A PROPOSAL TO SEQUENCE THE GENOME OF THE PLATYPUS, *ORNITHORHYNCHUS ANATINUS*.

Peter D.Temple-Smith[1] Jennifer A. Marshall Graves[2*], Frank Grützner[2] Janine Deakin[2], Marilyn B. Renfree[1], Katherine Belov[3], Robert Miller[4], Randy Jirtle[5], Kerstin Lindblad-Toh[6], Eric S. Lander[6] and Richard K. Wilson[7]

[1]The University of Melbourne, Melbourne; [2]Australian National University, Canberra; [3]The Australian Museum, Sydney; [4]University of New Mexico, NM [5]Duke University Medical Center, NC; [6]Whitehead Institute/MIT Center for Genome Research, Cambridge, MA; [6]Washington University School of Medicine, St. Louis, MO

*Corresponding author: graves@rsbs.anu.edu.au

I. Overview

The discovery of the egg-laying monotremes platypus and echidna in Australia more than 200 years ago caused great excitement and controversy in the zoological world. Their combination of reptilian, mammalian and unique characteristics seemed incongruous to 19th century scientists; and because of its strange and unique amalgam of characters, the platypus was initially thought to be a skilled taxidermist's hoax. However, monotremes were immediately identified as mammals because they bear fur and feed their young with milk. The evolutionary status of monotremes, either at the root of mammals or as offshoot of the marsupial lineage, remains controversial today although most monotreme biologists agree that the Prototherians (the monotremes) diverged before the Therian (marsupial and eutherian) mammals.

Platypus has been the subject of extensive anatomical and physiological studies because of its extraordinary features, including reproduction and development, neuroanatomy and special senses, as well as production (uniquely for a mammal) of venom. These studies have provided a considerable foundation of knowledge about this species that provides a solid basis for functional aspects of genomic comparison in mammals.

Equally extraordinary, unique and controversial is the genome of monotremes, which appears to combine typical mammalian features with reptilian characteristics (Grützner et al. in press). Although it is much the same size as other mammals, the size distribution of chromosomes is unusual for mammals, and has reptilian features. The sex chromosomes are unknown, and the therian sex determining gene *SRY* is apparently absent. Most peculiar of all is the presence of unpaired chromosomes that implies a system of translocation heterozygosity unique among mammals, whose effect on fertility is unknown.

The genome of the platypus therefore represents an extraordinary and unique resource for comparative genomics, offering power to identify genes and regulatory sequences by phylogenetic footprinting, dissect complex pathways of genetic control, and to determine how the mammalian genome evolved, and perhaps how it functions.

We therefore propose generating an 8-fold coverage high quality draft of the platypus genome.

II. Biological rationale for sequencing the platypus genome

A unique mammal for comparative studies

The unique suite of morphological and physiological characters of monotremes was recognized very soon after their discovery. The first serious study of monotremes by Caldwell (1884) demonstrated that their reproductive and developmental biology was a mixture of reptilian and mammalian features (Griffiths 1968, 1978, 1999, Temple-Smith 1974, Renfree 1993, 1995, Temple-Smith and Grant 2001, Lin and Jones 2001).

Platypus are, however, true mammals because they elaborate a complex milk under the control of the same hormones that control lactation in both other groups of

mammals (Griffiths 1978). The 210 million years of separate evolution of monotremes from other extant mammalian groups has led to many striking differences in their anatomy and physiology, some of which are highly specia lized to the group and reflect their early divergence from other mammals. Fundamental differences in modes of reproduction and development provide further unique opportunities to understand the genes and hormones that control mammalian reproduction and development. Despite this potential, and although they are relatively common animals in the wild, monotremes do not lend themselves readily to captive breeding for this reason. Only recently have researchers in Melbourne and Sydney been able to successfully breed these animals.

The most famous feature of monotremes is that they lay eggs. This is not to say that the embryo is not nourished in the uterus and they have an allantoic and yolk sac placenta (Griffiths 1968, 1978, Tyndale-Biscoe and Renfree 1987). The egg is laid when the embryo has reached the 19-20 somite stage, and in the echidna, but not the platypus, a pouch develops to receive the egg (Griffiths 1978, Renfree 1993). The young is dependent on milk for the first few months of life to sustain its development. Like all other mammals, the milk is produced by well-developed mammary glands. Unlike all other mammals, the females have no nipples and the milk is expressed by the female from two teatless areolae and the young suck the milk directly from the abdominal surface (Griffiths 1968, 1978). Monotreme milk composition, like that of marsupials, changes throughout lactation and the composition of platypus and echidna milks differ significantly.

Most organ systems are known to differentiate during pouch or nest life although much less is known about this process than in marsupials (Griffiths 1978). Two unique features of monotremes are the development of an electro-sensory mechanism in their snouts that assists them to locate their invertebrate prey (Scheich et al. 1986, Gregory et al. 1987, Proske et al. 1998, Pettigrew et al. 1998) and the presence in males of a keratinous canalized spur connected by a thin membranous duct to an associated venom gland (Temple-Smith 1974, Griffiths 1978). In the platypus this venom system is well developed and capable during the breeding season of inflicting wounds, but not apparently death, during territorial fighting between males (Temple-Smith 1974). It can also produce painful spur wounds in other mammals including man. In echidnas this system, while present, appears less functional and there are no records of intraspecific or interspecific aggression using the spurs. The presence of this system in all extant monotremes suggests a long history for the crural system in the group but little is known of its evolutionary derivation.

