Whole genomes reveal multiple candidate genes and pathways involved in the immune response of dolphins to a highly infectious viral disease

Running title: CeMV candidate immune genes

Kimberley C. Batley[1,2*], Jonathan Sandoval-Castillo[1], Catherine Kemper[3], Nikki Zanardo[1,2], Ikuko Tomo[3], Luciano B. Beheregaray[1], Luciana. M. Möller[1,2].

[1] Molecular Ecology Laboratory, College of Science and Engineering, Flinders University, Adelaide, South Australia 5042, Australia.

[2] Cetacean Ecology, Behaviour, and Evolution Laboratory, College of Science and Engineering, Flinders University, Adelaide, South Australia 5042, Australia.

[3] South Australian Museum, North Terrace, Adelaide, South Australia 5000, Australia.

kimberley.batley@flinders.edu.au

jonathan.sandoval-castillo@flinders.edu.au

catherine.kemper@samuseum.sa.gov.au

Nikki.Zanardo@sa.gov.au

ikuko.tomo@samuseum.sa.gov.au

luciano.beheregaray@flinders.edu.au

luciana.moller@flinders.edu.au

4.2 Abstract

Wildlife species are challenged and threatened by various infectious diseases that act as important selective forces and demographic drivers of populations. Yet, studies about host genetic factors and disease susceptibility are very limited. Cetacean morbillivirus (CeMV) has emerged as a major viral threat to cetacean populations worldwide, contributing to the death of tens of thousands of individuals of multiple dolphin and whale species. To understand the genomic basis of immune responses to CeMV, we generated and analysed whole genomes of 53 Indo-Pacific bottlenose dolphins (*Tursiops aduncus*) exposed to Australia's largest CeMV-related mortality event known to date. The genomic dataset consisted of 7,720,686 SNPs anchored onto 23 chromosome-length scaffolds and 77 short scaffolds. Allele frequency estimates between survivors and non-survivors of the outbreak revealed 11,009 candidate SNPs, of which 498 were annotated to 220 protein coding genes. These included 36 genes with functions related to innate and adaptive immune responses, and cytokine signalling pathways. The list also included genes known to be involved in immune responses to other morbilliviruses, such as measles in humans and the phocine distemper virus in pinnipeds. Our study characterised genomic regions and pathways that likely contribute to CeMV susceptibility and resistance in dolphins, representing a stride towards clarifying the complex interactions of the cetacean immune system. It also emphasises the relevance of whole genome datasets to study the genetics of wildlife diseases.

Keywords

Immune genes, whole genome sequencing, ecological genomics, wildlife disease, inshore dolphin, cetacean morbillivirus

1. Introduction

Climatic variations, natural and anthropogenic alterations to ecosystems, changes in host behaviour, and the movement of pathogens and vectors have all contributed to the emergence of infectious

diseases (EIDs) in wildlife populations (Williams et al. 2002, Morens et al. 2004, Cunningham et al. 2017, Titcomb et al. 2019). Infectious diseases have become a major conservation concern due to pathogens' ability to rapidly evolve, their short generation times and often complex transmission dynamics, as well as being able to cause swift and widespread mortality, diminish genetic diversity and contribute to population declines and extinctions (Altizer et al. 2003, Blanchong et al. 2016, Stejskalova et al. 2017).

Disease outbreaks in cetacean populations worldwide are also of concern, particularly for species that exhibit high social connectivity and gregarious behaviour, and for populations that are naïve, small, threatened, or immune-supressed (Gulland and Hall 2007, Van Bressem et al. 2009, Weiss et al. 2020). In recent years, the reporting of EIDs and strandings in cetaceans has increased, with one highly contagious and virulent pathogen emerging as a major threat to their populations; cetacean morbillivirus (CeMV) (Sacristán et al. 2015). CeMV belongs to the genus *Morbillivirus,* which affects both terrestrial mammals [humans (measles virus), canines (canine distemper virus), cattle, goats and sheep, and two novel morbilliviruses in cats and bats] and marine mammals [true seals (phocine distemper virus) and cetaceans] (Alfonso et al. 2016, Ohishi et al. 2019). These viral species are distinct, but share a common phylogenetic origin, similar genome structure, symptoms of infection and pathomorphology (da Fontoura Budaszewski and von Messling 2016, Diaz-Delgado et al. 2019). This, along with observed cross-species transmissions (Stejskalova et al. 2017, Jo et al. 2018, Padalino et al. 2019), suggests that knowledge gained on immune responses for one viral species may be applicable more generally to morbilliviruses.

Since its discovery in the late 1980's, CeMV has raised great concern given its rapid expansion in geographic distribution, increase in number of host species, incidents and number of mortalities, the discovery of new viral strains, and the possibility of spread through vertical, horizontal and cross-species transmission (reviewed in Van Bressem et al. 2014; Di Guardo & Mazzariol 2016). Susceptibility to CeMV is thought to be somewhat species-specific, with striped dolphins (*Stenella coeruleoalba*), common bottlenose dolphins (*Tursiops truncatus*) and Guiana dolphins (*Sotalia*

74 *guianensis*) being among the most impacted species (Diaz-Delgado et al. 2019). By contrast, other
75 delphinid species such as pilot whales (Globicephala spp.), dusky dolphins (*Lagenorhynchus*
76 *obscurus*), Fraser's dolphins (*Lagenodelphis hosei*) and melon-headed whales (*Peponocephala*
77 *electra*) are thought to act mainly as reservoirs and vectors (Van Bressem et al. 2014).

78 The CeMV has been implicated in the death of tens of thousands of cetaceans worldwide (Van
79 Bressem et al. 2014), but until recently had only been recognised as a contributing factor in the death
80 of four *Tursiops* individuals across Australia (Stone et al. 2011, Stone et al. 2012, Stephens et al.
81 2014). However, in 2013 CeMV was a major causative agent in the death of at least 50 dolphins in
82 South Australia, becoming the largest confirmed CeMV outbreak in Australia and the first recorded
83 deaths from CeMV in the state (Kemper et al. 2016). This unusual mortality event lasted
84 approximately six months (March to September), and the initial months of the outbreak coincided
85 with climatic anomalies that resulted in abnormally high sea surface temperatures (Kemper et al.
86 2016). This led to phytoplankton blooms and fish die-offs, particularly in Spencer Gulf and Gulf St
87 Vincent (GSV) (Kemper et al. 2016). The majority of the deaths, however, represented neonates and
88 calves from one genetic population of *T. aduncus* that inhabits GSV (Kemper et al. 2016). This
89 population is estimated at 700-1,200 dolphins (Bilgmann et al. 2019), exhibits high social
90 connectivity, shows relatively low genetic diversity and recent modelling suggests this population is
91 likely most vulnerable to epizootic events (Zanardo et al. 2016b, Pratt et al. 2018, Bilgmann et al.
92 2019, Reed et al. 2020). Therefore, it is important to understand why this population was particularly
93 susceptible to CeMV.

94 Host genetic factors are known to be key drivers in the plasticity of immune responses in natural
95 populations, being major determinants of an individual's susceptibility, resistance and tolerance to
96 infection (Karlsson et al. 2014, Stejskalova et al. 2017). Yet, studies investigating the role of genetic
97 factors of host defence against immune responses in wildlife populations are limited. In the case of
98 marine mammals, advances in understanding immune responses have generally come from studies of
99 freshly stranded individuals. These are not necessarily representative of the entire population, leaving

100 a knowledge gap of how immune responses are modulated in natural populations (Marsili et al. 2019).
101 To counteract this gap, association-based studies provide a favourable framework for identifying
102 associations between genomic locations, regions or genes, and complex traits in natural populations.
103 Studies addressing the role of host genetics in combating infection have generally targeted a small
104 number of genomic regions of known functional importance, and genes with strong effect. For
105 example, the Major Histocompatibility Complex (MHC) are among some of the most targeted and
106 well-studied immune associated genes in model and non-model species (Martin and Carrington 2005,
107 Acevedo-Whitehouse and Cunningham 2006, Cammen et al. 2015, Elbers et al. 2018, Pagan et al.
108 2018, Manlik et al. 2019). In the case of immune responses to the measles virus, specific alleles
109 within the human leukocyte antigen (HLA) genes class I and II (*B; DQA, DQB, DRB*) have been
110 associated with varying antibody titers following vaccination against the virus (Haralambieva et al.
111 2015). However, many other non-MHC genes have been proposed to be involved in host defence to
112 measles, including viral binding genes, cytokine receptor genes, pathogen-associated sensing genes
113 and antiviral genes (Hashiguchi et al. 2011, McCarthy et al. 2011, Haralambieva et al. 2015,
114 Stejskalova et al. 2017).

115 Advancements in next generation sequencing, computational power and improved availability of
116 genomic data has enabled the move from a targeted to a non-targeted approach of association-based
117 studies. This approach enables the search for multiple genetic variants across the genome under
118 selection and associated with a trait, without the need of prior knowledge. This framework is
119 frequently utilised in humans, model organisms and agricultural systems (Elbers et al. 2018), and
120 while still limited, advancing technologies have now enabled association studies in wildlife
121 populations. In particular, whole genome data has been utilised to investigate immune responses of
122 endangered and vulnerable Australian marsupials (Tasmanian devil, *Sarcophilus harrisii;* and the
123 koala*, Phascolarctos cinereus)* to two highly damaging diseases that continue to threaten populations
124 across their distribution (Wright et al. 2017, Johnson et al. 2018). The move to non-targeted
125 approaches and large genomic datasets improves our ability to address the genetic basis of adaptation

126 in wildlife populations, and allows us to understand the role that genetic variants play in the plasticity
127 of immune responses, and in the susceptibility of individuals and populations to infectious diseases.

128 In this study, we expand substantially on previous work based on reduced representation sequencing
129 (Batley et al. 2019) to investigate the genomic basis of resistance and susceptibility of bottlenose
130 dolphins to CeMV using whole genomes. Using a much larger dataset, we searched for regions of the
131 genome under selection between *case* (non-survivors) and *control* (survivors) from the viral outbreak
132 to identify genetic variants, genes and pathways associated with resistance and susceptibility to
133 CeMV. Our study provides the first whole genome-based information to enable the screening of other
134 cetaceans for potential genetic risk factors, ultimately enabling the identification of populations and
135 species particularly vulnerable to large-scale CeMV outbreaks.

136

137 **2. Materials and Methods**

138 **2.1 Study Species, Sites and Sample Collection**

139 Dolphins that died and stranded during an unusual mortality event throughout South Australia
140 between March and September 2013, were collected by the South Australian Museum for post-
141 mortem examinations. Histopathological examinations, Reverse-Transcription Polymerase Chain
142 Reaction (RT-PCR) and/or immunohistochemical assays confirmed that CeMV infection and related
143 pathologies were the main contributing factor in the dolphin deaths (see Kemper et al. 2016). Muscle
144 tissues from thirty-three *T. aduncus* from GSV and adjacent waters and a single individual from
145 Spencer Gulf were selected for DNA extractions. These *case* individuals are part of a set of 47
146 dolphins that tested positive for CeMV based on RT-PCR and/or immunohistochemical assays. *Case*
147 samples were classified into age classes, with the majority of the strandings being young dolphins
148 (neonates, calves, and juveniles) (young $n = 31$, adults $n = 3$). These samples were frozen at -80˚C and
149 kept at the South Australian Museum before being transferred to 90% ethanol and to the Molecular
150 Ecology Laboratory at Flinders University (MELFU).

151 *Case* samples were complemented with biopsy samples from free-ranging bottlenose dolphins from
152 GSV and adjacent waters collected between 2014 and 2015 (Bilgmann et al. 2007, Zanardo et al.
153 2016b, Pratt et al. 2018). These skin and blubber samples were collected using either the PAXARMS
154 biopsy system (Krützen et al. 2002) or a hand-held biopsy pole (Bilgmann et al. 2007). Age classes
155 (calves, juveniles, and adults) of sampled individuals were estimated in situ based on body size and
156 association with an adult dolphin (see Zanardo et al. 2016 for details). This resulted in a total of 34
157 *control* samples, with samples from young (n = 11) complemented with random adult samples from
158 GSV and close adjacency (n = 23). These *control* samples are considered putative survivors since they
159 belong to the same genetic and socially cohesive population (Zanardo et al. 2016, Pratt et al. 2018) as
160 the one most impacted during the unusual mortality event and were collected within 18 months of the
161 outbreak. Biopsy samples were preserved in a salt-saturated solution of 20% dimethyl sulphoxide
162 (DMSO) and stored at -80˚C at the MELFU. Dolphins were genetically sexed using the polymerase
163 chain reaction (PCR) (Banks et al. 1995). The phenotypic data for samples that passed quality controls
164 and were subsequently selected for whole genome sequencing is available in Table S1.

