The Mammalian Genome

Presented by: Heba M. Al Sherif Bashayer Jamal Shaikhah Al Shareefi Amna Dawoud Hana Hussain

Introduction to genomics (485) Fall 2019





Outline:

- The genome papers
- Introducing mammals
- Facts about the species
- Sequencing strategy
- Sequencing methods
- Assembly
- Annotation
- Interesting genomic fact
- Questions

The genome papers



ARTICLE

Received 30 Nov 2015 | Accepted 1 Apr 2016 | Published 17 May 2016

3/ncomms11519 OPEN

Giraffe genome sequence reveals clues to its unique morphology and physiology

Morris Agaba^{1,2,3}, Edson Ishengoma¹, Webb C. Miller³, Barbara C. McGrath³, Chelsea N. Hudson³, Oscar C. Bedoya Reina^{3,4}, Aakrosh Ratan^{3,5}, Rico Burhans³, Rayan Chikhi^{6,7}, Paul Medvedev^{6,7}, Craig A. Praul⁸, Lan Wu-Cavener³, Brendan Wood³, Heather Robertson⁹, Linda Penfold¹⁰ & Douglas R. Cavener^{1,3}

Vol 453 8 May 2008 doi:10.1038/nature06936

ARTICLES

nature

OPEN

Genome analysis of the platypus reveals unique signatures of evolution

ARTICLES https://doi.org/10.1038/s41588-018-0153-5	genetics
---	----------

Adaptation and conservation insights from the koala genome

Rebecca N. Johnson^{® 1,2,30,31*}, Denis O'Meally^{2,3,30}, Zhiliang Chen^{4,30}, Graham J. Etherington⁵, Simon Y. W. Ho^{®2}, Will J. Nash⁵, Catherine E. Grueber^{® 2,6}, Yuanyuan Cheng^{2,7}, Camilla M. Whittington⁸, Siobhan Dennison¹, Emma Peel², Wilfried Haerty⁵, Rachel J. O'Neill⁹, Don Colgan¹, Tonia L. Russell¹⁰, David E. Alquezar-Planas¹, Val Attenbrow¹, Jason G. Bragg^{11,12}, Parice A. Brandies², Amanda Yoon-Yee Chong^{5,13}, Janine E. Deakin¹⁴, Federica Di Palma^{5,15}, Zachary Duda⁹, Mark D. B. Eldridge¹, Kyle M. Ewart¹, Carolyn J. Hogg², Greta J. Frankham¹, Arthur Georges¹⁴, Amber K. Gillett¹⁶, Merran Govendir⁸, Alex D. Greenwood^{17,18}, Takashi Hayakawa^{19,20}, Kristofer M. Helgen^{1,21}, Matthew Hobbs^{© 1}, Clare E. Holleley²², Thomas N. Heider⁹, Elizabeth A. Jones⁸, Andrew King¹, Danielle Madden³, Jennifer A. Marshall Graves^{11,14,23}, Katrina M. Morris²⁴, Linda E. Neaves^{© 1,25}, Hardip R. Patel²⁶, Adam Polkinghorne³, Marilyn B. Renfree^{© 27}, Charles Robin^{© 27}, Ryan Salinas⁴, Kyriakos Tsangaras²⁸, Paul D. Waters⁴, Shafagh A. Waters⁴, Belinda Wright^{1,2}, Marc R. Wilkins^{4,10,30}, Peter Timms^{29,30} and Katherine Belov^{2,30,31} Renfree et al. Genome Biology 2011, **12**:R81 http://genomebiology.com/2011/12/8/R81



Open Access

RESEARCH

Genome sequence of an Australian kangaroo, *Macropus eugenii*, provides insight into the evolution of mammalian reproduction and development