Monotremes are seasonal breeders, and produce one to three young. The anatomy of the reproductive system is unequivocally ancestral to that of both marsupials and of eutherian mammals. Sperm structure has many reptilian characteristics, but the specialized uterine endometrium that nourishes the early embryo is mammalian. In echidnas both ovaries are functional, but in platypus ovulations occur from the left ovary only (as for many birds). Echidnas produce a single young. Platypus is capable of producing from one to three young in the annual reproductive cycle, but twins are the most common outcome. The ovary produces progesterone and estradiol, and the testes synthesise testosterone and dihydrotestosterone as in all other mammals (Jones 1998). All monotremes are testicond and differ in their accessory gland anatomy from other mammals (Griffiths 1968, 1978, Temple-Smith 1974, Carrick and Hughes 1978,

Jones and Djakiew 1978, Temple-Smith and Grant 2001, Griffiths in Walton and Richardson 1989), but their reproduction is otherwise characteristically mammalian (Tyndale-Biscoe and Renfree 1987).

Monotreme relationships and phylogeny

Generally classified in a single Order Monotremata of the mammalian Subclass Prototheria, only three species are usually recognised: the duck-billed platypus (*Ornithorhynchus anatinus*) and two echidna species, the Australian short beaked echidna (*Tachyglossus aculeatus*) and the long beaked Niugini echidna (*Zaglossus bruijni*). Recently, however, Flannery *et al.*(1998) reviewed long-beaked echidna taxonomy, erected two new species from museum specimens and suggested that three new subspecies of *Z.bruijni* be recognised.

The fossil record of monotremes comes almost exclusively from Australia and New Guinea. The oldest fossil dates from the Mesozoic 100 MYA (million years ago) (Archer et al. 1985 and Archer et al. 1992). Miocene and Pleistocene fossils of a giant echidna were also found in Australia. However, recent discovery of fossils of early Pleistocene platypus in southern Argentina show that the distribution of monotremes was wider and included eastern Gondwana during the late Cretaceous to early Paleocene (Pascual et al. 1992).

Although monotremes were immediately identified as mammals, their evolutionary relationship with other mammal groups is still controversial. Traditionally, they are considered the only extant order of the mammalian subclass Prototheria that diverged from the subclass Theria (marsupials and placental mammals) 150-200 MYA. This orthodox view is supported by a vast body of morphological, physiological and anatomical evidence (Carroll 1988, Rougier et al. 1996, Vaughn 1986). However, it has been challenged by sequence comparisons of mitochondrial DNA and nuclear 18S rRNA, which suggest that monotremes and marsupials form one group (Marsupionta) that diverged together from eutherians, and split later into the monotreme and marsupial lineages (Janke et al. 1997, Janke et al. 2002). However, more recent and more exhaustive analysis of nuclear genes (M6P/IGF2R and immunoglobulin genes) strongly supports the traditional idea of the Theria clade that includes marsupials and eutherian mammals (Killian et al. 2001b, Belov et al. 2002a,b,c), and the divergence is now dated from comparison of several gene sequences, at 210 MYA (Woodburne et al. 2003).

Cao et al. (1998) suggested that platypus and echidnas diverged between 20 to 45 MYR but evidence from mtDNA and immunoglobulin sequences and DNA-DNA hybridisation have narrowed this estimate to 21-25MYR (Westerman and Edwards 1992, Kirsch and Mayer 1998, Belov et al. 2003b).

Monotremes are therefore an outgroup to other mammal groups which can be used for many comparisons of genes and genomes within mammals, and with other vertebrates. Comparisons with the genomes of marsupial and eutherian mammals will enable us to establish the ancestral form of the mammal genome, and comparisons with the genomes of other vertebrates (eg chicken) will show us what features are mammal-specific and which have been present in the mammalian/sauropsid ancestor.

Unique features of the monotreme genome

Monotreme genomes, like the animals themselves, are an extraordinary mixture of typical mammalian, and reptilian characteristics.

The monotreme genome comprises about 3pg DNA (platypus 3.06pg, short-beaked echidna 2.98pg). This is slightly smaller than human (3.5pg) and well within the size range of other mammals (bats have1.8pg and some rodents up to 7pg).

The platypus karyotype consists of 52 chromosomes in both sexes, and includes six pairs of large autosomes, a large X chromosome, 17 pairs of smaller chromosomes that are hard to differentiate and four small elements with no pair (Bick and Sharman 1975, Wrigley and Graves 1988a, Graves et al. 2003). In males, there is a single X and five unpaired elements. The two members of homologous pairs frequently show differences in size and morphology. The heterochromatic and heteromorphic short arm of chromosome 6 bears the nucleolar organizer. The karyotype of the echidna is very similar, except that the diploid numbers are 63 in male and 64 in female. High resolution G-banding, R-banding and late replication banding stains (Wrigley and Graves 1988a), as well as comparative gene mapping (Watson et al. 1992) reveal a high degree of conservation of the larger chromosomes (including the X) between the echidna and platypus, and confirmed the identity between the karyotypes of the two echidna species. It has been difficult to map genes to the smaller elements, since they are hard to differentiate, but several genes apparently map to one of the small metacentrics.

The form of the monotreme karyotype is unusual for mammals, and forms a link between mammals and reptiles/birds in its distribution of very large and very small chromosomes. The small chromosomes bear interesting similarities to the gene rich microchromosomes of birds, occupying the central domain in cells where active chromatin is expected to lie (Grützner, unpublished).

A unique feature of the monotreme karyotype is the existence of several unpaired chromosomes. At meiosis, these unpaired elements form a chain of 8 elements (9 in echidna) with the X chromosome and other yet-to-be identified chromosomes (Bick and Sharman 1975, Murtagh 1977, Watson et al. 1992). This chain is interpreted as the result of heterozygosity for multiple translocations. The composition of the meiotic chain, the homologies between elements, and how these elements segregate to form balanced gametes and viable zygotes, is still a complete mystery. Questions of how the chain evolved, and why it was retained in monotremes, also remain unanswered and deeply puzzling, and are under active investigation (reviewed Grützner et al. in press).