165 **2.2 DNA Extractions and whole genome sequencing**

166 Genomic DNA was isolated from *control* samples following the salting out method (Sunnucks and
167 Hales 1996), while genomic DNA was extracted from *case* samples using the Qiagen DNeasy blood
168 and tissue kit following the manufacturer's protocol, or using the salting out method. The purity of
169 extractions was verified using a ND-1000 spectrophotometer (Nanodrop, Thermo Scientific), and
170 quantity assessed using a fluorometer (Qubit, Life Technologies). The DNA integrity was further
171 assessed by gel electrophoresis (2% agarose gels, produced in-house). All extractions were expected
172 to pass quality controls based on standards set by the Australian Genome Research Facility (AGRF),
173 where libraries were prepared and sequenced. Specifically, samples were required to have a quantity ≥
174 20 ng/µl, a high molecular weight (≥ 20 kb), free of RNA (assessed on agarose gels) and an A260/280
175 (protein contamination) ratio between 1.8-2.0. Extractions that did not pass quality controls were re-
176 extracted a maximum of three times, using the same method but with altering amounts of tissue to

potentially increase the concentration and improve the quality of DNA. As expected, extractions that failed the quality controls were typically of *case* samples, since these were obtained from carcasses rather than free-ranging dolphins. *Case* samples with a low concentration for all extractions were then combined, and concentrated using a centrifuge vacuum concentrator (Hetovac, Heto Lab). Extractions from 53 samples that passed all quality controls (*case*, n = 19 and *control*, n = 34) were subsequently selected for library preparation and whole genome sequencing at AGRF (Table S1). Libraries were prepared using the NEBNext Ultra II DNA library prep kit and sequenced on two lanes of the Illumina NovaSeq 6000 platform. Samples were sequenced at ~7x coverage, excluding one sample from GSV that was sequenced at a higher depth of coverage (~28x) to form a reference genome (Batley et al. unpublished data). While throughout this paper we refer to the species as *T. aduncus*, it has been previously suggested to represent a separate species, endemic to southern Australian waters, *T. australis* (Charlton-Robb et al. 2011). However, recent studies suggest this is more likely to be a subspecies of *T. aduncus* (Moura et al. 2020), and therefore we refer to the reference genome here as the southern Australian bottlenose dolphin (SABD). Details regarding the construction of this reference genome will be available in Batley et al. (in prep) but see Table S2 for the reference genomes' quality and statistics.

2.3 Read processing, SNP calling and filtering

Raw sequencing data was pre-processed following the pipeline adapted from GATK best practices, with modifications. Firstly, reads were trimmed if read quality was below 23 in a sliding window of five nucleotides, while adapters were removed using Trimmomatic v0.38. The remaining reads were mapped to the chromosome-length scaffolded *S*ABD reference genome using Bowtie2 v2.2.7. The resulting SAM files were then converted to BAM files, duplicate marked and sorted using Picard. Indels were then locally realigned to correct mapping errors using GATK before merging the replicate reads from different libraries with samtools.

SNPs were called from the mapped reads of all individuals using the SABD reference genome in a two-part process using bcftools. This involved generating genotype probabilities at each genomic

203 position before calling the SNPs. SNPs were then filtered with vcftools (Danecek et al. 2011) and
204 using parameters described in Brauer et al. (2016). In short, reads with a minor allele frequency <3%
205 and genotyped in <80% of the samples were excluded. Indels were removed and only SNPs with a
206 quality and depth ratio of 2%, mapping quality >30 and mean depth <12 were retained, while SNPs
207 not matching Hardy-Weinberg equilibrium expectations were excluded. Finally, SNPs were called
208 altogether, including from available whole genomes of *Delphinus delphis* and *T. truncatus* (data not
209 presented here), but as this study focused on *T. aduncus* from southern Australia, only SNPs that are
210 unique to this lineage were retained (see Table S3 for SNPs retained at each step).

211 **2.4 Whole genome association study**

212 *2.4a Potential effects of inbreeding, relatedness, sex, and age-classes.*

213 As inbreeding can reduce disease resistance due to the loss of genetic diversity (Acevedo-Whitehouse
214 et al. 2003), levels of inbreeding were calculated to test for potential effects of inbreeding. The
215 inbreeding coefficient, F, was calculated within and between *cases* and *controls* using the het
216 command in Plink v1.9, based on an unlinked SNP dataset (189,178 SNPs). The mean F of each
217 group was compared using an independent samples t-test.

218 Levels of relatedness within and between *cases* and *controls* as well as differences in the
219 representation of sexes and age classes between groups were also calculated to assess the potential
220 influence of these factors on the outcome of an individual. Pairwise relatedness between individuals
221 based on the unadjusted Ajk statistic method of Yang et al. (2010) was estimated using VCFtools.
222 Pairwise relatedness within and between groups, as well as the mean number of individuals of each
223 sex and age class between groups were then compared using an independent sample t-test.

224 *2.4b Identifying SNPs under selection*

225 Allele frequency differences between the two groups were calculated to identify SNPs potentially
226 involved in resistance or susceptibility. This analysis used 7,720,686 SNPs and was based on two

association tests implemented in Plink v1.9; the chi-square test and Fisher's exact test. SNPs with a highly significant *P*-value ($p \leq 0.001$) were selected as outlier SNPs, as per Batley et al. (2019). These two tests were complemented by the Weir and Cockerham's F_{ST} (Weir and Cockerham 1984), which estimates differentiation between groups based on allele frequency shifts. F_{ST} was calculated between *cases* and *controls* using the --weir-fst-pop command in VCFtools. SNPs with an F_{ST} value greater than five standard deviations from the mean (0.0019 ± 5SD) were selected as outlier SNPs (Axelsson et al. 2013, Kardos et al. 2015). Outlier SNPs from each of the three tests were compared, and those identified as outliers in at least two tests were selected as candidate SNPs to reduce selecting false positives. The two tests implemented in Plink also output Odds Ratios (OR), which were used to test the odds of the minor allele being in association with an outcome (i.e. non-survival).

2.4c Annotation of candidate SNPs

To annotate and explore the function of candidate SNPs, 600 bp flanking regions of the candidate SNPs were aligned to *T. truncatus* proteins (GCF_001922835.1) using blastx v2.2.28. This used an alignment length of above 30 amino acids, similarity above 50%, and an e-value threshold of 8e-07. Although the similarity cutoff was low, all alignments were validated using SnpEff (Cingolani et al. 2012) and with the SABD reference genome annotation. For all alignments to the proteins, the genomic region of the SNP (intronic or exonic) and their predicted functional effect (missense vs synonymous changes) were investigated using SnpEff. Specifically, a VCF file of all candidate SNPs with flanking regions that aligned to the protein database was generated and the SNPs were annotated against the SABD reference genome. All SNP annotations were cross-checked against the blastx sequence annotations to ensure annotations were correct between the two genome annotations. This was done as all SNPs were called from the SABD reference genome; however, the annotation of this genome is in its draft stages. Therefore, genes that were annotated against the *T. truncatus* proteins and not the SABD annotation were searched for within the SABD genome. If the gene was present but in a different location, the annotation was ignored. If, however, the gene was not present in the SABD genome, it was included as a correct annotation. Functions of the putative candidate genes

were explored using Gene Ontology (GO) terms provided by UniProtKB (UniProt 2019), and their involvement in immune pathways and gene interactions were explored with human Ensembl identifiers and Reactome (Fabregat et al. 2018).

2.4d Candidate immune gene approach

Several candidate genes for variation in CeMV immune responses have been previously proposed (Hashiguchi et al. 2011, McCarthy et al. 2011, Haralambieva et al. 2015, Stejskalova et al. 2017), yet some of these genes were not initially identified as candidate genes in this dataset. To investigate whether SNPs within these genes are in fact neutral between *cases* and *controls*, or alternatively are under selection but did not align to the protein coding regions, the allele frequency and genotype counts for each SNP within each gene were compared between *cases* and *controls*. To achieve this, each gene location was extracted from the SABD reference genome and SNPs within the specified regions were extracted using VCFtools. To confirm the gene regions from the SABD annotations were correct, gene sequences from either the *T. truncatus* (GCF_001922835.1) genome were downloaded from NCBI and mapped to the SABD reference genome using blast v2.2.28. All alignments from both algorithms were compared to ensure accurate alignments. Allele frequency differences were calculated using a chi-square test, and genotypes of the top performing SNP within each gene (i.e. SNP with the greatest differentiation between *cases* and *controls*) were counted using Plink.

3. Results

Whole genome sequencing produced a total of 4,274,472,237 reads for 53 *T. aduncus* individuals from South Australia (Figure 1). After quality filtering, 3,310,493,013 reads (mean = 31,231,066 ± 13,586,174) remained, of which an average of 96.81% of reads mapped to the SABD reference genome. Calling SNPs from the genome resulted in a total of 33,386,256 SNPs, of which 16,658,049 remained after quality filtering (Table S3). Of these SNPs, 7,720,686 (on 23 chromosome-length scaffolds and 77 smaller scaffolds, Figure S1) were unique to *T. aduncus*. The final dataset available

for analysis therefore, consisted of 7,720,686 SNPs for 53 individuals with an average of 1.22% missing data (SD ± 1.55%). Missing data did not differ significantly between *cases* and *controls* (*cases* = 1.15% ± 0.81%; controls = 1.25% ± 1.84%).

3.1 Potential effects of inbreeding, relatedness, sex, and age-classes.

There was no significant difference in the mean inbreeding coefficient between the two groups (*cases* = 0.0574 ± 0.049; *controls* = 0.0289 ± 0.061, $P = 0.087$), suggesting that genome-wide levels of inbreeding did not influence susceptibility of *case* dolphins to CeMV during the outbreak.

The mean relatedness of pairs of individuals within and between groups was not significantly different (*cases* = -0.0185 ± 0.056; *controls* = -0.0184 ± 0.051; *case-control* = -0.0226 0.036; all $P > 0.05$). Likewise, there was no significant difference between the sex composition between groups (*cases*: M = 10, F = 9; controls: M = 21, F = 9; $P = 0.527$). There was, however, a significant difference between the representation of different age classes in the two groups, but due to the limited number of adult *case* samples ($n = 2$), the influence of age could not be accounted for in the analysis.

3.2 Identifying SNPs under selection and annotation of candidate SNPs

Methods to detect SNPs under selection between *case* and *control* individuals identified outlier SNPs in all three tests, with the number of outlier SNPs ranging between 9,122 and 22,636 (Figure 2). Of these outliers, 3,870 SNPs were present in at least two tests, and a further 7,139 SNPs were present in all three tests (Figure 3). A total of 11,009 SNPs (< 0.15% of all SNPs) found on 22 chromosome-length scaffolds showed putative signatures of selection between *case* and *control* individuals and were considered candidate SNPs.

Of the 11,009 candidate SNPs and associated flanking regions, 498 aligned and annotated to the *T. truncatus* protein dataset and/or the SABD annotation. These SNPs annotated to 220 protein coding genes and seven uncharacterised proteins (Table S4). Investigation of all candidate genes and their involvement in different pathways found that 117 candidate genes were related to 654 different

biological sub-pathways that can be grouped into 23 pathways (Table S5). The key pathway of interest is the immune system (28 genes) (Figure 4, Table S6), however other pathways of interest include disease (15 genes), signal transduction (28 genes), and cell-cell communication (5 genes) (Table S5). The remaining 103 genes either did not have Ensembl identifiers or could not be characterised into pathways. However, inspection of GO terms suggests that a further eight genes could be involved in immune system pathways (Figure 4, Table S6).

Within the immune system pathways, Fc receptor proteins (FcRs) were well characterised with 7 genes (*DOCK1, MHY9, ACTR3, GRB2, NFATC2*, *CARD11* and *CALM1*), while F-box proteins that have previously been proposed to be involved in CeMV susceptibility and resistance (Batley et al. 2019) were also identified (*LMO7*, *FBXL7*, *FBX10 and FBXW11* [identified through the candidate gene approach]). Likewise, of the signaling pathway, the MAPK cascade was well characterised with three candidate genes (*FGF2*, *GRB2* and *CALM1*) and a further six genes with GO terms relating to the MAPK cascades (*PDE6H, NTRK2, KIDINS220, INHBA, NPSH1, PLCE1*). Finally, pathways and GO terms highlighted the importance of the regulation and expression of interleukins and T cells (see Table S7 for all pathways and sub-pathways).

