Transcriptomics of mussel transmissible cancer MtrBTN2 reveals accumulation of multiple cancerous traits and oncogenic pathways shared among bilaterians

Short title: **Transcriptomics of a mussel transmissible cancer**

E.A.V. Burioli[1], M. Hammel[1,2], E. Vignal[1], J. Vidal-Dupiol[1], G. Mitta[3], F. Thomas[4], N. Bierne[2], D. Destoumieux-Garzón[1], G.M. Charrière[1]*

[1]IHPE, Univ Montpellier, CNRS, IFREMER, Univ Perpignan Via Domitia, Montpellier, France;

[2]ISEM, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France;

[3]IFREMER, UMR 241 Écosystèmes Insulaires Océaniens, Labex Corail, Centre Ifremer du Pacifique, Tahiti, Polynésie française;

[4]CREEC/CANECEV (CREES), MIVEGEC, Unité Mixte de Recherches, IRD 224–CNRS 5290–Université de Montpellier, Montpellier, France;

*, erikaastrid.burioli@gmail.com

Abstract

Transmissible cancer cell lines are rare biological entities giving rise to diseases at the crossroads of cancer and parasitic diseases. These malignant cells have acquired the amazing capacity to spread from host to host. They have been described only in dogs, Tasmanian devils and marine bivalves. The *Mytilus trossulus* Bivalve Transmissible Neoplasia 2 (MtrBTN2) lineage has even acquired the capacity to spread inter-specifically between marine mussels of the *Mytilus edulis* complex worldwide. To identify the oncogenic processes underpinning the

biology of these atypical cancers we performed transcriptomics of MtrBTN2 cells. Differential gene expression, enrichment, protein-protein interaction network, and targeted analyses were used. Overall, our results suggests that the long-term evolution of MtrBTN2 has led to the accumulation of multiple cancerous traits. We also highlight that vertebrate and lophotrochozoan cancers share a large panel of common drivers, which supports the hypothesis of an ancient origin of oncogenic processes in bilaterians.

Teaser

Mussel transmissible cancers teach us that cancerous process has a remote origin and that their long-term evolution has made them super-cancers.

Abbreviations

BTN: Bivalve Transmissible Neoplasia

CTVT: Canine Transmissible Venereal Tumor

DEG: Differentially Expressed Gene

DFTD: Tasmanian Devil Facial Tumor Disease

ECM: Extra Cellular Matrix

FDR: False Discovery Rate

GO: Gene Ontology

HR: Homologous Recombination

Log2FC: Log2 value of Fold Change

MtrBTN2: *Mytilus trossulus* Bivalve Transmissible Neoplasia lineage 2

PPI: Protein-Protein Interaction

RLE: Relative Log Expression

43 SNV: Single Nucleotide Variation

44 TCA: TriCarboxylic acid cycle

45

46 **1. Introduction**

47 Transmissible cancers are rare and fascinating biological entities as they evolved the ability to
48 overcome host physical and immunological boundaries to become contagious and to spread
49 between animals through direct transfer of cancer cells, thereby behaving like parasites (1,2).
50 They have been reported in only two vertebrate species so far, in dogs and in Tasmanian
51 devils. The transmissibility of cancer cells was first demonstrated in Canine Transmissible
52 Veneral Tumor (CTVT) (3,4), which has probably emerged more than 4,000 years ago (5). CTVT
53 has then dispersed across continents, through coitus and oral contacts between dogs.
54 Tasmanian Devil Facial Tumor Disease (DFTD) is the other well-known transmissible cancer
55 that affects vertebrates. Infection occurs via bites, a common behavior in Tasmanian devil
56 social interactions (6-8). As a consequence, the cancer transmission has devastated a major
57 part of devil populations with some populations that dropped by 80% in the most DFTD-
58 affected geographical areas (9-12). However, the highest number of transmissible cancer
59 lineages identified so far has been found in bivalve mollusks and these multiple lineages have
60 been grouped under the term Bivalve Transmissible Neoplasia (BTN) (13-16). BTNs transmit
61 between individuals of the same species but some lineages have also crossed the species
62 barrier and are circulating in related bivalve host species (14-16). *Mytilus trossulus* Bivalve
63 Transmissible Neoplasia lineage 2 (MtrBTN2) is one of these lineages that transmit inter-
64 specifically. MtrBTN2 emerged originally in a *Mytilus trossulus* founder host and has since
65 spread to *M. trossulus*, *M. edulis*, *M. galloprovincialis*, and *M. chilensis* populations across
66 several continents —South America, Asia, and Europe (14-15,17-20).
67 Transmissible cancer lineages must have emerged from a first neoplastic transformation in a
68 founder host and then evolved specific phenotypes to become a new type of contagious

69 etiologic agents. Indeed, we know that non-transmissible cancers also occur in *Mytilus*
70 mussels (19). MtrBTN2 cells are found circulating in the hemolymph together with hemocytes,
71 the mussel immune cells. Similar to hemocytes, they are able to infiltrate connective tissues
72 of various organs (18,21). During disease progression, MtrBTN2 cells overgrow host cells and
73 progressively replace almost all the circulating cells present in the hemolymph and
74 disseminate in tissues. They show a characteristic morphology of rounded and basophilic cells
75 with a high nucleus-to-cell ratio and they are polyploid (18). Thus, like other hemolymphatic
76 cancers (19), they are easily diagnosed by histological/cytological observation or flow
77 cytometry of hemolymph (18,22). This diagnostic needs to be complemented with genetic
78 analyses to distinguish MtrBTN2 from other *Mytilus* transmissible cancers and regular
79 hemolymphatic neoplasia (15,23). MtrBTN2 cells are able to proliferate very quickly with a
80 mean doubling time of ~ 3 days (23). In addition, BTNs have the capacity to survive in the
81 outside-host environment long enough to infect a new host (23,26). In the case of MtrBTN2,
82 we showed that these cells were able to survive at least 3 days with no mortality and up to 8
83 days in seawater (23). Although MtrBTN2 cells harbor phenotypic traits of most neoplastic
84 cells including a high proliferating activity, genomic abnormalities such as aneuploidy, the
85 supermetastatic ability to colonize diverse tissue niches of multiple hosts, and extended cell
86 survival capabilities, oncogenesis molecular pathways in BTNs remain uncharacterized.