Monotremes are the only mammals known to maintain translocation heterozygosity for a number of chromosomes without obvious effects on fertility. Such a system is known to occur naturally in a few plants and invertebrate species, but was unprecedented in vertebrates. In mice heterozygous for translocations the chromosomes of these animals also form chains or rings, but this results in meiotic arrest and sterility of the heterozygous mice, or the production of high proportions of aneuploid sperm (Eaker et al. 2001) or embryos (Winking et al. 2000, Underkoffler et al. 2002). It is possible that in monotremes a balanced segregation is not always

achieved and that this may explain the relatively poor and apparently random individual breeding success in platypuses and short-beaked echidnas which have recently been reported from extensive field studies (Rismiller and Kelvey 2000, Temple-Smith and Grant 2001, Grant and Temple-Smith in press).

Monotremes have fibrillar spermheads like birds and reptiles. The chromosomes appear in a defined order in the monotreme spermhead, with the X at the apex (Watson et al. 1996. Greaves et al., 2002). Although chromosome position seems also

al. 1993), although no sign of a sex chromatin body could be found (McKay et al. 1987). Recently dosage compensation was demonstrated for three genes on the X by real-time PCR. It remains to be determined if this is achieved by X inactivation. This is especially interesting because all four genes that flank the X inactivation centre (*XIST*) in eutherians map to a platypus autosome (Delbridge, Waters, Deakin, unpublished results). If dosage compensation monotremes is achieved by X inactivation, it is therefore unlikely to be under the control of *XIST* as this region appears to be autosomal in the platypus.

The identity of the platypus Y chromosome (even the presence of a Y chromosome in males) is in doubt, because of the presence of other unpaired elements and the formation of the translocation chain. The question of chromosomal homologies is currently being addressed by comparisons of sequences and chromosome painting (Grützner et al. in press).

Many attempts over several years to clone the platypus homologue of the therian testis determining gene *SRY* have been unsuccessful, though several related autosomal *SOX* genes have been isolated and characterized (Kirby et al. 2002). The *SOX* gene from which *SRY* diverged appears to be autosomal in platypus. Therefore it seems likely that there is no male-specific *SRY* gene (Graves 2002), and platypus sex determination is accomplished by another gene on the X or a putative Y chromosome. Identification of this, perhaps ancestral, sex determining gene is important in understanding the evolution, and perhaps the function of the human sex-determining pathway.

The strange array of chromosomes in the monotremes provides unique classic model systems for studying the most basic rules for mammalian chromosome organization, function, stability and evolution. In the same way as fundamental differences in marsupial chromosome behaviour (e.g. the reversal of the sex difference in recombination rate in marsupials) have necessitated rewriting genetic rules for mammals, monotreme chromosome heteromorphy, translocation heterozygosity and mysterious sex chromosomes are likely to challenge our understanding of mammalian chromosomes.

Comparison of monotreme gene arrangements has also contributed to our understanding of the evolution of gene families. Thirty genes have been physically mapped by FISH in platypus and eight in the echidna (reviewed by Grützner et al. in press). This mapping has been the basis for new theories of sex chromosome evolution. In addition, the discovery of a conserved pair of *SOX* genes located together in the platypus (Kirby et al. 2002), provided insight into the evolution and homologies of the developmentally important *SOXB* gene family.

Thus gene mapping and sequencing in the platypus has revealed many unexpected features and challenged some of the rules of mammalian genomes (Graves 1996).

Informing the human sequence and bridging the gap between reptiles, birds and eutherians

Perhaps most significantly, monotremes occupy a unique phylogenetic position that will provide invaluable opportunities to explore mammalian genome organization and function.

Intergenomic comparison is an integral part of the analysis of the human genome sequence and has proved to be one the most effective techniques for identifying genes (Batzoglou et al. 2000; Roest Crollius et al. 2000). Such phylogenetic footprinting is also likely to be an efficient method to identify the non-coding conserved regulatory elements that are still largely unknown The power of this analysis depends on the richness and evolutionary depth of the species being compared. A huge international effort has been put into sequencing the human genome, and chimp, rodents and some domestic mammals, chicken and fish. These species provide unique and important comparative information in studies of vertebrate phylogeny.

However, sequencing only eutherian mammals leaves a large gap in the phylogeny between the divergence of e.g. carnivores from primates about 80 MYA, and the divergence of mammals from the bird-reptile lineage 350 MYA. For many important comparisons with human, other eutherians are too close, and chicken is too far away. For the same reason, a marsupial (divergence time 180 MYA) has now been given high-priority for genome sequencing.

An even stronger link between mammals and birds/reptiles wound be provided by adding the platypus genome as a link between marsupials and birds.

As for comparisons between marsupials and eutherians, sequence comparisons between monotreme and eutherian mammals should allow relatively straightforward alignment of conserved features but also display a high conservation signal to random noise ratio, reducing the extent and degree of homology required to infer functionally conserved sequences. Proof of principle has recently been supplied by the first large scale sequencing of a marsupial genome region (a BAC that encompasses four human and five mouse and marsupial genes including LYL1 (Chapman et al, 2003)). While mouse-human comparisons showed high conservation across the whole region, marsupial-human and marsupial-mouse comparisons showed reduced homology, and it was easy to identify all promoters and exons, as well as putative transcription factor binding sites consistent with those of the better studied paralogue SCL.

Comparisons between distantly related mammals can also reveal new human genes. There has not yet been an opportunity to compare human and monotreme genomes on a large scale. However, the discovery of new human genes including *RBMX*, a candidate for human X-linked mental retardation (Delbridge et al. 1999) and several related genes (Lingenfelter et al. 2000) that result from initial comparisons between human and marsupial genomes suggests that important discoveries will be made from similar comparisons between the monotreme genome and those of other mammals, especially human.