Further inspection of the annotated SNPs and the gene regions they fall within revealed that majority of the SNPs fell within introns (n = 362), while a further 42 SNPs were intergenic. In total, 48 SNPs were found in exonic regions, in which 18 caused a missense mutation and 22 SNPs resulted in synonymous substitutions (Table S8). Of the SNPs that annotated to immune genes, ten SNPs were found within exons, however just six of these within three genes (*CD300LF*, *NFATC2* and *NFKBIZ*) caused a missense change, while four SNPs within four genes resulted in no amino acid change (*DOCK1*, *FBXW10*, *MASP1*, *MHY9*) (Figure 4). The odds-ratio (OR) suggest that for the SNPs that caused a missense change, the minor allele increased the odds of succumbing to CeMV (Figure 4). Other genes of interest that were annotated include *IL4α*, which had an OR of 16.25 (Figure 4) and *PATJ*, both of which have been previously identified as potentially important in immune responses to morbilliviruses (McCarthy et al. 2011, Haralambieva et al. 2015, Batley et al. 2019).

328 **2.3. Candidate immune genes**

329 At least 29 genes have been previously identified as potentially playing a role in immune responses to
330 morbilliviruses in general (Hashiguchi et al. 2011, McCarthy et al. 2011, Haralambieva et al. 2015,
331 Stejskalova et al. 2017), but only two of these (*PATJ* and *IL4α*) were found to be under putative
332 selection between *cases* and *controls* in this study. For the remaining 27 genes, 23 aligned to the
333 SABD reference genome. Within the aligned genes, 7,271 SNPs were extracted, of which 29 SNPs on
334 five genes (*RARb*, *MAPK8*, *FBXW11, ANK3* and *ACOX3*) show significant allele frequency
335 differences ($P < 0.001$) between *cases* and *controls.* These SNPs were identified as outliers in the tests
336 for selection but did not align to the *T. truncatus* proteins and are therefore considered to be located
337 within intronic regions of the genes. The majority of the 23 genes were highly polymorphic (Table 1),
338 however inspection of genotype counts for the top performing SNP within each gene (i.e. SNP with
339 greatest allele frequency differences between *cases* and *controls*) highlighted a lack of heterozygosity
340 within 11 of the immune genes across all samples (Table 1). For these 11 SNPs, at least 84% of all
341 samples were homozygotes. For *TLR8* and *TLR3*, all *case* samples were homozygotes, while only four
342 control samples were heterozygotes. The genes *DQα*, *BSG* and *SLC11A1* also showed low levels of
343 variation, with only four samples being heterozygotes.

344

345 **4. Discussion**

346 We used whole genome datasets to characterise genomic regions underlying resistance and
347 susceptibility of dolphins to a highly contagious and fatal virus, CeMV. We found strong evidence of
348 selection across the genome in comparisons between *case* (non-survivors) and *control* (survivors)
349 bottlenose dolphins from the largest CeMV related outbreak in Australia (Kemper et al. 2016).
350 Significant allele frequency differences between *case* and *control* samples were identified in 11,009
351 SNPs, with a minimum of 36 immune-related genes apparently involved in CeMV resistance and
352 susceptibility.

353 Host genetic factors play an important role in mobilising immune responses to invading pathogens,
354 and may influence the outcome of an individual; yet relatively few studies have assessed the
355 importance of these factors and immunogenetic diversity in wildlife disease risk (Smith et al. 2009,
356 Bossart et al. 2019). Here, we not only uncovered host genetic variants and genes involved in CeMV
357 susceptibility and resistance but provide additional support for previously proposed candidate genes
358 associated with morbilliviruses. We also identified a lack of immunogenetic diversity in the studied
359 dolphin population within immune-related genes previously recognised as important in the fight
360 against pathogens more broadly. CeMV is of growing concern given ongoing climate change
361 threatening to lead to more stressful environments for populations and species, and altering host and
362 virus distributions (Burge et al. 2014). Since its discovery, CeMV has shown a rapid geographic
363 expansion and an increase in the number of host species and viral strains (Di Guardo and Mazzariol
364 2014, Van Bressem et al. 2014), which appears to be exacerbated by climate change associated
365 events. Through the identification of genes potentially involved in CeMV immune responses, our
366 work clarifies how host genetic factors drive CeMV outcomes and provides knowledge about the
367 diversity of immune responses, their interactions, and pathways in dolphins.

368 **4.1 Impact of inbreeding**

369 Levels of inbreeding within wildlife populations have been associated with disease emergence,
370 immunocompetence and increased disease susceptibility and severity (Smith et al. 2009). This has
371 been observed in several marine mammal populations. For example, high levels of inbreeding within
372 a population of Mediterranean striped dolphins (*Stenella coeruleoalba*) was correlated with
373 susceptibility and the outcome of individuals, with those stranded earlier in the outbreak exhibiting
374 higher levels of inbreeding than those that stranded later (Valsecchi et al. 2004). While the GSV
375 bottlenose dolphin population exhibits relatively low genetic diversity, it does not appear to exhibit
376 significant levels of inbreeding (Pratt et al. 2018). At the whole genome level, estimates of inbreeding
377 were not elevated in *case* samples compared to *controls*, supporting previous suggestions that the

outcome of an individual during this mortality event was not influenced by inbreeding (Batley et al. 2019).

4.2. Gene functions and immune system pathways

Gene annotations and investigation of associated gene pathways suggested that genes involved in the immune system, signal transduction, disease and cell-cell communication are likely to be involved in the susceptibility or resistance of dolphins to CeMV. Here, we focus on pathways and genes associated to immune functions, however, other pathways and gene functions were also disclosed (see Table S4 and S5 for details). A wide range of well characterised immune related pathways were found to be putatively under selection, including a similar number of genes from both the innate and adaptive immune systems, as well as cytokine signalling pathways. These pathways are distinct, but interconnected (Gelain and Bonsembiante 2019), reflecting the highly complex interactions and networks of the mammalian immune system.

Innate immune system

The innate immune system is rapid and non-specific in its response to pathogens, involving the interplay of the complement system, pattern recognition receptors (PRRs), cytokines and a diverse range of immune cells that detect and remove pathogens (Ohishi et al. 2011, Gui et al. 2013). Within this system, candidate genes identified were characterised into PRRs and toll-like receptor (TLR) cascades, the complement cascade, DAP12 interactions and neutrophil granulation (Table S7).

We found five candidate genes belonging to different PRR pathways (*ATF1*, *BCL2*, *MASP1*, *NFATC2* and *CARD11*). The PRRs are important in recognising Pathogen-Associated Molecular Patterns (PAMPs) to initiate a variety of different immune responses and inflammatory mediators (Amarante-Mendes et al. 2018, Gelain and Bonsembiante 2019). Interactions between the measles virus and PRRs, and in particular, TLRs, have been reported for humans (Bieback et al. 2002), and within cetaceans particular polymorphisms have been proposed to be of functional importance in CeMV

resistance (Stejskalova et al. 2017). This supports that TLRs and PRRs likely represent important genes and molecules for immune defence (McCarthy et al. 2011, Stejskalova et al. 2017).

The Fc receptor proteins (FcRs) are immunoglobulin receptors that have important functions in the activation and down-regulation of immune responses through their ability to bind to antibodies and stimulate cellular and humoral immune responses (Takai 2002, 2005). Morbilliviruses are known to cause immunosuppression that may last for long periods of time, leading to increased susceptibility to secondary infections (Kerdiles et al. 2006, Di Guardo and Mazzariol 2016). In humans, measles virus proteins have been reported to interact with FcRs on dentritic cells to generate immunosuppression through impairment of the cells' function, decreased production of interleukins, and the loss of antigen specific T cell proliferation (Marie et al. 2001). In this study, seven candidate genes were found in pathways relating to FcRs. In particular, *DOCK1, MHY9, ACTR3, GRB2* are part of the Fc gamma receptor dependent phagocytosis pathway, which recognise foreign pathogens and stimulate phagocytosis to engulf and eliminate infectious agents (Acevedo-Whitehouse and Cunningham 2006). Another three candidate genes, *NFATC2*, *CARD11* and *CALM1* are involved in the Fc-epsilon receptor (FcERI) signalling pathway, which is important for the activation of mast cells and basophils, and the release of inflammatory mediators (Turner and Kinet 1999). The *NFAT* is a family of transcription factors that appear to be key mediators of immune responses, specifically interleukin production (Fric et al. 2014). Within this pathway, it is thought that calcium in mast cells triggers the translocation of NFAT to the nucleus, where NFAT regulates the transcription of cytokine genes (Turner et al. 1998). *NFATC2* specifically is thought to regulate the expression of *TNF-a* and *IL-13* (Klein et al. 2006). *TNF-a* and interleukins are key cell signalling proteins and cytokines that fight infection, and their encoding genes have been associated with defence against the measles virus (Haralambieva et al. 2015). While these genes (*TNF-a* and interleukins) are known to be important for morbilliviruses' host defence, this finding highlights the need to look beyond these cytokine receptors, at activators and initiators of these proteins, as they may play a role in an individuals' outcome. Activation of FcERI also contributes to the recruitment of CARD11, which forms a signalling complex to activate NF-kB (Thome 2004). NF-kB is a transcription factor that is

responsible for the induction of cytokines and chemokines, which can initiate inflammation (Liu et al. 2017). CARD11 is critical for the activation of NF-kB, with the inactivation of CARD11 in mutant mice resulting in the block of cytokine production in T and B cells (Hara et al. 2003). Therefore, this gene may have been important in the activation of NF-kB and cytokine production, playing a role in susceptibility to CeMV. In summary, FcRs are important for the functioning of the innate immune system, and through the release of cytokines and phagocytes are essential for mounting an immune response to infection and preventing inflammation (Ben Mkaddem et al. 2019).

Neutrophils are among some of the most common white blood cells that circulate in the human body, participating in the inflammatory responses by releasing cytotoxic proteins during degranulation (Lacy 2006, Naegelen et al. 2015). They may also allow the presentation of antigens via MHC Class II, activating T cell proliferation (Wright et al. 2010). The presence of neutrophils in morbillivirus infected individuals are common signs of an acute inflammatory response and are commonly observed in CeMV cases (Duignan et al. 1992, Diaz-Delgado et al. 2017, Diaz-Delgado et al. 2019). Here, we found evidence of selection within three genes (*ADA2*, *RAB37* and *VAT1*) that are involved in neutrophil degranulation. These results suggest that variation within these three genes may play an important role in the release of cytotoxic proteins during neutrophil degranulation, and may be key contributors in the coordination of an inflammatory response, activating MHC or combating foreign pathogens, and thus influencing susceptibility or resistance to CeMV.

Adaptive immune system

The adaptive immune system, also known as the specific and non-rapid system, is mediated by B and T lymphocytes, and recognises pathogens by high affinity receptors (Werling and Jungi 2003). Within the adaptive immune system, seven of the candidate genes (*PDIA3*, *FBXW10*, *FBXL7*, *UBA5*, *SEC31A*, *AREL1*, *LMO7*) are known to be involved in the MHC class I pathway, which generally fight against intracellular pathogens such as viruses. Three other genes (*DCTN6*, *KIF5B*, *SEC31A*) are involved in the MHC class II pathways that fight against extracellular parasites and bacteria. The MHC complex is of known immune importance, being involved in resistance and susceptibility to

455 disease through antigen processing and presentation (de Sa et al. 2019). Specifically, five candidate
456 genes (*LMO7, AREL1, FBXL7, FBXL10, UBA5*) implicated in antigen processing of the MHC class I
457 pathway were found here. This process involves the ubiquitination and proteasome degradation,
458 whereby foreign proteins are degraded into short peptides for presentation to the MHC class I system
459 (Strehl et al. 2005). F-box proteins, in particular, are involved in the Skp, Cullin, F-box containing
460 complex (SCF complex) that catalyses the ubiquitination of proteins prior to proteasome degradation
461 (Kipreos and Pagano 2000). F-box proteins were previously suggested to be involved in the immune
462 response of bottlenose dolphins to CeMV (Batley et al. 2019), and likewise in this whole genome
463 study, allele frequency differences between *cases* and *controls* were observed at three additional
464 genes that encode F-box proteins (*LMO7*, *FBXL7* and *FBX10*). These results suggest that several F-
465 box proteins are likely to be important in host immune responses to CeMV, and genetic variants at
466 these genes may impact the likelihood of survival of an infected individual.

