore, we show that

the experiment works in three different amniotes, detecting similar amounts of S-phase cells across the species.

Keywords

Bromodeoxyuridine, BrdU, immunostain, Fluorescent *in situ* Hybridisation, FISH, directly-labelled FISH, chicken, mouse, platypus, FISH dot-assay.

Background

BrdU immunostaining to detect S-phase during the cell cycle

BrdU incorporation has long been used as a way to detect S-phase in interphase cells, and asynchronous replication across a karyotype in metaphase cells [1]. Its presence in DNA is detected by a monoclonal antibody which detects BrdU in single-stranded DNA [2]. It is important, however, to note that BrdU is not incorporated solely in S-phase; as a marker of synthesis, BrdU also detects events such DNA repair and abortive cell re-entry, although it has been shown that it is not incorporated into the DNA in high amounts through to these processes under normal growth conditions [3, 4, reviewed in 5]. In S-phase cells, DNA replication is characterised through BrdU staining due to the presence of a spattering of bright foci within the nuclear periphery coupled with an amount of diffuse uniform staining across the nucleus. However, in other types of BrdU incorporation, such as in DNA repair, only bright BrdU foci exist in the nucleus [3]. Thus, under normal growth conditions S-phase cells are quite easy to identify if immnostaining is working correctly.

FISH dot-assay to detect S-phase during the cell cycle

Many studies investigating asynchronous replication use the 'dot assay' FISH approach, whereby a fluorescently labelled single locus probe will indicate the number of allele copies (two and four before or after S-phase, three in S-Phase) which exist on hybridisation to the gene of interest [reviewed in 6]. A high frequency (above 25-30%) of cells showing three signals for a locus has been used as evidence of asynchronous replication in multiple species [7, 8, 9, 10, 11].

Simultaneous BrdU immunostaining and directly-labelled FISH allows dual detection of S-phase

Coupling BrdU immunostaining with a FISH based dot assay allows a cell in S-Phase to be identified in two ways: a general signature for DNA replication in a nucleus observed through BrdU incorporation, and replication in-process for a single locus (see **Figure 1**). Previous protocols for BrdU immunostaining have relied on acidic denaturation treatment (sometimes on top of a formamide treatment) [12, 13]. The low pH can have undesirable stringency affects on the in situ hybridisation, as well as causing around 20% DNA loss on average, which is not preferable when carrying out a BrdU immunostain coupled with FISH [14]. Moreover, more recent BrdU-FISH co-detection protocols, which use indirectly-labelled FISH, have only been tested in eutherian cells [15].

We have developed a protocol without an initial acidic denaturation for BrdU immunodetection, in combination with directly-labelled FISH. The directly-labelled FISH in this protocol includes RNase and pepsin treatments, formaldehyde re-fixation, and formamide denaturation. Our protocol also uses a series of blocking steps, carried out during BrdU immunodetection, that are coupled with Triton-X

permeabilisation to achieve less background after antibody detection. Finally, our optimised BrdU detection protocol has been shown to work across amniote species including mouse, platypus, and chicken.

Results

Incorporation of blocking enhances BrdU immnostaining

By incorporating two blocking steps when carrying out BrdU immunodetection (without any FISH), we found that the BrdU signals were much brighter and the slides had a reduction in the amount of background present, compared to when the slides underwent no blocking (**Figure 2**). These clearer BrdU signals would offer an advantage to any who are trying to easily identify S-phase cells, as the blocking made the contrast between heavily stained and lightly stained nuclei much greater. Furthermore, the brighter signals also make for easier identification of the amount of S-phase in the cell lines. All of our cell lines were grown to promote exponential growth, where cells were continuously passaged at 60-70% confluence, in order to decrease contact inhibition and the slowing of growth [16]. The cells were also not synchronised in any way as this can affect the natural replication profile of individual loci [17] . Under these growth conditions, we found that an average of 50-60% of interphase cells in all three species were positive for BrdU (**Table 2**).

Simultaneous detection of improved BrdU immunostaining with directly-labelled FISH

Our optimised protocol utilises directly-labelled FISH but requires more blocking steps and Triton X-100 treatment both before, and during, primary antibody detection (**Table 1**). However, it does not require harsh acidic treatments (which cause flurorophores to be rapidly quenched, and results in DNA loss) and only relies on a

formamide heat treatment to denature the DNA. Furthermore, our protocol utilises directly-labelled BAC-DNA probes, which negates the requirement for antibody detection during the FISH, and to our knowledge, has never been done in a BrdU co-detection experiment. In our own hands, our simultaneous detection protocol had a 100% success rate, with every slide tested in each of the species (mouse, platypus, and chicken) showing clear BrdU and FISH signals (**Figure 3**).

Across all three species the SD signal measured for the *Insulin-like growth factor II* (*Igf2*) gene in the fibroblasts was above 30%, and thus indicative of asynchronous replication [18, 19, 20]. Specifically, the mouse *Igf2* (GenBank:U71085) displayed around 39% SD signal, the platypus *Igf2* (GenBank:AY552324) around 33% SD signal, and the chicken *IGF2* (GenBank: AY267181) around 40% SD signal. When measuring the overlap of an SD signal at the *Igf2* locus with a positive BrdU incorporation status, we observe that an average of 22-23% of SD signals occur in cells positive for BrdU incorporation across the three species (**Table 3**).

Discussion

Detecting S-phase amount during exponential growth of cultured cells

The quick identification of cells that were shown to be in S-phase on two counts, due to the clear simultaneous BrdU and FISH signals achieved through our protocol, allowed us to measure how often an S-phase FISH signal (SD) coincided with an S-phase BrdU immunostained cell during exponential growth. When cells are undergoing exponential growth around 50% are observed to be in S-phase [16], compared to a non-exponential growth pattern, in which around only 25% of the cells will be in S-phase. Our BrdU incorporation and immunostaining experiments

indicated that around 50-60% of cells from each cell line tested positive for BrdU, indicative of the expected amount of S-phase in a cell line undergoing exponential growth (**Table 2**). We found that *Igf2* showed an average of around 33-40% SD signal in the three species, which is within the expected range of SD signal for an imprinted locus, or an imprinted ortholog, as measured in previous studies [7, 17, 20, 21 *manuscript in preparation*]. Intriguingly, only 22-23% of nuclei with an SD signal, for *Igf2* in all three species, also showed BrdU staining (**Table 3**). This observation has also been made in another study, which noted that some doublet FISH signals can be observed in the absence of S-phase markers, which was attributed to the separation of denatured chromatin strands or chromatin decondensation [11, 22]. Furthermore, some doublets may have not coincided with BrdU incorporation if they replicated earlier than the short BrdU pulse. It is interesting to speculate whether locus-specific DNA replication timing also comes into play here; the *Igf2* locus is known to undergo generally later replication timing in mouse, platypus and chicken cells [21 *manuscript in preparation*]. This suggests that by the time the *Igf2* gene is being replicated late in S-phase, most of the genome has already undergone replication, and thus BrdU incorporation is below a detectable limit at this point. Thus, only some of the nuclei with an SD signal for *Igf2* would also show BrdU immunostaining.

Conclusion

Replication timing and reorganisation of chromatin during S-phase is an important aspect of genome organisation in the nucleus. Our approach to improve global BrdU incorporation in combination with single-locus directly labelled FISH provides a powerful approach to investigate the evolution of replication timing in various species.

We have improved previous techniques detailing simultaneous detection of BrdU and FISH signals by optimising a technique that utilises directly-labelled FISH hybridisation, utilises additional blocking steps to aid in background suppression and boost BrdU immunostaining signal intensity, and finally, requiring no DNA-degrading acidic denaturation treatments during immunostaining. This optimised protocol has been tested in various species, namely mouse, platypus, and chicken, revealing that 22-23% of SD signals for the *Igf2* locus were present in cells with BrdU staining in all three species, indicating that not all SD signals are accompanied with a positive BrdU signal, perhaps due to the separation of denatured chromatin strands.

Methods

Cell culture

Female mouse (C57), chicken, and platypus skin-derived fibroblasts were grown in a culture medium containing Advanced DMEM (12491-023, Life Technologies), supplemented with 1% L-glutamine (25030-081, Invitrogen) and 10% Fetal Calf Serum (10099-441, Invitrogen). Cells were grown in 75 cm^2 TC flasks until they reached 60-70% confluence (in order to promote exponential growth and increase the amount of cells in S-phase). Cells were then removed from the flask with 3ml of 0.05% Trypsin (15400-054, Invitrogen) at 37^oC for 2 min with slight shaking. The Trypsin was then inactivated with 6 ml culture medium, and the solution was centrifuged at 1,200 rpm for 10 min to pellet the cells, before the supernatant was removed, and the cells resuspended in fresh culture medium. Cell lines underwent 5 to 7 passages before being harvested.

BrdU incorporation

When cells were at their final passage (before being harvested) they were pulse labelled for 45 min with 40 µg/ml BrdU (B9285, Sigma), then washed twice 1xPBS. Cells were removed from the flask with 3 ml of 0.05% Trypsin at 37 ^{o}C for 2 min with slight shaking, followed by Trypsin inactivation with 6 ml culture medium, and centrifugation of the solution at 1,200 rpm for 10 min to pellet the cells. Most supernatant was removed, leaving only 0.5 ml supernatant with the pellet in which the cells were carefully resuspended.

Harvesting and slide preparation

2 ml of 0.075M Potassium Chloride hypotonic solution at 37 ^{o}C was slowly added to the 0.5 ml of resuspended cell solution, and the mixture was carefully resuspended. Another 8 mls of hypotonic solution was added to the solution, and the cells carefully resuspended again, then the solution was incubated for 20 min at 37^{o}C. After incubation, 2 drops of ice-cold methanol/acetic acid (3:1) fixative were added to the solution and the mixture was again carefully resuspended. The cells were then pelleted by centrifugation (1,200 rpm for 10 min). Most supernatant was removed, leaving only 0.5 ml, to which 0.5 ml of ice-cold fixative was added, and the mixture carefully resuspended. The solution was then transferred to a 1.5 ml tube, and centrifuged for 30 sec at 13,000 rpm to pellet nuclei, before all supernatant was removed, and 1 ml of fresh fixative was added. The fixed cells were washed three times, following these washing steps: pellet resuspended in fresh fixative; solution centrifuged at 13,000 rpm for 10 min; removal of all supernatant; resuspended in fresh fixative. After being washed and resuspended in fresh fixative, cells were ready to be dropped onto freshly cleaned slides (slides cleaned with 1% HCl/Methanol), and dried at room-temperature for 1 hour under humid conditions (around 55-60%

humidity), then left for another hour at room-temperature. Slides were then ready to undergo the FISH slide preparation procedure.

Bacterial Artificial Chromosome (BAC) culture, DNA extraction, and identification of Igf2

BAC clones containing Mouse *Igf2* (RP23-51J21) and Chicken *IGF2* (CH261-122O12), were purchased from the Children's Hospital Oakland Research Institute (CHORI, Oakland, CA, USA), and grown following CHORI's instructions. The Platypus *Igf2* BAC clone (Oa_Bb-349H20) was purchased from Clemson University Genomics Institute (CUGI, Clemson, SC, USA), and grown following CUGI's instructions. BAC clone DNA was extracted using The Wizard Minipreps DNA Purification System (Promega, Madison, WI, USA). The presence of *Igf2* was confirmed on the BACs through PCR, using the following primers for Mouse and Chicken: Forward 5'- GACACACTTCAGTTCGTCTGT-3' and Reverse 5' AGCACTCCTCCACGATGCC-3' (yielding products sizes of 1303 bp and 6522 bp for Mouse and Chicken respectively); and the following primers in Platypus: Forward 5'-ATCGTGGAGGAGTGTTGCTTC-3', and Reverse 5'-CGTCACGGAAGAACTTTGTGC-3' (yield a product size of 1996 bp). Cycling parameters were as follows: 95^{o}C for 3 min; then 35 cycles of: 95^{o}C for 30 s, 56^{o}C for 30 s, 72^{o}C for 6 min; then a final cycle of 72^{o}C for 7 min, before being stored at 4^{o}C indefinitely.

BAC fluorescent labelling and precipitation for FISH

For a 50 µl labelling reaction: 5 µl BAC DNA (1µg total), 10 µl of 1 µg/µl random 9-mer primer (Geneworks), and 22 µl of water were added to a PCR tube, vortexed for

10 seconds, quick-spun, then incubated for 15 min at 95 ^{o}C, and cooled on ice for 5 minutes, before adding 9.2 ml of 5×dNTP supplemented NEB buffer 2 (supplemented with a 200 µM dNTP solution containing: 100 µM dACG-TPs and 40 µM dTTPs), 0.7 µl SpectrumGreen/Orange fluorescent dUTPs (Abbott Molecular), and 1 µl Exo-Klenow polymerase (MO221, New England Biolabs). The reaction was then incubated at 37^{o}C overnight. The labelled BAC DNA was then moved to a 1.5 ml tube, and salmon sperm (50 µg) as well as sonicated genomic DNA (10 µg; 300-600 bp fragments) were added to the tube, followed by the addition of 500 µl 100% ice-cold ethanol and precipitation for 20 min at −80^{o}C. The labelled probes were then centrifuged at 14,000 rpm for 30 min at 4 ^{o}C, before the supernatant was carefully removed, and the pellet dried for 8 min at 37^{o}C. The pellet was then dissolved in 5µl Deionised Formamide (F9037, Sigma) for 15 min at 37^{o}C and 5 µl hybridization mix (50% Deionised Formamide/10% dextran sulphate/2×SSC) at 37^{o}C for 15 min, before being denatured at 80^{o}C for 10 min, and pre-annealed at 37^{o}C for 40 min. Probes were then ready to be hybridised onto a slide.

BrdU immunostaining

Fresh slides were washed in 2xSSC for 5 min, then treated with 0.01% pepsin/10 mM HCl treatment at 37^{o}C for 10 min, washed twice in 1xPBS for 5 min, and equilibrated in 50 mM MgCl2/1xPBS, before being fixed in 1.3% Formaldehyde/1xPBS/50 mM MgCl2. The slides were then dehydrated in a 70%, 90% and 100% ethanol series (5 min for each ethanol solution), and denatured for 2.5 min in 70% (v/v) formamide/2xSSC at 70^{o}C, before being dehydrated in another ethanol series (as before) and air-dried for 10 min. Slides were then washed 1 x 5 min in 0.1% Triton X-100/1xPBS.

Blocking on slides: Slides were then blocked 3 x 5 min with 300 µl Blocking Solution 1 (0.5% (w/v) BSA/0.5% (w/v) milk powder/1xPBS) at 25^{o}C, before 200 µl of the BrdU primary antibody (Cat. B8434, Sigma-Aldrich) solution (1/100 dilution Anti-BrdU in 0.1% Triton X-100/10% (w/v) BSA/1xPBS) was spread across the slides and incubated under a parafilm coverslip in a moist chamber at 37^{o}C for 30 min. The slides were then washed 3 x 5 min in 1xPBS, before being blocked 3 x 5 min with 300 µl Blocking solution 2 (10% Goat serum, 5% (w/v) milk powder/1xPBS) at 25^{o}C, before 200 µl of the Anti-mouse Alexa-568 (A-11004, Life Technologies) secondary antibody solution (1/400 dilution of Anti-Mouse Alexa-568/Blocking solution 2) was spread across the slides and incubated under a parafilm coverslip in a moist chamber at 37 ^{o}C for 30 min. Slides were then washed 3 x 5 min in 1xPBS, counterstained with 1 µg/ml DAPI/2×SSC for 1 min, then air dried for 1 hour and mounted with Vectashield (Vector Laboratories).

No blocking on slides: 200 µl of the BrdU primary antibody (Cat. B8434, Sigma-Aldrich) solution (1/100 dilution Anti-BrdU in in 0.1% Triton X-100/10% (w/v) BSA/1xPBS) was spread across the slides and incubated under a parafilm coverslip in a moist chamber at 37^{o}C for 30 min. The slides were then washed 3 x 5 min in 1xPBS, before 200 µl of the Anti-mouse Alexa-568 (A-11004, Life Technologies) secondary antibody solution (1/400 dilution of Anti-Mouse Alexa-568 in 10% (w/v) BSA/1xPBS), was spread across the slides and incubated under a parafilm coverslip in a moist chamber at 37^{o}C for 30 min. The slides were then washed 3 x 5 min in 1xPBS, DAPI counterstained, mounted, and visualised as described previously.

Simultaneous detection of Directly-labelled FISH and BrdU immunoprecipitation

Fresh slides were washed in 2xSSC for 5 min, then treated with 100 µg/ml RNase/2xSSC at 37^{o}C for 30 min, and washed three times in 2xSSC for 5 min.