The cloning and characterisation of immunoglobulin and major histocompatibility complex (MHC) cDNAs from the platypus and echidna has enabled a comparison of these immunologically significant genes across all three mammalian lineages (Belov et al. 2003c). Recent studies have found that the immunological innovations that resulted in the complex 'mammalian' immune response occurred prior to the separation of the three mammalian lineages. All three groups use IgG and IgE and a compact IgA, while birds and lower vertebrates do not (Belov et al. 2003a). However, these studies are also finding many differences between the antigen receptors of the monotremes and the marsupials and eutherians. Preliminary studies on the MHC have shown that although the class II genes of the three mammalian lineages evolved from

a common ancestral gene, they are not orthologues. Instead the locus has been evolving dynamically, with a rapid turnover of genes (Belov et al. 2003c). Elucidation of these genes in monotremes will allow us to gain a better understanding of the evolutionary processes that lead to the eutherian (human) immune system.

In addition to its importance in comparing phylogenetic footprints, the sequenced monotreme genome will allow further testing of hypotheses of mammal gene organization and function, which may help to realign our knowledge and understanding of sex chromosomes, sex determination as well as the origin and evolution of X inactivation and genomic imprinting in mammals (Killian ey al. 2000 and Killian et al. 2001, Murphy and Jirtle, 2003). This has been an important outcome of genetic studies of marsupial where genomic comparisons have resulted in new interpretations of mammalian sex determination, recombination and X chromosome inactivation.

Evolution of regulatory systems and complex pathways Monotremes, like marsupials, are particularly valuable because although they share with eutherians many mammal-specific regulatory systems, genetic differences between these distantly related species can be exploited to analyze how pathways and complex regulatory systems were built up, and indirectly, how they work. Recent investigations of the sex-determination pathway, X-chromosome inactivation and genomic imprinting exemplify the insights to be gained.

Comparative studies of genes involved in testis determination have helped to establish the relationship of the steps they control. Isolation and characterization of the marsupial – and even the alligator – homologues of these genes has been particularly revealing (Pask et al. 2001, Western et al. 2000). Several genes involved in eutherian sex determination and spermatogenesis have already been isolated and characterized in platypus, including *ATRX*, *DMRT1*, *UBE1* and *RBMY*.

Both the X inactivation control region and imprinted domains of the human genome contain complex and enigmatic controls of gene expression, and involve unknown interactions of functional RNA molecules and chromatin. Comparative analysis of such complex regions will highlight conserved features for direct experimental investigation. X-chromosome inactivation will be a particularly fruitful area of investigation, since dosage compensation occurs apparently in the absence of the *XIST* controlling locus. Some molecular mechanisms are likely to be shared between human and platypus, whereas others will be specific to eutherians and may therefore be responsible for eutherian-specific features like randomness. Elucidation of the compensation pathway in monotremes may reveal the basic mechanism of dosage compensation from which the complex human X inactivation system (and perhaps also genomic imprinting) evolved.

In the same way, analysis of genomic imprinting in monotremes may show us how some autosomal genes are expressed only if they come from the mother, or only from the father in humans. Because monotremes lay eggs and therefore have minimal maternal investment in the embryo, they may provide an answer to the real question about imprinting – *why* did imprinting evolve?

Potential commercial and conservation value

Pharmaceutical Industry: Sequencing the platypus genome will identify highly conserved genes and regulatory signals. Identical expression profiles across human and platypus will indicate ancient and essential functions of interest to pharmaceutical companies, such as anchor genes to link in into processes such as drug metabolism and autoimmune reactions. It is also likely that the project will discover new human genes that would be of immediate interest to the pharmaceutical industry.

Vaccine Production: Investigators may also use monotreme molecules for producing vaccines for eutherians. The recombinant molecules are different enough for the eutherian (human) body to see them as foreign, but similar enough for them to be manipulated, refolded on the bench. The presence of a genome sequence would facilitate efficient cloning of recombinant molecules.

Reagent development: A significant issue for the many scientists working on monotremes worldwide is the poor availability of fundamental reagents that may be obtained over the counter for primates, rodents and farm animals. For example, reagents for localising cell surface markers could be developed once genomic sequence is available using anti-peptide antibodies. Similarly, immunological reagents for studying basic immune responses or responses to disease or pathogens, could be developed.

•*Monotreme venom:* The seasonal secretion from the crural gland in the platypus is one of very few mammalian venoms. The only other group to produce active venoms that have been partly characterised are three members of the shrew family. The venom is well known in humans to produce immediate local swelling and "insupportable pain" around the site of envenomation, loss of limb function and swelling of adjacent lymph nodes. Recent studies of platypus venom have identified various unique proteins and peptides, including four defensin-like polypeptides (Torres *et al.* 2000), a protein with hyaluronidase activity (Temple-Smith 1974, de Plater *et al.* 1995) and a C-type natriuretic peptide (Kourie 1999). The venom gene has been cloned in the Graves laboratory (Kirby, unpublished). These proteins and peptides have potential pharmaceutical significance, for example in the treatment management of pain – a significance cause of human morbidity. Discovery of the genes coding for these venom components and for the crural system components will not only have medical implications but also provide further markers to probe the evolution of this system in monotremes.

•*Monotreme conservation and management:* The platypus and short-beaked echidna are common throughout their ranges in Australia. However since the platypus is the last in a long lineage of ornithorhynchid monotremes it is important that their wide distribution and common status does not result in complacent attitudes to the continued survival of this species (Grant and Temple-Smith 2003). Water quality issues and disease are always a threat. Limited genetic studies of family and population structure have been useful (Gemmell et al. 1992, Gemmell et al. 1995), and the availability of sequence with which to design markers would be valuable.