467 *PDIA3* and *SEC31A* are involved in antigen presentation, folding, and loading of MHC class I
468 receptors. PDIA3 is an integral part of the peptide-loading complex (Santos et al. 2007) that
469 coordinates the movement of high-affinity peptides to MHC class I molecules, and is therefore
470 involved in the regulation of antigen presentation (Santos et al. 2007, Scholz and Tampe 2009). In
471 PDIA3 deficient mice, the MHC class I complex, and specifically the presentation of antigenic
472 peptides, was impaired, which may benefit pathogens survival and spread (Garbi et al. 2006,
473 Pressinotti et al. 2009). Another three candidate genes (*DCTN6*, *KIF5B* and *SEC31A*) were
474 characterised into antigen presentation of the MHC Class II pathway, while *NFX1* may be involved in
475 MHC Class II processes. It is within this Class II complex that cell surface glycoproteins recognise
476 and bind antigens, presenting them to T-lymphocytes that initiate an immune response (Moreno-
477 Santillan et al. 2016).

478 Across vertebrates, the MHC complex is one of the most well studied immune-related regions. It has
479 been implicated in responses to measles vaccination (Haralambieva et al. 2015), and suggested to be
480 functionally important in CeMV infection (Stejskalova et al. 2017). Although we found no evidence

of selection within key MHC genes (e.g. *DQA, DQB*), variation in several downstream genes suggests that the MHC Class I and II pathways may be involved in a dolphin's ability to fight CeMV infection.

Some cell receptors may also play a role in modifying the response of immune cells. We found significant variation between *cases* and *controls* within two genes (*ITGA4* and *CD300LF*) that may be involved in the regulation of immune functions. The CD300 family contains both inhibiting and activating receptors. The gene *CD300LF*, found to contain four SNPs that cause a missense mutation, may act as a negative regulator of TLR signalling and the FcR-mediated activation of mast cells (Izawa et al. 2014). By contrast, *CD300LF* may positively regulate the IL4-mediated signalling pathway by acting as a coreceptor for IL-4 (Moshkovits et al. 2015); a cytokine signalling gene that has been previously implicated in morbillivirus immune responses, and found to be putatively under selection in this study. *ITGA4* may also promote viral resistance by permitting T-lymphocytes to migrate to sites of inflammation. In particular, increased expression of *ITGA4* has been associated with elevated T-lymphocyte immune activity (Dedrick 2007).

In addition, several of the genes that were found to be putatively under selection here may also be involved in other adaptive pathways that relate to the signalling and differentiation of T and B cells. These pathways include TCR signalling (*CARD11*), co-stimulation by the CD28 family (GRB2) and signalling by B cell receptors (*NFAT*, *CARD11*, *GRB2*). The candidate gene *NFKBIZ*, which contained a SNP that caused a missense mutation, is also involved in the T cell receptor signalling pathway and the regulation of inflammatory responses. In regards to measles, TCR-mediated signals combined with the ligation of SLAM may trigger the downstream signalling of T helper 2 (Th2) cells, as well as key cytokines (IL-4 and IL-13) to fight infection (Sato et al. 2012). The co-stimulation of CD28 is also critical for T cell differentiation, with increased ligation of CD28 generally resulting in an increased production of IL-4 (Rulifson et al. 1997). A major immune response of humans to measles is controlled by T-lymphocytes that recognise measles antigens (Haralambieva et al. 2015). These T-lymphocytes also play a key role in immune responses of dolphins to CeMV, with seropositive dolphins showing a reduction in T cell proliferation in comparison to healthy dolphins

(Bossart et al. 2011, Bossart et al. 2019, Diaz-Delgado et al. 2019). Throughout the outbreak, all stranded dolphins showed clinical signs of lymphoid depletion (Kemper et al. 2016), suggesting that T cell proliferation may have been reduced, hampering the ability of an individual to fight the infection. Altogether, these results highlight the importance of T-lymphocytes in the ability of dolphins to mount an immune response to fight CeMV.

Cytokine signalling in the immune system

Cytokines and their receptors are very important in the modulation of immune responses and are key components of host defence. Given their role in combating pathogens, cytokines have been the focus of several vaccination efforts against measles (Haralambieva et al. 2015), and are proposed candidate genes for morbillivirus resistance and susceptibility. One of these cytokine signalling receptor genes, *IL4a*, was found to be under selection in this study. Likewise, McCarthy et al. (2011) identified variation within *IL4* of European harbour seals (*Phoca vitulina*) from populations exposed to phocine distemper virus. In our study, the importance of *IL4* has also been highlighted through its link to other immune pathways (i.e. TCR signalling, co-stimulation of CD28, and the regulation by *CD300LF*). These findings provide further support that variation within cytokine signalling genes such as interleukins, and particularly *IL4a,* play an important role in host immune responses to morbilliviruses in general. Four other candidate genes were found to be involved in the cytokine signalling pathway, including *CNTF*, *ATF1*, *GRB2* and *FGF2*. The ciliary neutrophic factor (*CNTF*) is a cytokine with immunomodulatory functions, and has been observed to increase the number of interferon gamma cytokines, resulting in the production of T cells (Tormo et al. 2012). These candidate genes further emphasise the importance of cytokines in mediating and regulating immune responses to morbillivirus infection.

4.4. Other pathways

Other candidate genes found under putative selection in dolphins are involved in multiple pathways that may indirectly be linked to immune responses. For example, signal transduction is an important

process where extracellular signals, such as hormones or growth factors change the cell state or activity (Nair et al. 2019). Here, three candidate genes (*FGF2*, *GRB2* and *CALM1*) are involved in the MAPK family signalling pathway, and another six (*PDE6H, NTRK2, KIDINS220, INHBA, NPSH1, PLCE1*) were associated to the MAPK cascades. These are involved in the initiation of the innate immune system, activation of the adaptive immune system, and cell death after infection (Dong et al. 2002). *MAPK8* was previously suggested to be involved in dolphin susceptibility and resistance to CeMV, and was related to a response to heat stress (Batley et al. 2019). Likewise, it was found to be under selection in this whole genome study (intron region, discussed in immune genes section). The MAPK signalling pathway has been suggested to be involved in functional changes of dendritic cells following contact with measles-hemagglutinin, and may be important in the immunosuppression of this virus (Romanets-Korbut et al. 2016). This suggests that MAPKs may be important not only in responses to stress, but also in immune responses to CeMV. Two other genes (*INHBA* and *BMPR1B*) are involved in signalling by the transforming growth factor (TGF) family members. TGFs have important functions in the regulation of inflammatory responses and in T cell regulation and differentiation (Li et al. 2006). TGFs have been associated with immune responses to a range of diseases (Akdis et al. 2016), but to the best of our knowledge, TGFs have not been implicated in immune responses to morbilliviruses.

4.5. Candidate immune genes

Numerous genes have been proposed to be involved in immune responses to morbilliviruses, including binding genes, pathogen-associated molecular pattern sensing genes, cytokine-cytokine receptor genes, antiviral genes, and vitamin A and D receptor genes (McCarthy et al. 2011, Haralambieva et al. 2015, Stejskalova et al. 2017). Due to the high number of gene annotations, we mainly focused on SNPs that annotated to protein coding regions and therefore may have missed variation between *cases* and *controls* in intronic regions of important immune genes. To counteract this, we assessed genetic variation and investigated putative signatures of selection in 23 genes previously proposed to be important in resistance and susceptibility to morbilliviruses. We found

558 significant differentiation in intronic regions in five of such candidate genes (*RARb*, *MAPK8*,
559 *FBXW11, ANK3* and *ACOX3)*. While the identified SNPs are within non-coding regions, the potential
560 role of these genes in fighting CeMV should not be discarded. A large proportion of the mammalian
561 genome is made up of introns (Chorev et al. 2017), and in the SABD reference genome less than 0.6%
562 of SNPs are within exons (Batley et al. unpublished data). While many introns act in a neutral manner
563 with apparently no function, intronic SNPs might indirectly influence gene function and immune
564 response genes through the alteration of splicing (Dhiman et al. 2008, Seoighe and Korir 2011, Singh
565 et al. 2018, Guigó and Ullrich 2020). Regarding measles, 18 intronic SNPs within TLRs, NFkB,
566 TRAF6 and IkK have been associated with measles-specific humoral and cellular immunity (Dhiman
567 et al. 2008). By contrast, inspection of the genotype counts highlighted a remarkable lack of
568 heterozygosity in important immune genes.

569 Genetic diversity is essential for natural populations to adapt to rapid and ongoing changes to their
570 environment (Manlik et al. 2019). Maintaining genetic diversity is particularly important for
571 populations to recognise and fight infectious diseases (Hendricks et al. 2017). Immune genes are
572 considered amongst some of the most polymorphic genes in wildlife populations (Morris et al. 2015,
573 Ruan et al. 2016, Dooley et al. 2018), with diversity suggested to be maintained through pathogen-
574 host balancing selection (Morris et al. 2015), and an excess of homozygous alleles likely impairing
575 an individual's ability to successfully fight pathogens (Smith et al. 2009, Shafer et al. 2012,
576 Blanchong et al. 2016). The dolphin population studied here has relatively low levels of standing
577 genetic variation compared to neighbouring populations (Pratt et al. 2018, Pratt et al. in prep), and this
578 may have negatively influenced their susceptibility to CeMV. While we observed a high level of
579 polymorphism in many immune genes, we found a lack of heterozygosity in some that are thought to
580 be functionally important. This lack of diversity was observed across *case* and *control* samples, and
581 therefore may not have led to *case* dolphins being more likely to succumb to CeMV, but the
582 population being more susceptible. Within the GSV, *T. aduncus* and *D. delphis* are considered
583 resident species (Kemper et al. 2008). *D. delphis* are very gregarious and form a larger population
584 (Zanardo et al. 2016a, Parra et al. in review) and although a small number of *D. delphis cases* were

585 recorded during the outbreak, the virus did not seem to have a similar impact in this population
586 (Kemper et al. 2016). *D. delphis* from GSV are more genetically diverse than *T. aduncus* from the
587 same bioregion (Bilgmann et al. 2014, Pratt et al. 2018), and this difference in diversity may have
588 influenced their ability to fight and survive CeMV infection.

589

590 **5. Conclusions**

591 This whole-genome association study disclosed the importance of key immune response genes and
592 pathways in the susceptibility and resistance of dolphins to the highly infectious and fatal CeMV. We
593 have uncovered novel genes and pathways that have not previously been the target of morbillivirus
594 immune response studies. In particular, the genes *CD300LF*, *NFATC2* and *NFKBIZ* may be important
595 in the regulation and expression of interleukins and T cells, while the gene pathways FcRs and MAPK
596 cascade may be important for recognising pathogens and activating immune responses, and the
597 initiation and activation of the immune system, respectively. In addition, we found evidence for
598 putative selection in genes previously suggested to be involved in responses to morbilliviruses, adding
599 evidence that knowledge gained on immune responses by one species can be more broadly applied to
600 other morbilliviruses. The results highlighted the importance of cytokines, T cells (particularly Th2)
601 and *IL4*, in fighting infection by these viruses. Overall, our work highlights the complex interactions
602 between the innate, adaptive, and signalling processes of the mammalian immune system in fighting
603 infection by viruses and adds to our understanding of major marine mammal immune responses. The
604 unravelled interactions of the immune systems emphasise the significance of whole genome studies to
605 characterise the interplay of immune responses and genes involved in combating infections.
606 Additional whole genome studies of larger CeMV outbreaks should clarify the role of these genes and
607 pathways across virus strains, and cetacean populations and species. These results add weight that
608 knowledge gained on immune responses by one species may be more broadly applied to other species
609 of morbilliviruses.

610 **Acknowledgements**

611 Fieldwork was conducted under permits from the South Australian Department of Environment,
612 Water and Natural Resources (DEWNR), permit #K25761-6, and under Ministerial Exemption from
613 Primary Industries Resources South Australia (PIRSA), exemption #9902648. Ethics approval was
614 obtained from the Flinders University Animal Welfare Committee, project #E375. We thank the many
615 people who assisted with the collection of stranded animals during the unusual mortality event, and
616 the South Australian Museum (SAM) for their contribution of samples. We also thank members of the
617 Cetacean Ecology Behaviour and Evolution Lab (CEBEL) members, particularly E. Pratt and K.
618 Bilgmann who assisted with fieldwork for the collection of biopsy samples. Financial support was
619 provided by the Sea World Research and Rescue Foundation Inc (SWRRFI), Flinders University and
620 the Australian Government Research Training Program Scholarship to K. B.