Slides were then treated with 0.01% pepsin/10 mM HCl treatment at 37^{o}C for 10 min, washed twice in 1xPBS for 5 min, and equilibrated in 50 mM MgCl2/1xPBS, before being fixed in 1.3% Formaldehyde/1xPBS/50 mM MgCl2. The slides were then dehydrated in a 70%, 90% and 100% ethanol series (5 min for each ethanol solution), and denatured for 2.5 min in 70% (v/v) formamide/2×SSC at 70^{o}C, before being dehydrated in another ethanol series (as before) and air-dried for 30 min. Half a slide (18 mm x 18 mm area) was then hybridised with 10 µl labelled BAC/formamide/hybridization mix, capillary action spread under a half-slide coverslip and left in a water moist chamber at 37^{o}C overnight. The following day, slides were washed 3 x 5 min in 50% Formamide/2xSSC at 42^{o}C, washed in 2xSSC at 42^{o}C for 5 min and 0.1xSSC at 60^{o}C for 5 min. Slides were then washed 1 x 5 min in 0.1% Triton X-100/1xPBS, blocked 3 x 5 min with 300 µl Blocking Solution 1 (0.5% (w/v) BSA/0.5% (w/v) milk powder/1xPBS) at 25^{o}C, before 200 µl of the BrdU primary antibody (Cat. B8434, Sigma-Aldrich) solution (1/100 dilution Anti-BrdU in 0.1% Triton X-100/10% (w/v) BSA/1xPBS) was spread across the slides and incubated under a parafilm coverslip in a moist chamber at 37^{o}C for 30 min. The slides were then washed 3 x 5 min in 1xPBS, before being blocked 3 x 5 min with 300 µl Blocking solution 2 (10% Goat serum, 5% (w/v) milk powder/1xPBS) at 25^{o}C, before 200 µl of the Anti-mouse Alexa-568 (A-11004, Life Technologies) secondary antibody solution (1/400 dilution of Anti-Mouse Alexa-568/Blocking solution 2) was spread across the slides and incubated under a parafilm coverslip in a moist chamber at 37^{o}C for 30 min. Slides were then washed 3 x 5 min in 1xPBS, counterstained with 1 µg/ml DAPI/2×SSC for 1 min, rinsed twice in MQ water, then air dried for 1 hour and mounted with Vectashield (Vector Laboratories).

Microscopy and data collection

Slides were viewed using a Zeiss AxioImagerZ.1 epifluorescence microscope equipped with a CCD camera, and images were taken using Zeiss Axiovision software. Three replicates were taken with combined BrdU and FISH data (ie. each individual nucleus had its combined BrdU status and FISH dot-assay replication status recorded) for the mouse (3 replicates of 40 nuclei; Total N = 120), platypus (2 replicates of 37, and 1 replicate of 38; Total N = 112), and chicken (3 replicates of 40 nuclei; Total N = 120).

List of abbreviations

BrdU: bromodeoxyuridine; FISH: fluorescent *in situ* hybridization; S-phase: DNA synthesis phase of the cell-cycle; SS: single-single dot FISH signal; SD: single-double dot FISH signal; DD: double-double dot FISH signal; *Igf2*: *Insulin-like growth factor II*; BAC: bacterial artificial chromosome.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Megan Wright carried out all experimental and optimization work (including BAC identification, immunostaining procedures, and FISH experiments), all data collection and statistics (taking microscope images and BrdU + FISH dot-assay counting/statistics), wrote the manuscript draft and compiled all figures and tables. Frank Grützner supervised all research and manuscript drafting, as well as acting as primary correspondent during the submission process.

Authors' information

Frank Grützner is an Australian Research Council (ARC) research fellow.

Megan Wright's PhD is funded through an Australian Postgraduate Award (bestowed by the Australian Government).

Acknowledgments

The authors would like to acknowledge Aaron Casey for his advice with immunostaining procedures.

References

1. Latt SA: **Microfluorometric detection of deoxyribonucleic acid replication in human metaphase chromosomes.** *Proc Natl Acad Sci U S A* 1973, **70:**3395-3399.
2. Gratzner HG: **Monoclonal antibody to 5-bromo- and 5-iododeoxyuridine: A new reagent for detection of DNA replication.** *Science* 1982, **218:**474-475.
3. Kao GD, McKenna WG, Yen TJ: **Detection of repair activity during the DNA damage-induced G2 delay in human cancer cells.** *Oncogene* 2001, **20:**3486-3496.
4. Bauer S, Patterson PH: **The cell cycleâ€“apoptosis connection revisited in the adult brain.** *The Journal of Cell Biology* 2005, **171:**641-650.
5. Taupin P: **BrdU immunohistochemistry for studying adult neurogenesis: paradigms, pitfalls, limitations, and validation.** *Brain Res Rev* 2007, **53:**198-214.
6. Wright ML, Grutzner F: **Replication Timing: Evolution, Nuclear Organization and Relevance for Human Disease.** In *DNA Replication - Current Advances.* Edited by Seligmann H: InTech; 2011: 437-466
7. Dünzinger U, Nanda I, Schmid M, Haaf T, Zechner U: **Chicken orthologues of mammalian imprinted genes are clustered on macrochromosomes and replicate asynchronously.** *Trends in Genetics* 2005, **21:**488-492.
8. Ho KKK, Deakin JE, Wright ML, Graves JAM, Grutzner F: **Replication asynchrony and differential condensation of X chromosomes in female platypus (Ornithorhynchus anatinus).** *Reproduction, Fertility and Development* 2009, **21:**952-963.

9. Kitsberg D, Selig S, Brandeis M, Simon I, Keshet I, Driscoll DJ, Nicholls RD, Cedar H: **Allele-specific replication timing of imprinted gene regions.** *Nature* 1993, **364:**459-463.

10. Bentley L, Nakabayashi K, Monk D, Beechey C, Peters J, Birjandi Z, Khayat FE, Patel M, Preece MA, Stanier P, et al: **The imprinted region on human chromosome 7q32 extends to the carboxypeptidase A gene cluster: an imprinted candidate for Silver-Russell syndrome.** *J Med Genet* 2003, **40:**249-256.

11. Bickmore WA, Carothers AD: **Factors affecting the timing and imprinting of replication on a mammalian chromosome.** *J Cell Sci* 1995, **108 (Pt 8):**2801-2809.

12. Boggs BA, Chinault AC: **Analysis of DNA replication by fluorescence in situ hybridization.** *Methods* 1997, **13:**259-270.

13. De Sario A, De Carli L, Raimondi E: **An improved method for chromosome banding after fluorescence in situ hybridization: use of a tetramethylrhodhamine-conjugated BrdU antibody.** *Cytotechnology* 1991, **7:**55-58.

14. Raap AK, Marijnen JGJ, Vrolijk J, van der Ploeg M: **Denaturation, renaturation, and loss of DNA during in situ hybridization procedures.** *Cytometry* 1986, **7:**235-242.

15. Moralli D, Monaco ZL: **Simultaneous detection of FISH signals and bromo-deoxyuridine incorporation in fixed tissue cultured cells.** *PLoS One* 2009, **4:**e4483.

16. Clausen OP: **DNA synthesis rate and unlabeled cells with S-phase DNA content in P388 leukemic cells in vivo studied by flow cytometry and cell sorting.** *Cytometry* 1987, **8:**612-617.

17. Bergstrom R, Whitehead J, Kurukuti S, Ohlsson R: **CTCF regulates asynchronous replication of the imprinted H19/Igf2 domain.** *Cell Cycle* 2007, **6:**450-454.

18. Baumer A, Riegel M, Schinzel A: **Non-random asynchronous replication at 22q11.2 favours unequal meiotic crossovers leading to the human 22q11.2 deletion.** *Journal of Medical Genetics* 2004, **41:**413-420.

19. Wilson ND, Ross LJ, Close J, Mott R, Crow TJ, Volpi EV: **Replication profile of PCDH11X and PCDH11Y, a gene pair located in the non-pseudoautosomal homologous region Xq21.3/Yp11.2.** *Chromosome Res* 2007, **15:**485-498.

20. 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:**759-773.

21. Wright ML, Zechner U, Haaf T, Grutzner F: **Conservation of replication asynchrony and S-phase entry predates evolution of genomic imprinting in mammals.** *manuscript in preparation*.

22. Houseal TW, Klinger KW: **What's in a spot?** *Hum Mol Genet* 1994, **3:**1215-1216.

Figures

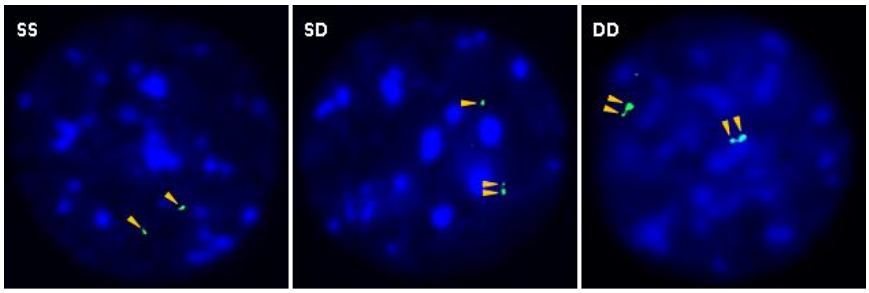


Figure 1: FISH dot assay

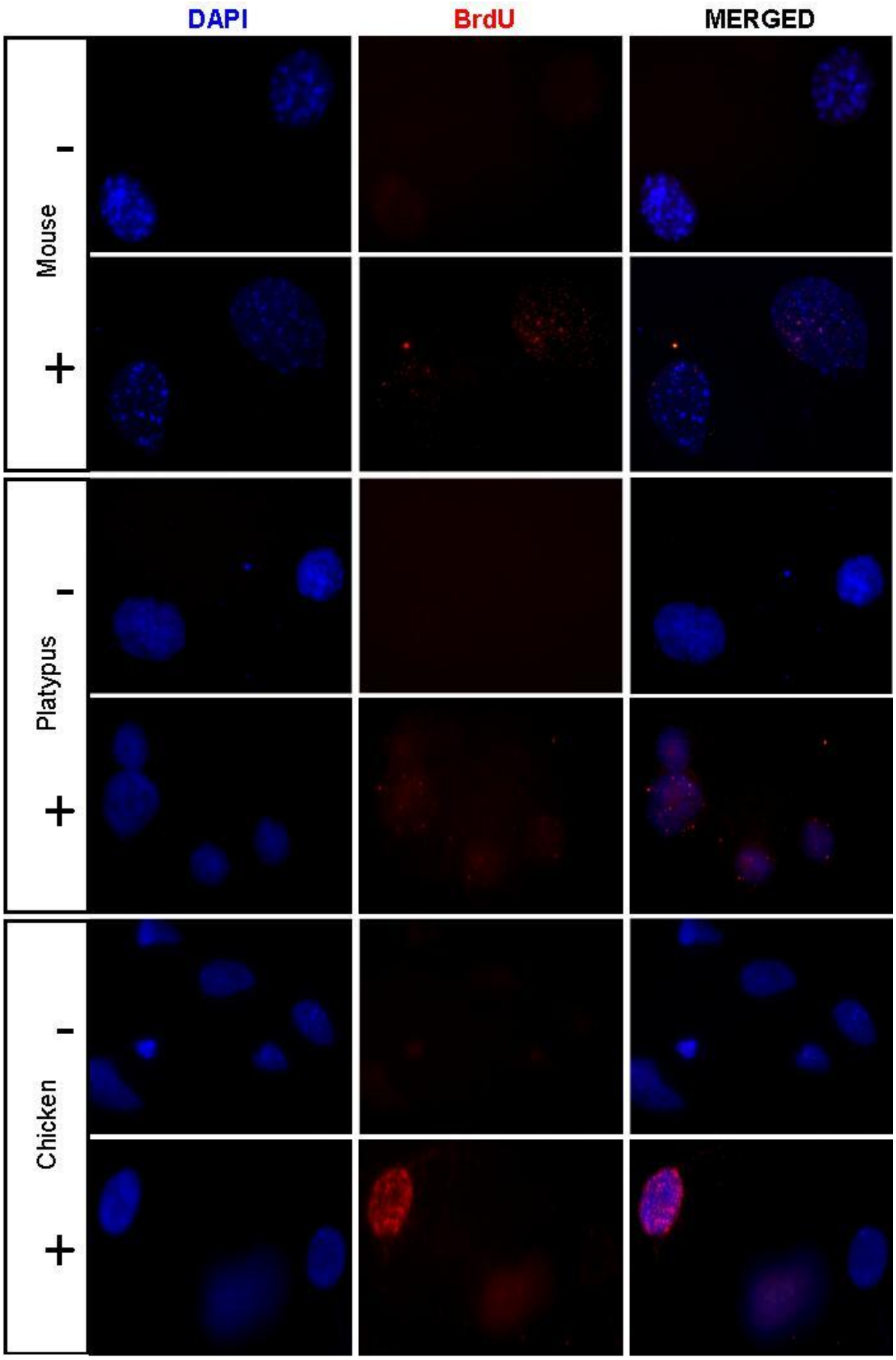


Figure 2: BrdU immunostaining experiments with enhanced blocking

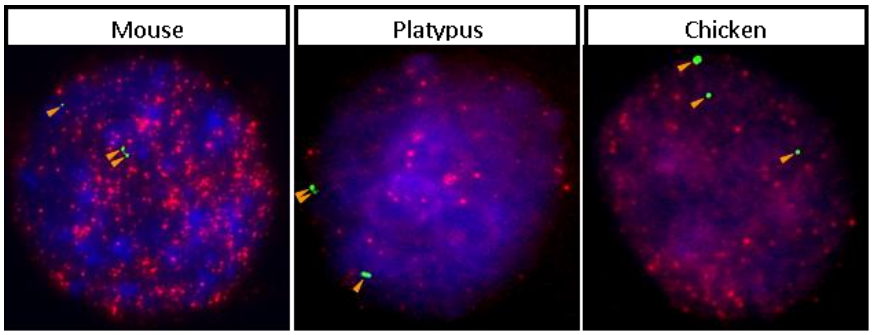


Figure 3: BrdU Immunostaining coupled with directly-labelled FISH

Figure legends

Figure 1: FISH dot assay

In these individual mouse nuclei the *Igf2* locus has a replication status in one of three states; SS: two single dot signals show that locus is unreplicated; SD: a single-double signal pair show that the locus is replicating and the cell is in S-phase (this is counted in the dot assay as asynchronous replication, as one allele has replicated whilst the other has not yet done so); DD: a double-double dot signal showing that the locus has replicated.

Figure 2: BrdU immunostaining experiments with enhanced blocking

Mouse, platypus, and chicken slides after BrdU immunostaining with (+) and without (-) blocking (with blocking solutions 1 and 2) for the primary (anti-BrdU) and secondary (anti-mouse Alexa-568) antibodies.

Figure 3: BrdU Immunostaining coupled with directly-labelled FISH

BrdU immunostaining is shown in red. The FISH hybridisation signals, achieved using a species-specific (mouse, platypus and chicken) *Igf2* BAC (labelled with green fluorescence) in each case, are indicated with the orange arrows.

Tables

Table 1: Optimised steps of the improved directly-labelled FISH + BrdU immunostaining protocol

Steps in Method	Included in our Protocol?
3 Methanol: 1 Acetic Acid fixation of nuclei	Y
Acidic Denaturation	N
Direct-DNA probe labelling with fluorophore	Y
Pepsin treatment/and formaldehyde re-fixation	Y
Slide formamide treatment (70%, 70 ^{o}C for 10 min)	Y
Triton X-100 permeabilisation	Y
Blocking before 1^{o}Ab	Y
Permeabilization with 1^{o}Ab (with 0.1% Triton X-100)	Y
Blocking before 2^{o}Ab	Y
Immunodetection of FISH signals	N

Table 2: BrdU positive nuclei in each cell line, and the average amount of SD% measured for all loci

Species	% BrdU positive cells
Mouse	52 ± 7
Platypus	59 ± 7
Chicken	58 ± 17

Table 3: Total %SD for *Igf2* in each species, and % of SD signals which exist within BrdU stained nuclei

Species	Total % SD signal for Igf2	% nuclei with SD signal for *Igf2* and are BrdU positive
Mouse	39 ± 12	23 ± 4
Platypus	33 ± 4	23 ± 6
Chicken	40 ± 10	22 ± 14

CHAPTER 4: Monoallelic expression predates genomic imprinting

This chapter consists of one publication style manuscript.

Chapter overview

The manuscript (unpublished) presented in this chapter is formatted for submission to the journal *Genome Biology*, and includes research looking at the monoallelic expression, and allele expression ratios, at platypus orthologs of eutherian imprinted genes. The results indicate that these genes may have evolved random monoallelic expression and long-range DNA interactions before becoming subject to parent-of-origin dependent (i.e. genomic imprinting) expression in therian mammals.