The long-beaked echidna in New Guinea is, by contrast, highly endangered and

currently threatened with extinction. The three extant monotreme species are remnants of a much wider monotreme radiation that extended across Gondwanaland with a range of unusual genera and species that are now represented only in the fossil record. Knowledge of the genome of an extant monotreme will provide important evolutionary information for their conservation and also for the management of the remaining populations, especially the Niugini long-beaked echidna. Recently, Flannery *et al.* (1998) reviewed long-beaked echidna taxonomy, erected two new species from museum specimens and suggested that three new subspecies of *Z. bruijni* be recognised. A detailed genetic analysis of *Zaglossus* using genomic sequences from the platypus genome project would provide more definition for this specific analysis – a critically important outcome for long beaked echidna conservation strategies since the single species *Z. bruijni* was already regarded as critically endangered.

Choice of the platypus as model monotreme

There is little choice in adopting a model monotreme, since only two (of the grand

However, the easy culture of apparently immortal diploid lines of platypus fibroblasts makes it a simple matter to provide living cells to overseas investigators, and increasing numbers have been provided with platypus cell cultures, tissues, DNA or gene sequences over the last few years.

The monotreme genetics community

Not surprisingly, most of the significant work on monotreme biology, including genetics and genomics, has been done in Australia. Leaders in the field (authors of this proposal, Temple-Smith, Renfree, Graves) operate or are aligned with other major research groups with significant interests in monotreme and marsupial biology. There is an extremely active monotreme group that operates within the Australian Mammal Society, meeting regularly and publishing up-to-date books and journal issues on monotreme biology (e.g. Augee 1992 and *in preparation*, Nichol (2003).

Monotreme tissues, DNA, clones and sequences 2 11.2 23y sought byg nuegros.

intenrational and Australiancollaboaratioes,fjorinstaencn the ollowhing(intenrational) Tj 0 -13.5 TD -0.2038 Tc .71103 T
-Ggenva),

Monotreme cell hybrids. Cell hybrids between monotremes and other mammalian species would provide a valuable gene mapping resource. Five platypus hybrids bearing regions of platypus chromosomes were produced for early gene mapping studies (Wrigley and Graves, 1987).

Flow-sorted and microdissected chromosomes, and regional libraries. Platypus chromosomes have been individually sorted in collaboration with Prof Malcolm Ferguson-Smith (Cambridge). Six of the small chromosomes, and chromosome arms have been microdissected. DOP-PCR amplified chromosome paints have been used to explore homologies between unpaired platypus chromosomes (Grützner et al. unpublished), between platypus and echidna chromosomes, and to discover the position of chromosomes in sperm and in interphase cells (Greaves et al. 2002).

Physical maps. Thirty genes have been physically mapped by FISH in platypus and eight in the echidna (reviewed by Grützner et al. in press). An updated standard karyotype has been assembled for platypus as well as both echidna species (Graves et al. 2003). This provides a skeleton on which to hang a more detailed physical map.

DNA and libraries. Several platypus genomic and four cDNA libraries (including spleen) have been prepared. Genomic libraries are routinely prepared. A BAC library is available through Dr R, Jirtle (personal communication) and a 10x BAC library has been recently completed at the University of Arizona and is publicly accessible from Clemson labs in the US.

Monotreme genome sequencing strategy and cost

Experience and insights gained while recently mapping and sequencing the mouse and chimpanzee genomes lead us to propose a whole genome shotgun (WGS) strategy with the potential to turn into a combined clone-based strategy for the platypus genome.

This choice is based on the fact that the platypus genome is known to have some peculiar characteristics, including unpaired chromosomes containing roughly 13% of the genome (6.5% in female). An 8.0-fold whole genome shotgun component consisting of both small and large insert clones (see Table 1) will provide the community of users with rapid access to most of the platypus genome sequence. The relatively high coverage of 8.0x is chosen based on the fact that we wish to generate at least 4x coverage of the unpaired platypus genome regions. We expect the large insert clones to provide good linkage for a well ordered and oriented assembly.

To anchor the major sequence supercontigs to platypus chromosomes we propose to use FISH, mapping roughly 2 clones for each of the largest 200 supercontigs as well as a small number of clones to resolve questions surrounding the unpaired chromosomes and their heterozygous translocations. A BAC clone-based physical map generated by restriction fingerprinting, will serve to validate the assembly as well as to provide clone contigs for walking into regions where the whole genome coverage might be low. We expect that the cost of generating the physical map and 8.0x sequence coverage would be approximately $43M. As with the mouse genome sequence, we would expect that the proposed approach would result in anchored

supercontigs (sequence contigs connected by at least two read-pair links) of greater than 10 Mb in average length (MGSC, 2002).

Table 1. Proposed whole genome shotgun sequencing of the platypus genome.

Clone type	Insert size	No. of attempted reads	Seq coverage
plasmid	4 kb	28.0M	5.8x
plasmid	8 kb	8.1M	1.6x
Fosmid	40 kb	2.0M	0.4x
BAC	150 kb	0.6M	0.2x
TOTALS	---	38.7M	8.0x

Note: based on a genome size of 3.0Gb and an average sequence read length of 680 bp.

Since the platypus genome is known to have peculiar characteristics, it would be prudent to design a sequencing approach that provides a good deal of flexibility. For example, we would expect to perform initial assemblies of the genome sequence once approximately 2-fold and 4-fold coverage has been generated. If the 4-fold assembly suggested problems with a WGS approach, the remainder of the sequence coverage could be generated by constructing subclone libraries from mapped BAC clones at a minimal increase in cost. To ensure the availability of mapped clones anchored in the genome assembly all the Fosmids and BACs would be generated within the first 4x coverage.