621 **Data Accessibility**

622 The sequence data will be made available on the Sequence Read Archive (SRA) upon acceptance of
623 the manuscript.

624 **Author Contributions**

625 The study was designed by KB and LM, in conjunction with LB. Tissue samples from stranded
626 dolphins and information about samples and the outbreak were provided by CK and IT. Biopsy
627 samples of live dolphins were collected by LM and NZ. Laboratory work and bioinformatics were
628 primarily conducted by KB, with guidance and assistance from JSC. Data analysis and interpretation
629 was conducted by KB, with guidance from LM, JSC and LB. KB and LM wrote the paper, with
630 critical revisions made by LB, JSC, CK, IT and NZ.

631 **References**

632 Acevedo-Whitehouse, K., and A. A. Cunningham. 2006. Is MHC enough for understanding
633 wildlife immunogenetics? Trends in Ecology & Evolution **21**:433-438.

634 Acevedo-Whitehouse, K., F. Gulland, D. Greig, and W. Amos. 2003. Disease susceptibility in
635 California sea lions. Nature **422**:35.
636 Akdis, M., A. Aab, C. Altunbulakli, K. Azkur, R. A. Costa, R. Crameri, S. Duan, T. Eiwegger,
637 A. Eljaszewicz, R. Ferstl, R. Frei, M. Garbani, A. Globinska, L. Hess, C. Huitema, T.
638 Kubo, Z. Komlosi, P. Konieczna, N. Kovacs, U. C. Kucuksezer, N. Meyer, H. Morita,
639 J. Olzhausen, L. O'Mahony, M. Pezer, M. Prati, A. Rebane, C. Rhyner, A. Rinaldi, M.
640 Sokolowska, B. Stanic, K. Sugita, A. Treis, W. van de Veen, K. Wanke, M.
641 Wawrzyniak, P. Wawrzyniak, O. F. Wirz, J. S. Zakzuk, and C. A. Akdis. 2016.
642 Interleukins (from IL-1 to IL-38), interferons, transforming growth factor beta, and
643 TNF-alpha: Receptors, functions, and roles in diseases. J Allergy Clin Immunol
644 **138**:984-1010.
645 Alfonso, C. L., G. K. Amarasinghe, K. Bányai, Y. Bào, C. F. Basler, S. Bavari, and A.
646 Bukreyev. 2016. Taxonomy of the order Mononegavirales. Archives of Virology
647 **161**:2351-2360.
648 Altizer, S., D. Harvell, and E. Friedle. 2003. Rapid evolutionary dynamics and disease
649 threats to biodiversity. Trends in Ecology & Evolution **18**:589-596.
650 Amarante-Mendes, G. P., S. Adjemian, L. M. Branco, L. C. Zanetti, R. Weinlich, and K. R.
651 Bortoluci. 2018. Pattern Recognition Receptors and the Host Cell Death Molecular
652 Machinery. Front Immunol **9**:2379.
653 Axelsson, E., A. Ratnakumar, M. L. Arendt, K. Maqbool, M. T. Webster, M. Perloski, O.
654 Liberg, J. M. Arnemo, A. Hedhammar, and K. Lindblad-Toh. 2013. The genomic
655 signature of dog domestication reveals adaptation to a starch-rich diet. Nature
656 **495**:360-364.
657 Banks, J., K. Levine, M. Syvanen, J. Theis, and A. Gilson. 1995. DNA tissue gender typing.
658 Standard operating procedure of the Wildlife Forensic Laboratory of the California
659 Department of Fish and Game. Sacramento, CA: California Department of Fish and
660 Game.
661 Batley, K. C., J. Sandoval-Castillo, L. Beheregaray, and L. Moller. in prep. The southern
662 Australian bottlenose dolphin genome.
663 Batley, K. C., J. Sandoval-Castillo, C. M. Kemper, C. R. M. Attard, N. Zanardo, I. Tomo, L. B.
664 Beheregaray, and L. M. Möller. 2019. Genome-wide association study of an unusual
665 dolphin mortality event reveals candidate genes for susceptibility and resistance to
666 cetacean morbillivirus. Evolutionary Applications **12**:718-732.
667 Ben Mkaddem, S., M. Benhamou, and R. C. Monteiro. 2019. Understanding Fc Receptor
668 Involvement in Inflammatory Diseases: From Mechanisms to New Therapeutic Tools.
669 Front Immunol **10**:811.
670 Bieback, K., E. Lien, I. M. Klagge, E. Avota, J. Schneider-Schaulies, W. P. Duprex, H.
671 Wagner, C. J. Kirschning, V. Ter Meulen, and S. Schneider-Schaulies. 2002.
672 Hemagglutinin protein of wild-type measles virus activates toll-like receptor 2
673 signaling. J Virol **76**:8729-8736.
674 Bilgmann, K., O. J. Griffiths, S. J. Allen, and L. M. Möller. 2007. A Biopsy Pole System for
675 Bow-Riding Dolphins: Sampling Success, Behavioral Responses, and Test for
676 Sampling Bias. Marine Mammal Science **23**:218-225.
677 Bilgmann, K., G. J. Parra, L. Holmes, K. J. Peters, I. D. Jonsen, and L. M. Moller. 2019.
678 Abundance estimates and habitat preferences of bottlenose dolphins reveal the
679 importance of two gulfs in South Australia. Sci Rep **9**:8044.
680 Bilgmann, K., G. J. Parra, N. Zanardo, L. B. Beheregaray, and L. M. Möller. 2014. Multiple
681 management units of short-beaked common dolphins subject to fisheries bycatch off
682 southern and southeastern Australia. Marine Ecology Progress Series **500**:265-279.
683 Blanchong, J. A., S. J. Robinson, M. D. Samuel, and J. T. Foster. 2016. Application of
684 genetics and genomics to wildlife epidemiology. The Journal of Wildlife Management
685 **80**:593-608.
686 Bossart, G. D., T. A. Romano, M. M. Peden-Adams, A. Schaefer, S. McCulloch, J. D.
687 Goldstein, C. D. Rice, J. T. Saliki, P. A. Fair, and J. S. Reif. 2011.
688 Clinicoimmunopathologic findings in Atlantic bottlenose dolphins Tursiops truncatus

689 with positive cetacean morbillivirus antibody titers. Diseases of Aquatic Organisms
690 **97**:103-112.
691 Bossart, G. D., T. A. Romano, M. M. Peden-Adams, A. M. Schaefer, C. D. Rice, P. A. Fair,
692 and J. S. Reif. 2019. Comparative Innate and Adaptive Immune Responses in
693 Atlantic Bottlenose Dolphins (Tursiops truncatus) With Viral, Bacterial, and Fungal
694 Infections. Frontiers in Immunology **10**.
695 Brauer, C. J., M. P. Hammer, and L. B. Beheregaray. 2016. Riverscape genomics of a
696 threatened fish across a hydroclimatically heterogeneous river basin. Molecular
697 Ecology **25**:5093-5113.
698 Burge, C. A., C. Mark Eakin, C. S. Friedman, B. Froelich, P. K. Hershberger, E. E. Hofmann,
699 L. E. Petes, K. C. Prager, E. Weil, B. L. Willis, S. E. Ford, and C. D. Harvell. 2014.
700 Climate change influences on marine infectious diseases: implications for
701 management and society. Ann Rev Mar Sci **6**:249-277.
702 Cammen, K. M., L. A. Wilcox, P. E. Rosel, R. S. Wells, and A. J. Read. 2015. From genome-
703 wide to candidate gene: an investigation of variation at the major histocompatibility
704 complex in common bottlenose dolphins exposed to harmful algal blooms.
705 Immunogenetics **67**:125-133.
706 Charlton-Robb, K., L. A. Gershwin, R. Thompson, J. Austin, K. Owen, and S. McKechnie.
707 2011. A new dolphin species, the Burrunan Dolphin Tursiops australis sp. nov.,
708 endemic to southern Australian coastal waters. PLoS ONE **6**:e24047.
709 Chorev, M., A. Joseph Bekker, J. Goldberger, and L. Carmel. 2017. Identification of introns
710 harboring functional sequence elements through positional conservation. Sci Rep
711 **7**:4201.
712 Cingolani, P., A. Platts, L. Wang le, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, and D.
713 M. Ruden. 2012. A program for annotating and predicting the effects of single
714 nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster
715 strain w1118; iso-2; iso-3. Fly (Austin) **6**:80-92.
716 Cunningham, A. A., P. Daszak, and J. L. N. Wood. 2017. One Health, emerging infectious
717 diseases and wildlife: two decades of progress? Philosophical Transactions of the
718 Royal Society B **372**:10.1098/rstb.2016.0167.
719 da Fontoura Budaszewski, R., and V. von Messling. 2016. Morbillivirus Experimental Animal
720 Models: Measles Virus Pathogenesis Insights from Canine Distemper Virus. Viruses
721 **8**:10.3390/v8100274.
722 Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E.
723 Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, and G.
724 Genomes Project Analysis. 2011. The variant call format and VCFtools.
725 Bioinformatics **27**:2156-2158.
726 de Sa, A. L. A., B. Breaux, T. C. T. Burlamaqui, T. C. Deiss, L. Sena, M. F. Criscitiello, and
727 M. P. C. Schneider. 2019. The Marine Mammal Class II Major Histocompatibility
728 Complex Organization. Front Immunol **10**:696.
729 Dedrick, R. L. 2007. Understanding gene expression patterns in immune-mediated
730 disorders. J Immunotoxicol **4**:201-207.
731 Dhiman, N., I. G. Ovsyannikova, R. A. Vierkant, J. E. Ryan, V. S. Pankratz, R. M. Jacobson,
732 and G. A. Poland. 2008. Associations between SNPs in toll-like receptors and related
733 intracellular signaling molecules and immune responses to measles vaccine:
734 preliminary results. Vaccine **26**:1731-1736.
735 Di Guardo, G., and S. Mazzariol. 2014. Cetacean morbillivirus in Northern and Southern
736 Hemispheres. Frontiers in Microbiology **5**.
737 Di Guardo, G., and S. Mazzariol. 2016. Cetacean Morbillivirus-associated pathology:
738 Knowns and unknowns. Frontiers in Microbiology **7**:10.3389/fmicb.2016.00112.
739 Diaz-Delgado, J., K. R. Groch, E. Sierra, S. Sacchini, D. Zucca, O. Quesada-Canales, M.
740 Arbelo, A. Fernandez, E. Santos, J. Ikeda, R. Carvalho, A. F. Azevedo, J. Lailson-
741 Brito, Jr., L. Flach, R. Ressio, C. T. Kanamura, M. Sansone, C. Favero, B. F. Porter,
742 C. Centelleghe, S. Mazzariol, L. Di Renzo, G. Di Francesco, G. Di Guardo, and J. L.
743 Catao-Dias. 2019. Comparative histopathologic and viral immunohistochemical