87 Among the most frequent functional capabilities observed in human cancerous cells, which
88 have been the most studied, the core hallmarks are the sustaining of proliferative signaling,
89 the evasion from growth suppressors, the resistance to cell death, the replicative immortality,
90 the activation of invasion and metastasis, the reprogramming of cellular metabolism, and the
91 avoidance of immune destruction (24). The acquisition of these capabilities is ensured through

92 the activation or inactivation of specific signaling pathways. Among them, Sanchez-Vega et al.
93 (25) highlighted ten oncogenic signaling pathways as the most frequently altered in most
94 human cancers (HIPPO, MYC, NOTCH, NRF2, PI3K, RTK/RAS, TGFβ, TP53, WNT, and cell cycle),
95 and mainly involved in promotion of cell proliferation. These conclusions were drawn from
96 massive data over hundreds of human cancers acquired by transcriptomics, genomics and
97 epigenomics analyses.

98 Measuring relative levels of gene expression has been a key approach to identify genes and
99 biological pathways associated with the cancerous process and cancer functional adaptations
100 (27). In recent years, RNA sequencing (RNA-seq) has emerged as a fast, cost-effective, and
101 robust technology to address various biological questions (28,29). Transcriptome analyses
102 enable to link cellular phenotypes and their molecular underpinnings. In the context of
103 cancers, these links represent an opportunity to dissect the complexity of the cancer biological
104 adaptations. For non-model organisms and in the absence of a suitable reference genome,
105 which is the case for *M. trossulus*, RNA-seq is used to reconstruct and quantify *de novo*
106 transcriptomes (30). Differential gene expression analysis (DGE) is then carried out to compare
107 the effect of treatments or conditions on gene expression. To date, even if several
108 transmissible neoplasia have been described in marine bivalves, transcriptome-wide studies
109 are still lacking.

110 Here, we performed a deep sequencing of cancerous cell transcriptomes to investigate the
111 gene expression profile of MtrBTN2. We sequenced mRNA from circulating cells in both
112 healthy and MtrBTN2-affected *M. edulis*, and considering that the founder host species was
113 *M. trossulus*, we included hemocytes from this species as control. Analysis of differentially

114 expressed genes unveiled some functional characteristics of this uncommon cancer and
115 highlighted potential cancer drivers shared by all bilaterians.

116

117 **2. Results and discussion**

118 MtrBTN2 cells were sampled from the hemolymph of *M. edulis* mussels grown in the English
119 Channel (Normandy, France). Only individuals with more than 95% of circulating neoplastic
120 cells in their hemolymph were used for MtrBTN2 cell collection. *M. edulis* hemocytes were
121 collected from individuals from the same geographical site but tested negative for MtrBTN2
122 by cytology and qPCRs targeting a sequence in the EF1 gene specific of the *M. trossulus* species
123 and a MtrBTN2 specific sequence in the mitochondrial COI gene (15,23). *M. trossulus*
124 hemocytes were collected from the hemolymph of wild *M. trossulus* present in the Barents
125 Sea (Kola Bay, Russia) as this species is absent along the French coasts. All *M. trossulus* mussels
126 were positive for the qPCR targeting the *M. trossulus* specific sequence but were negative for
127 MtrBTN2 by cytology and the qPCR. We did not observe any other pathogen by histology in
128 the nine individuals used for RNA-Seq analysis. Transcripts from hemolymph samples were
129 sequenced for each condition: MtrBTN2-positive (i.e. cancerous) *M. edulis* (CAN 1-3),
130 MtrBTN2-negative *M. edulis* (EDU 1-3) and MtrBTN2-negative *M. trossulus* (TRO 1-3). In the
131 absence of reference genomes for our species of interest and for MtrBTN2, all RNA-seq reads
132 were used to assemble *de novo* a pantranscriptome from the 9 samples (MtrBTN2 cells, *M.*
133 *edulis* hemocytes, and *M. trossulus* hemocytes). The graphical representation of the
134 experimental set up and bioinformatic analyses is shown in Fig. 1.