Marilyn B Renfree^{1,2*†}, Anthony T Papenfuss^{1,3,4*†}, Janine E Deakin^{1,5}, James Lindsay⁶, Thomas Heider⁶, Katherine Belov^{1,7}, Willem Rens⁸, Paul D Waters^{1,5}, Elizabeth A Pharo², Geoff Shaw^{1,2}, Emily SW Wong^{1,7}, Christophe M Lefèvre⁹, Kevin R Nicholas⁹, Yoko Kuroki¹⁰, Matthew J Wakefield^{1,3}, Kyall R Zenger^{1,7,11}, Chenwei Wang^{1,7}, Malcolm Ferguson-Smith⁸, Frank W Nicholas⁷, Danielle Hickford^{1,2}, Hongshi Yu^{1,2}, Kirsty R Short^{1,2}, Hannah V Siddle^{1,7} Stephen R Frankenberg^{1,2}, Keng Yih Chew^{1,2}, Brandon R Menzies^{1,2,13}, Jessica M Stringer^{1,2}, Shunsuke Suzuki^{1,2}, Timothy A Hore^{1,14}, Margaret L Delbridge^{1,5}, Amir Mohammadi^{1,5}, Nanette Y Schneider^{1,2,15}, Yangiu Hu^{1,2}, William O'Hara⁶, Shafagh Al Nadaf^{1,5}, Chen Wu⁷, Zhi-Ping Feng^{3,16}, Benjamin G Cocks¹⁷, Jianghui Wang¹⁷, Paul Flicek¹⁸, Stephen MJ Searle¹⁹, Susan Fairley¹⁹, Kathryn Beal¹⁸, Javier Herrero¹⁸, Dawn M Carone^{6,20}, Yutaka Suzuki²¹, Sumio Sugano²¹, Atsushi Toyoda²², Yoshiyuki Sakaki¹⁰, Shinji Kondo¹⁰, Yuichiro Nishida¹⁰, Shoji Tatsumoto¹⁰, Ion Mandiou²³, Arthur Hsu^{3,16}, Kaighin A McColl³, Benjamin Lansdell³, George Weinstock²⁴, Elizabeth Kuczek^{1,25,26}, Annette McGrath²⁵, Peter Wilson²⁵, Artem Men²⁵, Mehlika Hazar-Rethinam²⁵, Allison Hall²⁵, John Davis²⁵, David Wood²⁵, Sarah Williams²⁵, Yogi Sundaravadanam²⁵, Donna M Muzny²⁴, Shalini N Jhangiani²⁴, Lora R Lewis²⁴ Margaret B Morgan²⁴, Geoffrey O Okwuonu²⁴, San Juana Ruiz²⁴, Jireh Santibanez²⁴, Lynne Nazareth²⁴, Andrew Cree²⁴, Gerald Fowler²⁴, Christie L Kovar²⁴, Huyen H Dinh²⁴, Vandita Joshi²⁴, Chyn Jing²⁴, Fremiet Lara²⁴, Rebecca Thornton²⁴, Lei Chen²⁴, Jixin Deng²⁴, Yue Liu²⁴, Joshua Y Shen²⁴, Xing-Zhi Song²⁴, Janette Edson²⁵, Carmen Troon²⁵, Daniel Thomas²⁵, Amber Stephens²⁵, Lankesha Yapa²⁵, Tanya Levchenko²⁵, Richard A Gibbs²⁴, Desmond W Cooper^{1,28} Terence P Speed^{1,3}, Asao Fujiyama^{22,27}, Jennifer A M Graves^{1,5}, Rachel J O'Neill⁶, Andrew J Pask^{1,26}, Susan M Forrest^{1,25} and Kim C Worley²⁴