studies on CeMV infection among Western Mediterranean, Northeast-Central, and Southwestern Atlantic cetaceans. PLoS ONE **14**:e0213363.
Diaz-Delgado, J., E. Sierra, A. I. Vela, M. Arbelo, D. Zucca, K. R. Groch, and A. Fernandez. 2017. Coinfection by Streptococcus phocae and cetacean morbillivirus in a short-beaked common dolphin Delphinus delphis. Dis Aquat Organ **124**:247-252.
Dong, C., R. J. Davis, and R. A. Flavell. 2002. MAP kinases in the immune response. Annu Rev Immunol **20**:55-72.
Dooley, C. T., T. Ferrer, H. Pagan, and G. M. O'Corry-Crowe. 2018. Bridging immunogenetics and immunoproteomics: Model positional scanning library analysis for Major Histocompatibility Complex class II DQ in Tursiops truncatus. PLoS ONE **13**:e0201299.
Duignan, P. J., J. R. Geraci, J. A. Raga, and N. Calzada. 1992. Pathology of Morbillivirus Infection in Striped Dolphins (Stenella coeruleoalba) from Valencia and Murcia, Spain. Canadian Journal of Veterinary Research **56**:242-248.
Elbers, J. P., M. B. Brown, and S. S. Taylor. 2018. Identifying genome-wide immune gene variation underlying infectious disease in wildlife populations - a next generation sequencing approach in the gopher tortoise. BMC Genomics **19**:64.
Fabregat, A., S. Jupe, L. Matthews, K. Sidiropoulos, M. Gillespie, P. Garapati, R. Haw, B. Jassal, F. Korninger, B. May, M. Milacic, C. D. Roca, K. Rothfels, C. Sevilla, V. Shamovsky, S. Shorser, T. Varusai, G. Viteri, J. Weiser, G. Wu, L. Stein, H. Hermjakob, and P. D'Eustachio. 2018. The Reactome Pathway Knowledgebase. Nucleic Acids Res **46**:D649-D655.
Fric, J., T. Zelante, and P. Ricciardi-Castagnoli. 2014. Phagocytosis of Particulate Antigens - All Roads Lead to Calcineurin/NFAT Signaling Pathway. Front Immunol **4**:513.
Garbi, N., S. Tanaka, F. Momburg, and G. J. Hammerling. 2006. Impaired assembly of the major histocompatibility complex class I peptide-loading complex in mice deficient in the oxidoreductase ERp57. Nat Immunol **7**:93-102.
Gelain, M. E., and F. Bonsembiante. 2019. Acute Phase Proteins in Marine Mammals: State of Art, Perspectives and Challenges. Front Immunol **10**:1220.
Gui, D., K. Jia, J. Xia, L. Yang, J. Chen, Y. Wu, and M. Yi. 2013. De novo assembly of the Indo-Pacific humpback dolphin leucocyte transcriptome to identify putative genes involved in the aquatic adaptation and immune response. PLoS ONE **8**:e72417.
Guigó, R., and S. Ullrich. 2020. Dynamic changes in intron retention are tightly associated with regulation of splicing factors and proliferative activity during B-cell development. Nucleic Acids Research **48**:1327-1340.
Gulland, F. M. D., and A. J. Hall. 2007. Is Marine Mammal Health Deteriorating? Trends in the Clobal Reporting of Marine Mammal Disease. EcoHealth **4**:135-150.
Hara, H., T. Wada, C. Bakal, I. Kozieradzki, S. Suzuki, N. Suzuki, M. Nghiem, E. K. Griffiths, C. Krawczyk, B. Bauer, F. D'Acquisto, S. Ghosh, W.-C. Yeh, G. Baier, R. Rottapel, and J. M. Penninger. 2003. The MAGUK Family Protein CARD11 Is Essential for Lymphocyte Activation. Immunity **18**:763-775.
Haralambieva, I. H., R. B. Kennedy, I. G. Ovsyannikova, J. A. Whitaker, and G. A. Poland. 2015. Variability in Humoral Immunity to Measles Vaccine: New Developments. Trends in Molecular Medicine **21**:789-801.
Hashiguchi, T., T. Ose, M. Kubota, N. Maita, J. Kamishikiryo, K. Maenaka, and Y. Yanagi. 2011. Structure of the measles virus hemagglutinin bound to its cellular receptor SLAM. Nature Structural & Molecular Biology **18**:135-141.
Hendricks, S., B. Epstein, B. Schonfeld, C. Wiench, R. Hamede, M. Jones, A. Storfer, and P. Hohenlohe. 2017. Conservation implications of limited genetic diversity and population structure in Tasmanian devils (Sarcophilus harrisii). Conserv Genet **18**:977-982.
Izawa, K., M. Isobe, T. Matsukawa, S. Ito, A. Maehara, M. Takahashi, Y. Yamanishi, A. Kaitani, T. Oki, K. Okumura, T. Kitamura, and J. Kitaura. 2014. Sphingomyelin and ceramide are physiological ligands for human LMIR3/CD300f, inhibiting FcepsilonRI-mediated mast cell activation. J Allergy Clin Immunol **133**:270-273 e271-277.

Jo, W. K., J. Kruppa, A. Habierski, M. van de Bildt, S. Mazzariol, G. Di Guardo, U. Siebert, T. Kuiken, K. Jung, A. Osterhaus, and M. Ludlow. 2018. Evolutionary evidence for multi-host transmission of cetacean morbillivirus. Emerg Microbes Infect **7**:201.

Johnson, R. N., D. O'Meally, Z. Chen, G. J. Etherington, S. Y. W. Ho, W. J. Nash, C. E. Grueber, Y. Cheng, C. M. Whittington, S. Dennison, E. Peel, W. Haerty, R. J. O'Neill, D. Colgan, T. L. Russell, D. E. Alquezar-Planas, V. Attenbrow, J. G. Bragg, P. A. Brandies, A. Y. Chong, J. E. Deakin, F. Di Palma, Z. Duda, M. D. B. Eldridge, K. M. Ewart, C. J. Hogg, G. J. Frankham, A. Georges, A. K. Gillett, M. Govendir, A. D. Greenwood, T. Hayakawa, K. M. Helgen, M. Hobbs, C. E. Holleley, T. N. Heider, E. A. Jones, A. King, D. Madden, J. A. M. Graves, K. M. Morris, L. E. Neaves, H. R. Patel, A. Polkinghorne, M. B. Renfree, C. Robin, R. Salinas, K. Tsangaras, P. D. Waters, S. A. Waters, B. Wright, M. R. Wilkins, P. Timms, and K. Belov. 2018. Adaptation and conservation insights from the koala genome. Nat Genet **50**:1102-1111.

Kardos, M., G. Luikart, R. Bunch, S. Dewey, W. Edwards, S. McWilliam, J. Stephenson, F. W. Allendorf, J. T. Hogg, and J. Kijas. 2015. Whole-genome resequencing uncovers molecular signatures of natural and sexual selection in wild bighorn sheep. Mol Ecol **24**:5616-5632.

Karlsson, E. K., D. P. Kwiatkowski, and P. C. Sabeti. 2014. Natural selection and infectious disease in human populations. Nature Reviews Genetics **15**:379-393.

Kemper, C. M., M. Bossley, and P. Shaughnessy. 2008. Marine mammals of Gulf St Vincent, investigator strait and backstairs passage.339–352.

Kemper, C. M., I. Tomo, J. Bingham, S. S. Bastianello, J. Wang, S. E. Gibbs, L. Woolford, C. Dickason, and D. Kelly. 2016. Morbillivirus-associated unusual mortality event in South Australian bottlenose dolphins is largest reported for the Southern Hemisphere. Royal Society Open Science **3**:10.1098/rsos.160838.

Kerdiles, Y. M., B. Cherif, J. C. Marie, N. Tremillion, B. Blanquier, G. Libeau, A. Diallo, T. F. Wild, M. B. Villiers, and B. Horvat. 2006. Immunomodulatory properties of morbillivirus nucleoproteins. Viral immunology **19**:324-334.

Kipreos, E. T., and M. Pagano. 2000. The F-box protein family. Genome Biology **1**:3002.3001-3002.3007.

Klein, M., S. Klein-Hessling, A. Palmetshofer, E. Serfling, C. Tertilt, T. Bopp, V. Heib, M. Becker, C. Taube, H. Schild, E. Schmitt, and M. Stassen. 2006. Specific and redundant roles for NFAT transcription factors in the expression of mast cell-derived cytokines. J Immunol **177**:6667-6674.

Krützen, M., L. Barre, L. M. Möller, M. Heithaus, C. Simms, and W. Sherwin. 2002. A biopsy system for small cetaceans: Darting success and wound healing in Tursiops spp. . Marine Mammal Science **18**:863-878.

Lacy, P. 2006. Mechanisms of Degranulation in Neutrophils. Allergy, Asthma & Clinical Immunology **2**.

Li, M. O., Y. Y. Wan, S. Sanjabi, A. K. Robertson, and R. A. Flavell. 2006. Transforming growth factor-beta regulation of immune responses. Annu Rev Immunol **24**:99-146.

Liu, T., L. Zhang, D. Joo, and S. C. Sun. 2017. NF-kappaB signaling in inflammation. Signal Transduct Target Ther **2**.

Manlik, O., M. Krützen, A. M. Kopps, J. Mann, L. Bejder, S. J. Allen, C. Frère, R. C. Connor, and W. B. Sherwin. 2019. Is MHC diversity a better marker for conservation than neutral genetic diversity? A case study of two contrasting dolphin populations. Ecology and Evolution.

Marie, J. C., J. Kehren, M. C. Trescol-Biemont, A. Evlashev, H. Valentin, T. Walzer, R. Tedone, B. Loveland, J. F. Nicolas, C. Rabourdin-Combe, and B. Horvat. 2001. Mechanism of Measles Virus–Induced Suppression of Inflammatory Immune Responses. Immunity **14**.

Marsili, L., G. Di Guardo, S. Mazzariol, and S. Casini. 2019. Insights Into Cetacean Immunology: Do Ecological and Biological Factors Make the Difference? Front Immunol **10**:1219.

Martin, M. P., and M. Carrington. 2005. Immunogenetics of viral infections. Current Opinion in Immunology **17**:510-516.

McCarthy, A. J., M. A. Shaw, P. D. Jepson, S. M. J. M. Brasseur, P. J. H. Reijnders, and S. J. Goodman. 2011. Variation in European harbour seal immune response genes and susceptibility to phocine distemper virus (PDV). Infection, Genetics and Evolution **11**:1616-1623.

Moreno-Santillan, D. D., E. A. Lacey, D. Gendron, and J. Ortega. 2016. Genetic Variation at Exon 2 of the MHC Class II DQB Locus in Blue Whale (Balaenoptera musculus) from the Gulf of California. PLoS ONE **11**:e0141296.

Morens, D. M., G. K. Folkers, and A. S. Fauci. 2004. The challenge of emerging and re-emerging infectious diseases. Nature **430**:242-249.

Morris, K. M., B. Wright, C. E. Grueber, C. Hogg, and K. Belov. 2015. Lack of genetic diversity across diverse immune genes in an endangered mammal, the Tasmanian devil (Sarcophilus harrisii). Mol Ecol **24**:3860-3872.

Moshkovits, I., D. Karo-Atar, M. Itan, H. Reichman, P. Rozenberg, N. Morgenstern-Ben-Baruch, D. Shik, A. Ejarque-Ortiz, A. Y. Hershko, L. Tian, J. E. Coligan, J. Sayos, and A. Munitz. 2015. CD300f associates with IL-4 receptor alpha and amplifies IL-4-induced immune cell responses. Proc Natl Acad Sci U S A **112**:8708-8713.

Moura, A. E., K. Shreves, M. Pilot, K. R. Andrews, D. M. Moore, T. Kishida, L. Möller, A. Natoli, S. Gaspari, M. McGowen, I. Chen, H. Gray, M. Gore, R. M. Culloch, M. S. Kiani, M. Sarrouf Willson, A. Bulushi, T. Collins, R. Baldwin, A. Willson, G. Minton, L. Ponnampalam, and A. Rus Hoelzel. 2020. Phylogenomics of the genus Tursiops and closely related Delphininae reveals extensive reticulation among lineages and provides inference about eco-evolutionary drivers. Molecular Phylogenetics and Evolution.

Naegelen, I., N. Beaume, S. Plancon, V. Schenten, E. J. Tschirhart, and S. Brechard. 2015. Regulation of Neutrophil Degranulation and Cytokine Secretion: A Novel Model Approach Based on Linear Fitting. J Immunol Res **2015**:817038.

Nair, A., P. Chauhan, B. Saha, and K. F. Kubatzky. 2019. Conceptual Evolution of Cell Signaling. Int J Mol Sci **20**.

Ohishi, K., T. Maruyama, F. Seki, and M. Takeda. 2019. Marine Morbilliviruses: Diversity and Interaction with Signaling Lymphocyte Activation Molecules. Viruses **11**.

Ohishi, K., R. Shishido, Y. Iwata, M. Saitoh, R. Takenaka, D. Ohtsu, K. Okutsu, and T. Maruyama. 2011. Lipopolysaccharide-induced innate immune factors in the bottlenose dolphin (Tursiops truncatus) detected in expression sequence tag analysis. Microbiol Immunol **55**:790-797.

Padalino, I., G. Di Guardo, A. Carbone, P. Troiano, A. Parisi, D. Galante, M. A. Cafiero, M. Caruso, L. Palazzo, L. Guarino, L. De Riso, C. Centelleghe, S. Mazzariol, and A. Petrella. 2019. Dolphin Morbillivirus in Eurasian Otters, Italy. Emerging Infectious Diseases **25**:372-374.

Pagan, H. J. T., T. Ferrer, and G. O'Corry-Crowe. 2018. Positive selection in coding regions and motif duplication in regulatory regions of bottlenose dolphin MHC class II genes. PLoS ONE **13**:e0203450.

Parra, G., K. Bilgmann, K. Peters, and L. Moller. in review. Abundance and potential biological removal of common dolphins subject to fishery-impacts in South Australian Waters.

Pratt, E. A. L., L. B. Beheregaray, K. Bilgmann, N. Zanardo, F. Diaz-Aguirre, and L. M. Möller. 2018. Hierarchical metapopulation structure in a highly mobile marine predator: the southern Australian coastal bottlenose dolphin (Tursiops cf. australis). Conservation genetics **19**:637-654.