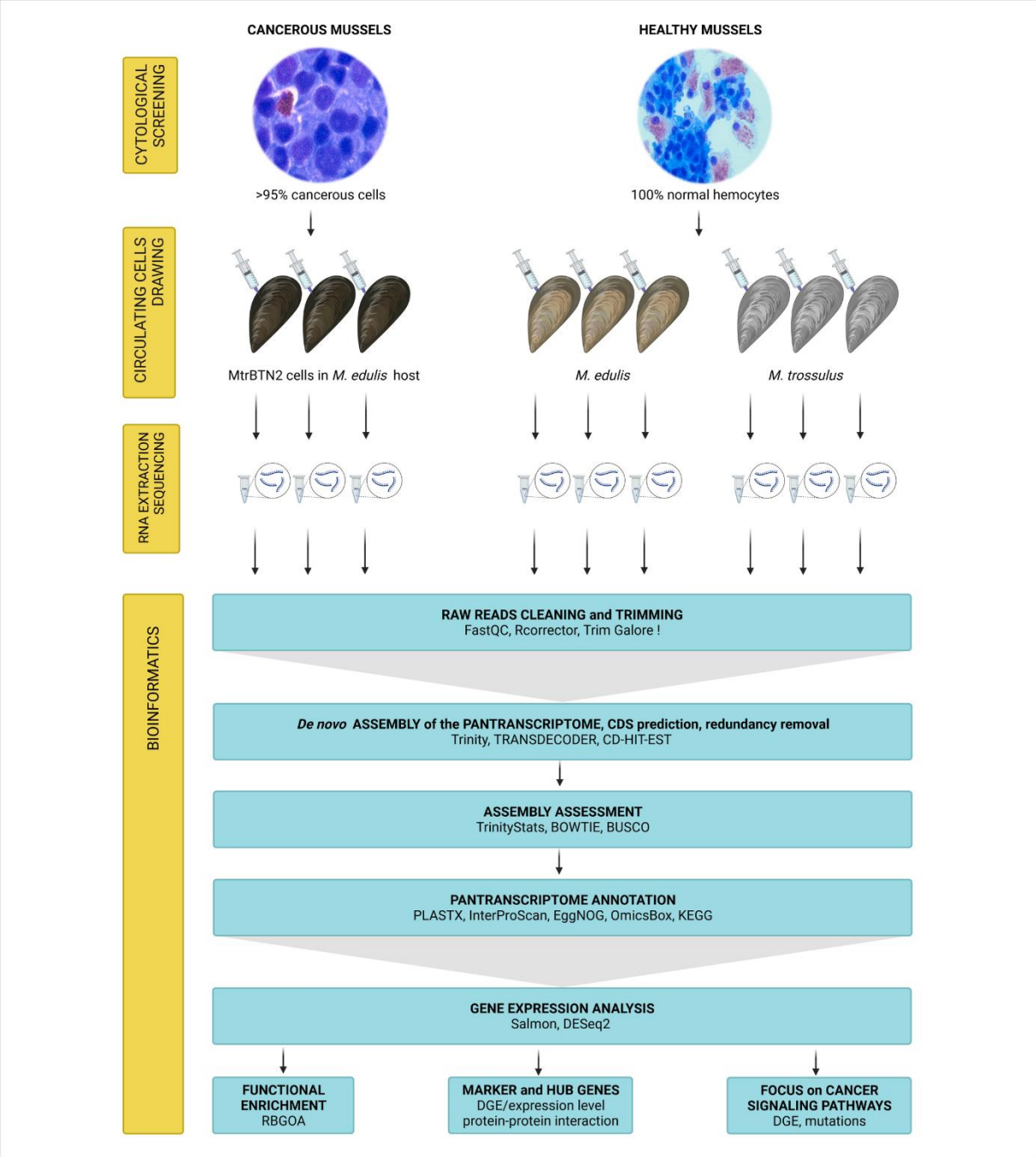


Figure 1: Graphical representation of the experimental design and RNA-Seq data analysis. DGE: differential gene expression analysis. Microphotographs of cancerous and healthy circulating cells are on the same scale.

2.1 Pantranscriptome assembly and annotation

139 The raw assembly contained 2,149,788 transcripts. After filtration steps, we retained a total
140 of 138,079 transcripts. The pantranscriptome completeness evaluation indicated that 95.6%
141 of the highly conserved single-copy metazoan genes (n = 954) were present in full length as a
142 single (81.2%) or duplicated copy (14.4%). Mean re-mapping rate reached 53.38 ± 0.58% with
143 negligible differences among samples (Kruskal-Wallis, H(2)=1.69, p=0.43). Only 5826
144 transcripts did not show a blast similarity or a conserved domain. OmicsBox program v2.0
145 (31,32) associated GO terms to 131,346 transcripts but finally, 41.43% of transcript
146 annotations were retained after computing a filtration step based on a minimum annotation
147 score. Assembly and annotation metrics are reported in Table 1 and supplementary Table S1.
148 Raw read sequences coming from RNA-seq were deposited with links to BioProject accession
149 number PRJNA749800 in the NCBI BioProject database
150 (https://www.ncbi.nlm.nih.gov/bioproject/). Individual SRA numbers are displayed in
151 supplementary Table S2.

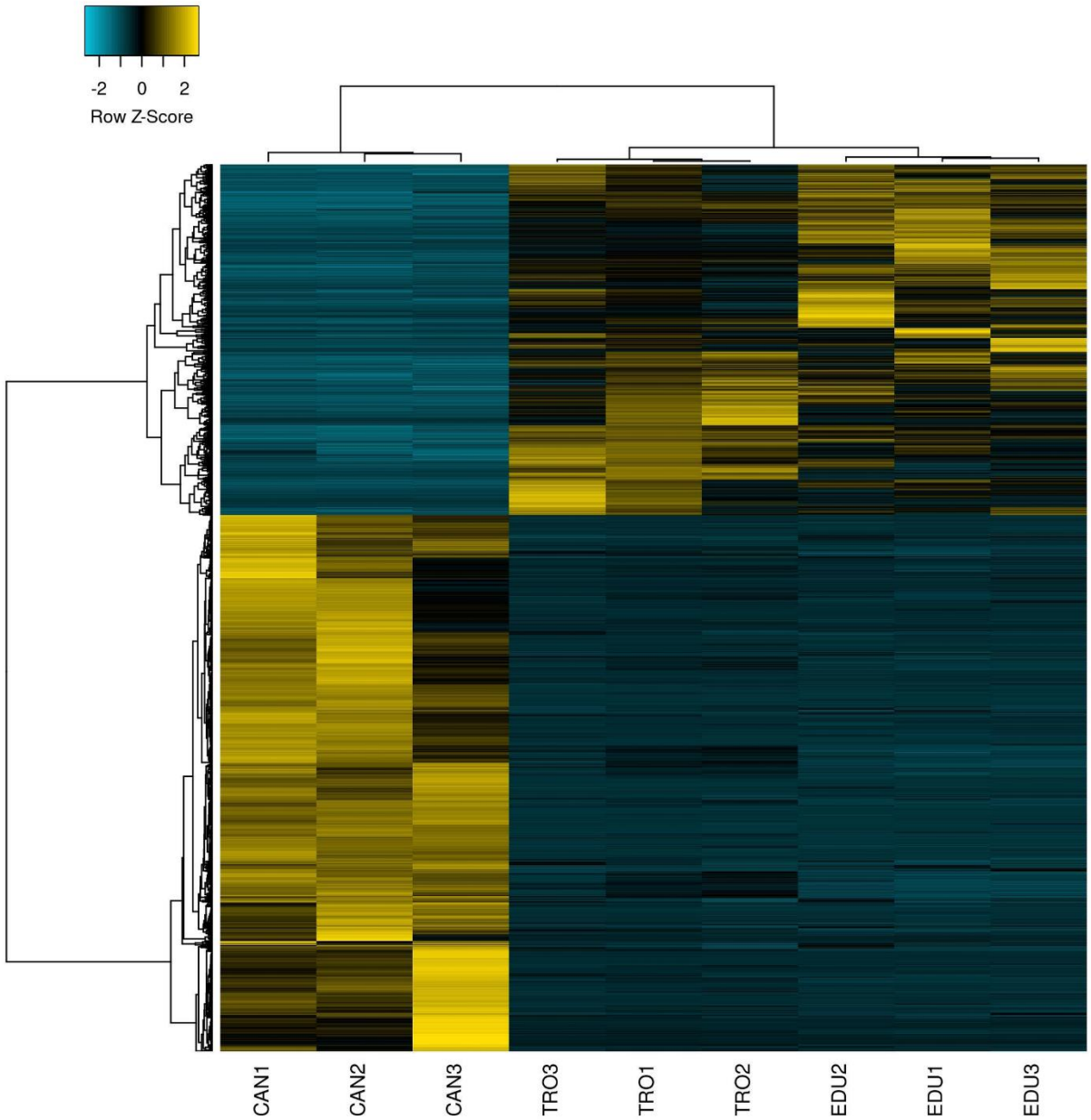
Attributes	Cancerous individuals			Healthy individuals					
	MtrBTN2			*Mytilus edulis*			*Mytilus trossulus*		
	CAN1	CAN2	CAN3	EDU1	EDU2	EDU3	TRO1	TRO2	TRO3
Clean reads (2x) x10^6	83.69	82.83	82.05	83.88	81.93	98.33	81.57	76.67	84.82
GC%	36	36	36	37	37	37	36	36	36
Transcriptome statistics									
	Raw number of transcripts						2,149,788		
	Final number of transcripts						138,079		
	GC%						38		
	Total assembled bases						97,665,057		
	Contigs N50 (bp)						1,038		
	Median contig length (bp)						439		
	Average contig length (bp)						745		
	BUSCO % (*Metazoa, n=954*)						95.6		
	Complete single-copy						167		
	Duplicated						745		
	Fragmented						34		
	Missing						8		
Back alignment on pantranscriptome %	53.53	53.92	53.48	53.88	53.83	52.83	53.70	52.18	53.14
Annotation statitics									
	Transcripts with GO associated						131,346		
	Transcrips annotated*						57,214		
	No-blast, no Interproscan						5,826		