Vol 463|21 January 2010|doi:10.1038/nature08696

ARTICLES

nature

The sequence and *de novo* assembly of the giant panda genome

Ruiqiang Li^{1,2}*, Wei Fan¹*, Geng Tian^{1,3}*, Hongmei Zhu¹*, Lin He^{4,5}*, Jing Cai^{3,6}*, Quanfei Huang¹, Qingle Cai^{1,7}, Bo Li¹, Yinqi Bai¹, Zhihe Zhang⁸, Yaping Zhang⁶, Wen Wang⁶, Jun Li¹, Fuwen Wei⁹, Heng Li¹⁰, Min Jian¹, Jianwen Li¹, Zhaolei Zhang¹¹, Rasmus Nielsen¹², Dawei Li¹, Wanjun Gu¹³, Zhentao Yang¹, Zhaoling Xuan¹, Oliver A. Ryder¹⁴, Frederick Chi-Ching Leung¹⁵, Yan Zhou¹, Jianjun Cao¹, Xiao Sun¹⁶, Yonggui Fu¹⁷, Xiaodong Fang¹, Xiaosen Guo¹, Bo Wang¹, Rong Hou⁸, Fujun Shen⁸, Bo Mu¹, Peixiang Ni¹, Runmao Lin¹, Wubin Qian¹, Guodong Wang^{3,6}, Chang Yu¹, Wenhui Nie⁶, Jinhuan Wang⁶, Zhigang Wu¹, Huiqing Liang¹, Jiumeng Min^{1,7}, Qi Wu⁹, Shifeng Cheng^{1,7}, Jue Ruan^{1,3}, Mingwei Wang¹, Zhongbin Shi¹, Ming Wen¹, Binghang Liu¹, Xiaoli Ren¹, Huisong Zheng¹, Dong Dong¹¹, Kathleen Cook¹¹, Gao Shan¹, Hao Zhang¹, Carolin Kosiol¹⁸, Xueying Xie¹³, Zuhong Lu¹³, Hancheng Zheng¹, Yingrui Li^{1,3}, Cynthia C. Steiner¹⁴, Tommy Tsan-Yuk Lam¹⁵, Siyuan Lin¹, Qinghui Zhang¹, Guoqing Li¹, Jing Tian¹, Timing Gong¹, Hongde Liu¹⁶, Dejin Zhang¹⁶, Lin Fang¹, Chen Ye¹, Juanbin Zhang¹, Wenbo Hu¹⁷, Anlong Xu¹⁷, Yuanyuan Ren¹, Guojie Zhang^{1,3,6}, Michael W. Bruford¹⁹, Qibin Li^{1,3}, Jing Thao¹, Ning Qu^{1,7}, Shancen Zhao¹, Feng Tian¹, Xiaoling Yua², Shiping Lu²³, Hemin Zhang²⁴, Desheng Li²⁴, Yan Huang²⁴, Xia Wang¹, Guohua Yang¹, Zhi Yan¹, Jiang¹, Jiunyi Wang¹, Nan Qin¹, Li ¹, Jingxiang Li¹, Lizhi Xu¹, Xiao Liu¹, Tomas Vinar²⁰, Yajun Wang²¹, Tak-Wah Lam²², Junyi Wang¹, Nan Qin¹, Li Li¹, Jingxiang Li¹, Lizri Solund¹, Karsten Kristiansen^{1,2}, Gane Ka-Shu Wong^{1,25}, Maynard Olson²⁶, Xiuqing Zhang¹, Songgang Li¹, Huanming Yang¹, Jian Wang¹ & Jun Wang^{1,2}

Introducing Mammals

- Mammals are endothermic animals.
- Special Features: Hair or fur and mammary glands.
- Most mammals give birth to live young.
- Have the most genetic similarity to human.



Giraffa camelopardalis tippelskirchi

Masai Giraffe

By: Heba M. Al Sherif



- Giraffa camelopardalis tippelskirchi.
- Okapia johnstoni.
- Tallest living animal.
- Neck length reaches up to 2 meters.
- Affects the cardiovascular system, the musculoskeletal system and the

nervous system.



- 70 genes exhibit multiple signs of adaptation.
- HOX, NOTCH and FGF signaling pathways
- FGFRL1 is the strongest candidate for unique growth.





- Liver samples for two female giraffes and one male okapi.
- Sequencing strategy: WGS
- Sequencing method: Illumina
- Assembly: SOAPdenovo2 and GapCloser
- Chromosomes: Giraffe 30 and Okapi 44-46
- Predicted genes: 17K



	Genome size	Coverage	GC%	# Contigs	N50 contig	# Scaffolds	N50 Scaffold
Giraffe	2.9 Gb	30x	41.5%	2 M	47 Kb	2 M	339.6 Kb
Okapi	3.3 Gb	30x	42.3%	3.8 M	33.7 Kb	3.7 M	104.6 Kb



Interesting Facts

- Giraffe has to spread it's front legs to be able to
 - reach the ground.
- Giraffe spots are like
 - human fingerprints.





Questions

What are the most distinctive features of mammals?

- 1. Covered in hair or fur
- 2.Mammary glands

What are the systems most effected by the giraffe height?

- 1. Cardiovascular system
- 2. Musculoskeletal system
- 3. Nervous system



Ailuropoda melanoleura.

"Giant panda"

By: Hana Hussain



- The *Ailuropoda melanoleura* simply known as pandas; is a species of bear that have a distinctive black and white fur.
- Commonly found in the mountains of central and western China.
- They are good at climbing trees; swimming and they spend 14 hours eating per day.
 - They are one of the rarest animal and is under immense threat from habitat loss.
 - The panda genome contains 20 pairs of autosomes and one pair of sex chromosomes (2n=42).