To improve on the quality and continuity of the sequence, we would propose a subsequent round of computer-directed "pre-finishing" in which oligonucleotides are algorithmically selected to extend sequence contigs into gap and other low coverage regions. BAC and Fosmid clones would serve as the templates for the necessary sequencing reactions. The methods, computational tools and laboratory pipelines for automated finishing are already in place at the Washington University Genome Sequencing Center and the Whitehead Institiute/MIT Center for Genome Research. We would expect that with ~8.0x sequence coverage from the plasmids and large insert clones, a directed read would be required at 4 kb average intervals. The cost of this pre-finishing activity would be approximately $3M.

At this point, manual finishing could be employed for targeted regions (or the whole genome, although it does not seem necessary) to further improve contiguity and sequence accuracy. Our experience with finishing other genomes using this approach suggests that most of this work would involve using PCR to sequence and/or size regions that were missing, ambiguous or repetitive in content and might be considerably more expensive than the pre-finishing proposed above.

Choice of platypus to sequence

Despite the presence of unpaired regions in the genome, the X chromosome is present in two copies in a female platypus, suggesting that a female should be sequenced to give appropriate coverage to the important X chromosome. If the sequence of the Y chromosome is desired, male-specific element(s) can be sequenced separately by a clone-based approach.

The heterozygosity of the platypus is unknown at this point and may even be influenced by the translocational heterozygosity seen in the genome. To address this question WUGSC has started heterozygosity testing of platypus from the Graves and Temple-Smith laboratories.

In addition, the varying amount of heterochromatin seem between different platypus suggests that karyotypic examination to select an individual with low levels of heterochromatin would be a reasonable idea. Ideally, to make use of the existing BAC library, this particular individual should be examined for both heterozygosity rate and heterochromatin extent and chosen if possible.

SNPs
Several thousand SNPs that could be used both for mapping and conservation efforts could be generated both for platypus and the two echidna species by generating ~ 10,000 WGS reads for each species and aligning them to the platypus assembly. For the platypus, SNPs would be immediately discovered, whereas for the echidna species a smaller number of loci would be chosen and SNP discovery performed by re-sequencing a few individuals.

IV Conclusion

Just as these extraordinary mammals provide us with the last anatomical link to the bird-reptile lineage that spawned us all, their unique genomes will provide a molecular link to our past. As the earliest offshoot within the mammalian lineage, monotremes provide an invaluable outgroup to therian mammals. The structure of the monotreme genome has and will provide us with crucial information about ancestral features of our own genome and those of other therians. The comparison of genome sequence and gene maps will reveal ancient features of the mammalian genome (Graves et al, 1990).

Comparative mapping of genes and DNA sequences between monotreme, marsupial and eutherian species – and also between birds and monotremes – will indicate which parts of the genome have been conserved and which are ancestral, and which are unique to monotremes. Comparison of the molecular mechanisms that control sex determination, genomic imprinting and X chromosome inactivation will allow us to reconstruct the assembly of these complex systems, and to provide information on how they operate in all mammals.

In addition, the developmental system unique to the egg-laying mammals provides excellent opportunities to study the evolution of genomic imprinting and X chromosome inactivation, which is believed to have evolved in close relation to placental development and the differentiation of heteromorphic sex chromosomes.

The platypus genome represents a treasure trove of comparative genomics data. Analysis of individual genes, and of gene arrangement, has already contributed significantly to our understanding of human biology and genetics. Analysis on a genome-wide scale will provide sequences for identifying conserved genes, functional domains and regulatory elements. The unique value of platypus sequences for such comparisons will motivate the complete sequencing the platypus genome.

References

Augee, M.L. (1978) Ed: Monotreme Biology *Australian Zoologist* 20:1-257.

Augee, M.L. (1992) Ed: Platypus and Echidna *Proceedings of the Royal Society of New South Wales* pp1-198.

Baldwin, J. and Temple-Smith, P.D. (1973) Distribution of LDHX in mammals: presence in marsupials and absence in the monotremes platypus and echidna. *Comp Biochem Physiol.* 46B: 805-811.

Belov, K., Hellman, L., Cooper, D.W. (2002a) Characterisation of echidna IgM provides insights into the time of divergence of extant mammals. *Dev Comp Immunol* 26: 831-839

Belov, K., Hellman, L., Cooper, D.W. (2002b) Characterisation of IgG1 from a monotreme, *Tachyglossus aculeatus*. *Immunogenetics* 53: 1065-1071.

Belov, K., Zenger, K.R., Hellman, L., Cooper, D.W. (2002c) Echidna IgA supports mammalian unity and traditional Therian relationship. *Mamm Genome* 13:656-63.

Belov, K., Hellman, L. (2003a) Immunoglobulin genetics of *Ornithorhynchus anatinus* (platypus) and *Tachyglossus aculeatus* (short-beaked echidna). *Comparative Biochemistry and Physiology Part A* (in press).

Belov, K., Hellman, L. (2003b) Platypus IgM and the divergence of the two extant monotreme lineages. *Australian Mammalogy* 25: 87-94.

Belov, K., Lam, M.K., Hellman, L., Colgan, D.J. (2003c) Evolution of the major histocompatibility complex: Isolation of class II b cDNAs from two monotremes, the platypus and the short-beaked echidna. *Immunogenetics* 55: 402-11.

Bick, Y.A.E. and Sharman, G.B. (1975). The chromosomes of the platypus (*Ornithorhynchus*: Monotremata). *Cytobios* 14: 17-28.

Caldwell, H (1887) The embryology of Monotremata and Marsupialia Pt I. *Phil. Tran.s R. Soc.* B 178: 463-486.

Carrick FN and RL Hughes (1978) Reproduction in male monotremes. *Australian Zoologist* 20:211-232

Chapman M. A., *et al.* (2003) Comparative and functional analyses of LYL1 loci establish marsupial sequences as a model for phylogenetic footprinting. *Genomics* 81: 249-59.