Pressinotti, N. C., H. Klocker, G. Schafer, V. D. Luu, M. Ruschhaupt, R. Kuner, E. Steiner, A. Poustka, G. Bartsch, and H. Sultmann. 2009. Differential expression of apoptotic genes PDIA3 and MAP3K5 distinguishes between low- and high-risk prostate cancer. Mol Cancer **8**:130.

908 Reed, J., R. Harcourt, L. New, and K. Bilgmann. 2020. Extreme Effects of Extreme
909 Disturbances: A Simulation Approach to Assess Population Specific Responses.
910 Frontiers in Marine Science **7**.
911 Romanets-Korbut, O., L. M. Kovalevska, T. Seya, S. P. Sidorenko, and B. Horvat. 2016.
912 Measles virus hemagglutinin triggers intracellular signaling in CD150-expressing
913 dendritic cells and inhibits immune response. Cell Mol Immunol **13**:828-838.
914 Ruan, R., J. Ruan, X. L. Wan, Y. Zheng, M. M. Chen, J. S. Zheng, and D. Wang. 2016.
915 Organization and characteristics of the major histocompatibility complex class II
916 region in the Yangtze finless porpoise (Neophocaena asiaeorientalis asiaeorientalis).
917 Sci Rep **6**:22471.
918 Rulifson, I. C., A. Sperling, P. E. Fields, F. W. Fitch, and J. A. Bluestone. 1997. CD28
919 Costimulation Promotes the Production of Th2 Cytokines. The Journal of
920 Immunology **158**:658-665.
921 Sacristán, C., M. Carballo, M. J. Muñoz, E. N. Bellière, E. Neves, V. Nogal, and F. Esperón.
922 2015. Diagnosis of Cetacean morbillivirus: A sensitive one step real time RT fast-
923 PCR method based on SYBR® Green. Journal of Virological Methods **226**:25-30.
924 Santos, S. G., E. C. Campbell, S. Lynch, V. Wong, A. N. Antoniou, and S. J. Powis. 2007.
925 Major histocompatibility complex class I-ERp57-tapasin interactions within the
926 peptide-loading complex. J Biol Chem **282**:17587-17593.
927 Sato, H., M. Yoneda, T. Honda, and C. Kai. 2012. Morbillivirus receptors and tropism:
928 multiple pathways for infection. Front Microbiol **3**:75.
929 Scholz, C., and R. Tampe. 2009. The peptide-loading complex--antigen translocation and
930 MHC class I loading. Biol Chem **390**:783-794.
931 Seoighe, C., and P. K. Korir. 2011. Evidence for intron length conservation in a set of
932 mammalian genes associated with embryonic development. BMC Bioinformatics **12**
933 **Suppl 9**:S16.
934 Shafer, A. B., C. W. Fan, S. D. Cote, and D. W. Coltman. 2012. (Lack of) genetic diversity in
935 immune genes predates glacial isolation in the North American mountain goat
936 (Oreamnos americanus). J Hered **103**:371-379.
937 Singh, I., S.-H. Lee, A. S. Sperling, M. K. Samur, Y.-T. Tai, M. Fulciniti, N. C. Munshi, C.
938 Mayr, and C. S. Leslie. 2018. Widespread intronic polyadenylation diversifies
939 immune cell transcriptomes. Nature Communications **9**.
940 Smith, K. F., K. Acevedo-Whitehouse, and A. B. Pedersen. 2009. The role of infectious
941 diseases in biological conservation. Animal Conservation **12**:1-12.
942 Stejskalova, K., Z. Bayerova, J. Futas, K. Hrazdilova, M. Klumplerova, J. Oppelt, P.
943 Splichalova, G. Di Guardo, S. Mazzariol, C. E. Di Francesco, G. Di Francesco, G.
944 Terracciano, R. M. Paiu, T. D. Ursache, D. Modry, and P. Horin. 2017. Candidate
945 gene molecular markers as tools for analyzing genetic susceptibility to Morbillivirus
946 infection in stranded Cetaceans. HLA:10.1111/tan.13146.
947 Stephens, N., P. J. Duignan, J. Wang, J. Bingham, H. Finn, L. Bejder, I. A. P. Patterson, and
948 C. Holyoake. 2014. Cetacean morbillivirus in coastal indo-pacific bottlenose dolphins,
949 Western Australia. Emerging Infectious Diseases **20**:666-670.
950 Stone, B. M., D. J. Blyde, J. T. Saliki, U. Blas-Machado, J. Bingham, A. Hyatt, J. Wang, J.
951 Payne, and S. Crameri. 2011. Fatal cetacean morbillivirus infection in an Australian
952 offshore bottlenose dolphin (Tursiops truncatus). Australian Veterinary Journal
953 **89**:452-457.
954 Stone, B. M., D. J. Blyde, J. T. Saliki, and J. M. Morton. 2012. Morbillivirus infection in live
955 stranded, injured, trapped, and captive cetaceans in southeastern Queensland and
956 Northern New South Wales, Australia. Journal of Wildlife Diseases **48**:47-55.
957 Strehl, B., U. Seifert, E. Krüger, S. Heink, U. Kuckelkorn, and P. M. Kloetzel. 2005.
958 Interferon-g, the functional plasticity of the ubiquitin–proteasome system, and MHC
959 class I antigen processing. Immunological Reviews **207**:19-30.
960 Sunnucks, P., and D. F. Hales. 1996. Numerous transposed sequences of mitochondrial
961 cytochrome oxidase I-II in aphids of the genus Sitobion (Hemiptera: Aphididae).
962 Molecular Biology and Evolution **13**:510-524.

963 Takai, T. 2002. Roles of Fc receptors in autoimmunity. Nature Reviews Immunology **2**:580–
964 592.
965 Takai, T. 2005. Fc Receptors and Their Role in Immune Regulation and Autoimmunity.
966 Journal of Clinical Immunology **25**.
967 Thome, M. 2004. CARMA1, BCL-10 and MALT1 in lymphocyte development and activation.
968 Nature Reviews Immunology **4**:348-359.
969 Titcomb, G. C., C. L. Jerde, and H. S. Young. 2019. High-Throughput Sequencing for
970 Understanding the Ecology of Emerging Infectious Diseases at the Wildlife-Human
971 Interface. Frontiers in Ecology and Evolution **7**.
972 Tormo, A. J., M. C. Letellier, R. Lissilaa, L. A. Batraville, M. Sharma, W. Ferlin, G. Elson, S.
973 Crabe, and J. F. Gauchat. 2012. The cytokines cardiotrophin-like cytokine/cytokine-
974 like factor-1 (CLC/CLF) and ciliary neurotrophic factor (CNTF) differ in their receptor
975 specificities. Cytokine **60**:653-660.
976 Turner, H., M. Gomez, E. McKenzie, A. Kirchem, A. Lennard, and D. A. Cantrell. 1998. Rac-
977 1 Regulates Nuclear Factor of Activated T Cells (NFAT) C1 Nuclear Translocation in
978 Response to Fc ⇵ Receptor Type 1 Stimulation of Mast Cells. Journal of
979 Experimental Medicine **188**:527-537.
980 Turner, H., and J. P. Kinet. 1999. Signalling through the high-afinity IgE receptor FceRI.
981 Nature **402**.
982 UniProt, C. 2019. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Research
983 **47**:D506-D515.
984 Valsecchi, E., W. Amos, J. A. Raga, M. Podestà, and W. Sherwin. 2004. The effects of
985 inbreeding on mortality during a morbillivirus outbreak in the Mediterranean striped
986 dolphin (Stenella coeruleoalba). Animal Conservation **7**:139-146.
987 Van Bressem, M. F., P. J. Duignan, A. Banyard, M. Barbieri, K. M. Colegrove, S. de Guise,
988 G. di Guardo, A. Dobson, M. Domingo, D. Fauquier, A. Fernandez, T. Goldstein, B.
989 Grenfell, K. R. Groch, F. Gulland, B. A. Jensen, P. D. Jepson, A. Hall, T. Kuiken, S.
990 Mazzariol, S. E. Morris, O. Nielsen, J. A. Raga, T. K. Rowles, J. Saliki, E. Sierra, N.
991 Stephens, B. Stone, I. Tomo, J. Wang, T. Waltzek, and J. F. X. Wellehan. 2014.
992 Cetacean morbillivirus: Current knowledge and future directions. Viruses **6**:5145-
993 5181.
994 Van Bressem, M. F., J. A. Raga, G. Di Guardo, P. D. Jepson, P. J. Duignan, U. Siebert, T.
995 Barrett, M. C. De Oliveira Santos, I. B. Moreno, S. Siciliano, A. Aguilar, and K. Van
996 Waerebeek. 2009. Emerging infectious diseases in cetaceans worldwide and the
997 possible role of environmental stressors. Diseases of Aquatic Organisms **86**:143-157.
998 Weir, B. S., and C. C. Cockerham. 1984. Estimating F-Statistics for the Analysis of
999 Population Structure. Evolution **38**:1358-1370.
1000 Weiss, M. N., D. W. Franks, K. C. Balcomb, D. K. Ellifrit, M. J. Silk, M. A. Cant, and D. P.
1001 Croft. 2020. Modelling cetacean morbillivirus outbreaks in an endangered killer whale
1002 population. Biological Conservation **242**.
1003 Werling, D., and T. W. Jungi. 2003. TOLL-like receptors linking innate and adaptive immune
1004 response. Veterinary Immunology and Immunopathology **91**:1-12.
1005 Williams, E. S., T. Yuill, M. Artois, J. Fischer, and S. A. Haigh. 2002. Emerging infectious
1006 diseases in wildlife. Revue Scientifique et Technique de l'OIE **21**:139-157.
1007 Wright, B., C. E. Willet, R. Hamede, M. Jones, K. Belov, and C. M. Wade. 2017. Variants in
1008 the host genome may inhibit tumour growth in devil facial tumours: evidence from
1009 genome-wide association. Scientific Reports **7**:423.
1010 Wright, H. L., R. J. Moots, R. C. Bucknall, and S. W. Edwards. 2010. Neutrophil function in
1011 inflammation and inflammatory diseases. Rheumatology (Oxford) **49**:1618-1631.
1012 Yang, J., B. Benyamin, B. P. McEvoy, S. Gordon, A. K. Henders, D. R. Nyholt, P. A.
1013 Madden, A. C. Heath, N. G. Martin, G. W. Montgomery, M. E. Goddard, and P. M.
1014 Visscher. 2010. Common SNPs explain a large proportion of the heritability for
1015 human height. Nat Genet **42**:565-569.
1016 Zanardo, N., K. Bilgmann, G. J. Parra, and L. M. Möller. 2016a. Socio-genetic structure of
1017 short-beaked common dolphins in southern Australia. Journal of Zoology **299**:89-97.

Zanardo, N., G. J. Parra, and L. M. Möller. 2016b. Site fidelity, residency, and abundance of bottlenose dolphins (Tursiops sp.) in Adelaide's coastal waters, South Australia. Marine Mammal Science **32**:1381-1401.

Tables and Figures

Table 1: Genotype counts for the most differentiated SNP between *case* and *control* bottlenose dolphins for 23 previously identified genes as being involved in morbillivirus immune responses.