MANUSCRIPT:

Monoallelic expression predates imprinting: platypus imprinted orthologs display an amount of monoallelic expression

Statement of Authorship

Title of paper	**Monoallelic expression predates imprinting: platypus imprinted orthologs display an amount of monoallelic expression**
Publication status	Unpublished
Publication details	Wright, M.L., Casey, A., Necsulea, A., Julien, P., Kaessmann, H., and Grutzner, F. 2014. Monoallelic expression predates imprinting: platypus imprinted orthologs display an amount of monoallelic expression. *Manuscript in preparation*. For planned submission in 2015 to *Genome Biology.*

Author Contributions

By signing the Statement of Authorship, each author certifies that their stated contribution to the publication is accurate and that permission is granted for the publication to be included in the candidate's thesis.

Name of Principal Author (Candidate)	Megan Wright
Contribution to the Paper	Identifcation and analysis of clones, cytological analysis. With support from Kaessmann group analysis of RNA-seq data. Performed statistical analyses. Wrote the paper.
Signature	19.12.2014

Name of Co-Author	Anamaria Necsulea
Contribution to the Paper	Analysed original RNA-seq datasets.
Signature	19.12.2014

Name of Co-Author	Aaron Casey
Contribution to the Paper	Assisted with RNA –seq analysis
Signature	20.12.2014

Name of Co-Author	Philippe Julien
Contribution to the Paper	Completed some of the data analysis using the ‘R’ program.
Signature	19.12.2014

Name of Co-Author	Henrik Kaessman
Contribution to the Paper	Supervised the RNA-seq research, and aided in the manuscript construction.
Signature	19.12.2014

Name of Co-Author	Frank Grützner
Contribution to the Paper	Supervised the RNA-FISH and cDNA-seq research, and aided in the manuscript construction/outlining.
Signature	19.12.2014

Manuscript Title: Monoallelic expression predates imprinting: platypus imprinted orthologs display an amount of monoallelic expression

Manuscript prepared for submission to the journal *Genome Biology* (Research Paper)

Full names of authors

Megan Lynne Wright[1], Aaron Casey[1], Anamaria Necsulea[2,3], Philippe Julien[2,3], Henrik Kaessmann[2,3], and Frank Grützner[1]*

*all correspondence to this author.

Affiliations

[1]The Robinson Research Institute, School of Molecular & Biomedical Sciences, The University of Adelaide, Australia

[2]Center for Integrative Genomics, University of Lausanne, Switzerland

[3]Swiss Institute of Bioinformatics, Lausanne, Switzerland

Corresponding author information:

frank.grutzner@adelaide.edu.au

Phone: + 61 8 8331 4812

Fax: + 61 8 8313 4362

The Robinson Research Institute

School of Molecular & Biomedical Sciences

The University of Adelaide

Adelaide

S.A. 5005

AUSTRALIA

Emails for all non-corresponding authors:

Megan Wright: megan.wright@adelaide.edu.au
Aaron Casey: aaron.casey@adelaide.edu.au
Anamaria Necsulea: anamaria.necsulea@epfl.ch
Philippe Julien: philippe.julien@crg.eu
Henrik Kaessmann: Henrik.Kaessmann@unil.ch

Running head: Monoallelic expression platypus

Keywords: Monoallelic, expression, platypus, genomic imprinting, RNA-FISH

Abstract

Parent-of-origin dependent monoallelic expression (genomic imprinting) has so far been shown to exist in marsupials and eutherian mammals but not in monotremes and birds. Surprisingly, replication asynchrony, a signature of monoallelically expressed genes has been observed in monotremes and birds for genes that are imprinted in therian mammals. In this study we first investigated whether an allele expression bias exists at 24 orthologs of eutherian imprinted genes in platypus, and eight control genes, using RNA-seq data. We find that three orthologs of imprinted genes show an allele expression bias, where one allele is expressed in higher amounts than the other, which could indicate the presence of distinct *cis*-regulatory mechanisms at each allele. Using RNA FISH we also find that three of four genes imprinted in other mammals showed evidence of monoallelic expression. These findings raise the possibility that random monoallelic expression may have been the first step towards the evolution of genomic imprinting in mammals.

Introduction

Genomic imprinting is a form of monoallelic expression conferring parent of origin dependent monoallelic expression of an autosomal gene, and is controlled by parent-of-origin methylation of CpGs inherited at an allele. Imprinting is believed to have evolved as a result of an evolutionary conflict between parent and offspring genes involved in controlling nutrient allocation, due to the fact that many imprinted genes are involved in interuterine foetal growth and nourishment, as well as behaviour [1]. In earlier diverged amniote clades, such as egg-laying monotremes and birds, there is no evidence of genomic imprinting, indicating that it evolved sometime after prototherian divergence from the mammalian common ancestor [reviewed in both 2,

3]. Interestingly, whilst all of the aforementioned monoallelic expression control processes are achieved in slightly different ways, there are some shared characteristics among monoallelically expressed loci, including common chromatin marks and asynchronous replication [4, reviewed in 5]. Furthermore, many monoallelically expressed loci also undergo long-range interactions as part of their expression control mechanism, indicating that DNA interactions may have also played a role in an ancestral form of monoallelic expression [reviewed in 6].

Random monoallelic expression occurs when one allele at an autosomal locus is epigenetically silenced, causing transcription products from the other 'active' allele to be present, and is dissimilar to imprinting in the way that expression is not determined in a parentally-inherited manner. Monoallelic expression is controlled in many ways in the therian genome, from chromosome-wide allelic shutdown processes, such as X-chromosome inactivation and the regional control processes observed in imprinted gene clusters, to single-locus allelic inactivation processes for autosomal monoallelically expressed genes, genes under solitary imprinted control, and genes which undergo allelic exclusion [reviewed in 5]. Whilst most of the genes on the eutherian X-chromosome are monoallelically expressed (with only a few escaping), 10-24% of autosomal eutherian genes display random monoallelic expression and less than 1% display imprinted expression [7, 8, 9, 10, 11, 12].

Monoallelic expression is an extreme effect of *cis*-regulated differential allele expression, and thus imprinted genes, with their parentally inherited monoallelic expression pattern, are an interesting example of 'tight' *cis*-regulatory control. This is not to say that all monoallellic expression is *cis*-regulated, as there are known examples of monoallelic expression regulated in *trans* [reviewed in 13], however, the

cis-regulation of monoallelic expression becomes particularly pertinent when studying imprinting in therian mammals, whereby one imprinting control region (ICR) can enact positive or negative control over expression of many alleles within an imprinted cluster. In therian mammals, differential CpG methylation, specifically at Imprinted Control Regions (ICRs), occurs at the two alleles, and the allele with the methylated ICR will be shut down. Thus, the installed differential imprinted marks allow control *in cis* of the protein-DNA interactions that occur at each allele, and are then maintained (according to requirement) in a tissue-specific manner. This can result in differential long-range DNA interactions occurring at each allele, where the two alleles will interact with different regions of the genome, as observed with genes such as the eutherian imprinted, *Igf2* [14, 15].

Previous studies investigating the expression of monotreme orthologs of eutherian-imprinted genes have not identified any form of monoallelic expression [16, 17]. However, the observation that asynchronous replication, often seen as a hallmark of monoallelic expression, occurs at eutherian-imprinted orthologs of platypus, raises the question as to why this characteristic would still be conserved at these loci in monotremes? Notably, studies have found that no differential methylation patterns exist at eutherian-imprinted orthologs in monotremes, ruling out this well-understood form of *cis*-regulatory control [17]. However, this does not rule out other forms of *cis*-regulatory control, such as the presence of differential chromatin states, or differential chromatin interactions [18]. Furthermore, asynchronous replication at imprinted loci is reportedly consistent with the organisation of alleles into separate subnuclear compartments, and independent of differential DNA methylation [19]. These studies indicate that differential chromatin environments and long-range DNA interactions occurring in *cis* could also be responsible for the conservation of asynchronous

replication at orthologous regions in the platypus. However, the presence of bivalent histone marks and other long-range DNA interactions, remain to be seen in the platypus genome.

Differential allele expression has been analysed in order to determine any *cis*-acting affects on expression, with differences in the amount of expression between alleles of autosomal genes being observed in different species [12, 20, 21]. Allele expression bias, where one allele has a measurably higher expression than the other allele, is observed in 20-50% of human genes [20, 22], however, this characteristic has not been studied extensively outside of eutherians.

In this study we firstly sought to observe allele expression biases at eutherian-imprinted platypus orthologs (EIPOs) that could be characteristic of *cis*-regulatory processes. We find that although three EIPOs do show some allele expression bias, indicative of *cis*-regulatory effects, the majority of EIPOs do not. We then investigated whether any monoallelic expression occurred at EIPOs, and find that all the tested EIPOs showed an amount of monoallelic expression when measured cytogenetically. Finally, we report preliminary evidence of a conserved long-range DNA interaction in platypus and mouse, indicative that such interactions may be conserved in all mammalian lineages, including monotremes.

Results

Imprinted genes are expressed from both alleles in platypus

Firstly, we analysed RNA-seq data from 8 tissues (brain, cerebellum, fibroblasts, heart, kidney, liver, ovary, testes) from five adult platypus animals (four males and one female) [23], to look for evidence of biased Allele Expression Ratios (AERs) at EIPOs, comparing them to control genes that are classified as non-imprinted in eutherians. We were able to observe biallelic expression for 21 platypus orthologs of genes imprinted in therian mammals (see **Table 1 and Supplementary Table 1**), in different tissues (**Supplementary Table 2**). Notably, we found that the spread of AERs observed between the imprinted orthologs and non-imprinted orthologs were significantly different (p-value = 0.031) (**Figure 1**), with the non-imprinted orthologs showing higher allele expression ratios than the imprinted orthologs, indicating a larger expression bias between the alleles of non-imprinted genes. Furthermore, we measured how many Single Nucleotide Polymorphisms (SNPs) showed an expression bias towards one allele in all tissues (**Table 1**). It has previously been stipulated in human studies, that any observed allele expression ratio that is lower than a 60:40 ratio (represented by the value 0.6 in our AER data) is due to experimental noise, thus only genes that show more that 60% of their expression from one allele can be considered to be undergoing cis-regulatory effects [22]. We found that two imprinted orthologs (*Mcts2* and *Zim1*) and two non-imprinted (*Actb* and *Wsb1*) orthologs possessed an allele expression bias under these criteria, with one allele being expressed in consistently higher amounts than the other.

Next, the allele expression ratios were measured for a subset of genes by cDNA sequencing and measuring the relative signal strength of each nucleotide at an SNP

on a sequencing chromatogram (**Supplementary Figure 1**), as performed in previous studies [16]. We chose to specifically test highly expressed SNPs that were detected in the RNA-seq (**Table 2**). These ratios were similar to the averages observed in RNA-seq for all genes except *Diras3*, which showed a ratio 0.79, indicative of a marked expression bias in kidney tissue.

Monoallelic expression of EIPOs

Next we wanted to investigate, at the single cell level, whether both alleles are expressed at EIPOs. Using RNA-FISH we investigated the expression of *Copg2*, *Igf2*, *Igf2r*, *Ube3a* (imprinted in eutheria), and *Wsb1* (not imprinted) in platypus fibroblast cells (**Table 3** and **Figure 2A**). Overall we found that *Wsb1* showed expression in all nuclei, while the EIPOs were expressed in lower amounts of nuceli (**Table 3**).

A combination of monoallelic and biallelic expression was observed for the EIPOs *Copg2*, *Igf2*, and *Ube3a*, which all showed between 65-84% monoallelic expression, and 16-35% biallelic expression (**Figure 2B**). The *Igf2r* gene showed monoallelic expression in all cells, however, it was also expressed in very low amounts (**Figure 2B**). Interestingly, the amount of monoallelic expression observed at the EIPOs (65-100%) was at least close to double (or higher) the amount of biallelic expression (0-35%) observed. Conversely, at the *Wsb1* non-imprinted ortholog, the expression was overwhelmingly biallelic (98%) (**Figure 2B**). T-tests confirmed that the amount of monoallelic expression observed at the EIPOs is significantly different to the amount observed at *Wsb1*.

Long-range interaction Igf2 and the Wsb1 are conserved in amniotes

An interchromosomal interaction observed between the mouse genomic regions containing the imprinted *Igf2* and *Wsb1* loci was previously measured through DNA FISH to occur at a frequency of around 30-42% depending on the cell line [14]. We used DNA FISH *Igf2* and *Wsb1* BAC clones to measure whether the interaction between these genes occurs in platypus (**Figure 3A**). As a control we also measured the same interaction in mouse cultured skin fibroblast, so that we could compare our results to previous studies [14]. We measured around 19% of the *Igf2-Wsb1* interaction in mouse, which was significantly different to the amount of interaction measured between the non-interaction control loci (in mouse these loci contained the genes *Mest* and *B3galt1*), but also lower than what was observed in previous literature (30-42%) [14] . We found that the *Igf2-Wsb1* interaction was conserved in platypus with 8% of the interaction occurring in platypus skin fibroblasts, and that this interaction was occurring in significantly higher amounts than the 0.5% average interaction occurring between non-interacting control loci (in platypus these loci were *Mest* and *Sox1*) (**Figure 3**).

Discussion

Differential allele expression is an indicator of *cis*-acting affects on expression of autosomal genes [12, 20, 21]. For the EIPOs tested, the lack of consensus for an AER bias indicates that allele expression at EIPOs occurs mainly from both alleles. Only three EIPOs, *Mcts2*, *Zim1*, and *Diras3*, as well as two non-imprinted orthologs, *Actb* and *Wsb1*, show ratios that indicate a significant allele expression bias in the RNA-seq and cDNA-seq data analysis (**Table 1** and **2**). Thus, in light of the RNA FISH evidence for partial monoallelic expression at EIPOs, these ratios suggest that

monoallelic expression may not be exclusive to one allele, and instead hint at mutually exclusive monoallelic expression from both alleles. Notably, previous studies on monoallelic expression and biased allele expression in humans has found that many monoallelically expressed genes are preferentially expressed from one allele, with our results suggesting that this could be a conserved characteristic across mammals [12].

Our observation of partial monoallelic expression of the platypus *Igf2*, *Copg2*, *Igf2r* and *Ube3a* genes, provides the first evidence that monoallelic expression at these loci pre-dates therian imprinting, with results from the transcriptome data suggesting that both alleles are expressed in these cells. A recent study on random monallelic expression in humans found that it varied across different cell-types, with some cell-types showing strict monoallelic expression for a specific locus, whilst other cell types showed biallelic expression for the same locus [12]. It may be that such a random monoallelic expression pattern facilitates the evolution of parent-of-origin dependent expression in therian mammals; with the mixture of monoallelic and biallelic expression eventually shifting into parentally-inherited imprinted expression. Perhaps more importantly, the presence of partial monoallelic expression at EIPOs could suggest that gene dosage may have also remained more similar at these genes across mammalian evolution than previously thought.

The conservation of the long-range interaction between the *Igf2-Wsb1* genomic regions, from platypus to mouse (**Figure 4**), shows that such interactions may be conserved across mammalian clades. Previous research has shown that DNA-interactions are conserved across eutherian species, with mouse and human interactome maps showing conservation of interaction frequency [24]. Notably, we

found that the *Igf2-Wsb1* DNA interaction occured in a higher proportion of cells in mouse compared to platypus (**Figure 3**), indicating the interaction may have increased in frequency during mammalian evolution. The conservation of interactions at imprinted genes and their orthologs aids in explaining the conservation of other epigenetic characteristics, such as replication asynchrony [19], and also adds to evidence showing the conservation of epigenetic regulation over allele specific expression [25]. Future work will need to investigate the interplay of long-range DNA interactions in the evolution of imprinted loci and their orthologs, particularly in terms of asynchronous replication and monoallelic expression.

Conclusion

Here we show, for the first time, data that suggests specific genes were monoallelically expressed before they became imprinted. Previous work has suggested that imprinting evolved after the divergence of monotreme mammals. However, asynchronous replication, a hallmark of monoallelically expressed genes, is conserved in all mammalian lineages as well as birds, which display no form genomic imprinting. In the present work we show that another vital aspect of *Igf2* monoallelic activity, the interchromosomal interaction with the *Wsb1* genomic region, is also conserved in monotremes. We have addressed the question of monoallelic expression using RNA-seq and RNA-FISH and found that while many cells express only one allele transcript, both alleles are present in the transcriptome. This suggests that random monoallelic expression and other associated characteristics, including replication timing and chromosomal interaction, were established before the evolution of genomic imprinting in therian mammals.

Methods

Tissue and RNA preparation

Sample treatment and RNA sequencing were performed as previously described, with total RNA being extracted with Trizol (Invitrogen) or RNAeasy/RNAeasy Lipid/miRNeasy (Qiagen) column purification kits [26]. The sequencing and read-alignment were performed using previously established protocols [26].