Cooper, D.W., Johnston, P.G., Graves. J.A.M., and Watson, J.M. (1993) X-inactivation in marsupials and monotremes. *Sem Dev Biol* 4: 117-128.

Flannery, T.F., Groves, C.P. (1998) A revision of the genus Zaglossus (Monotremata, Tachyglossidae), with description of new species and subspecies. *Mammalia* 62(3): 367-396.

Gemmell, N.J., Grant, T.R., Western, P.S., Watson, J.M., Murray, N.D. and Graves, J.A.M. (1992) Preliminary molecular studies of platypus family and population structure. In *"Platypus and Echidnas"*, M. Augee (ed), Royal Soc. of NSW pp 277-284.

Gemmell, N.J., Grant, T.R., Western, P.S., Watson, J.M., Murray, N.D. and Graves, J.A.M. (1995). Determining platypus relationships. *Aust J Zool* 43: 283-291.

Grant, T.R. (1989) The Platypus. UNSW Press:Kensington 73pp.

Grant, T.R. and Temple-Smith, P.D. (1998) Field biology of the platypus: current and historical perspectives. *Phil.Trans.Roy.Soc.Lond.* B 353: 1081-1090.

Grant, T.R and Temple-Smith, P.D. (2003) Conservation of the platypus: threats and challenges. *Aquatic.Ecosystem Health and Management* 6: 5-18

Grant, T.R., Griffiths, M., Temple-Smith, P.D. (in press) Breeding in a Free-ranging population of Platypuses, *Ornithorhynchus anatinus*, in the upper Shoalhaven River, New South Wales - a 26 Year Study. *J Linn Soc NSW.*

Graves, J.A.M., Hope, R.M. and Cooper, D.W. (eds) (1990) *"Mammals from Pouches and Eggs: Genetics, Breeding and Evolution of Marsupials and Monotremes"*. CSIRO Press, Melbourne, 337pp.

Graves, J.A.M. and Shetty, S. (2001) Sex from W to Z – evolution of vertebrate sex chromosomes and sex determining genes. *J. Exptl. Zool.* 281: 472-481.

Graves, J.A.M. (1996). Mammals that break the rules: Genetics of marsupials and monotremes. *Annual Rev. Genet.* 30:233–260.

Graves, J.A.M. (1998) Gene maps of monotreme mammals (mammalian subclass Prototheria). *Inst Lab Animal Res J* 39: 225-228.

Graves, J.A.M. (2002) The rise and fall of *SRY. Trends in Genetics* 18: 259-264.

Graves, J.A.M., Watson, J.M., Grützner, F. (accepted) Mitotic chromosomes of the platypus (*Ornithorhynchus anatinus*, Order Monotremata of Mammalian Subclass Prototheria). Mammalian Chromosome Atlas, S.J. O'Brien (ed).

Greaves, I.K., Rens, W., Ferguson-Smith, M.A., Griffin, D., Graves. J.A.M. (2003) Conservation of

chromosome arrangement and position of the X in mammalian sperm suggests functional significance. *Chromosome Research* 11: 503-512.

Gregory, J.E., Iggo, A., McIntyre, A.K., Proske, U. (1987) Electroreceptors in the platypus. Nature. 6111: 386-387.Griffiths, M. (1968) *Echidnas* Pergamon.

Griffiths, M. (1978) *The Biology of the Monotremes* Academic Press 367pp.

Grützner, F., Deakin J., Rens, W., El-Mogharbel, N., Graves, J.A.M. The monotreme genome: a patchwork of reptile, mammal and unique features? *Comp Biochem Physiol,* In press

Jones, R.C. and Djakiew, D. (1978) The role of the excurrent ducts from the testes of tesicond mammals *Australian Zoologist* 20: 201-210.

Killian, J.K., Byrd, J.C, Jirtle, J.V., Munday, B.L., Stoskopf, M.K., and Jirtle RL (2000). M6P/IGF2R imprinting evolution in mammals. *Mol. Cell* 5: 707-716.

Killian, J.K., Nolan, C.M., Stewart, N., Munday, B.L. Andersen, N.A., Stewart Nicol, S., and Jirtle RL (2001). Monotreme IGF2 expression and ancestral origin of genomic imprinting. *J. Exp. Zoology* 291:205-212.

Kirby, P.J., Waters, P.D, Delbridge, M.L., Svartman, M., Stewart, A.N., Nagai, K. and Graves, J.A.M. (2002) Mapping *SOX2* and *SOX14* in the platypus. *Cytogenetics and Genome Research* 98: 96-100

Kourie, J.I. (1999) Characterization of a C-type natriuretic peptide (CNP-39)-formed cation-selective channel from platypus (Ornithorhynchus anatinus) venom. *J Physiol* 518: 359-369

Lin, M. and Jones, R.C. (2000) Spermiogenesis and spermiation in a monotreme mammal, the platypus, Ornithorhynchus anatinus. *J Anat.* 196: 217-32

McKay, L.M., Wrigley, J.M. and Graves, J.A.M. (1988) Evolution of mammalian X chromosome inactivation: sex chromatin in monotremes and marsupials. *Aust J Biol Sci.* 40: 397-404

Miska, K.B., Harrison, G.H., Hellman, L., Miller, R.D. (2002) The Major Histocompatibility Complex in monotremes: An analysis of the evolution of Mhc class I genes across all three mammalian subclasses. *Immunogenetics* 54: 381-393

Murtagh, C.E. (1977) A unique cytogenetic system in monotremes. *Chromosoma* 65: 37-57

Murphy, S.K., and Jirtle, R.L. (2003)Imprinting Evolution and the Price of Silence. *BioEssays* 25: 577-588.

Nicholl, S. (ed) (2003) Special issue on Monotreme Biology, *Comp Biochem Physiol* (in press)

Pettigrew, J.D., Manger, P.R. and Fine, S.L.B. (1998) The sensory world of the platypus. *Phil Trans R Soc Lond* B 353: 1199-1210.