Table 1: Assembly and annotation metrics. The pantranscriptome was obtained from the mRNA sequencing of circulating cells from three *Mytilus edulis*, three *M. trossulus*, and three MtrBTN2-affected *M. edulis* at a late stage of the disease (>95% of cancerous cells).*after applying the OmicsBox annotation rule (e-value hit filter of 1.10^{-3}, annotation cutoff of 55, a HSP-Hit coverage of 60%).

2.2 The specific transcriptomic profile of MtrBTN2 cells

We investigated differentially expressed genes in MtrBTN2 cells (CAN 1-3) by comparing their transcript abundance to the *M. edulis* (EDU 1-3) and *M. trossulus* hemocytes (TRO 1-3) using DESeq2. A total of 4358 transcripts were significantly differentially expressed; 2632 were more expressed and 1726 were less expressed in MtrBTN2 cells than in hemocytes (see Log2 fold change (Log2FC) in supplementary Table S3). The three groups of mussels were well discriminated through hierarchical clustering on a heatmap (log2 centered) (Fig. 2).

164 Remarkably, the transcriptome of MtrBTN2 cells (CAN 1-3) was very specific and distinct from
165 that of hemocyte samples of both healthy *M. edulis* (EDU 1-3) and *M. trossulus* (TRO 1-3),
166 which clustered together. These first results indicate that MtrBTN2 transcriptomic profile
167 significantly differs from the transcriptomic profile of healthy *Mytilus* hemocytes,
168 independently of their species or geographical origin. This biological material was therefore
169 used to identify the specificities of MtrBNT2 cell transcriptome.



170

171 **Figure 2:** Heatmap (log2 centered) with hierarchical clustering of the 4358 differentially expressed genes at a

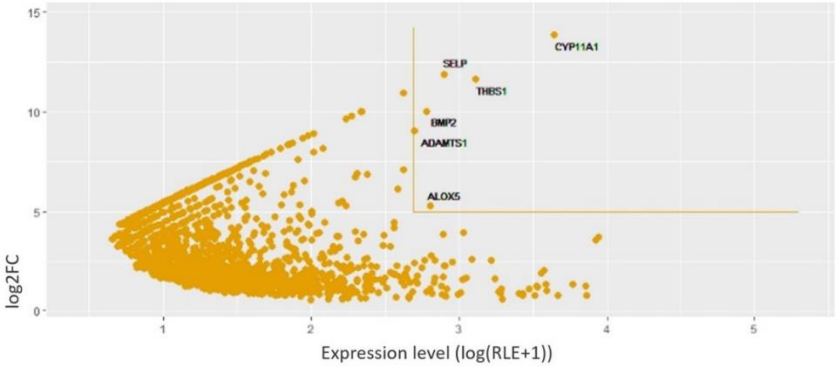
172 FDR <0.05.

173 Among the most differentially expressed genes (DEG) (with log2FC > 5), we identified six top-

174 expressed genes (RLE > 500) in MtrBTN2 cells (*CYP11A1*, *THBS1*, *SELP*, *ALOX5*, *BMP2*, and

175 *ADAMTS1*) (Fig. 3). These genes may be considered as specific of the malignant state and have

176 been described as implicated in tumorigenesis and metastasis in human (Table 2).



177

178 **Figure 3:** Scatter plot representing y=log2FC, x=log(RLE+1). Lines limit the areas above log2FC values > 5 and RLE

179 > 500. Six genes (*CYP11A1*, *THBS1*, *SELP*, *ALOX5*, *BMP2*, and *ADAMTS1)* exceeded these values and may be

180 considered as specific of the malignant state.

gene	Product	Function	Biological processes	References
CYP11A1	Cholesterol side-chain cleavage enzyme	- catalysis of cholesterol to pregnenolone convertion	- modulation of immune response (glucocorticoids) - sexual development and gametogenesis (androgens and estrogens)	Hu et al., 2010 [33]
THBS1	Thrombospondin-1	- cell-to-cell and cell-to-matrix interactions - major activator of TGF-β to its mature form	- chemotactic gradient for immune cells participating in inflammatory response during acute phase - induction macrophage polarization to the M2 anti-inflammatory phenotype	Murphy-Ullrich and Poczatek, 2000 [34] Letterio and Roberts, 1998 [35] Ashcroft, 1999 [36]
SELP	P-selectin	- vascular adhesion molecules	- extravasation of circulating metastatic cells - inflammation - induction macrophage polarization to the M2 anti-inflammatory phenotype	Läubli and Borsig, 2010 [37] Fabricius et al., 2021 [38]
ALOX5	5-Lipoxygenase	- synthesis of leukotrienes	- inflammation - proliferation of cancerous cells - apoptosis inhibition of cancerous cells	Anderson et al., 1998 [39] Bishayee et al., 2013 [40]
BMP2	Bone morphogenetic protein 2	- growth factor - cell differentiation	- stemness maintenance - induction macrophage polarization to the M2 anti-inflammatory phenotype	Lee et al., 2013 [41]
ADAMTS1	A disintegrin and metalloproteinase with thrombospondin motifs 1	- matrix proteolytic enzyme	- ECM dismantling - induction macrophage polarization to the M2 anti-inflammatory phenotype	Redondo-García et al., 2021 [42] Rucci et al., 2011 [43]

181

182 **Table 2: Top-expressed genes in MtrBTN2 cells (log2FC > 5 and RLE > 500).** Most are involved in inflammation

183 and immune response.

