

Genomic characterisation of
Bettongia penicillata
papillomavirus type 1

Mark David Bennett

BVSc(Hons) MACVSc PhD

This thesis is presented for the degree of Master
of Veterinary Science of Murdoch University

2011

Declaration

I declare that this thesis is my own account of my research and contains as its main content work which has not previously been submitted for a degree at any tertiary education institution.

.....
Mark David Bennett

August 2011

Abstract

The first fully sequenced papillomavirus from a marsupial host, *Bettongia penicillata* papillomavirus type 1 (BpPV1), was detected in facial papillomas from an adult male woylie (*Bettongia penicillata ogilbyi*). Woylies are critically endangered Australian diprotodont marsupials that now occur only in parts of south-west Western Australia and on a few islands off the coast of South Australia. The gross and microscopic pathology was typical of papillomavirus infections in a wide range of animal species. Papillomavirus L1 protein was demonstrated using indirect immunohistochemistry within keratinocyte nuclei, as was BpPV1 DNA using *in situ* hybridization. The circular, double-stranded DNA genome contained 7737 base pairs and encoded 7 open reading frames: *E6*, *E7*, *E1*, *E2*, *E4*, *L2* and *L1* in typical papillomavirus conformation. Phylogenetic analysis revealed that BpPV1 is a close-to-root papillomavirus within the beta-gamma-pi-xi ($\beta + \gamma + \pi + \xi$) supertaxon of the family *Papillomaviridae*. It appears to be most closely related to *Erinaceus europaeus* papillomavirus type 1 – a papillomavirus detected in the European hedgehog. A phylogenetic tree constructed using the *L1* and *L2* open reading frames of BpPV1, 62 other papillomavirus types and the bandicoot papillomatosis-carcinomatosis viruses types 1 and 2 (BPCVs), revealed possible common ancestry between the late protein-encoding open reading frames of the BPCVs and BpPV1. This result is important for determining the likely evolutionary history of the BPCVs, and lends credence to the hypothesis that the BPCVs arose due to a recombination event between an ancient papillomavirus and an ancient polyomavirus, several tens of millions of years before present.

Contents

| | |
|---|----|
| Declaration | 2 |
| Abstract | 3 |
| Contents | 4 |
| Acknowledgements | 9 |
| Abbreviations | 10 |
| | |
| 1. Introduction | |
| 1.1 <i>Bettongia penicillata ogilbyi</i> | 13 |
| 1.1.1 Classification and taxonomy | 13 |
| 1.1.2 The family Potoroidae | 14 |
| 1.1.3 Evolution of potoroids | 16 |
| 1.1.4 Anatomical and physiological characteristics | 18 |
| 1.1.5 Habits and habitat | 22 |
| 1.1.6 Ecological interactions | 26 |
| 1.1.7 Conservation status of <i>Bettongia penicillata ogilbyi</i> | 28 |
| 1.1.8 Infectious diseases | 31 |
| 1.1.8.1 Parasites | 31 |
| 1.1.8.2 Bacteria and viruses | 34 |
| 1.2 Papillomaviruses and related viruses of animals | 34 |
| 1.2.1 The family <i>Papillomaviridae</i> | 34 |
| 1.2.2 Papillomaviruses & related viruses from marsupials and monotremes | 37 |
| 1.2.3 Significance of papillomavirus infections | 39 |

| | |
|---|----|
| 1.2.4 Evolution of papillomaviruses and related viruses | 44 |
| 2. Materials and methods | |
| 2.1 Clinical examination and biopsy collection | 47 |
| 2.2 Diagnostic testing | 48 |
| 2.2.1 Histopathology | 48 |
| 2.2.2 Immunohistochemistry | 48 |
| 2.2.3 Transmission electron microscopy | 49 |
| 2.2.4 DNA extraction | 50 |
| 2.2.5 Multiply primed rolling circle amplification | 50 |
| 2.2.6 Restriction enzyme analysis | 50 |
| 2.2.7 PCR testing | 51 |
| 2.2.8 Preliminary DNA sequencing | 51 |
| 2.3 Cloning | 52 |
| 2.4 Genome sequencing | 53 |
| 2.4.1 Sequencing primers | 53 |
| 2.4.2 Sequencing reactions | 54 |
| 2.5 Sequence analysis | 54 |
| 2.6 Phylogenetics | 55 |
| 2.7 <i>In situ</i> hybridization experiments | 56 |
| 2.8 Estimating the time of divergence | 58 |