Pask, A. and Graves, J.A.M. (2001) Sex chromosomes and sex determining genes – insights from marsupials and monotremes. In G. Scherer (ed*) "Genes and mechanisms in vertebrate sex determination*." Birkenhauser, Berlin pp 71-95

Proske, U., Gregory, J.E. Iggo, A. (1998) Sensory receptors in monotremes. *Phil Trans R Soc Lond* B 353: 1187-1198.

Renfree, M.B. (1993) Ontogeny, genetic control and phylogeny of female reproduction in monotreme and therian mammals. In: *Mammal Phylogeny* Eds. F.S.

Renfree, M.B. (1995) Monotreme and Marsupial Reproduction. *Reprod Fert Dev* 7: 1003-1020.

Spencer, J.A., Watson, J.M., Lubahn, D.B., Joseph, D.R., French, F.S., Wilson, E.M. and Graves, J.A.M. (1991) The androgen receptor gene is located on a highly conserved region of the X chromosome of marsupial and monotreme, as well as eutherian marsupials. *Journal of Heredity* 82: 134-139

Szalay, J.J. Novacek and M.C. McKenna, Springer-Verlag, New York. pp 4-20

Scheich, H., Langner, G., Tidemann, C., Coles, R.B., Guppy, A. (1986) Electroreception and electrolocation in platypus. *Nature* 319: 401-402

Rismiller, P.D. and McElvey, M.W. (2000) Frequency of breeding and recruitment in the short-beaked echidna Tachyglossus aculeatus J Mammal 81: 1-17

Temple-Smith, P.D. (1973) Seasonal breeding biology of the platypus, *Ornithorhynchus anatinus* (Shaw1799), with special reference to the male. PhD Thesis, Australian National University, 367pp.

Temple-Smith, P.D. and Grant, T.R. (2001) Uncertain breeding: a short history of reproduction in monotremes. *Reprod Fertil Dev* 13: 487-497.

Torres, A.M., de Plater, G,M., Doverskog, M., Birinyi-Strachan, L. C., Nicholson, G.M., Gallagher, C. H., Kuchel, P.W. (2000) Defensin-like peptide-2 from platypus venom: member of a class of peptides with a distinct structural fold. *Biochemical Journal* 348: 649-656

Tyndale-Biscoe, C.H. and Renfree, M.B. (1987) *Reproductive Physiology of Marsupials*. Cambridge University Press, Cambridge.

Walton, D. and Richardson, B. (1989) *Fauna of Australia.* Australian Government Printer, Canberra, Australia.

Waters, P., Duffy, B., Frost, C.J., Delbridge, M.L. and Graves, J.A.M. (2001) The human Y chromosome derives largely from a single autosomal region added 80-130 million years ago. *Cytogenet Cell Genet* 92: 74–79

Watson, J.M., Spencer, J.A., Riggs, A.D. and Graves, J.A.M. (1990) The X chromosome of monotremes shares a highly conserved region with the eutherian and marsupial X chromosomes, despite the absence of X chromosome inactivation. *Proc Natl Acad Sci U.S.A.* 87: 7125-7129

Watson, J.M., Spencer, J.A., Riggs, A.D. and Graves, J.A.M. (1991). Sex chromosome evolution: Platypus gene mapping suggests that part of the human X chromosome was originally autosomal. *Proc Natl Acad Sci U.S.A.* 88: 11256-11260

Watson, J.M., Spencer, J.A., Graves, J.A.M., Snead, M.L. and Lau, E.C. (1992a) Autosomal localization of the amelogenin gene in monotremes and marsupials: implications for mammalian sex chromosome evolution. *Genomics* 14: 785-789

Watson, J.M., Riggs, A.D. and Graves, J.A.M. (1992b) Gene mapping studies confirm the homology between the platypus X and echidna X_1 chromosomes and identify a conserved ancestral monotreme X. *Chromosoma* 101: 596-601

Watson, J.M., Meyne, J. and Graves, J.A.M. (1992c) Studies of the chromosomes of the echidna meiotic translocation chain. In *"Platypus and Echidnas"*, M. Augee (ed), Royal Soc. NSW pp53-63

Watson, J.M., Frost, C., Spencer, J.A. and Graves, J.A.M. (1993) Sequences homologous to the human X-and Y-borne zinc finger protein genes (*ZFX/Y*) are autosomal in monotreme mammals. *Genomics* 15: 317-322

Watson, J.M., Meyne, J. and Graves, J.A.M. (1996) Ordered tandem arrangement of chromosomes in the sperm heads of monotreme mammals. *Proc Natl Acad Sci USA* 93: 10200-10205

Western, P.S., Harry, J.L. Graves, J.A.M. and Sinclair, A.H. (2000) Temperature dependent sex determination: the expression of SF1, WT1 and DAX1 during gonadogenesis. *Genes* 241: 223-232.

Woodburne, M.O., Rich, T.H., Springer, M.S. (2003) The evolution of tribospheny and the antiquity of mammalian clades. Molec Phylogenet Evol 28: 360-385.

Wrigley, J.M. and Graves, J.A.M. (1984) Two monotreme cell lines derived from female platypuses. (*Ornithorhynchus anatinus*; Monotremata, Mammalia). *In Vitro.* 20: 321-328

Wrigley, J.M. and Graves, J.A.M. (1988a) Karyotypic conservation in the mammalian Order Monotremata (subclass Prototheria). *Chromosoma* 96: 231-247

Wrigley, J.M. and Graves, J.A.M. (1988b) Sex chromosome homology and incomplete, tissue-specific inactivation suggest that monotremes represent an intermediate stage of mammalian sex chromosome evolution. *J. Heredity* 79: 115-118 (cover story)