184

185 They are linked to inflammation, immune processes, and extra-cellular matrix (ECM)
186 disruption (33-43). The exact role of some of these genes (*THBS1, SELP, ALOX5,* and *BMP2*)
187 during tumorigenesis is still debated mainly for their implication in inflammatory process
188 (44,45). However, *THBS1* expression is increased in numerous cancers promoting invasion and
189 metastasis (46-49). *SELP* and *ALOX5* are highly expressed in some tumors and cancer cell lines,
190 inducing proliferation of cancerous cells (50-53) and inhibiting apoptosis (39,40). Conflicting
191 data also exist with regard to the effect of *BMP2* on cancer (54). Nevertheless, most studies
192 suggest that *BMP2* enhances metastatic progression and tumorigenesis. Although mollusk
193 immunity against cancer has never been studied so far, most of these genes could be involved
194 in the modulation of host immune response. Some of them like *THBS1*, *SELP*, *BMP2*, and
195 *ADAMTS1* are involved in the induction of macrophage polarization to the M2 anti-
196 inflammatory phenotype in mammals and their roles in the interactions of MtrBTN2 with host
197 immune response deserve further attention to better understand the host invasion process.

198 **2.3 Dysregulated biological processes in MtrBTN2**

199 To identify the main biological processes that are specific of the MtrBTN2 phenotype, we
200 performed a GO_MWU enrichment analysis on DEGs (for protein descriptions and gene IDs
201 see supplementary Table S3). The complete list of genes clustered by Gene Ontology (GO)
202 term is provided in the supplementary Table S4. The GO_MWU analysis revealed several
203 biological functions differentially regulated in MtrBTN2 cells compared to normal hemocytes
204 (Fig. 4).

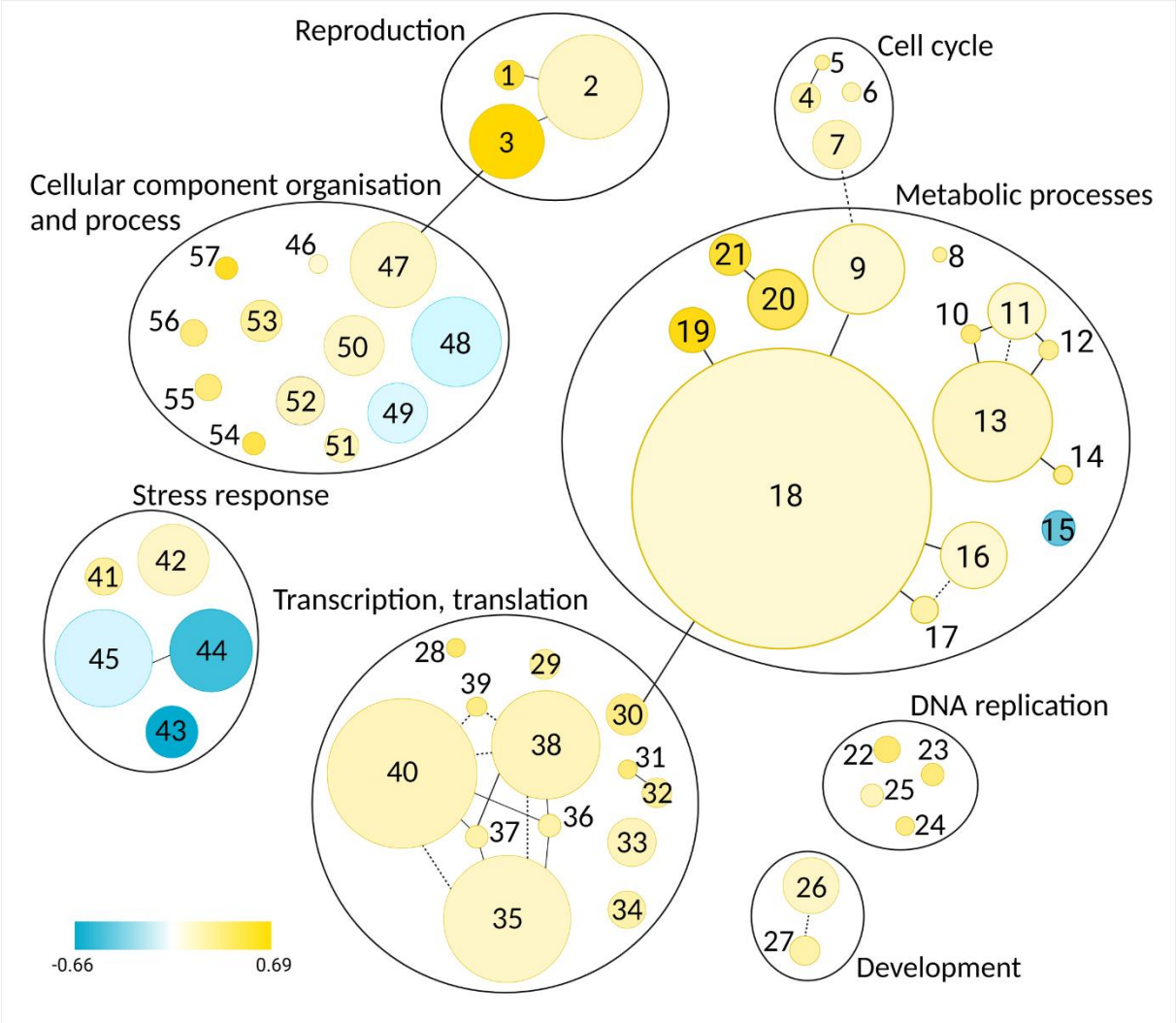


Figure 4: Biological processes dysregulated in transmissible neoplasia: **1.**Meiosis I; **2.**Cellular process involved in reproduction; **3.**Sperm motility; **4.**Cytokenetic process; **5.**Contractil ring contraction; **6.**Mitotic spindle elongation; **7.**Regulation of cell cycle process; **8.**Protein deacylation; **9.**Proteolysis; **10.**Adenosine metabolic process; **11.**Nucleoside phosphate biosynthetic process; **12.**IMP metabolic process; **13.**Carbohydrate derivative metabolic process; **14.**Dolichol-linked oligosaccharide biosynthetic process; **15.**Tricarboxylic acid cycle; **16.**Peptidyl-amino acid modification; **17.**Histone modification; **18.**Protein metabolic process; **19.**Protein deubiquitination; **20.**Nucleobase metabolic process; **21.**Pyrimidine nucleobase metabolic process; **22.**DNA replication initiation; **23.**DNA strand elongation; **24.**DNA replication, synthesis of RNA primer; **25.**DNA-dependent DNA replication; **26.**Regulation of growth; **27.**Cell fate determination; **28.**Cis assembly of pre-catalytic spliceosome; **29**.RNA phosphodiester bond hydrolysis; **30.**Mitochondrial translation; **31.**Transcription initiation

