Chromosomal sex via epigenetic modification?

Zhang, X.[†], Wagner. S.[†], Holleley, C.E., Graves, J.A.M[+] and Georges, A.[+]
Institute for Applied Ecology, University of Canberra, School of Life Sciences, La Trobe University

Abstract:
The central bearded dragon (*Pogona vitticeps*) has a ZZ male/ZW female sex chromosome system. The Z and W are tiny microchromosomes, but they are easily distinguishable microscopically by the regions of repetitive sequence on the larger W. We searched the sex chromosome sequences for candidate sex determining genes, but we detected no single copy W-specific genomic sequence. The most promising candidate gene was *nr5a1*, which codes for steroidogenic factor SF1, a protein essential for mammalian sex determination. This gene has alleles on both the Z and W (*Z-nr5a1* and *W-nr5a1*), which have the same genomic sequence, and are subject to recombination. Three isoforms were detected in *Z-nr5a1* transcripts from gonads of adult ZZ males, two of which would translate into intact protein. However, the *Z-nr5a1* and *W-nr5a1* alleles of ZW females produced sixteen isoforms, most of which contained chain terminating sequences. This suggests that the W-borne allele produces transcripts that are differentially spliced and produce truncated polypeptides. Structures of these truncated polypeptides confirm that the DNA binding domain is intact, so that they could act as competitive inhibitors of the full length intact protein. We propose that an altered configuration of the W chromosomes affects the splicing to generate inhibitory W-borne isoforms that suppress testis determination. Thus this GSD system may be controlled, not genetically but epigenetically, by the sex chromosomes.

1. Mapping *nr5a1* sequences in Z and W microchromosomes in Pogona

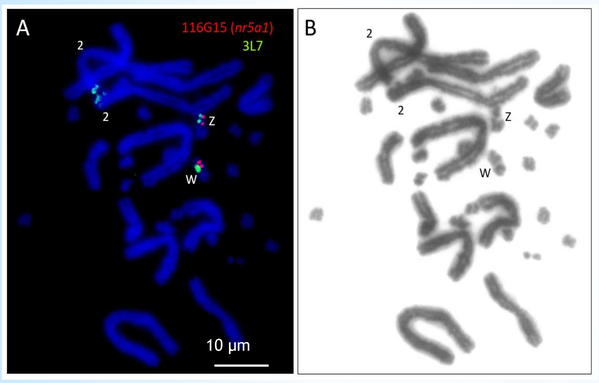


Figure 1
A: *nr5a1*-containing BAC (BAC 116G15, red) maps to the Z and W chromosomes (Marker BAC 3L7, green) in a metaphase cell of female *P. vitticeps*. B: Inverted DAPI image of the same metaphase highlighting Z, W and chromosomes.

2. The hinge region of Pogona *nr5a1* gene contains GC-rich microsatellite repeats (STR)

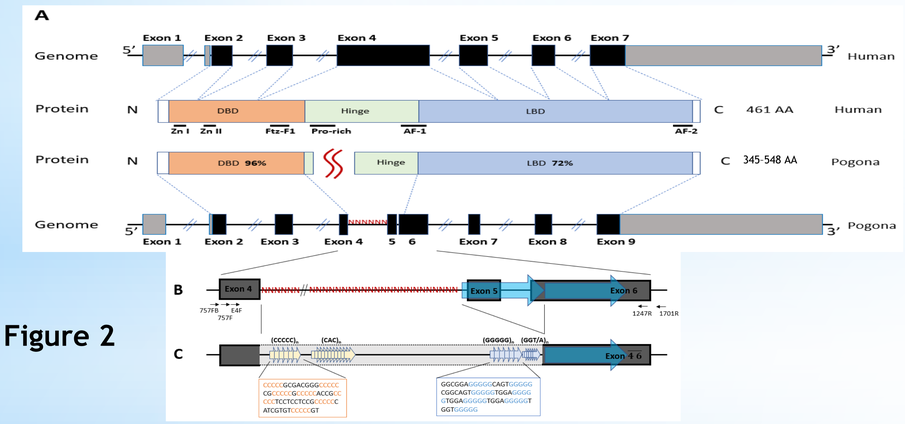


Figure 2

The protein encoded by *nr5a1* is SF1, commonly contain a N-terminal DNA-binding domain (DBD) and a C-terminal Ligand binding domain (LBD), linked by a hinge region (**Figure 2A**). The DBD functions in binding the target genes, and the LBD in ligand-mediated transcription activity, both are highly conserved. The hinge region displays high diversity between species. For most species, the hinge region is encoded within a large exon in the middle of the gene, but the model RefSeq transcript generated by NCBI for Pogona splits it into three exons (Exon 4 to 6) and defines an unresolved gap (**Figure 2B**). We resolve the genomic sequence of the region by Sanger sequencing. The region has a high GC content harboring four microsatellites (**Figure 2C**), including a monomeric $(CCCCC)_{6\text{-}7}$, a trinucleotide $(CAC)_{13\text{-}18}$, a monomeric$(GGGGG)_{2\text{-}6}$ and trinucleotide $(GGT/A)_{4\text{-}7}$ repeats. The monomeric C and G STRs were absent from other species, so appear to be Pogona specific.