RNA-seq

Mapped reads, aligning to annotated Ensembl genes, were then selected if they showed a heterozygous SNP, with at least two alleles at the nucleotide. Due to a large amount of SNPs with more than 2 alleles in the data, SNPs with more than 2 alleles (ie. more than 2 nucleotides present at the single SNP) were discarded if they did not meet this requirement: where the sum of the read counts of the additional alleles (whether there was one or two of them, meaning that total alleles present would be 3 and 4 respectively) did not add up to more than 30% of the read count of the SNP with the second highest read-count. If the SNP did pass this requirement, but had more than two alleles present, the counts for the additional alleles were discarded, and only the counts of the two alleles of highest count were used to calculate the AER. AERs were calculated by dividing the allele with the highest read count, by the total read count observed at the SNP (for both alleles). Thus to minimise the effects of read-errors on observed AERs, (caused by one of the two alleles mapping to a single SNP with a very low read-count that was more likely due to read-errors during sequencing) we discarded AERs if they had values lying between 0.90 and 1.0.

RT-PCR and sequencing

RNA was extracted at a concentration of 250-1000 ng/ul from platypus kidney samples, using TRIzol (Invitrogen) following the manufacturer's instructions, then treated with DNase I (New England Biolabs). The RNA was then reverse transcribed using a SuperScript III reverse transcriptase kit (Invitrogen) according to the manufacturer's instructions. Primers (see **Supplementary table 3**) were then designed to span an intron in exonic sequence for each gene (except *Diras3* which contains no intron), and the cDNA expression was tested using the following PCR cycles: 4 min at 95^{o}C; 35 cycles of: 45 sec at 95^{o}C, 30 sec at 55^{o}C, 2 min 20 sec at 72^{o}C; followed by 7 mins at 72 ^{o}C, before being stored at 4^{o}C. The PCR products (and SNPs) were confirmed by sequencing, using the Big Dye Terminator version 3.1 cycle sequencing kit (Applied Biosystems). Allele expression ratios were analysed via PCR using previously established methods [16], with peak values measured for each nucleotide at the SNP, and the allele expression ratio found by dividing the larger peak value by the total of both values.

BAC DNA extraction, characterisation, and labelling

Platypus BAC clones (**Supplementary table 1**) were purchased from the Clemson University Genomics Institute (CUGI), and grown following the company's instructions. The Mouse BAC clones, which correspond to the *Igf2* (RP23-51J21), *Wsb1* (RP24-256H2), *Mest* (RP24-211G11) and *B3galt1* (RP23-346P15) genomic regions, were purchased from Children's Hospital Oakland Research Institute (CHORI, Oakland, CA, USA) and grown following the company's instructions. BAC clones were extracted using The Wizard Minipreps DNA Purification System (Promega). BACs were characterised as described previously [27 *manuscript in preparation*], and the platypus *Ube3a* BAC PCR was run with primers: For 5'-

ACCCTCCCTCTTCATATTGAC-3' and Rev 5'-CGTACTAGATGCTTGGGAGAG-3'. For a 25 ul labelling reaction: 3 µl BAC DNA (500 ng total), 5 µl of 1 µg/µl random 9-mer primer (Geneworks), and 12.5 µl of water were incubated for 15 min at 95^{o}C, and cooled on ice for 5 min, before adding 4.6 ml of 5×dNTP supplemented NEB buffer 2 (supplemented with a 200 µM dNTP solution containing: 100 µM dACG-TPs and 40 µM dTTPs), 0.4 µl SpectrumGreen/Orange fluorescent dUTPs (Abbott Molecular), and 0.5 µl Exo-Klenow polymerase (MO221, New England Biolabs). The reaction was then incubated at 37^{o}C overnight.

Tissue Culture

Skin-derived platypus and mouse fibroblasts, were cultured in Advanced DMEM (12491-023, Life Technologies), and supplemented with 10% Fetal Calf Serum and 1% L-glutamine. Fibroblasts were cultured to around passages 3-6, before they were removed from the flasks and grown directly on glass slides for 3-5 hours (until most cells had attached to the glass), directly followed by the RNA FISH experiment.

RNA FISH

RNA FISH protocol is a modified version of a previous protocol [28]. All solutions were prepared with DEPC-treated and autoclaved Milli-Q water. Slide preparation: Slides were washed in 1xPBS for 5 minutes, before being permeabilised on ice for 10 min in freshly prepared CSK buffer (100 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$, 10 mM PIPES pH 6.8)/ 0.5% Triton X 100/ 2 mM Vanadyl Ribonucleoside complex (VRC). Slides were then fixed in freshly prepared 3% paraformaldehye/1xPBS at room-temperature for 10 min, washed twice in 70% ethanol, then dehydrated in an ethanol series (80%, 95% and 100%) for 3 min each. The slides were then allowed to dry for 10 min before the BAC probes were hybridized.

Probe preparation: 10 ul of labelled 20 ng/µl BAC DNA (total of 200 ng of labelled DNA used for hybridisation of half a slide) was added to 1 µg Cot-1 suppressor, 10 ug of Salmon Sperm, and 200 µl ice-cold 100% ethanol. The mixture precipitated for 20 min at -80^{o}C, centrifuged at 14,000 rpm for 30 min, before the ethanol was removed and the pellet left to dry for 10 min at 37^{o}C. The probe was then denatured for 10 min at 80^{o}C, placed on ice for 5 min, before 5 µl of deionized formamide was added and the mixture shaken for 15 min at 37^{o}C, then 5 µl of hybridization buffer (4 x SSC, 40% dextran sulphate, 2mg/ml BSA, 10 mM Vanadyl Ribonucleoside Complex) was added and the mixture was once again shaken for 15 min at 37^{o}C. The probe was pre-annealed for 20 min at 37^{o}C, before being deposited onto the prepared slide, sealed with rubber cement, and hybridised overnight 37^{o}C in a humid chamber.

Detection: Slides were washed 2 x 5 min in 50% formamide/2xSSC (adjusted to pH 7.2) at 42^{o}C, then washed 2 x 5 min in 3xSSC at 42^{o}C. The slides were then counterstained with 1 µg/ml DAPI/2×SSC for 1 min, rinsed twice in 2xSSC water, then air dried for 30 min and mounted with Vectashield (Vector Laboratories).

DNA FISH

Slide preparation: Slides were washed 3 x 5 min in 2xSSC (until all mounting Vectashield was removed), then treated with 100 µg/ml RNase/2xSSC at 37^{o}C for 30 min, then washed again 3 x 5 min in 2xSSC. Slides were treated with 0.01% pepsin/10 mM HCl at 37^{o}C for 10 min, washed 2 x 5 min in 1xPBS, and equilibrated in 50 mM MgCl2/1xPBS for 5 min, before being fixed in 1.3% Formaldehyde/1xPBS/50 mM MgCl2 for 10 min. The slides were then dehydrated for 3 min each in an ethanol series (70%, 90% and 100%) and denatured for 2.5 min in 70% (v/v) formamide/2×SSC at 70^{o}C. The slides were once again dehydrated in an

ethanol series (as before), air-dried for 30 min, before 10 µl labelled BAC/formamide/hybridization mix was deposited on each half-slide, covered with a coverslip (18 mm x 18 mm area) and sealed with rubber cement, before being hybridised overnight in a water moist chamber at 37^{o}C.

Probe preparation: 25 ul of fluorescently labelled 20 ng/µl BAC DNA in DNA FISH following RNA FISH (total of 500 ng of labelled DNA used for hybridisation of half a slide) was added to 10 µg platypus competitor DNA (300-500 bp fragments), 50 ug of Salmon Sperm, and 500 µl ice-cold 100% ethanol. For probe preparation of two co-precipitated BACs for DNA FISH 25 ul of each fluorescently labelled 20 ng/µl BAC DNA (total of 1000 ng of labelled DNA used for hybridisation of half a slide) was added to 15 µg competitor DNA (300-500 bp fragments), 50 ug of Salmon Sperm, and 500 µl ice-cold 100% ethanol. Probe preparation was then carried out following the same steps as the RNA FISH probe preparation.

Detection: Slides were washed 3 x 5 min in 50% Formamide/2xSSC at 42^{o}C, washed in 2xSSC at 42^{o}C for 5 min, 0.1xSSC at 60^{o}C for 5 min, and again in 2xSSC at 42^{o}C for 5 min. Slides were counterstained with 1 µg/ml DAPI/2×SSC for 1 min, rinsed twice in MQ water, then air dried for 1 hour and mounted with Vectashield (Vector Laboratories).

Microscopy and Images

Slides were visualised with a Zeiss AxioImagerZ.1 epifluorescence microscope equipped with a CCD camera, and images acquired using Zeiss Axiovision software.

List of abbreviations

EIPO - Eutherian-imprinted platypus ortholog

AER - Allele expression ratio

ICR - Imprinting control region

Competing interests

The author(s) declare that they have no competing interests

Funding

Megan Wright and Aaron Casey were funded through Australian Postgraduate Awards (bestowed by the Australian Government). Anamaria Necsulea was supported by a long-term FEBS postdoctoral fellowship. Henrik Kaessman and Philippe Julien were supported by grants from the European Research Council (Starting Independent Grant) and the Swiss National Science Foundation. Frank Grützner is an Australian Research Council (ARC) research fellow.

Acknowledgements

The authors would like to acknowledge Tasman Daish for his help and guidance with the RNA FISH experiments.

References

1. Moore T, Haig D: **Genomic imprinting in mammalian development: a parental tug-of-war.** *Trends in Genetics* 1991, **7:**45-49.
2. Renfree MB, Papenfuss AT, Shaw G, Pask AJ: **Eggs, embryos and the evolution of imprinting: insights from the platypus genome.** *Reproduction, Fertility and Development* 2009, **21:**935-942.
3. Renfree MB, Hore TA, Shaw G, Graves JA, Pask AJ: **Evolution of genomic imprinting: insights from marsupials and monotremes.** *Annu Rev Genomics Hum Genet* 2009, **10:**241-262.
4. Rougeulle C, Navarro P, Avner P: **Promoter-restricted H3 Lys 4 di-methylation is an epigenetic mark for monoallelic expression.** *Hum Mol Genet* 2003, **12:**3343-3348.
5. Goldmit M, Bergman Y: **Monoallelic gene expression: a repertoire of recurrent themes.** *Immunol Rev* 2004, **200:**197-214.
6. Ling JQ, Hoffman AR: **Epigenetics of Long-Range Chromatin Interactions.** *Pediatric Research* 2007, **61:**11R-16R 10.1203/pdr.1200b1013e31804575db.
7. Gimelbrant A, Hutchinson JN, Thompson BR, Chess A: **Widespread monoallelic expression on human autosomes.** *Science* 2007, **318:**1136-1140.
8. Deng Q, Ramskold D, Reinius B, Sandberg R: **Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells.** *Science* 2014, **343:**193-196.
9. Filippova GN, Cheng MK, Moore JM, Truong JP, Hu YJ, Nguyen DK, Tsuchiya KD, Disteche CM: **Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development.** *Dev Cell* 2005, **8:**31-42.

10. DeVeale B, van der Kooy D, Babak T: **Critical Evaluation of Imprinted Gene Expression by RNA-Seq: A New Perspective.** *PLoS Genet* 2012, **8:**e1002600.

11. Savova V, Vigneau S, Gimelbrant AA: **Autosomal monoallelic expression: genetics of epigenetic diversity?** *Curr Opin Genet Dev* 2013, **23:**642-648.

12. Gendrel AV, Attia M, Chen CJ, Diabangouaya P, Servant N, Barillot E, Heard E: **Developmental dynamics and disease potential of random monoallelic gene expression.** *Dev Cell* 2014, **28:**366-380.

13. Zakharova IS, Shevchenko AI, Zakian SM: **Monoallelic gene expression in mammals.** *Chromosoma* 2009, **118:**279-290.

14. Ling JQ, Li T, Hu JF, Vu TH, Chen HL, Qiu XW, Cherry AM, Hoffman AR: **CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1.** *Science* 2006, **312:**269-272.

15. Kurukuti S, Tiwari VK, Tavoosidana G, Pugacheva E, Murrell A, Zhao Z, Lobanenkov V, Reik W, Ohlsson R: **CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2.** *Proc Natl Acad Sci U S A* 2006, **103:**10684-10689.

16. Killian JK, Nolan CM, Stewart N, Munday BL, Andersen NA, Nicol S, Jirtle RL: **Monotreme IGF2 expression and ancestral origin of genomic imprinting.** *J Exp Zool* 2001, **291:**205-212.

17. Weidman JR, Murphy SK, Nolan CM, Dietrich FS, Jirtle RL: **Phylogenetic footprint analysis of IGF2 in extant mammals.** *Genome Res* 2004, **14:**1726-1732.

18. McEwen KR, Ferguson-Smith AC: **Distinguishing epigenetic marks of developmental and imprinting regulation.** *Epigenetics Chromatin* 2010, **3:**2.

19. 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:**759-773.

20. Lo HS, Wang Z, Hu Y, Yang HH, Gere S, Buetow KH, Lee MP: **Allelic variation in gene expression is common in the human genome.** *Genome Res* 2003, **13:**1855-1862.

21. Goncalves A, Leigh-Brown S, Thybert D, Stefflova K, Turro E, Flicek P, Brazma A, Odom DT, Marioni JC: **Extensive compensatory cis-trans regulation in the evolution of mouse gene expression.** *Genome Res* 2012, **22:**2376-2384.

22. Serre D, Gurd S, Ge B, Sladek R, Sinnett D, Harmsen E, Bibikova M, Chudin E, Barker DL, Dickinson T, et al: **Differential allelic expression in the human genome: a robust approach to identify genetic and epigenetic cis-acting mechanisms regulating gene expression.** *PLoS Genet* 2008, **4:**e1000006.

23. Julien P, Brawand D, Soumillon M, Necsulea A, Liechti A, Schutz F, Daish T, Grutzner F, Kaessmann H: **Mechanisms and evolutionary patterns of mammalian and avian dosage compensation.** *PLoS Biol* 2012, **10:**e1001328.

24. Ryba T, Hiratani I, Lu J, Itoh M, Kulik M, Zhang J, Schulz TC, Robins AJ, Dalton S, Gilbert DM: **Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types.** *Genome Res* 2010, **20:**761-770.

25. Zwemer LM, Zak A, Thompson BR, Kirby A, Daly MJ, Chess A, Gimelbrant AA: **Autosomal monoallelic expression in the mouse.** *Genome Biol* 2012, **13:**R10.

26. Brawand D, Soumillon M, Necsulea A, Julien P, Csardi G, Harrigan P, Weier M, Liechti A, Aximu-Petri A, Kircher M, et al: **The evolution of gene expression levels in mammalian organs.** *Nature* 2011, **478:**343-348.

27. Wright ML, Zechner U, Haaf T, Grutzner F: **Conservation of replication asynchrony and S-phase entry predates evolution of genomic imprinting in mammals.** *manuscript in preparation*.

28. Deakin JE, Hore TA, Koina E, Marshall Graves JA: **The status of dosage compensation in the multiple X chromosomes of the platypus.** *PLoS Genet* 2008, **4:**e1000140.

29. Morison IM, Ramsay JP, Spencer HG: **A census of mammalian imprinting.** *Trends Genet* 2005, **21:**457-465.

Figure legends

Figure 1: Allele expression ratio range for all platypus orthologs measured from RNA-seq data

Ratios were determined for the platypus orthologs of loci that are imprinted and non-imprinted in eutherians. This graph encompasses all of the data for every imprinted ortholog SNP tested, and every non-imprinted ortholog SNP tested. *p-value <0.05.

Figure 2: Allele expression status measured by RNA FISH in platypus fibroblast cells

A. RNA (top row) and DNA (bottom row) FISH experiments were carried out on the same nucleus to observe the amount of monoallelic or biallelic expression for each locus. The RNA (green) and DNA (red) FISH signals are indicated by yellow arrows. **B.** A graph of the relative amount of mono vs biallelic expression for each locus tested. T-tests results, testing for a significant difference observed between the amount of monoallelic expression observed at an EIPOs compared to *Wsb1*, are indicated by asterisk: *p-value <0.05; **p-value<0.01.

Figure 3: *Igf2-Wsb1* long-range interaction measured in platypus by DNA FISH

A. An example of a long-range DNA interaction between the platypus *Igf2* (green) and *Wsb1* (red) loci as measured by DNA FISH. Allele interaction was counted if two of the DNA FISH signals were observed to reside in the same plane, and touch (as indicated by the yellow arrow). **B.** An example of a non-interaction, as no FISH signal overlap is observed in this case for the platypus *Igf2* (green) and *Wsb1* (red) loci. **C.** A graph of the percentage of *Igf2-Wsb1* long-range interaction, and the amount of interaction in a control experiment (between two genes that are not expected to undergo any required DNA interaction), measured in platypus and mouse skin cultured fibroblasts (N = 50). In platypus the non-interacting control genes were *Mest* and *Sox1*, and in mouse, *Mest* and *B3galt1*. *P-value <0.05; **P-value<0.01.