216 from RNA polymerase I promoter; **32.**DNA-templated transcription, initiation; **33.**Translational elongation;
217 **34.**Spliceosomal snRNP assembly; **35.**mRNA processing; **36.**tRNA splicing, via endonucleolytic cleavage and
218 ligation; **37.**tRNA wobble base modification; **38.**tRNA metabolic process; **39.**Deadenylation-dependent
219 decapping of nuclear-transcribed mRNA; **40.**ncRNA metabolic process; **41.**Response to oxidative stress;
220 **42.**Double-strand break repair; **43.**Cellular response to amino acid starvation; **44.**Response to bacterium;
221 **45.**Defense response; **46.**Actin filament based process; **47.**Microtubule-based movement; **48.**Cell surface
222 receptor signaling pathway; **49.**Anion transport; **50.**Golgi vesicle transport; **51.**Ribosomal small subunit
223 assembly; **52.**Ribosomal large subunit assembly; **53.**rRNA-contain RNP complex. The disk diameter is
224 proportional to the number of DEGs linked to the process. The color intensity is proportional to the enrichment
225 intensity, expressed as the ratio between the number of genes that were significantly over- (yellow) or down-
226 expressed (blue) compared with the total number of genes in the process. Lines connecting two functions
227 represent shared genes (solid line means that all genes in the smaller family are part of the largest family; dotted
228 line means that only some genes are shared).

229 ***Proliferation-related functions are enriched***

230 Most of the functions over-represented in MtrBTN2 cells are consistent with a high
231 proliferative activity (4.Cytokenetic process; 5.Contractil ring contraction; 6.Mitotic spindle
232 elongation; 7.Regulation of cell cycle process, in Fig. 4). Many genes linked to "DNA
233 replication" and "transcription and translation" underpin this enrichment (supplementary
234 Table S4). Among them, we found *TERT,* which is essential for the replication of chromosome
235 telomeres in most eukaryotes and is particularly active in progenitor and cancer cells (55). By
236 maintaining telomere length, telomerase relieves a main barrier on cellular lifespan, enabling
237 limitless proliferation driven by oncogenes (56). In addition, we found in the "development"
238 family a higher expression of transcripts involved in the promotion of cell growth such as *TTK*
239 (57), *METAP2* (58), *THBS1* (activator of TGF-β) (34), *MEAK7* (59), and *LAMTOR1* (60). Several

240 gene transcripts acting as growth inhibitors, especially in cancer contexts, were less expressed
241 in MtrBTN2, such as *FAM107A* (61), and *LTBP3* (62), which keeps TGF-β in a latent state.

242 The "cell cycle" family (4.Cytokenetic process; 5.Contractil ring contraction; 6.Mitotic spindle
243 elongation; 7.Regulation of cell cycle process, in Fig. 4) was also enriched in MtrBTN2 cells.
244 The negative regulation of the cell cycle progression appeared less expressed: *ASAH2,* which
245 inhibits the mediation of the cell cycle arrest by ceramide (63), was more expressed, whereas
246 *PPP2R5B,* which inhibits the cycle progression (64), was less expressed in neoplasia. Moreover,
247 genes promoting the cell cycle progression were more expressed in MtrBTN2 cells, particularly
248 five genes forming the Anaphase Promoting Complex (*ANAPC1, ANAPC4, ANAPC5, CDC26, and*
249 *CDC23*) and *CCND1* encoding cyclin D1 that promotes progression through the G1-S phase of
250 the cell cycle (65).

251 ***Regulation of metabolic pathways are modified***

252 Many metabolic processes were shown to be enriched (Fig. 4 and Suppl. Table S4). The
253 nucleotide metabolism was characterized by a higher expression in MtrBTN2 cells of more
254 than 25 genes involved in the *de novo* and salvage biosynthetic pathways of purines and
255 pyrimidines and in the homeostasis of cellular nucleotides. Protein and carbohydrate
256 derivative metabolisms were also over-represented in neoplasia. Enrichment concerned
257 genes encoding ribosomal proteins, protein turnover and localization (especially
258 deubiquitination), proteolysis, peptidyl amino-acid modification, histone acetylation, and lipid
259 and protein glycosylation. In addition to *MEAK7* and *LAMTOR1*, eight genes (*BMTP2*, *C12orf66*,
260 *ITFG2*, *MIOS*, *NPRL2*, *RagA, RagD, and SEC13*) involved in the mTORC1 amino-acid sensing
261 pathway (43.Cellular response to amino acid starvation, in Fig. 4) (66) were differentially

262 expressed. Interestingly, beside the role of mTORC1 in the regulation of autophagy, mTORC1
263 plays a pivotal role as a master regulator of protein, lipid, and nucleic acid metabolism
264 modifications reported in many different cancer cells (67-70). Several transcripts involved in
265 organelle activities and linked to protein metabolism (anabolism, catabolism, post-
266 translational modifications, folding, and secretion) were also more expressed in MtrBTN2.

267 The tricarboxylic acid cycle (TCA) appeared also highly differentially regulated in MtrBTN2 with
268 a lower expression of *DLST* transcripts (a component of 2-oxoglutarate dehydrogenase
269 complex), *FH*, *SDHD* and *NNT*, and a higher expression of *SUCLA2*. TCA cycle is crucial for
270 generating cellular energy and precursors for biosynthetic pathways (71). By screening the
271 expression of 16 glycolysis-related genes, we found *PFKP*, *PKM2*, and *SLC2A* more expressed;
272 however, *PGK1*, *TPI1*, and *ALDOA* were less expressed than in control hemocytes. These
273 conflicting results did not allow us to conclude on how the regulation of glycolytic activities in
274 transmissible neoplasia cells is modified. However, these results suggest major modifications
275 in nucleotides, amino-acid and energy metabolism in MtrBTN cells as often seen in other
276 cancers (72).