| | |
|--|----|
| 3. Results | |
| 3.1 Initial diagnostics | 59 |
| 3.2 RCA and restriction enzyme analysis | 62 |
| 3.3 Preliminary PCR results | 62 |
| 3.4 <i>In situ</i> hybridization | 62 |
| 3.5 Sequencing | 65 |
| 3.5.1 Non-coding regions | 65 |
| 3.5.2 Open reading frames | 67 |
| 3.5.2.1 The <i>L2</i> open reading frame | 67 |
| 3.5.2.2 The <i>L1</i> open reading frame | 67 |
| 3.5.2.3 The <i>E6</i> open reading frame | 68 |
| 3.5.2.4 The <i>E7</i> open reading frame | 68 |
| 3.5.2.5 The <i>E1</i> open reading frame | 69 |
| 3.5.2.6 The <i>E2</i> open reading frame | 69 |
| 3.5.2.7 The <i>E4</i> open reading frame | 69 |
| 3.6 Phylogenetic trees | 71 |
| 3.7 Time since divergence | 71 |
| 3.8 Archives containing BpPV1 materials | 72 |
| | |
| 4. Discussion and conclusions | |
| 4.1 <i>Bettongia penicillata</i> papillomavirus type 1 | 76 |
| 4.1.1 Notable features of the virus genome | 76 |
| 4.1.2 Interpretation of <i>in situ</i> hybridization results | 78 |
| 4.1.3 Phylogenetic trees | 80 |
| 4.2 Taxonomy and phylogeny | 81 |

| | |
|---|-----|
| 4.2.1 Formal classification of BpPV1 | 81 |
| 4.2.2 Implications for the BPCVs | 82 |
| 4.2.3 Molecular clocks and mutation rates | 84 |
| 4.3 Implications for <i>Bettongia penicillata ogilbyi</i> conservation | 85 |
| 4.4 Conclusions | 88 |
| 5. Bibliography | 90 |
| 6. Appendices | |
| 6.1 <i>Bettongia penicillata</i> papillomavirus type 1: complete genome | 124 |
| 6.2 Predicted amino acid sequences of BpPV1 open reading frames | 130 |
| 6.3 Animal PVs and PV-like viruses, excluding those infecting humans | 132 |
| 6.3.1 Mammalia: Carnivora | 134 |
| 6.3.2 Mammalia: Rodentia | 135 |
| 6.3.3 Mammalia: Insectivora | 135 |
| 6.3.4 Mammalia: Chiroptera | 135 |
| 6.3.5 Mammalia: Lagomorpha | 135 |
| 6.3.6 Mammalia: Perrisodactyla | 136 |
| 6.3.7 Mammalia: Cetacea | 136 |
| 6.3.8 Mammalia: Sirenia | 136 |
| 6.3.9 Mammalia: Artiodactyla | 137 |
| 6.3.10 Mammalia: Primates (excluding humans) | 138 |
| 6.3.11 Mammalia: Monotremata | 142 |
| 6.3.12 Mammalia: Didelphimorphia | 142 |
| 6.3.13 Mammalia: Peramelemorphia | 142 |

| | |
|---|-----|
| 6.3.14 Mammalia: Diprotodontia | 142 |
| 6.3.15 Reptilia: Testudines | 143 |
| 6.3.16 Aves: Galliformes | 143 |
| 6.3.17 Aves: Passeriformes | 143 |
| 6.3.18 Aves: Psittaciformes | 143 |
| 6.4 Similarity matrices | |
| 6.4.1 Comparison of the KoAA1 partial L1 sequence to those of selected PV and PV-like viruses | 144 |
| 6.4.2 Comparison of the <i>T. vulpecula</i> partial L1 sequence to those of selected PV and PV-like viruses | 145 |

Acknowledgements

This research project received funding from Murdoch University, the Department of Environment and Conservation of Western Australia and the Weston-Fernie Wildlife Research Fund.