Figures

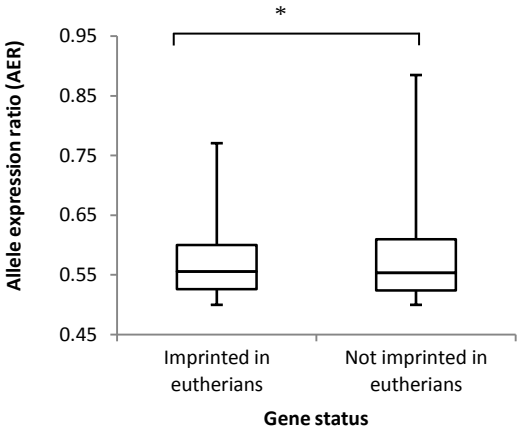


Figure 1: Allele expression ratio range for all platypus orthologs measured from RNA-seq data

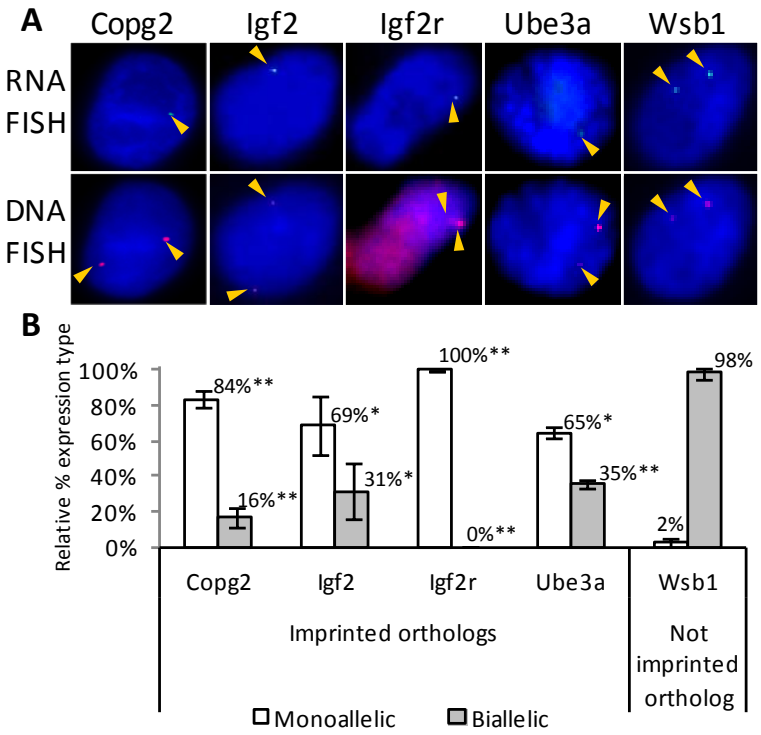


Figure 2: Allele expression status measured by RNA FISH in platypus fibroblast cells

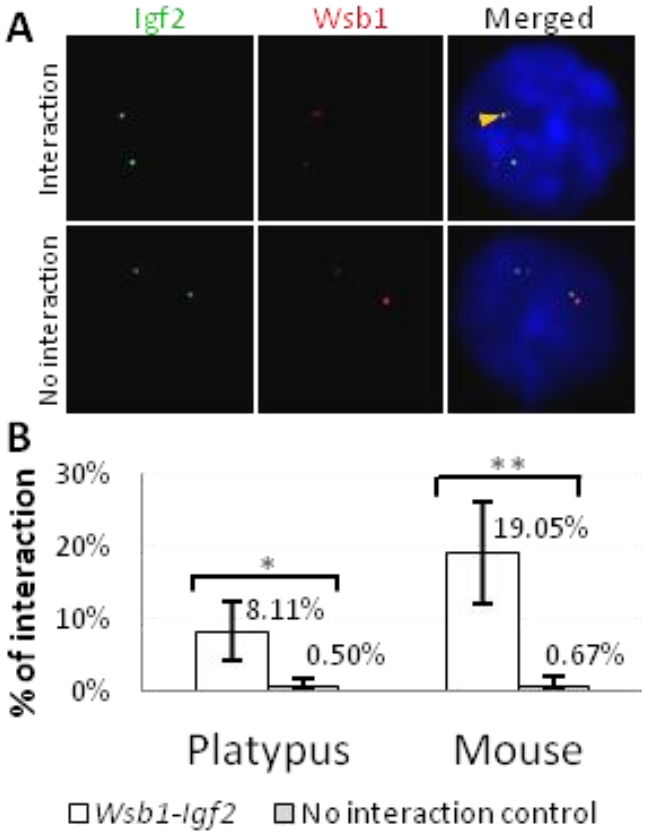


Figure 3: *Igf2-Wsb1* long-range interaction measured in platypus by DNA FISH

Tables

Table 1: Allele expression ratios in platypus for genes that have at least one detectable SNP

Status	Platypus orthologs	No. of unique SNPs tested	Avg. no. of reads/SNP	SNP no. with allele expression bias one way only in all tissues tested	Avg. allele ratio
Imprinted in eutherians	*Atp10a*	2	28.5	NA	0.55
	Commd1	2	27.5	0	0.56
	Copg2	10	77.6	1	0.56
	Dcn	1	24	NA	0.50
	Dhcr7	5	50.9	1	0.57
	Diras3	1	89.4	0	0.57
	Dlgap2	3	38.3	NA	0.54
	Gabra5	1	43	NA	0.51
	Gatm	2	103.4	0	0.56
	Igf2r	7	42.2	1	0.55
	Impact	1	31	NA	0.52
	Kcnk9	2	18	NA	0.60
	Lrrtm1	1	36	NA	0.53
	Magel2	3	48.5	0	0.60
	Mcts2	1	48	NA	0.77
	Phlda2	1	18	NA	0.56
	Plagl1	2	17.5	NA	0.54
	Sgce	2	55.9	0	0.56
	Slc38a4	1	240	NA	0.54
	Zim1	2	30	NA	0.64
	Zrsr1	2	105.5	1	0.60
AVERAGE IMPRINTED IN EUTHERIANS					**0.57**
Not imprinted in eutherians	*Actb*	3	1587.2	2	0.62
	B2m	9	326.7	0	0.59
	Gapdh	3	643.8	0	0.58
	Hprt	1	25	NA	0.56
	Tpb	1	26.5	0	0.58
	Wsb1	2	85	1	0.62
	Yhaz	2	1584.3	0	0.58
AVERAGE NON-IMPRINTED IN EUTHERIANS					**0.59**

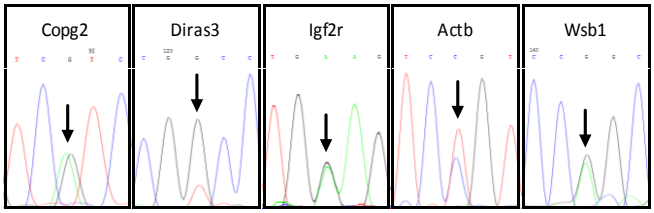
Table 2: Gene expression ratios as measured by cDNA-seq chromatogram peak intensity (PCR sequencing)

Gene status	Gene	SNP position	Published base	cDNA seq		RNA seq	
				Highest peak base	Allele ratio	Base with more reads	Avg. allele ratio
Imprinted in eutherians	*Copg2*	G/A Chr10: 5020040	A	A	0.50	A	0.56
	Diras3	G/T Chr18: 897932	T	G	0.79	T	0.57
	Igf2r	G/A Ultra61: 99114	G	G	0.52	A	0.55
Not imprinted in eutherians	*Actb*	T/C Chr2: 24871651	C	T	0.61	T	0.62
	Wsb1	G/A Ultra17: 703706	A	G	0.54	A	0.62

Table 3: Number of cells with monoallelic expression in platypus fibroblast line

Gene status	Gene	% cells with one RNA-FISH signal	N
Imprinted in eutherians	*Copg2*	35%	48
	Igf2	24%	50
	Igf2r	4%	50
	Ube3a	30%	47
Not imprinted in eutherians	*Wsb1*	100%	51

Supplemental Material



Supplementary Figure 1: Chromatograms of SNPs detected in cDNA-seq

Genes with SNPs tested in cDNA-seq are labelled accordingly. Black arrows point to the chromatogram signal that corresponds to the SNP. AERs were measured from the cDNA-seq chromatograms, where relative nucleotide intensities at the SNP were measured to determine the expression level of each allele.

Supplementary table 1: Information about genes used in this study

Gene	Ensembl/EMBL annotation	Protein Name or Description	Platypus chromosome/contig region	Platypus BAC	Imprinted allele in eutherians
Atp10a	ENSOANG00000009526	ATPase, Class V	Ultra292, 95.2 Kb	-	P[HSA]
Commd1	ENSOANG00000012420	Copper metabolism gene Murr1	Ultra56, 4.47 Mb	-	P[MUS]
Copg2	ENSOANG00000011471	Coatomer protein complex subunit	Chr.10p, 5.01 Mb	Oo_Bb-59D5	P[HSA], M[MUS]
Dcn	ENSOANG00000006630	Proteoglycan	Ultra443, 10.3 Mb	-	P[MUS]
Dhcr7	ENSOANG00000011295	7-dehydrocholesterol reductase	Chr.3p, 7.06 Mb	-	P[MUS]
Diras3	ENSOANG00000007765	Ras homolog	Chr.18q, 897 Kb	-	M[OV]
Dlgap2	ENSOANG00000014126	Membrane associated guanylate kinase	Chr.X1p, 9.06 Mb	-	M[HSA]
Gabra5	ENSOANG00000009524	Gamma-aminobutyric acid receptor	Ultra292, 416 Kb	-	M[HSA]
Gatm	ENSOANG00000006872	Glycine amidinotransferase	Chr.2p, 1.67 Mb	-	P[MUS]
Igf2	AF225876.2	Insulin-like growth factor 2	Contig27935, 2.5 Kb	Oa_Bb-349H20	P[HSA & MUS]
Igf2r	ENSOANG00000007296	Insulin-like growth factor receptor 2	Ultra61, 137 Kb	Oo_Bb-300J22	P[HSA & MUS]
Impact	ENSOANG00000006577	Imprinted and ancient	Contig20839, 9.66 Kb	-	M[MUS & RAT]
Kcnk9	ENSOANG00000007121	Potassium channel	Chr.4p, 5.73 Mb	-	P[HSA & MUS]
Lrrtm1	ENSOANG00000002561	leucine rich repeat transmembrane	Contig8748, 0.96 Kb	-	M[HSA]
Magel2	ENSOANG00000020339	MAGE-like protein	Ultra403, 645 Kb	-	M[HSA & MUS]
Mcts2	ENSOANG00000002577	RNA binding protein	Ultra390, 611 Kb	-	M[HSA & MUS]
Mest	ENSOANG00000011469	Alpha/beta hydrolase fold family	Chr.10p, 5.01 Mb	Oo_Bb-491E23	M[HSA & MUS]
Phlda2	ENSOANG00000010715	Pleckstrin homology-like domain	Contig17670, 0.53 Kb	-	P[HSA & MUS]
Plagl1	ENSOANG00000014866	Zinc finger protein	Chr.2p, 17.2 Mb	-	M[HSA, MUS & OV]
Sgce	ENSOANG00000004259	Sarcoglycan, epsilon	Contig1947, 60.6 Kb	-	M[HSA & MUS]
Slc38a	ENSOANG00000006348	Amino acid transporter	Chr.2q, 53.5 Mb	-	M[MUS]
Ube3a	ENSOANG00000007102	Ubiquitin protein ligase	Congtig13421, 16.0 Kb	Oa_Bb-358D20	P[HSA & MUS]
Zim1	ENSOANG00000004401	Zinc finger protein	Contig1021, 271 Kb	-	P[MUS]
Zrsr1	ENSOANG00000002571	U2 small nuclear RNP auxiliary factor	Ultra462, 8.48 Mb	-	M[MUS]
Actb	ENSOANG00000007716	Beta-actin	Chr.2p, 24.8 Mb	-	-
B2m	ENSOANG00000001539	Beta-2-microglobulin	Contig3082, 40.0 Kb	-	-
Gapdh	ENSOANG00000004492	Glyceraldehyde-3-phosphate dehydrogenase	Ultra687, 3.0 Kb	-	-
Hprt	ENSOANG00000011689	Hypoxanthine phosphoribosyltransferase 1	Chr.6p, 8.42 Mb	-	-
Sox1	ENSOANG00000011188	Sox-1 transcription factor	Contig126960, 0.20 Kb	Oa_Bb-117J7	-
Tpb	ENSOANG00000012858	TATA box binding protein	Contig11766, 12.6 Kb	-	-
Wsb1	ENSOANG00000005668	WD repeat and SOCS box containing 1	Ultra17, 703 Kb	Oa_Bb-155A12	-
Ywhaz	ENSOANG00000007305	Tyrosine 3-monooxygenase/tryptophan 5-	Ultra432, 916 Kb	-	-

M - maternally imprinted gene (paternal allele expressed); P - paternally imprinted gene (maternal allele expressed); HSA - Human, MUS - Mouse, OV – Pig. Locus and species imprinting information from Morison IM, Ramsay JP and Spencer HG [29].

Supplementary Table 2: AERs measured through RNA-seq in individual tissues

Status	Gene	Brain	Cerebellum	Fibroblast	Heart	Kidney	Liver	Ovary	Testis
Imprinted in eutherians	*Atp10a*	0.54	0.55	-	-	-	-	E	E
	Commd1	0.54	E	0.58	0.57	0.56	E	E	0.57
	Copg2	0.56	0.55	0.56	E	E	E	E	0.56
	Dcn	E	E	E	E	E	E	E	0.50
	Dhcr7	0.53	E	0.53	E	E	0.64	0.56	0.61
	Diras3	0.57	E	0.60	E	-	-	0.51	E
	Dlgap2	0.54	E	-	-	-	-	E	-
	Gabra5	0.51	E	E	-	-	-	-	E
	Gatm	E	0.59	E	E	0.56	0.55	E	0.56
	Igf2r	0.54	0.54	0.57	E	0.57	E	0.55	0.53
	Impact	E	E	E	E	E	E	E	0.52
	Kcnk9	E	0.60	-	E	E	E	E	E
	Lrrtm1	0.53	E	E	E	-	-	-	-
	Magel2	0.60	0.62	0.59	0.70	0.58	E	0.58	0.54
	Mcts2	E	E	E	E	E	E	E	0.77
	Phlda2	-	-	0.56	-	E	E	E	-
	Plagl1	E	-	E	E	E	E	E	0.54
	Sgce	0.57	0.52	0.63	E	E	E	E	0.54
	Slc38a4	E	-	-	-	-	0.54	-	E
	Zim1	E	0.64	E	E	E	E	E	E
	Zrsr1	0.52	E	0.60	E	E	E	0.62	0.60
	AVERAGE for TISSUE	**0.55**	**0.58**	**0.58**	**0.63**	**0.57**	**0.57**	**0.56**	**0.57**
Not imprinted in eutherians	*Actb*	0.67	0.54	0.55	E	0.67	0.78	E	0.55
	B2m	0.54	0.54	0.58	0.59	0.54	0.57	0.73	0.56
	Gapdh	0.52	0.55	0.53	0.64	0.64	0.61	0.60	0.60
	Hprt	0.56	E	E	E	E	E	E	E
	Tpb	E	E	E	E	E	E	E	0.58
	Wsb1	0.71	E	0.62	0.50	E	0.50	0.71	0.68
	Ywhaz	0.55	0.59	0.61	0.50	0.60	0.56	0.73	0.54
	AVERAGE for TISSUE	**0.59**	**0.55**	**0.58**	**0.56**	**0.61**	**0.60**	**0.69**	**0.58**

E – expression measured in the tissue (but no observable SNP/data not deep enough for analysis). A dash (-) represents no measured expression in that tissue.

Supplementary Table 3: Primers for SNP-detection in platypus tissue cDNA-libraries, and fragment sizes

Gene	SNP position (Ensembl)	Forward 5'-3'	Reverse 5'-3'	T_m °C	DNA bp	RNA bp
Copg2	Chr10:5020040	ATTCGTCATCGTATCCATCC	AGGGACCTGTTACACTCTTG	54	702	143
Diras3	Chr18:897932	TTCTTGCCATCGACCTGTAG	GCCTTCATCCTGGTCTACTC	55	302	302
Igf2r	Ultra61:99114	TTCCCTCTCCATTCCTTCTG	TGATGCACAAGATCACGAAC	55	2045	431
Actb	Chr2:24871651	ATCCGTAAGGACCTGTATGC	CCGGACTCATCGTACTCTTG	55	420	233
Wsb1	Ultra17:703706	GTGAGATCACGGACTATAGGA	ATTGGTGACATTCTTGGTGG	54	758	204

CHAPTER 5: Conclusion

Chapter Overview

Replication timing, chromatin interaction and chromosome condensation are an integral part of genome biology and signatures of gene activity. Work presented in this thesis reveals novel insights into the evolution of these aspects of genome biology. The results are relevant to our understanding of the evolution of X chromosome inactivation and genomic imprinting. The first area of research investigates the asynchronous replication of sex-specific loci in the platypus and chicken, and also differential chromosome condensation at platypus homologs, to address unanswered questions in regards to sex chromosome inactivation in these species. The second area of research looks at eutherian imprinted orthologs in platypus, and how their observed characteristics, such as asynchronous replication, conserved replication-timing entry into S-phase, long-range DNA interactions, and partial monoallelic expression, help piece together the unique evolution of the epigenetic changes that affect the activity of these genes. These results lead to an alternative hypothesis about the evolution of autosomal monoallelic expression and genomic imprinting. Finally, I discuss the current limitations of the study and the future directions of the research.