3. No difference in gDNA sequences nor expression levels detected between male and female individuals

Figure 3. Genetic instability

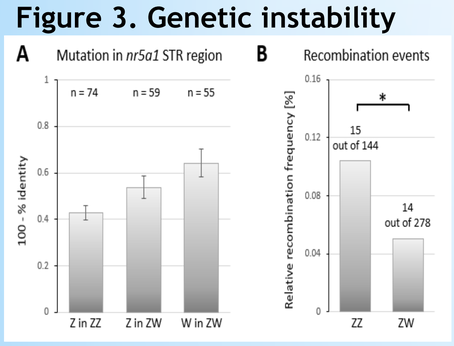
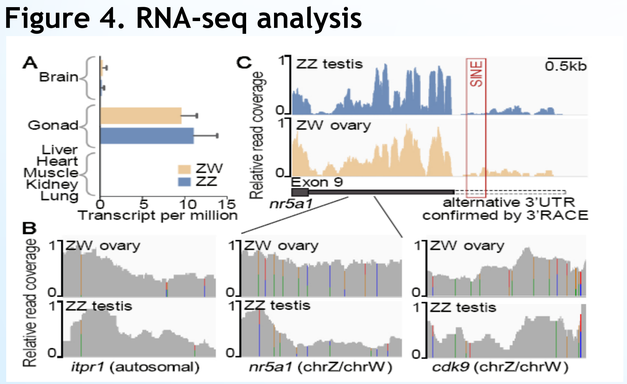


Figure 4. RNA-seq analysis



We sequenced STR region of *nr5a1* in 33 individuals of *P. vitticeps* and assigned alleles to the Z or W chromosome by pedigree analysis. The STR region appears genetically instable and the *W-nr5a1* has a higher mutation rate (**Figure 3A).** Mitotic recombination still occurs between the Z and W chromosome in *nr5a1* although the frequency of recombination showed lower than that between two Z in ZZ individuals (**Figure 3B**).
RNA-seq data showed that high levels of *nr5a1* mRNA were only found in gonads, consistent with its significant role in gonadal development and maintenance (**Figure 4A**). We observed no significant difference in *nr5a1* expression level between adult ZZ testis and ZW ovaries (**Figure 4B**). RNA-seq and 3'RACE both detected a long 3'UTR tail of 3,895 nt in addition to a 2,210 bp 3'UTR (**Figure 4C).**

4. Sex differential alternative splicing in *nr5a1* STR region

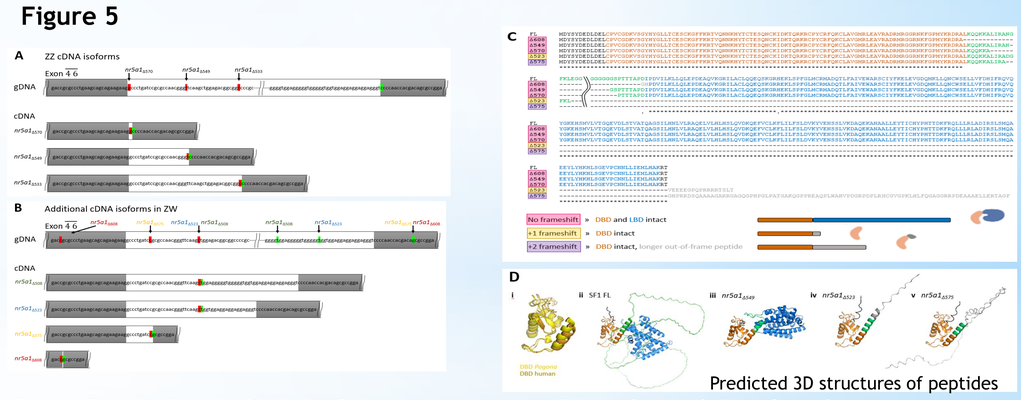
Table 1. *nr5a*1 cDNA variants identified in male and female gonads