Genome facts:

- Taste is an important factor in the development of dietary habits.
- The five components of basic taste one of them is umami, which can be tasted in foods like meats, cheese, broth, stock which contain a high level of the amino acid and glutamate.
- Lack of *T1R1* gene prevent the panda from expressing a functional umami taste receptor, which may partly explain why the panda diet is primarily herbivorous despite its taxonomic classification as a carnivore.
- The panda appears to have all the genes needed for a carnivorous digestive system but lacks digestive cellulase genes, therefore they depend on its gut microbiome to digest the bamboo.



≻Sources of DNA:

 Peripheral venous blood was collected 3-year-old female giant panda.

≻The Sequence stargates:

• Whole-genome shotgun sequencing strategy was used.

≻The Sequence methods:

• Illumina Genome Analyser sequencing technology was used.

≻The Assembly:

• The genome sequence was assembled with short reads using SOAP de-novo software.

≻The Annotation:

- Genome size = 2.4 Gb
- % of the genome that contains genes = 3.2%
- GC content= Between 20% and 80%.







Contig N50	Scaffold N50	Genome size	GC%	% of the genome that contains genes	T.E.	LTR	LINE	SINE	# SNP
40 kb	1.3 Mb	2.4 Gb	Between 20% and 80%.	3.2 %	36.2%	5.6%	18.2%	7.9%	2.7 million heterozygous SNPs.



Question:

Why the panda diet is primarily herbivorous despite its taxonomic classification as a carnivore?

 \checkmark The panda lack *T1R1* gene that prevent them from expressing a functional umami taste receptor.







" Platypus" Ornithorhynchus anatinus

By : Bashayer Jamal.



Platypus

Order: Monotremata

Common name: Platypus (*Platy:* flat, *Pus:* foot)

Scientific name:

Ornithorhyncus anatinus. (Ornitho: bird, rhyncus: snout. Anatinus: duck-like).

Habitat: Fresh water (Australia).



Biology of Platypus



Platypus Genome





Interesting Fact

Hunting

- Under the water, platypus's eyes & ears are closed. They locate their food using the electroreceptors (in the bill).
- Platypus and Cormorant feeding together, the cormorant pecks the platypus to make it dive for food.

 The X-chromosome is very similar to the Zchromosome in the chicken, but not similar to human X.







Assembly & Annotation

Assembly	Annotation
Genome size: 2.7 Gb	LTR: 13,723
GC content: 45.5%	LINE: 2,491,710
Sequencing strategy: BAC & WGS	SINE: 2,529,582
Assembly method: PCAP & FISH	Low complexity: 52,125
Sequencing coverage: 6X	
Contig N50: 39.589 kb	
Scaffold N50: 298 kb	
Number of contigs: 177.028	
Number of scaffold: 61.239	



Questions

- 1) Why it is a <u>semi-aquatic</u> (neither terrestrial nor aquatic)?
- Olfactory receptors & the males venomous.
- Webbed feet.
- 2) Can you eat platypus?
- No, the platypus is poisonous so it's not edible.
- 3) What about their egg?
- You can, but they are only <u>11mm in diameter</u>, it would take a lot of them to make a meal.



"Australian kangaroo" Macropus eugenii.

by: Amna Dawoud



Facts about the specie:

Common name: Tammar wallaby

Order: Diprotodontia **Scientific name**: macropus eugenii

• "Ile Eugene island" South Australia

Distribution:

- Humans : kangaroos
- National symbol of Australia

Characteristics:

- Food Habits: herbivore
- Use hopping as their primary method of locomotion.
- Can jump 3m high and 7.6m long.
- 2n =16





Sequencing:





Genome Assembly:

Meug_1.0

Sanger
 sequencing

• 2x

Meug_1.1

- ABI SOLID
- additional5.9x

Meug_2

- Illumina
- 5x

Genome Assembly:

• Genome size \approx 2.9 Gb

version	Contigs no.	Scaffolds no.	N50/contig	N50/ scaffold
Meug_1	1.211 million	616,418	2.5 kb	-
Meug_1.1	1.174 million	277,711	2.6 kb	41.8 kb
Meug_2	1.111 million	379,858	2.91 kb	34.3 kb





Genome Annotation:

- RepeatMasker & RepeatModeler programs.
- Repbase database

total Interspersed repeats	Transposons	LTR	LINE	SINEs
52.8%	2.9%	3.91%	28.6%	11.7%
	(3)	(25)	(4)	(1)