			Case			Control			All samples		
Gene	**# SNPs**	**SNP**	**AA**	**AB**	**BB**	**AA**	**AB**	**BB**	**AA**	**AB**	**BB**
TLR3	23	chr_unlocalized.8:21644753	0	0	19	0	4	30	0	4	49
BSG	37	chr_unlocalized.3:129929304	1	3	15	0	1	33	1	4	48
DQa	319	chr_unlocalized.10:24317329	1	3	15	0	1	33	1	4	48
SLC11A1	12	chr_unlocalized.13:19817092	1	2	16	0	2	32	1	4	48
TLR8	18	chr_unlocalized.3:10277523	0	0	18	0	4	28	0	4	46
TLR7	10	chr_unlocalized.3:10232052	0	5	14	0	1	32	0	6	46
DRb	146	chr_unlocalized.10:24261165	2	2	15	0	4	29	2	6	44
SLAMF1	45	chr_unlocalized.8:49373642	0	0	18	2	6	25	2	6	43
CD209	10	chr_unlocalized.3:124318575	0	5	14	0	2	31	0	7	45
IL10	13	chr_unlocalized.10:97296013	1	4	14	0	4	30	1	8	44
NCR1	55	chr_unlocalized.1:55733562	2	4	12	0	4	30	2	8	42
TRIM6	9	chr_unlocalized.2:60993451	0	2	17	2	11	21	2	13	38
DRa	68	chr_unlocalized.10:24244730	2	5	12	3	12	19	5	17	31
ADAR	20	chr_unlocalized.8:44739127	5	5	9	1	12	21	6	17	30
IL2	3	chr_unlocalized.7:19991473	1	6	12	4	12	18	5	18	30
MAPK8	230	chr_unlocalized.20:20510091	1	2	16	8	16	10	9	18	26
RARb	1160	chr_unlocalized.4:147546454	9	8	2	5	11	18	14	19	20
TNFa	2	chr_unlocalized.10:23544844	4	4	11	5	16	13	9	20	24
MX2	109	chr_unlocalized.4:4369741	8	5	6	4	15	15	12	20	21
FBXW11	2688	chr_unlocalized.11:3389906	10	7	2	3	16	15	13	23	17
NECTIN4	32	chr_unlocalized.6:87201732	6	9	4	5	16	13	11	25	17
ACOX3	837	chr_unlocalized.7:3120566	5	6	8	0	7	26	5	34	13
ANK3	1425	chr_unlocalized.11:3626693	0	0	19	2	14	17	2	36	14

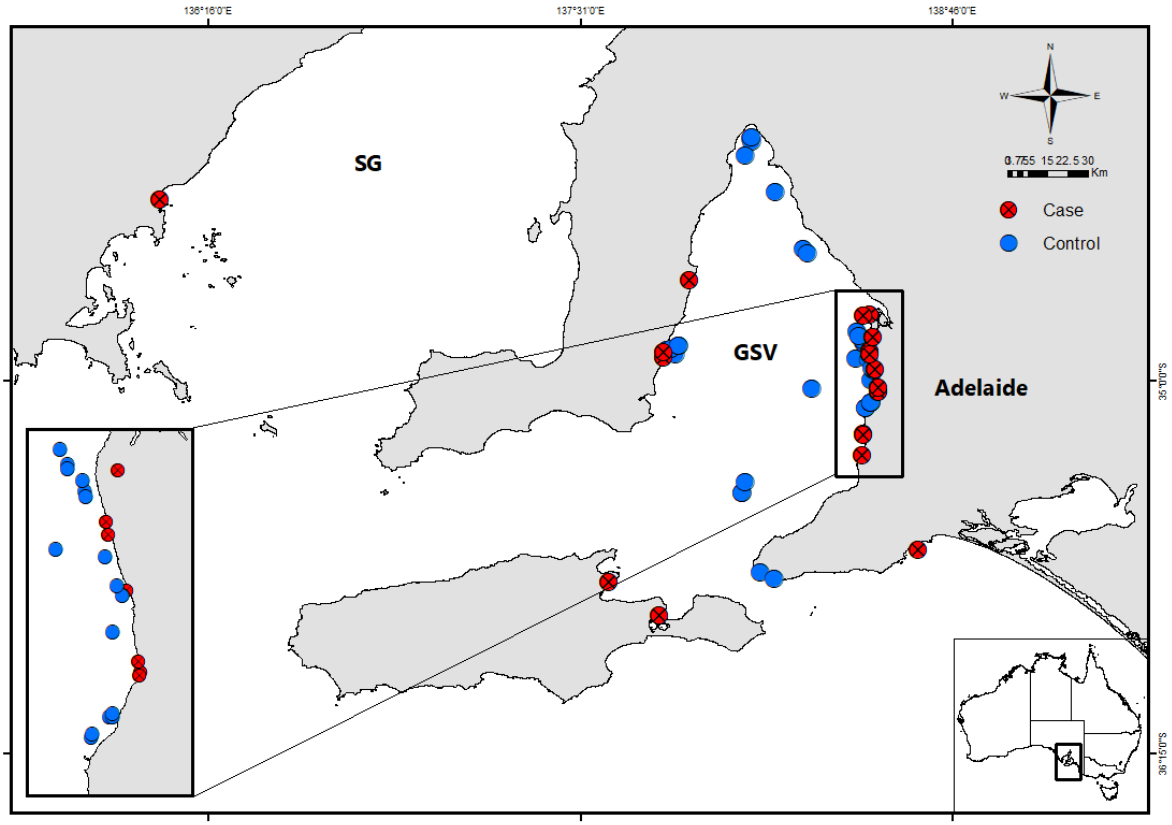


Figure 1: Sampling locations of southern Australian bottlenose dolphins, *Tursiops aduncus*. *case* (n= 19) are non-survivors and *control* (n= 34) are survivors from the 2013, Gulf St. Vincent outbreak used in the whole-genome association study of resistance and susceptibility to CeMV.

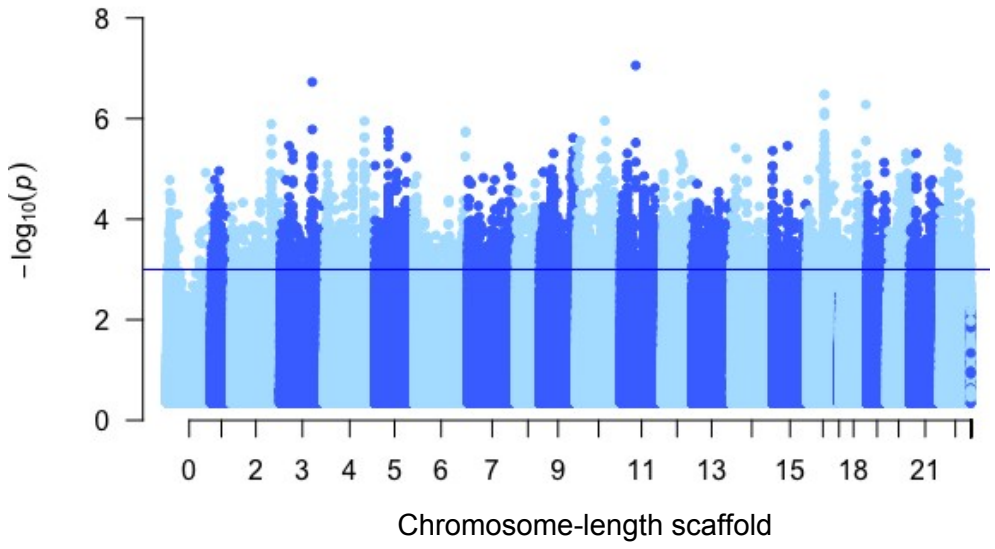


Figure 2: Manhattan plot of the Chi-square test for differences in allele frequencies of 7,720,686 SNPs between *case* and *control* bottlenose dolphins. SNPs above the blue line ($P < 0.001$) represent candidate SNPs putatively associated with cetacean morbillivirus resistance and susceptibility.

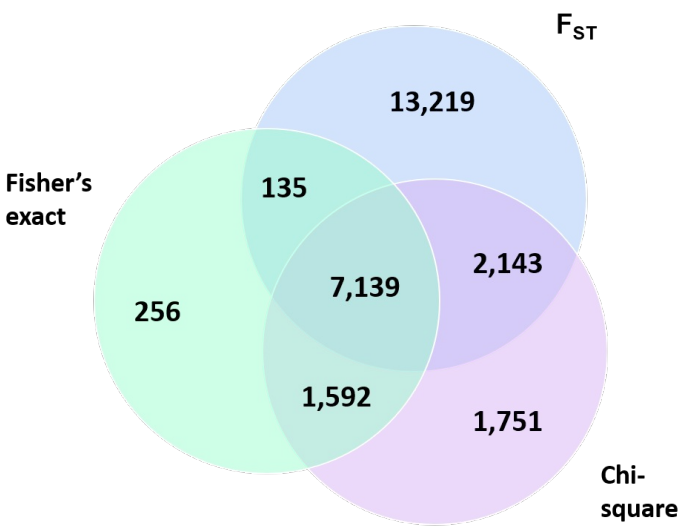
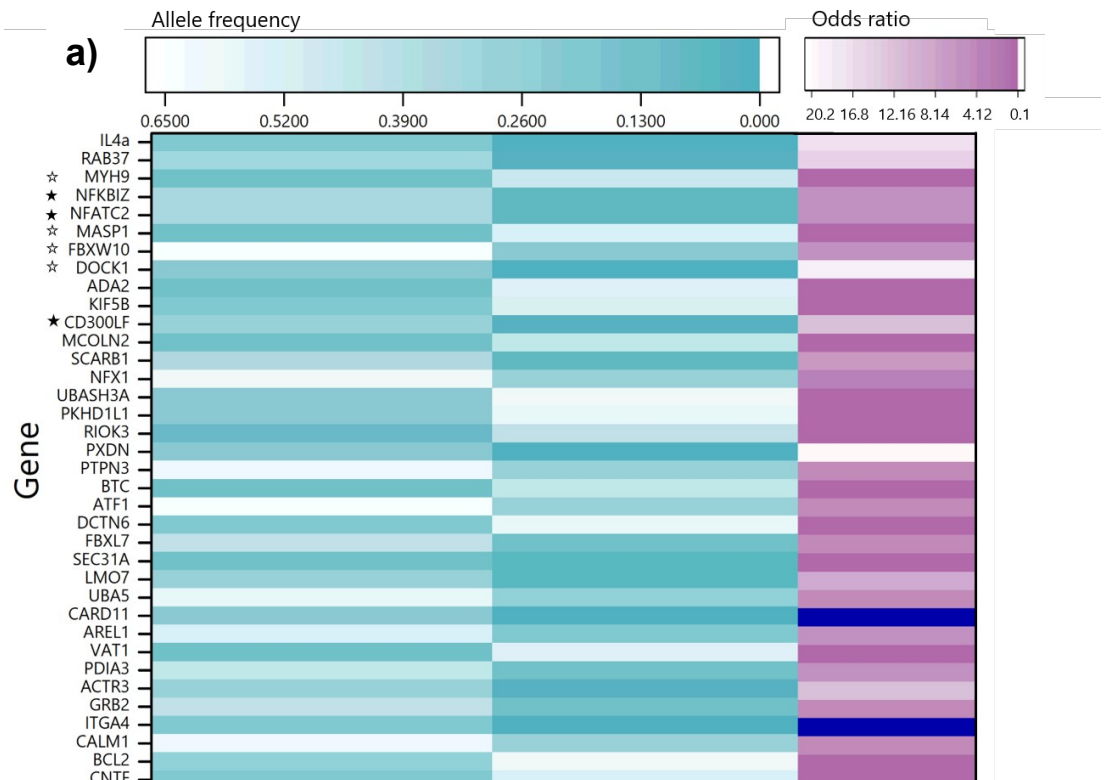


Figure 3: Comparison of the number of candidate SNPs identified in three tests of differentiation between *case* and *control* bottlenose dolphins. Chi-square test (n = 12,625), Fisher's exact test (n = 9,122) and Weir and Cockerham's F_{ST} (n = 22,636).



Cases Controls Odds Ratio

b)

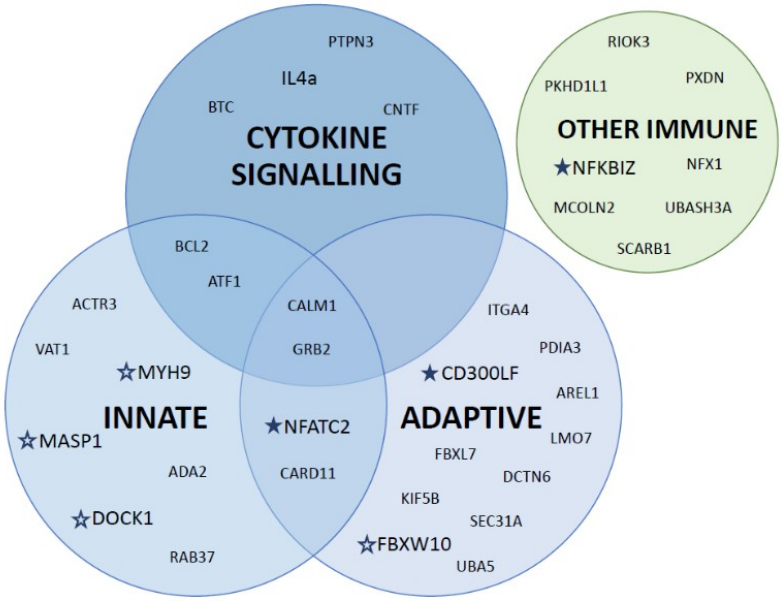


Figure 4: 36 immune-related genes putatively associated with cetacean morbillivirus resistance and susceptibility; **a)** allele frequency differences between *case* and *control* individuals and their corresponding odds ratio. Blue odds ratios represent non-applicable odds ratios (as allele frequency in *controls* = 0); **b)** Immune sub-pathways of the candidate immune-related genes. Other immune refer to the genes that had GO terms relating to immune functions. Stars represent the seven genes that include exonic SNPs (filled = missense change; non-filled = synonymous change).