277 ***DNA repair systems are dysregulated***

278 MtrBTN2 neoplasia were found to express at a higher level a significant panel of genes
279 involved in the DNA Homologous Recombination (HR) (42.Double-strand break repair, in Fig.
280 4), such as *BABAM1*, *ZSWIM7*, *RAD51L*, *RAD54L*, *MRE11*, *MCM9*, and *TONSL*. The protein DNA
281 helicase MCM9 is also related to the repair of inter-strand crosslinks (73), as well as MUS81
282 (74). This kind of damage is a double-edged sword as DSB can induce cell death unless repaired
283 efficiently, but inefficient or inappropriate repair can lead to mutation, gene translocation and

284 cancer (75). Due to their high proliferation rate and enhanced metabolic activities, cancer cells
285 suffer a huge replication stress (76), which emphasizes the need for cancer cells to activate
286 DNA repair. Cells are able to repair DSB through two major pathways: nonhomologous end
287 joining and HR (77). HR is critical for reestablishing replication forks at the sites of damage
288 during S and G2 phases of the cell cycle. However when combined with depletion of cell cycle
289 checkpoints and apoptosis this mechanism can result in genomic instability. As MtrBTN2
290 ploidy appears highly abnormal (varying between 8N to 18N, (18)), further investigations will
291 be needed to determine whether the dysregulation of these DNA repair pathways favors
292 recombination between homologous chromosomes and chromosomal abnormalities
293 accumulation, or whether these pathways are activated in order to limit DNA damages.
294 Interestingly, up-regulation of double-stranded DNA break repair via HR pathway was also
295 found in DFTD (78).

296 ***Interaction with ECM is affected***

297 Seven genes encoding integrin sub-units (48.Cell surface receptor signaling pathway, in Fig. 4),
298 a major class of transmembrane glycoproteins that mediate cell-matrix and cell-cell adhesion
299 were dysregulated (Suppl. Table S4). As integrins are mostly less expressed in MtrBTN2 cells
300 this suggests that the capacities of these cells to interact with ECM components are
301 profoundly modified which is often involved at multiple stages of the cancerous processes
302 (79).

303 ***Cell fate determination pathways are modified and soma-to-germline transition may have***
304 ***occurred***

305 Major pathways of cell fate determination appear modified (Fig. 4 and Suppl. Table S4). The
306 expression of six genes involved in the Notch and Wnt pathways were altered. These
307 modifications may drive dedifferentiation processes often seen in aggressive cancer cells or
308 cancer stem cells (80). We report in paragraph 2.6 the results of the focused analysis carried
309 out on these two pathways. Additionally three biological processes clustered under GO terms
310 linked to the "reproduction" were enriched (1.Meiosis I; 2.Cellular process involved in
311 reproduction; 3.Sperm motility, in Fig. 4). Nineteen transcripts more expressed in MtrBTN2
312 cells gathered under the "sperm motility" process and encode dynein assembly factors, dynein
313 chains, and cilia and flagella associated proteins. Moreover, transcripts of genes involved in
314 meiosis such as *MSH5* (81), *SYCP3* (82), and *TEX11* (83), were more expressed in MtrBTN2.
315 These unexpected results could reflect a soma-to-germline transition during the oncogenesis
316 process that was reported in a diversity of cancers and reinforce the hypothesis of profound
317 modifications of the differentiation state of MtrBTN2 cells (84,85). Interestingly when
318 abnormally produced in mitotically dividing cells, the DEG *SYCP3*, may impair recombination
319 and drives ploidy changes influencing chromosomal segregation in cancer cells, which could
320 be one of the mechanisms that led to the hyperploidy of MtrBTN2 cells (85).

321 ***Immune response is underrepresented***

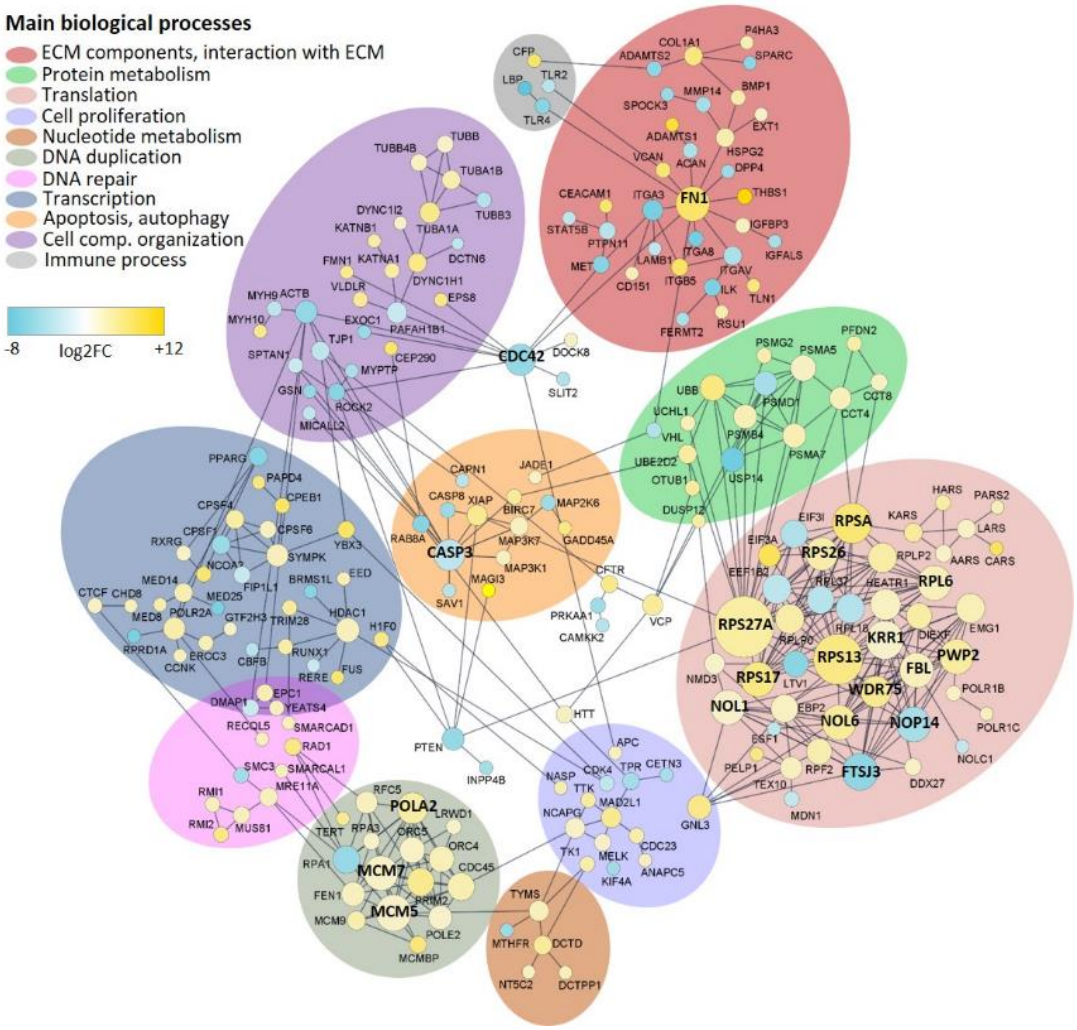
322 One of the striking differences between MtrBTN2 cells and hemocytes is that biological
323 processes belonging to innate immunity (44.Response to bacterium; 45.Defense response; in
324 Fig. 4) are significantly underrepresented in MtrBTN2 cells. The genes with significant lower
325 expression in MtrBTN2 cells encode major mussel antimicrobial peptides and bactericidal
326 proteins such as *Mytilin B*, *Myticin A*, *Myticin B*, *MGD1*, and *BPI* (86-88). They also encode
327 multiple pattern recognition receptors: *TLR2* and *TLR4,* which belong to the Toll-Like-Receptor