Identified genes	Protein coding	Pseudo-genes	GC %	CpG
18,258	15,290	1,496	38.8%	3.55

Genomic Facts :

life cycle:

• Short gestation & long lactation

Immunity:

• Immune factors (antimicrobial peptides)

Concurrent asynchronous lactation:

- lactation genes:
- Milk production: (CSN1,CSN2,CSN3,LALBA &LGB)
- -Marsupial Phase- specific: (LLPA, LLPB & WAP)





Questions:

1. Which one do you think is the best version of -Meug assemblies & why?

(Meug_2); smallest number of contigs covers 50% of the assembled genome

2. What is the unique feature in marsupials life cycle:

Complete their development inside the pouch



"Koala" *Phascolarctos cinereus.*

By: Shaikhah Al-Shareefi.



Facts about the species :

- The koala is a marsupial mammal (pouched mammals).
- The word 'koala' is thought to mean 'no drink' or 'no water' in the Aboriginal language.
- It sleeps 22h, They live and sleep in the eucalyptus trees.
- Koalas do not build nests.
- Pear-shaped body provides stability while the koala sits in trees.







Facts about the species :

- The koala's young are born while still at the embryonic stage, It crawls into the mother's posteriorly opening pouch and attaches to a teat, where it remains for 6–7 months. It continues to suck after it has left the pouch until about a year old.
 - The species was heavily exploited by a pelt trade (1870s to late 1920s), which harvested millions of animals.
 - threats are primarily due to loss and fragmentation of habitat, climate change and disease.





Interesting Facts:

Ability to tolerate a highly toxic diet :

- koala's diet of *eucalyptus* leaves contains high levels of plant secondary metabolites that would be lethal to most other mammals.
- Koalas experience little competition for food resources.
- Adaptive expansion of CYP2C and maintenance of duplicates appear to have worked in concert, resulting in higher enzyme levels for detoxification.
- Due to the low caloric content of this diet, the koala rests and sleeps up to 22 h a day.
- meloxicam, a nonsteroidal anti-inflammatory drug (NSAID) known to be metabolized by the protein product of CYP2C in humans. Is so rapidly metabolized in the koala and a handful of other eucalypt-eating marsupials.
- Anti-chlamydia antibiotics such as chloramphenicol are degraded rapidly by koalas.





Interesting Facts:

Taste, smell and food choice :

- expansion of one lineage of vomeronasal receptor type 1 (V1R) genes associated with the detection of nonvolatile odorants. There are six such genes in koala, and none found in, human, mouse, dog, platypus or chicken.
- The expansion of one lineage of *V1R* genes is consistent with the koala's ability to discriminate among diverse plant secondary metabolites.
- genomic evidence of expansions within the taste receptor families that would enable the koala to optimize ingestion of leaves with a higher moisture and nutrient content in concert with the concentration of toxic plant secondary metabolites in their food plants.
- functional duplication of the aquaporin 5 gene, which plays a role in the generation of saliva, tears and pulmonary secretions.
- The *TAS2R* family has a role in 'bitter' taste, enabling recognition of structural toxins such as terpenes, phenols and glycosides.

- DNA source: tissue biopsy.
- Sequencing strategy: hierarchical
- Sequencing method : illumina.
- Assembly: *de novo*
- Annotation: structural annotation (automated genome annotation).



Questions:

- 1. How does the koala genome force the organism on just one type of food?
 - 2. koala is born with no immunity, how do they gain it?



	Genome size	Coverage	GC%	# Contigs	N50 contig	# Scaffolds	N50 Scaffold
koala	3.42Gbb	57.3 ×	39%	1906	11.6Mb	-	11,589 (contig)

	Genome size	Coverage	GC%	# Contigs	N50 contig	# Scaffolds	N50 Scaffold	chromosome s
Giraffe	2.9 Gb	30x	41.5%	2 M	47 Kb	2 M	339.6 Kb	30
platyus	2.7 Gb	6X	45.5%	177.028	39.589 kb	61.239	298 kb	52 (26 pairs).
Giant panda	2.4 Gb	56x	Between 20% and 80%.		40 kb		1.3 Mb	
koala	3.42 Gb	57.3 ×	39%	1906	11.6Mb	-	-	16
kangaroo	≈ 2.9 Gb	2x 5.9x 5x	38.8%	1.211 M 1.174 M 1.111 M	2.5 kb 2.6 kb 2.91 kb	616,418 277,711 379,858	- 41.8 kb 34.3 kb	16