PART 1: The evolution of platypus and chicken sex chromosomes: the conservation of asynchronous replication and other epigenetic characteristics

Summary of results

Asynchronous replication of sex-specific regions in platypus and chicken

Asynchronous replication occurs at sex-linked loci in the non-therian amniote species of platypus and chicken (Ho et al. 2009; Wright and Grutzner 2014 *manuscript in revision*). In platypus females, only loci within the X-specific regions underwent asynchronous replication (Ho et al. 2009) (**Table 1**). In male birds, locus-specific asynchronous replication was observed to be dynamic across development at specific Z-borne loci (Wright and Grutzner 2014 *manuscript in revision*). The *DMRT1* region on the chicken Z-chromosome, however, replicated synchronously in males across development (**Table 1**).

Platypus X-chromosome condensation

Out of the X1, X3, and X5 homologous pairs of X-chromosomes measured in female platypus fibroblast cells, differential chromosome condensation was only observed at the X3 homologues (Ho et al. 2009). A statistically significant difference in size ($p = 0.002$) was measured for the platypus X3 homologues, whilst the other homologous pairs, X1 and X5, showed no statistical difference in size (Ho et al. 2009).

Discussion

Sex-chromosome inactivation in platypus and chicken

Monotremes feature a complex sex chromosome system with homology to chicken sex chromosomes (Grutzner et al. 2004). This raises questions about existence of

chromosome wide inactivation in monotremes, which is observed in other mammals but not birds. Replication asynchrony and chromosome condensation were used to address this in platypus. Previous chromosome replication banding-studies have shown that both female monotremes and male (ZZ) chickens appear to show no asynchronous replication at the chromosome-wide level (Wrigley and Graves 1988; Schmid et al. 1989). However, on the gene and region specific level this research has shown that locus-specific asynchronous replication is observed to occur on X- and Z-chromosomes (Wright and Grutzner 2014 *manuscript in revision;* Wright et al. 2014b *manuscript in preparation*). Interestingly, only one Z-specific region measured in this study remained synchronously replicating across development, and contained the *DMRT1* gene (Wright and Grutzner 2014 *manuscript in revision*). *DMRT1* gene expression in male chicken is about twice as high as in female chicken, and there is evidence suggesting that this gene may undergo a form of RNA-mediated silencing in females, a mechanism akin to eutherian *Xist* and marsupial *Rsx* RNA-mediated X-chromosome silencing (Brown et al. 1991; Raymond et al. 1999; Teranishi et al. 2001; Grant et al. 2012). Thus, the sex-specific dosage effect observed for *DMRT1*, where the homogametic males display twice the amount of expression compared to heterogametic females, correlates with synchronous replication of this region in males.

The results in platypus and chicken support previous gene expression profiling studies, which show region-specific expression changes on the X- and Z-chromosomes (Arnold et al. 2008; Julien et al. 2012). In therian females, the vast majority of X-specific loci are observed to asynchronously replicate, with the expressed allele replicating earlier than the X-inactivated silenced allele (Boggs and Chinault 1994; Gribnau et al. 2005; Dutta et al. 2009). The few genes on the

eutherian X-chromosome that do escape the X-inactivation process tend to replicate synchronously (Boggs and Chinault 1994). In platypus and chicken there is a partial upregulation of the expression of some genes on the sex chromosomes of the heterogametic sex, which explains why no chromosome-wide inactivation mechanisms appear to exist in the homogametic sex (Julien et al. 2012). However, some X-borne genes that are monoallelically expressed in platypus females indicate that inhibitory expression mechanisms, which contribute to gene dosage, exist at the gene-specific level in this species (Deakin et al. 2008; Julien et al. 2012).

The presence of differential chromosome condensation in monotremes, observed at the platypus X3 homologues, indicates heterochromatisation of large parts of one of the X3 homologs (Ho et al. 2009). X-chromosome condensation and inactivation, which occurs in therian female somatic cells, is a key aspect of the dosage-compensation mechanism. A heterochromatised condensed chromosome, coinciding with epigenetic transcriptional silencing for dosage compensation purposes and late S-phase replication, are hallmarks of the silent X-homolog in therians (Sharman 1971; Latt 1973; Takagi and Martin 1984; Al Nadaf et al. 2010; Julien et al. 2012). However, the presence of differentially condensed chromosome for X3 of the three (X1, X3 and X5) homologous pairs tested, also suggests that the presence of expression inhibitory effects may be region-specific in the platypus.

Overall, the presence of asynchronous replication, monoallelic expression and chromosome condensation, on female platypus X-chromosomes, indicate that the platypus mechanisms governing X chromosome inactivation may be similar to those in therians, except they are region-specific and not chromosome-wide (Deakin et al. 2008; Ho et al. 2009). Previous studies have indicated that dosage compensatory

mechanisms appear to be at work at most female platypus X-specific genes (Julien et al. 2012), however, very few genes have been mapped to the platypus X3 chromosome, meaning that their expression pattern still remains elusive. However, chromosome condensation results suggest their expression may be quite different to the genes on other well-studied sex chromosomes, such as X1.

Table 1: Overview of Asynchronous replication profiles of specific loci measured through FISH dot-assay in amniote species

Gene Category	Placentals				Monotremes (Platypus)				Birds (Chicken)			
	Gene/ Region	% SD	N	Ref.	Gene/ Region	% SD	N	Ref.		% SD	N	Ref.
Sex specific regions	*Xist*	39% [Mus]	138	(Gribnau et al. 2005)	*Ogn*	22%	587	(Ho et al. 2009)	*DMRT1*	19%	200	(Wright and Grutzner 2014 *manuscript in revision*)
	Mecp2	33% [Mus]	108		*Apc*	29%	420		*GPR98*	25.5%	200	
	Smcx	38% [Mus]	157		X5q (UltraContig304 1.9Mb)	28%	493		*PALM2*	32%	200	
Eutherian imprinted genes and their orthologs	*Igf2*	23% [Mus]	>100	(Kitsberg et al. 1993)	*Igf2*	38%	339	(Wright et al. 2014b *manuscript in preparation*)	*IGF2*	25%	258	(Dünzinger et al. 2005)
	Igf2R	35% [Mus]	>100		*Igf2R*	36%	251		*IGF2R*	22%	279	
	MEST/ COPG2	25% [HSA]	>200	(Bentley et al. 2003)	*Mest/ Copg2*	39%	240		*MEST/ COPG2*	18%	299	
Allelic exclusion	*TCRβ*	46% [Mus]	100-300	(Mostoslavsky et al. 2001)	ND							
	B-cell receptor (κ)	48% [Mus]	100-300									
	IL-2	68% [Mus]	100	(Hollander et al. 1998)								
	Olfactory receptor	31% [Mus]	> 99	(Simon et al. 1999)								
Auto-somal genes (biallelic)	*α globin*	17% [Mus]	131	(Gribnau et al. 2005)	*Wsb1*	26%	339	(Wright et al. 2014b *manuscript in preparation*)	*WSB1*	26%	257	(Wright and Grutzner 2014 *manuscript in revision*)
	L23mrp	13% [Mus]	135		*Sox1*	25%	328		*GAPDH*	10%	267	(Dünzinger et al. 2005)
	CFTR	10% [HSA]	>200	(Bentley et al. 2003)	*Insm1*	9%	431	(Ho et al. 2009)	*SLC6A8*	4%	274	

ND – no data exists. MUS – measured in mouse; HSA – measured in human. Areas in grey represent findings from work done in this thesis. Table adapted from Wright and Grutzner 2011.

While mammalian clades appear to share some characteristics that coincide with X-inactivation, other parts of the X-inactivation mechanism are not shared. For example, individual therian lineages have evolved their own mechanisms for controlling chromosome-wide dosage-compensation, with lineage specific long non-coding RNAs, such as eutherian *Xist* and marsupial *Rsx*, playing roles in the X-inactivation process in their respective clades (Brown et al. 1991; Grant et al. 2012).

Implications of the research

The observation that asynchronous replication corresponds to X-inactivated and dosage-compensated regions in therians, suggests that the presence of locus-specific asynchronous replication could reflect regional inactivation effects at specific loci in the homogametic sexes of platypus and chicken. The presence of asynchronous replication of X-borne platypus genes and Z-borne chicken genes indicates that this characteristic is not specific to the X-chromosomes of therian mammals (Ho et al. 2009; Wright and Grutzner 2014 *manuscript in revision*). Replication asynchrony is known to be characteristic of monoallelically expressed regions in eutherians, and this also appears to be true in platypus (Deakin et al. 2008; Ho et al. 2009). Thus, the locus-specific presence of asynchronous replication at sex-borne genes in chicken likely represents a conserved sex-chromosome inactivation process, perhaps also resulting in monoallelic expression (Gimelbrant and Chess 2006).

PART 2: Platypus imprinted orthologs: the conservation of asynchronous replication, s-phase entry timing, monoallelic expression, and long-range interaction

Summary of results

Asynchronous replication and replication timeline of S-phase entry of imprinted genes and their orthologs

This study has shown, for the first time, that the asynchronously replicating characteristic of genes that undergo imprinting is conserved in monotreme orthologs, which lack imprinting. The eutherian imprinted orthologs (EIOs) in platypus, *Igf2*, *Igf2R*, and *Mest*, were observed to asynchronously replicate (Wright et al. 2014b *manuscript in preparation*) (**Table 1**). The EIO, *Wt1*, in platypus did not asynchronously replicate. This confirms and extends previous work which showed the conservation of ansynchronous replication between therian mammals and birds (Dünzinger et al. 2005; Wright et al. 2014b *manuscript in preparation*) (**Table 1**).

In addition, this research has shown, for the first time, that the four EIOs in platypus and chicken, and their corresponding eutherian imprinted genes in mouse, underwent S-phase entry in relatively the same order (i.e. *Wt1*, *Mest*, *Igf2R*, and *Igf2*) (Wright et al. 2014b *manuscript in preparation*).

Allele expression ratios and monoallelic expression of imprinted orthologs in the platypus

Allele expression ratio (AER) measurements through RNA-seq revealed that 6 out of 21 platypus EIOs, and 5 out of 7 platypus control genes had AERs above 0.6, indicative of differential *cis*-regulatory effects at each allele (Wright et al. 2014a

manuscript in preparation). The cDNA seq measurements for AER were much the same as the RNA-seq, with only one in four EIOs possessing an AER above 0.6, and only one control gene, *Actb*, possessing an AER above 0.6 (Wright et al. 2014a *manuscript in preparation*). Overall, the RNA and cDNA-seq results suggest that very few platypus genes had allele-specific *cis*-regulatory effects.

The RNA-FISH results indicated that monoallelic expression occurred at the platypus EIOs: *Copg2, Igf2*, *Igf2R* and *Ube3a*. All four genes displayed between 65-100% monoallelic expression and between 0-35% biallelic expression (note that this total is out of cells which were showing expression) (**Table 2**).

Table 2: Monoallelic expression measured in different species through RNA FISH

Gene Status	Eutherians		Marsupials		Monotremes (Platypus)		Birds (Chicken)	
	%	Reference	%	Reference	%	Reference	%	Reference
Sex chromosome loci (measured in homogametic sex)	19-99%	(Al Nadaf et al. 2012)	94-99%	(Al Nadaf et al. 2010)	16-80%	(Deakin et al. 2008)	16-25%	(Kuroda et al. 2001)
Imprinted and imprinted ortholog loci	47-79%	(Kohda et al. 2001; Herzing et al. 2002)	77.7%	(Lawton et al. 2008)	65-100%	(Wright et al. 2014a *manuscript in preparation*)	ND	
Allelic exclusion loci	30-60%	(Gendrel et al. 2014)	ND		ND		ND	
Autosomal (biallelic) loci	0-20%	(Al Nadaf et al. 2012; Gendrel et al. 2014)	0% approx	(Al Nadaf et al. 2010)	3-4%	(Deakin et al. 2008)	15%	(Kuroda et al. 2001)

ND – no data exists. Areas in grey represent findings from work done in this thesis.

Conserved long-range DNA interaction between Igf2 and Wsb1 in amniotes

We found that the platypus had a conserved interaction between the *Igf2* and *Wsb1* regions, with the interaction occurring in around 8.11% of nuclei, which was significantly different to the non-interaction control (0.50%). The interaction was previously characterised in mouse, so as a control we also tested this interaction in mouse fibroblasts, and found that the interaction occurred in around 19% of nuclei (which was significantly different to the non-interaction control amount of 0.67%). Notably, this was a reduction in the amount of interaction observed in mouse fibroblasts compared to the previous study, which observed around 30-42% interaction (Ling et al. 2006).

Discussion

Replication asynchrony, replication timing, and monoallelic expression at platypus orthologs of imprinted genes

This research has shown that replication asynchrony and monoallelic expression are both conserved characteristics at imprinted orthologs in platypus (Wright et al. 2014a *manuscript in preparation;* Wright et al. 2014b *manuscript in preparation*). The asynchronous replication result in platypus makes sense in light of the finding that some imprinted orthologs in the platypus undergo amounts of monoallelic expression, with previous studies in eutherians using the asynchronously replicating characteristic as a marker for monoallelic expression (Gimelbrant and Chess 2006). However, if the conserved presence of asynchronous replication is a marker for (at least) partial monoallelic expression at these loci in amniotes, it is intriguing to speculate as to why they have evolved to express in this manner in the first place.

There is evidence to show that ncRNAs aid in the establishment of differential histone marks at eutherian imprinted loci. For example, the human *KCNQ1OT1* ncRNA associates with imprinted genes from the *KCNQ1* cluster, and mouse *Air* ncRNA expression plays a role in the imprinted expression of the *insulin-like growth factor receptor* (*Igf2r*) gene (Murakami et al., 2007; Yamasaki et al., 2005). Generally, the ncRNA association leads to recruitment of histone modifiers at the inactive allele, including methyltransferases, which aid in the establishment of differential chromatin environments at the alleles of the ncRNA-associated genes (Delaval and Feil, 2004; Terranova et al., 2008; Yamasaki et al., 2005; reviewed in Zakharova et al. 2009). However, marsupials do not appear to require ncRNAs in the establishment of imprinting at a locus (Weidman et al., 2006), suggesting that this silencing mechanism may have only evolved in eutherian mammals.

Replication asynchrony is not random for imprinted genes, and is determined in a parent of origin specific manner, unlike random monoallelically expressed genes (Simon et al. 1999). This asynchronous replication timing pattern is established in the gametes for each individual imprinted locus, and is upheld throughout development (Simon et al. 1999). In terms of the relationship between early/late allele replication and imprinted expression, it is generally accepted that the expressed allele is replicated earlier during S-phase than the silenced allele at imprinted genes (Greally et al. 1998; Zakharova et al. 2009). The amount of asynchronous replication that we observed at platypus imprinted orthologs is compared to therian imprinted genes and chicken orthologs in **Table 2**, and generally reflects the amount observed at other monoallelically expressed genes. It is understood that the differential methylation pattern, that guides the parentally-determined expression pattern at imprinted loci, is not required to set up asynchronous replication at these loci (Gribnau et al. 2003). In

fact, there is evidence to show that DNA-binding proteins, akin to that observed for the random monoallelic expressed genes, are controlling replication asynchrony at imprinted loci (Bergstrom et al. 2007).

The presence of the conserved interaction between the *Igf2* and *Wsb1* regions in platypus and mouse, indicates that interactome maps in mammals may be more conserved than previously thought (Wright et al. 2014a *manuscript in preparation*). A recent genome-wide study performed in mouse and human has shown a conservation of replication timing and chromatin interaction for orthologous regions between the species, with late replicating regions displaying a higher amount of interaction (Ryba et al. 2010). The conserved S-phase entry replication timing, observed at the eutherian imprinted genes *Igf2*, *Igf2R*, *Wt1* and *Mest* in mouse, and their orthologs in platypus and chicken, may therefore reflect the presence of conserved long-range DNA interactions in amniotes. Furthermore, there is evidence to suggest that DNA-binding proteins which mediate long-range interactions also control replication asynchrony at a locus (Bergstrom et al. 2007). The CTCF protein, which interacts with a region adjacent to the *Igf2* locus in mouse, mediates the DNA interaction between *Igf2* and *Wsb1*, and is also known to control asynchronous replication in this region (Ling et al. 2006; Bergstrom et al. 2007). Thus, the conserved presence of long-range interactions at imprinted genes, and their orthologous regions in other amniotes, may explain the conservation of asynchronous replication at these genomic regions.