Gender	Individual ID	cDNA variants
ZZ	Pit_001003379128	$nr5a1_{\Delta549}$ $nr5a1_{\Delta533}$
	Pit_001003344020	$nr5a1_{\Delta549}$ $nr5a1_{\Delta533}$ $nr5a1_{\Delta570}$
	Pit_001003351790	$nr5a1_{\Delta549}$ $nr5a1_{\Delta533}$
	Pit_001003344962	$nr5a1_{\Delta549}$ $nr5a1_{\Delta533}$
	Pit_005005003628	$nr5a1_{\Delta549}$ $nr5a1_{\Delta533}$
ZW	Pit_005005003514	$nr5a1_{\Delta549}$ $nr5a1_{\Delta520}$ $nr5a1_{\Delta526}$ $nr5a1_{\Delta568}$ $nr5a1_{\Delta703}$ $nr5a1_{\Delta881}$
	Pit_005005003443	$nr5a1_{\Delta533}$ $nr5a1_{\Delta537}$ $nr5a1_{\Delta604}$ $nr5a1_{\Delta640}$ $nr5a1_{\Delta679}$
	Pit_001003182571	$nr5a1_{\Delta549}$ $nr5a1_{\Delta533}$ $nr5a1_{\Delta570}$ $nr5a1_{\Delta523}$ $nr5a1_{\Delta608}$
	Pit_001003342982	$nr5a1_{\Delta549}$ $nr5a1_{\Delta533}$ $nr5a1_{\Delta570}$ $nr5a1_{\Delta508}$ $nr5a1_{\Delta523}$ $nr5a1_{\Delta575}$ $nr5a1_{\Delta608}$

RT-PCR detected three *nr5a1* cDNA variants from five ZZ male individuals and sixteen *nr5a1* cDNA variants from four ZW female individuals **(Table 1).**
The cDNA variants result from alternative splicing in the STR region (**Figure 5A and B).**

5. Putative function of polypeptides encoded by alternative transcripts

Figure 5



Two of the three cDNA isoforms common to ZZ male and ZW female ($nr5a1_{\Delta549}$ and $nr5a1_{\Delta570}$) and one only from ZW females ($nr5a1_{\Delta608}$) would translate into intact protein (**Figure 5C and D**). Thirteen of cDNA isoforms in ZW female contain a translation frame shift that introduces a premature stop codon. They could translate into polypeptide containing the DBD but not the LBD (**Figure 5C and D**), and retain the function of binding to the target genes.

6. Hypothesis

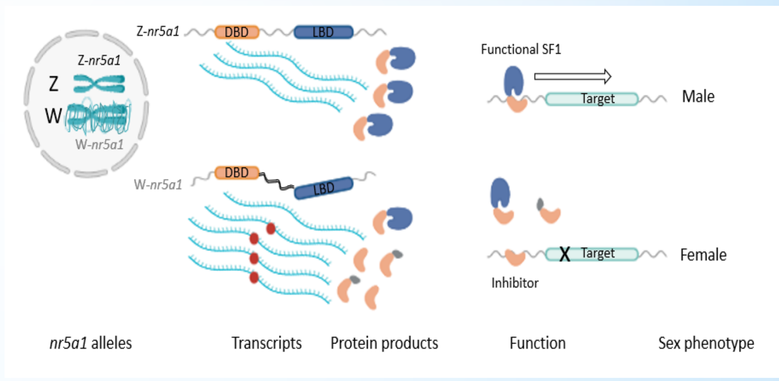


Figure 6. Hypothesis of the epigenetic control of sex determination by post-transcriptional control of *nr5a1*. The Z and W-borne *nr5a1* have the same base sequence, but the W is held in an altered conformation. This makes the STR region in the middle of W-*nr5a1* genetically unstable, and affects transcription and secondary structure of the pre-mRNA. This affects splicing, leading to many isoforms that are translated into truncated SF1 polypeptides that act as competitive inhibitors of SF1 function. Suppression of SF1 leads to female development of ZW animals.

Future work: It will be exciting to test this hypothesis by phased long-read sequencing of Z and W chromosomes, Hi-C and recombination analysis across the entirety of both sex chromosomes, genome wide CHiP as well as detailed studies of transcripts early in embryonic development, and the effects of temperature on differential splicing.

Acknowledgement: We thank Janine Deakin, Kazumi Matsubara, Ira W. Deveson, Denis O'Meally, Hardip R. Patel,[7] Tariq Ezaz, Zhao Li, Chexu Wang, Melanie Edwards,Wendy Ruscoe, Jacqui Richardson, Andrew Sinclair, Sarah Whiteley, Meghan Castelli, Shayer Alam and Foyez Shams for all sorts of helps.
This work is funded by an Australian Research Council Discovery Grant (DP170101147) awarded to AG (lead), CEH, JWD, JAMG, TE, SS, LS, and PW.