Thanks to Sabrina Trocini, Andrea Reiss and Adrian Wayne for allowing me to reproduce their photographs of woylies in this thesis.

I am also very grateful for the kind assistance of the staff at the Laboratory of Clinical and Epidemiological Virology, Rega Institute for Medical Research, at the University of Leuven. Special thanks are owed to Hans Stevens for his work on the phylogenetic analyses of BpPV1.

Thanks to my academic supervisors for supporting this project and to my family, friends and fellow residents in veterinary anatomic and clinical pathology at Murdoch University's School of Veterinary and Biomedical Sciences for their encouragement.

Abbreviations

| | |
|---------|---|
| aa | Amino acid |
| AIDS | Acquired immune deficiency syndrome |
| ARWH | Australian registry of wildlife health |
| ATPase | Adenosine triphosphatase |
| BLAST | Basic local alignment search tool |
| BPCV | Bandicoot papillomatosis carcinomatosis virus |
| BPCV1 | Bandicoot papillomatosis carcinomatosis virus type 1 |
| BPCV2 | Bandicoot papillomatosis carcinomatosis virus type 2 |
| BPV | Bovine papillomavirus |
| BPV1 | Bovine papillomavirus type 1 |
| cDNA | Complementary DNA |
| CI | Confidence interval |
| CPV1 | Canine oral papillomavirus (formerly abbreviated COPV) |
| DEC | Department of Environment and Conservation of Western Australia |
| DNA | Deoxyribonucleic acid |
| dNTP | Deoxyribonucleotide triphosphate |
| DPX | Distyrene-plasticiser-xylene |
| DSEWP&C | Dept. of Sustainability, Environment, Water, Population & Communities |
| EDTA | Ethylenediaminetetraacetic acid |
| EePV1 | <i>Erinaceus europaeus</i> papillomavirus type 1 (formerly EHPV) |
| EHPV | European hedgehog papillomavirus (now EePV1) |
| EV | Epidermodysplasia verruciformis |
| Fab | Fragment antigen-binding (antigen-binding parts of Ig molecules) |

| | |
|-------|--|
| FFPE | Formalin fixed, paraffin embedded |
| HaOPV | Hamster oral papillomavirus (now MaPV1) |
| HPV | Human papillomavirus |
| HRP | Horseradish peroxidase |
| Ig | Immunoglobulin |
| ISH | <i>In situ</i> hybridization |
| ISIS | International species information system |
| IUCN | International Union for the Conservation of Nature and Natural Resources |
| LB | Luria-Bertani |
| LCR | Long control region |
| MaPV1 | Hamster oral papillomavirus (formerly HaOPV) |
| mRNA | Messenger RNA |
| mya | Million years ago |
| NCBI | National center for biotechnology information |
| NCR | Non-coding region |
| nt | Nucleotide |
| ORF | Open reading frame |
| PBS | Phosphate buffered saline |
| PBST | Phosphate buffered saline with 0.05% (v/v) Tween-20 |
| PCR | Polymerase chain reaction |
| pRb | Retinoblastoma protein |
| PV | Papillomavirus |
| RCA | Multiply primed rolling circle amplification |
| RNA | Ribonucleic acid |
| RNase | Ribonuclease |

| | |
|-------|---|
| SSC | Sodium chloride-sodium citrate |
| TBST | Tris buffered saline with 0.05% (v/v) Tween-20 |
| Tris | Tris(hydroxymethyl)aminomethane |
| URR | Upstream regulatory region |
| UV | Ultra-violet |
| VLPs | Virus-like particles |
| WHIMS | Warts hypogammaglobulinaemia infection & myelokathexis syndrome |
| X-gal | Bromo-chloro-indolyl-galactopyranoside |
| XSCID | X-linked severe combined immunodeficiency |