A theory on the evolution of monoallelic expression in amniotes

There is much evidence to support the theory that monoallelic expression did in fact arise due to the presence of pre-existing regulatory mechanisms, such as long-range interactions, in specific clustered regions of the amniote ancestral genome. Firstly, imprinted orthologs in monotremes show an amount of monoallelic expression, combined with biallelic expression, indicating that differential *cis*-regulation at the alleles of these genes was present before imprinting evolved (Wright et al. 2014a *manuscript in preparation*). Secondly, regulatory elements that mediate long-range DNA interactions, such as LCRs and ICRs, are conserved in amniote species with and without imprinting (Weidman et al. 2004; Paulsen et al. 2005). Thirdly, the presence of asynchronous replication and conserved S-phase entry replication timing at imprinted genes and their orthologs, suggests that DNA interactions at these regions may be highly conserved in amniotes (Dünzinger et al. 2005; Ryba et al. 2010; Wright et al. 2014a *manuscript in preparation;* Wright et al. 2014b *manuscript in preparation*). Fourthly, the linked arrangement of imprinted genes, and their non-imprinted paralogs, to other imprinted clusters, indicates that these genes were likely previously organised in 'pre-imprinted' clusters present only on one or a few chromosomes in the common ancestor (Walter and Paulsen 2003). This theory also coincides with the observation that conserved clustered arrangements of imprinted orthologs exist in bird species, and that general clustering is observed at many autosomal random monoallelically expressed genes in eutherians (Young et al. 2002; Dünzinger et al. 2005; Gimelbrant and Chess 2006). Duplication, transposition, and translocation events would have spread genes from these ancestral clusters throughout the genome, and may have caused them to acquire new expression control mechanisms in their new positions.

A recent study has shown that imprinted expression of a non-imprinted *Hox* cluster transgene occurred after the transgene underwent genomic insertion in a position-dependent manner (Lonfat et al. 2013). The imprinted transgene had established *cis*-regulatory long-range contacts with a nearby locus in its new genomic position, indicating that this process may have helped to establish genomic imprinting at the inserted transgene (Lonfat et al. 2013). Notably, highly conserved and biallelicly expressed *Hox* cluster genes undergo multiple chromatin interactions as part of their normal regulatory network (**Table 3**), much like those observed at imprinted and random monoallelically expressed gene clusters (Fraser et al. 2009; Ferraiuolo et al. 2010; Wang et al. 2011). Thus, it is feasible to assume that the *Hox* transgene in the study already possessed some of the regulatory elements required to mediate long-range DNA interactions. If the clustered ancestral orthologs of imprinted genes also underwent many long-range DNA interactions as part of their expression control mechanism, perhaps it was the breaking-up of clusters, and subsequent insertions into new genomic regions, that lead to the 'pre-imprinted' monoallelic expression that we observe at platypus imprinted orthologs today. Notably, many imprinted genes are still found in clusters in birds and mammals, and thus it is interesting to speculate as to whether the insertion of a smaller cluster (and not an individual gene) would lead to a 'halfway' monoallelically expressing state and not full imprinted expression. Furthermore, unlike eutherian monoallelically expressed genes, *Hox* cluster genes are resistant to tandem duplication and transposable element (TE) insertions, such as short and long interspersed elements (SINEs and LINEs respectively) (Kim et al. 2000; Wagner et al. 2003; Simons et al. 2006) (**Table 3**). Whilst TE insertions could seriously jeopardise the imperative *cis*-regulatory elements at collinearly arranged gene groups, such as the *Hox* cluster, it has previously been speculated that they

Table 3: Characteristics of autosomal monoallelically expressed and biallelicly expressed *Hox* cluster genes

Characteristic	Random monoallelic genes		Imprinted genes		*Hox* cluster genes	
Expression	Monoallelic		Monoallelic		Biallelic	
Gene status examples	*Olfactory Receptors* and *T and B cell receptors*	(Godfrey et al. 2004; Malnic et al. 2004)	*Beckwidth Wideman region, Prader-Willi/ Angelman locus*	(Paulsen et al. 2005; reviewed in Kantor et al. 2006)	*Hox* clusters	(reviewed in Duboule 1998)
Large scale regional duplication events	Yes	(Amadou et al. 2003; reviewed in Market and Papavasiliou 2003)	Yes	(Walter and Paulsen 2003)	Yes	(Wagner et al. 2003)
Tandem gene duplication	Yes	(Amadou et al. 2003; Gimelbrant and Chess 2006; reviewed in Kambere and Lane 2007)	Yes	(Olinski et al. 2006; Rapkins et al. 2006)	No	(Wagner et al. 2003)
Conserved gene order and intergenic distances (collinearity)	No	(Godfrey et al. 2004; Malnic et al. 2004)	No	(Walter and Paulsen 2003)	Yes	(reviewed in Duboule 1998)
LINE insertions (compared to autosomal biallelicly expressed genes)	Increased around majority	(Allen et al. 2003)	Increased around majority	(Allen et al. 2003)	Depleted	(Kim et al. 2000; reviewed in Wagner et al. 2003; Simons et al. 2006)
SINE insertion (compared to autosomal biallelic expressed genes)	Depleted around majority	(Allen et al. 2003)	Depleted around majority	(Greally 2002; Allen et al. 2003)	Depleted	(Kim et al. 2000; reviewed in Wagner et al. 2003; Simons et al. 2006)
Long-range interactions	Yes	(reviewed in Jhunjhunwala et al. 2009; Ribeiro de Almeida et al. 2012)	Yes	(Kurukuti et al. 2006; Ling et al. 2006; Zhao et al. 2006; Rabinovitz et al. 2012; Wright et al. 2014a *manuscript in preparation*)	Yes	(reviewed in Ferraiuolo et al. 2010; Montavon and Duboule 2013)

Areas in grey represent fields to which some findings from this thesis have contributed.

may also act as an adaptive force in expression control at other, more dynamic, regions of the genome, such as at imprinted clusters (Feschotte 2008). Thus, it is likely that ancestral regional duplication and TE insertions have also shaped the 'pre-imprinted' and imprinted status of certain genes, with research indicating that both events have aided in the evolution of monoallelic and imprinted expression that we observe in amniote genomes today (Barlow 1993; Ono et al. 2001; Walter and Paulsen 2003; Feschotte 2008; Pask et al. 2009).

Implications of the research

Previous theories have stipulated TE insertions around imprinted genes could cause them to become imprinted (Barlow 1993; McDonald et al. 2005), however, they do not take into account the evidence that suggests clustered *cis*-regulatory chromatin interactions may also play a role in the final imprinted adaption of these genes. This work, in combination with other current research, allows a new model to be purported for the evolution of random monoallelic and imprinted gene evolution. The 'pre-imprinted' status at imprinted orthologs in earlier diverged amniotes most likely reflects the presence of regional *cis*-regulation, in the form of chromatin environment and interaction, and has played a role in the movement and adaption of monoallelically expressed genes, with particular TE insertions and tandem duplications aiding in the creation of new cis-regulatory mechanisms. The final imprinting step for some loci may be mediated by the pre-existing regional *cis*-regulation, with the adaptive selection of new and favourable maternal or paternal allele-specific regional *cis*-regulatory interactions that driving the evolution of favourable parentally-inherited imprinted interactions.

Limitations of this study

The cytogenetic approach taken in this study has allowed us to measure replication and expression at the single-nucleus level, and thus is a powerful tool when studying monoallelic and biallelic expression occurring in mutually exclusive states. We have combined this cytogenetic approach with some molecular technologies, including RNA-seq, in order to get an overall picture of what locus-specific and allele-specific expression levels are present within certain tissues. However, in the platypus it is particularly hard to assess the expression of imprinted genes, not only because of observed mutually exclusive mono- and biallelic expression status, but also because of the type of tissues available (i.e. from adult animals). Many imprinted genes are expressed in high levels during fetal development, and expression often reduces post-birth, making expression studies (with a large subset of imprinted orthologs) difficult.

Furthermore, there are still limitations to the platypus genome build. Whilst the assembly has improved over the course of this research, there are still many areas of the platypus genome that remain elusive, in terms of both sequenced content and syntenic alignment. This limited the investigation of allele specific expression in the available deep sequencing datasets. It also limited our ability to perform chromosome confirmation capture in the platypus, and so we decided to use a cytogenetic approach as an initial way to measure the conservation of long-range interactions in mammalian clades.

Future Directions

This study has successfully identified the presence of asynchronous replication at specific sex-borne genes in platypus and chicken, showing that this is not just a characteristic of loci on therian female sex-chromosomes, but a characteristic of loci on amniote sex chromosomes. As an extension of this research, it would be relevant to cytogenetically study the expression status of individual sex-borne loci on the male chicken Z-chromosome, to see if mutually exclusive monoallelic and biallelic expression occurs, akin to what occurs at loci on the platypus female sex chromosomes (Deakin et al. 2008). Future findings on the expression status of chicken Z-borne loci will aid in theories about the dosage compensatory processes witnessed at amniote sex chromosomes.

Furthermore, it would be particularly interesting to see if the any of the platypus X-chromosomes mediate X-inactivation through a long non-coding RNA, and contain a gene similar to the eutherian X-inactivation gene, *Xist*, or the marsupial X-inactivation gene, *Rsx*. As marsupials and eutherians have evolved separate X-inactivation mechanisms, it would be of particular importance to investigate whether the platypus has evolved its own form of dosage compensation, or has an ancestral form of one of the therian X-inactivation processes, with recent work suggesting that this could be a possibility (Necsulea et al. 2014). While there has been some work on identifying heterochromatic regions on the platypus sex chromosomes (Rens et al. 2004), future work will need to look at the histone modifications involved in the monotreme X-inactivation process.

This study also supports the hypothesis that monoallelic expression exists at the eutherian imprinted orthologs of earlier-diverged amniotes. It has shown that replication asynchrony is upheld at platypus imprinted orthologs, as previously

observed in chicken, which could indicate that some form of monoallelic expression occurs at these regions across amniote clades (Dünzinger et al. 2005; Wright et al. 2014b *manuscript in preparation*). Furthermore, it has shown that partial monoallelic expression, interspersed with biallelic expression, occurs at imprinted orthologs in the platypus (Wright et al. 2014a *manuscript in preparation*). Future studies should include non-mammalian species to investigate whether there is an evolutionary trajectory from monoallelic expression which then evolved into imprinted expression in therian mammals.

References

Al Nadaf S, Deakin JE, Gilbert C, Robinson TJ, Graves JA, Waters PD. 2012. A cross-species comparison of escape from X inactivation in Eutheria: implications for evolution of X chromosome inactivation. *Chromosoma* **121**: 71-78.

Al Nadaf S, Waters PD, Koina E, Deakin JE, Jordan KS, Graves JA. 2010. Activity map of the tammar X chromosome shows that marsupial X inactivation is incomplete and escape is stochastic. *Genome Biol* **11**: R122.

Allen E, Horvath S, Tong F, Kraft P, Spiteri E, Riggs AD, Marahrens Y. 2003. High concentrations of long interspersed nuclear element sequence distinguish monoallelically expressed genes. *Proc Natl Acad Sci U S A* **100**: 9940-9945.

Amadou C, Younger RM, Sims S, Matthews LH, Rogers J, Kumanovics A, Ziegler A, Beck S, Lindahl KF. 2003. Co-duplication of olfactory receptor and MHC class I genes in the mouse major histocompatibility complex. *Hum Mol Genet* **12**: 3025-3040.

Arnold AP, Itoh Y, Melamed E. 2008. A bird's-eye view of sex chromosome dosage compensation. *Annu Rev Genomics Hum Genet* **9**: 109-127.

Barlow DP. 1993. Methylation and imprinting: from host defense to gene regulation? *Science* **260**: 309-310.

Bentley L, Nakabayashi K, Monk D, Beechey C, Peters J, Birjandi Z, Khayat FE, Patel M, Preece MA, Stanier P et al. 2003. The imprinted region on human chromosome 7q32 extends to the carboxypeptidase A gene cluster: an imprinted candidate for Silver-Russell syndrome. *J Med Genet* **40**: 249-256.

Bergstrom R, Whitehead J, Kurukuti S, Ohlsson R. 2007. CTCF regulates asynchronous replication of the imprinted H19/Igf2 domain. *Cell Cycle* **6**: 450-454.

Boggs BA, Chinault AC. 1994. Analysis of replication timing properties of human X-chromosomal loci by fluorescence in situ hybridization. *Proc Natl Acad Sci U S A* **91**: 6083-6087.

Brown CJ, Ballabio A, Rupert JL, Lafreniere RG, Grompe M, Tonlorenzi R, Willard HF. 1991. A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* **349**: 38-44.

Deakin JE, Hore TA, Koina E, Marshall Graves JA. 2008. The status of dosage compensation in the multiple X chromosomes of the platypus. *PLoS Genet* **4**: e1000140.

Delaval K, Feil R. 2004. Epigenetic regulation of mammalian genomic imprinting. *Current Opinion in Genetics & Development* **14**: 188-195.

Duboule D. 1998. Vertebrate hox gene regulation: clustering and/or colinearity? *Curr Opin Genet Dev* **8**: 514-518.

Dünzinger U, Nanda I, Schmid M, Haaf T, Zechner U. 2005. Chicken orthologues of mammalian imprinted genes are clustered on macrochromosomes and replicate asynchronously. *Trends in Genetics* **21**: 488-492.

Dutta D, Ensminger AW, Zucker JP, Chess A. 2009. Asynchronous replication and autosome-pair non-equivalence in human embryonic stem cells. *PLoS One* **4**: e4970.

Ferraiuolo MA, Rousseau M, Miyamoto C, Shenker S, Wang XQ, Nadler M, Blanchette M, Dostie J. 2010. The three-dimensional architecture of Hox cluster silencing. *Nucleic Acids Res* **38**: 7472-7484.

Feschotte C. 2008. Transposable elements and the evolution of regulatory networks. *Nat Rev Genet* **9**: 397-405.

Fraser J, Rousseau M, Shenker S, Ferraiuolo MA, Hayashizaki Y, Blanchette M, Dostie J. 2009. Chromatin conformation signatures of cellular differentiation. *Genome Biol* **10**: R37.

Gendrel AV, Attia M, Chen CJ, Diabangouaya P, Servant N, Barillot E, Heard E. 2014. Developmental dynamics and disease potential of random monoallelic gene expression. *Dev Cell* **28**: 366-380.

Gimelbrant AA, Chess A. 2006. An epigenetic state associated with areas of gene duplication. *Genome Res* **16**: 723-729.

Godfrey PA, Malnic B, Buck LB. 2004. The mouse olfactory receptor gene family. *Proc Natl Acad Sci U S A* **101**: 2156-2161.

Grant J, Mahadevaiah SK, Khil P, Sangrithi MN, Royo H, Duckworth J, McCarrey JR, VandeBerg JL, Renfree MB, Taylor W et al. 2012. Rsx is a metatherian RNA with Xist-like properties in X-chromosome inactivation. *Nature* **487**: 254-258.

Greally JM. 2002. Short interspersed transposable elements (SINEs) are excluded from imprinted regions in the human genome. *Proceedings of the National Academy of Sciences* **99**: 327-332.

Greally JM, Starr DJ, Hwang S, Song L, Jaarola M, Zemel S. 1998. The mouse H19 locus mediates a transition between imprinted and non-imprinted DNA replication patterns. *Hum Mol Genet* **7**: 91-95.

Gribnau J, Hochedlinger K, Hata K, Li E, Jaenisch R. 2003. Asynchronous replication timing of imprinted loci is independent of DNA methylation, but consistent with differential subnuclear localization. *Genes Dev* **17**: 759-773.

Gribnau J, Luikenhuis S, Hochedlinger K, Monkhorst K, Jaenisch R. 2005. X chromosome choice occurs independently of asynchronous replication timing. *J Cell Biol* **168**: 365-373.

Grutzner F, Rens W, Tsend-Ayush E, El-Mogharbel N, O'Brien PC, Jones RC, Ferguson-Smith MA, Marshall Graves JA. 2004. In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes. *Nature* **432**: 913-917.

Herzing LB, Cook EH, Jr., Ledbetter DH. 2002. Allele-specific expression analysis by RNA-FISH demonstrates preferential maternal expression of UBE3A and imprint maintenance within 15q11- q13 duplications. *Hum Mol Genet* **11**: 1707-1718.

Ho KKK, Deakin JE, Wright ML, Graves JAM, Grutzner F. 2009. Replication asynchrony and differential condensation of X chromosomes in female platypus (Ornithorhynchus anatinus). *Reproduction, Fertility and Development* **21**: 952-963.

Hollander GA, Zuklys S, Morel C, Mizoguchi E, Mobisson K, Simpson S, Terhorst C, Wishart W, Golan DE, Bhan AK et al. 1998. Monoallelic expression of the interleukin-2 locus. *Science* **279**: 2118-2121.

Jhunjhunwala S, van Zelm MC, Peak MM, Murre C. 2009. Chromatin architecture and the generation of antigen receptor diversity. *Cell* **138**: 435-448.

Julien P, Brawand D, Soumillon M, Necsulea A, Liechti A, Schutz F, Daish T, Grutzner F, Kaessmann H. 2012. Mechanisms and evolutionary patterns of mammalian and avian dosage compensation. *PLoS Biol* **10**: e1001328.

Kambere MB, Lane RP. 2007. Co-regulation of a large and rapidly evolving repertoire of odorant receptor genes. *BMC Neurosci* **8 Suppl 3**: S2.

Kantor B, Shemer R, Razin A. 2006. The Prader-Willi/Angelman imprinted domain and its control center. *Cytogenet Genome Res* **113**: 300-305.

Kim CB, Amemiya C, Bailey W, Kawasaki K, Mezey J, Miller W, Minoshima S, Shimizu N, Wagner G, Ruddle F. 2000. Hox cluster genomics in the horn shark, Heterodontus francisci. *Proc Natl Acad Sci U S A* **97**: 1655-1660.

Kitsberg D, Selig S, Brandeis M, Simon I, Keshet I, Driscoll DJ, Nicholls RD, Cedar H. 1993. Allele-specific replication timing of imprinted gene regions. *Nature* **364**: 459-463.

Kohda A, Taguchi H, Okumura K. 2001. Visualization of biallelic expression of the imprinted SNRPN gene induced by inhibitors of DNA methylation and histone deacetylation. *Biosci Biotechnol Biochem* **65**: 1236-1239.

Kuroda Y, Arai N, Arita M, Teranishi M, Hori T, Harata M, Mizuno S. 2001. Absence of Z-chromosome inactivation for five genes in male chickens. *Chromosome Res* **9**: 457-468.

Kurukuti S, Tiwari VK, Tavoosidana G, Pugacheva E, Murrell A, Zhao Z, Lobanenkov V, Reik W, Ohlsson R. 2006. CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proc Natl Acad Sci U S A* **103**: 10684-10689.

Latt SA. 1973. Microfluorometric detection of deoxyribonucleic acid replication in human metaphase chromosomes. *Proc Natl Acad Sci U S A* **70**: 3395-3399.

Lawton B, Carone B, Obergfell C, Ferreri G, Gondolphi C, VandeBerg J, Imumorin I, O'Neill R, O'Neill M. 2008. Genomic imprinting of IGF2 in marsupials is methylation dependent. *BMC Genomics* **9**: 205.

Ling JQ, Li T, Hu JF, Vu TH, Chen HL, Qiu XW, Cherry AM, Hoffman AR. 2006. CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. *Science* **312**: 269-272.

Lonfat N, Montavon T, Jebb D, Tschopp P, Nguyen Huynh TH, Zakany J, Duboule D. 2013. Transgene- and locus-dependent imprinting reveals allele-specific chromosome conformations. *Proc Natl Acad Sci U S A* **110**: 11946-11951.

Malnic B, Godfrey PA, Buck LB. 2004. The human olfactory receptor gene family. *Proc Natl Acad Sci U S A* **101**: 2584-2589.

Market E, Papavasiliou FN. 2003. V(D)J recombination and the evolution of the adaptive immune system. *PLoS Biol* **1**: E16.

McDonald JF, Matzke MA, Matzke AJ. 2005. Host defenses to transposable elements and the evolution of genomic imprinting. *Cytogenet Genome Res* **110**: 242-249.

Montavon T, Duboule D. 2013. Chromatin organization and global regulation of Hox gene clusters. *Philos Trans R Soc Lond B Biol Sci* **368**: 20120367.

Mostoslavsky R, Singh N, Tenzen T, Goldmit M, Gabay C, Elizur S, Qi P, Reubinoff BE, Chess A, Cedar H et al. 2001. Asynchronous replication and allelic exclusion in the immune system. *Nature* **414**: 221-225.

Murakami K, Oshimura M, Kugoh H. 2007. Suggestive evidence for chromosomal localization of non-coding RNA from imprinted LIT1. *Journal of human genetics* **52**: 926-933.

Necsulea A, Soumillon M, Warnefors M, Liechti A, Daish T, Zeller U, Baker JC, Grutzner F, Kaessmann H. 2014. The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature* **505**: 635-640.

Olinski RP, Lundin LG, Hallbook F. 2006. Conserved synteny between the Ciona genome and human paralogons identifies large duplication events in the molecular evolution of the insulin-relaxin gene family. *Mol Biol Evol* **23**: 10-22.

Ono R, Kobayashi S, Wagatsuma H, Aisaka K, Kohda T, Kaneko-Ishino T, Ishino F. 2001. A retrotransposon-derived gene, PEG10, is a novel imprinted gene located on human chromosome 7q21. *Genomics* **73**: 232-237.

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

Paulsen M, Khare T, Burgard C, Tierling S, Walter J. 2005. Evolution of the Beckwith-Wiedemann syndrome region in vertebrates. *Genome Res* **15**: 146-153.

Rabinovitz S, Kaufman Y, Ludwig G, Razin A, Shemer R. 2012. Mechanisms of activation of the paternally expressed genes by the Prader-Willi imprinting center in the Prader-Willi/Angelman syndromes domains. *Proc Natl Acad Sci U S A* **109**: 7403-7408.

Rapkins RW, Hore T, Smithwick M, Ager E, Pask AJ, Renfree MB, Kohn M, Hameister H, Nicholls RD, Deakin JE et al. 2006. Recent assembly of an imprinted domain from non-imprinted components. *PLoS Genet* **2**: e182.

Raymond CS, Kettlewell JR, Hirsch B, Bardwell VJ, Zarkower D. 1999. Expression of Dmrt1 in the genital ridge of mouse and chicken embryos suggests a role in vertebrate sexual development. *Dev Biol* **215**: 208-220.

Rens W, Grutzner F, O'Brien P C, Fairclough H, Graves JA, Ferguson-Smith MA. 2004. Resolution and evolution of the duck-billed platypus karyotype with an X1Y1X2Y2X3Y3X4Y4X5Y5 male sex chromosome constitution. *Proc Natl Acad Sci U S A* **101**: 16257-16261.

Ribeiro de Almeida C, Stadhouders R, Thongjuea S, Soler E, Hendriks RW. 2012. DNA-binding factor CTCF and long-range gene interactions in V(D)J recombination and oncogene activation. *Blood* **119**: 6209-6218.

Ryba T, Hiratani I, Lu J, Itoh M, Kulik M, Zhang J, Schulz TC, Robins AJ, Dalton S, Gilbert DM. 2010. Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Res* **20**: 761-770.

Schmid M, Enderle E, Schindler D, Schempp W. 1989. Chromosome banding and DNA replication patterns in bird karyotypes. *Cytogenet Cell Genet* **52**: 139-146.

Sharman GB. 1971. Late DNA replication in the paternally derived X chromosome of female kangaroos. *Nature* **230**: 231-232.

Simon I, Tenzen T, Reubinoff BE, Hillman D, McCarrey JR, Cedar H. 1999. Asynchronous replication of imprinted genes is established in the gametes and maintained during development. *Nature* **401**: 929-932.

Simons C, Pheasant M, Makunin IV, Mattick JS. 2006. Transposon-free regions in mammalian genomes. *Genome Res* **16**: 164-172.

Takagi N, Martin GR. 1984. Studies of the temporal relationship between the cytogenetic and biochemical manifestations of X-chromosome inactivation during the differentiation of LT-1 teratocarcinoma stem cells. *Dev Biol* **103**: 425-433.

Teranishi M, Shimada Y, Hori T, Nakabayashi O, Kikuchi T, Macleod T, Pym R, Sheldon B, Solovei I, Macgregor H et al. 2001. Transcripts of the MHM region on the chicken Z chromosome accumulate as non-coding RNA in the nucleus of female cells adjacent to the DMRT1 locus. *Chromosome Res* **9**: 147-165.

Terranova R, Yokobayashi S, Stadler MB, Otte AP, van Lohuizen M, Orkin SH, Peters AH. 2008. Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos. *Dev Cell* **15**: 668-679.

Wagner GP, Amemiya C, Ruddle F. 2003. Hox cluster duplications and the opportunity for evolutionary novelties. *Proc Natl Acad Sci U S A* **100**: 14603-14606.

Walter J, Paulsen M. 2003. The potential role of gene duplications in the evolution of imprinting mechanisms. *Hum Mol Genet* **12 Spec No 2**: R215-220.

Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y, Lajoie BR, Protacio A, Flynn RA, Gupta RA et al. 2011. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472**: 120-124.

Weidman JR, Dolinoy DC, Maloney KA, Cheng JF, Jirtle RL. 2006. Imprinting of opossum Igf2r in the absence of differential methylation and air. *Epigenetics* **1**: 49-54.

Weidman JR, Murphy SK, Nolan CM, Dietrich FS, Jirtle RL. 2004. Phylogenetic footprint analysis of IGF2 in extant mammals. *Genome Res* **14**: 1726-1732.

Wright ML, Casey A, Necsulea A, Julien P, Kaessmann H, Grutzner F. 2014a. Monoallelic expression predates imprinting: platypus imprinted orthologs display an amount of monoallelic expression. *manuscript in preparation*.

Wright ML, Grutzner F. 2014. Asynchronous replication timing of loci on the Z-chromosome in male chicken *manuscript in revision*.

Wright ML, Zechner U, Haaf T, Grutzner F. 2014b. Conservation of replication asynchrony and S-phase entry predates evolution of genomic imprinting in mammals. *manuscript in preparation*.

Wrigley JM, Graves JA. 1988. Sex chromosome homology and incomplete, tissue-specific X-inactivation suggest that monotremes represent an intermediate stage of mammalian sex chromosome evolution. *J Hered* **79**: 115-118.

Yamasaki Y, Kayashima T, Soejima H, Kinoshita A, Yoshiura K, Matsumoto N, Ohta T, Urano T, Masuzaki H, Ishimaru T et al. 2005. Neuron-specific relaxation of Igf2r imprinting is associated with neuron-specific histone modifications and lack of its antisense transcript Air. *Hum Mol Genet* **14**: 2511-2520.

Young JM, Friedman C, Williams EM, Ross JA, Tonnes-Priddy L, Trask BJ. 2002. Different evolutionary processes shaped the mouse and human olfactory receptor gene families. *Hum Mol Genet* **11**: 535-546.

Zakharova IS, Shevchenko AI, Zakian SM. 2009. Monoallelic gene expression in mammals. *Chromosoma* **118**: 279-290.

Zhao Z, Tavoosidana G, Sjolinder M, Gondor A, Mariano P, Wang S, Kanduri C, Lezcano M, Sandhu KS, Singh U et al. 2006. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat Genet* **38**: 1341-1347.

THESIS SUPPLEMENTARY

Other published work by author of thesis

Long HK, Sims D, Heger A, Blackledge NP, Kutter C, Wright ML, Grutzner F, Odom DT, Patient R, Ponting CP et al. 2013. Epigenetic conservation at gene regulatory elements revealed by non-methylated DNA profiling in seven vertebrates. *Elife* **2**: e00348.

AMENDMENTS

1. Page XVII, paragraph 2, line 2: replaced “albeit” with “although”

2. Added comma around the word ‘however’ at the following places (not included on these pages due to pre-existing publication formatting):

   Page 6, paragraph 1, line 4 (after); page 6, paragraph 3, line 4 (after); page 8, paragraph 1, line 5 (after); page 9, paragraph 3, line 5 (before and after); page 10, paragraph 2, line 6 (before and after); page 14, paragraph 4, line 3 (after); page 15, paragraph 1, line 4 (after); page 15, paragraph 4, line 6 (after); page 16, paragraph 2, line 11 (before); page 19, paragraph 3, line 6 (before) page 20, paragraph 2, line 13 (after); page 22, paragraph 2, line 6 (after).

3. Page 8, after the end of paragraph 1: The following paragraph should be included (not included on page 8 due to pre-existing publication formatting):

   There is evidence to suggest that both non-coding RNAs and DNA-binding proteins also have roles in coordinating replication timing. The long ncRNA gene present on human chromosome 6, *ASAR6*, has been shown to affect asynchronous replication of chromosome 6 loci both in cis, and also in trans when opposite coordination is observed across the centromere (Donley et al. 2013). Furthermore, asynchronous replication on the eutherian X-chromosome is established soon after the accumulation of *Xist* ncRNA, however, deletion of the *Xist* locus after inactivation has been established still results in maintained inactivation and late replication, suggesting that replication timing may play a role in preserving the inactivated state (Csankovszki et al. 1999; Keohane et al. 1996; Takagi and Martin 1984). In terms of DNA-binding proteins, the Polycomb Group protein EED, a subunit of the Polycomb repressive complex 2 (PRC2) with

histone H3 lysine 27 methyltransferase activity, is observed to coordinate replication asynchrony, through DNA-binding, at some monoallelicly expressed genes in differentiated cells (Alexander et al., 2007). Furthermore, the DNA-binding protein CTCF, which mediates imprinting at the *Igf2-H19* locus, has also been found to regulate asynchronous replication at this locus (Bergstrom et al., 2007).

Added references to Chapter 1:

Csankovszki G, Panning B, Bates B, Pehrson JR, Jaenisch R. 1999. Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation. Nat Genet 22: 323-324.

Donley N, Stoffregen EP, Smith L, Montagna C, Thayer MJ. 2013. Asynchronous replication, mono-allelic expression, and long range Cis-effects of ASAR6. PLoS Genet 9: e1003423.

Takagi N, Martin GR. 1984. Studies of the temporal relationship between the cytogenetic and biochemical manifestations of X-chromosome inactivation during the differentiation of LT-1 teratocarcinoma stem cells. Dev Biol 103: 425-433.

Keohane AM, O'Neill L P, Belyaev ND, Lavender JS, Turner BM. 1996. X-Inactivation and histone H4 acetylation in embryonic stem cells. Dev Biol 180: 618-630.

4. Page 9, paragraph 1, line 6: After the sentence ends, the following sentence should be included (not included on page 9 due to pre-existing publication formatting):

It should be added here that replication timing does not always fall into an early/late pattern, indeed there are many areas of the genome that follow intermediate patterns of replication timing. Furthermore, replication asynchrony doesn't mean early or late replication timing; it only means that replication occurs at the two homologues at different times.

5. Page 11, paragraph 2, line 5: Add to the sentence after the words "including cancer" (not included on page 11 due to pre-existing publication formatting):

   and neuronal diseases, such as familial Alzheimer's disease (AD1), and SOD1 associated with familial amyotrophic lateral sclerosis.

6. Page 12, paragraph 1, line 2: After the sentence ends, the following sentence should be included (not included on page 12 due to pre-existing publication formatting):

   It has been noted that larger probes can actually result in the appearance of doublet signals in the absence of replication, meaning smaller sized probes are often preferable to larger ones when measuring asynchronous replication through DNA FISH (Carothers and Bickmore, 1995).

   Added reference to Chapter 1:

   Carothers AD, Bickmore WA. 1995. Models of DNA replication timing in interphase nuclei: an exercise in inferring process from state. Biometrics 51: 750-755.

7. Page 13, paragraph 1, lines 8-9: Change the end of the sentence from "and the final epigenetic mark is the late replication status of the inactive-X during the S-

phase" to read (change not included on page 13 due to pre-existing publication formatting):

and the inactive-X undergoes late replication during the S-phase.

8. Page 46: text formatting at the bottom of the page was adjusted.

9. Page 116, paragraph 1, line 10: after previous sentence ends, additional sentence was inserted which reads:

Furthermore, some doublets may have not coincided with BrdU incorportation if they replicated earlier than the short BrdU pulse.

10. Page 146, paragraph 2, line 8: replaced "facilities" with "facilitates"

11. Page 180, paragraph 1: inserted paragraph reads as follows:

There is evidence to show that ncRNAs aid in the establishment of differential histone marks at eutherian imprinted loci. For example, the human *KCNQ1OT1* ncRNA associates with imprinted genes from the *KCNQ1* cluster, and mouse *Air* ncRNA expression plays a role in the imprinted expression of the *insulin-like growth factor receptor* (*Igf2r*) gene (Murakami et al., 2007; Yamasaki et al., 2005). Generally, the ncRNA association leads to recruitment of histone modifiers at the inactive allele, including methyltransferases, which aid in the establishment of differential chromatin environments at the alleles of the ncRNA-associated genes (Delaval and Feil, 2004; Terranova et al., 2008; Yamasaki et al., 2005; reviewed in Zakharova et al. 2009). However, marsupials do not appear to require ncRNAs in the establishment of imprinting at a locus (Weidman et al., 2006), suggesting that this silencing mechanism may have only evolved in eutherian mammals.

12. Page 187, paragraph 2, line 1: deleted "has"