328 family essential to innate immunity against pathogens (89), and the lectin-related pattern
329 recognition molecules *Ficolin-2*, *FIBCD1*, *CLEC4E*, and *CLEC4M* involved in the recognition of
330 pathogens and apoptotic or necrotic cells (90-92). This under-representation of biological
331 functions related to host defense suggests that MtrBTN2 cells are immunologically
332 incompetent. During MtrBTN2 disease progression, we also observe a drastic decrease in the
333 number of host hemocytes, progressively replaced by circulating MtrBTN2 cells, which can
334 represent over 95% of the circulating cells in cancerous mussels (18,21). To date, it remains to
335 be determined whether MtrBTN2 cells outcompete regular hemocytes in the hemolymph or
336 whether they interfere with hematopoiesis, as seen for some human leukemias (93). Still, the
337 combination of fewer circulating hemocytes and the immune incompetence of MtrBTN2 cells
338 could have dramatic consequences for the host health and could lead to lethal opportunistic
339 systemic infections (94).

340 **2.4 CASP3, FN1, and CDC42 are hub genes in the interaction network of MtrBTN2 DEGs**

341 Protein-protein interaction (PPI) networks of DEGs were constructed to evaluate their
342 connectivity and identify hub genes that could play a critical role in the MtrBTN2 phenotype.
343 The complete network of PPI is represented in Fig. 5. The top 20 genes in connectivity ranking
344 were found in functions already identified as enriched with GO_MWU such as translation
345 (most involved in ribosome biogenesis) and DNA replication. Outputs of NetworkAnalyzer
346 v4.4.8 (95) are reported in Supplementary table S5. Interestingly, this analysis highlighted
347 three DEGs as major hubs in the PPI networks that did not stand out in the GO_MWU analysis,
348 which are *FN1, CDC42,* and *CASP3*. Fibronectin 1 encoded by *FN1*, is a major component of
349 ECM and plays an important role in cell adhesion, migration, growth and differentiation (96);
350 the number of transcripts was substantially higher in MtrBTN2 cells (log2FC > 5). CDC42*,* is a

351 member of the Rho GTPase family and plays an important role in cell- cell and cell-matrix
352 adhesion, actin cytoskeletal architecture, cell migration, and cell cycle regulation (97), the
353 number of transcripts was significantly lower in MtrBTN2 cells. As FN1 and CDC42 are involved
354 in complementary functions and signaling pathways (like cell-ECM interactions and cell
355 migration) both hub genes were found interconnected in the PPI network, moreover a large
356 number of integrin sub-units were found dysregulated. Altogether these results strongly
357 suggest that cell-ECM interactions and cell migration functions are profoundly modified in
358 MtrBTN2 cells in line with their observed rounded morphology, low adhesion properties and
359 f-actin modifications (18,20). The master effector of apoptosis, Caspase 3 encoded by *CASP3*
360 is less expressed as well in MtrBTN2 cells. It plays a central role in the execution-phase of cell
361 apoptosis (98) that is commonly inhibited in cancerous cells allowing them to evade cell death
362 despite DNA damage accumulation and dysfunctions of the cell cycle.



363

364 **Figure 5:** Network of protein-protein interactions (PPI). The disk size is proportional to the number of gene
365 connections. Top 20 Hub genes are in bold. In yellow, up-regulated genes; in blue, down-regulated genes. The
366 color intensity is proportional to the Log2FC value.

367 **2.6 Most major oncogenic signaling pathways are altered in MtrBTN2**

368 Based on the Cancer Genome Atlas project, Sanchez-Vega et al. (25) highlighted that 10
369 signaling pathways are very frequently altered in most cancers. We found that seven of these
370 10 oncogenic signaling pathways were altered in MtrBTN2 neoplasia; six were highlighted by
371 differential gene expression analysis and one by SNV analysis. These pathways are involved in
372 cell proliferation and differentiation.

373 Key DEGs were found in the “Hippo”, “Notch”, “Wnt”, “Myc”, “PI3K”, and “Cell Cycle” signaling

374 pathways (Table 3).

Pathway	Main functions	Gene	Action on pathway	MtrBTN2/hemocytes expression
Hippo	anti-proliferative anti-stem cell self renewal pro-apoptotic	SAV1	activation	Lower
Notch	cell proliferation cell fate	JAG1 FBXW7	activation repression	Higher Higher
Wnt	cell proliferation stemness maintenance	SFRP1 SFRP3 APC CTNNB1	repression repression repression /	Lower Higher Higher Higher
Myc	cell proliferation apoptosis	MXD3	repression	Lower
PI3K	cell proliferation cell survival	PTEN PIK3R1	repression repression	Lower Lower
Cell Cycle		CCND1	activation	Higher

375

376 **Table 3: Genes differentially expressed in MtrBTN2 among the most frequently altered genes in the oncogenic**

377 **signaling pathways**. We report their action on the corresponding pathway and their state of expression.

378 Hippo signaling pathway plays a role in the inhibition of cell proliferation, control of cell fate,

379 and apoptosis promotion through the phosphorylation of YAP/TAZ transcription coactivators,

380 their cytoplasmic retention and degradation. *SAV1* is a key activation factor of the Hippo

381 pathway participating to LATS1/2 phosphorylation, a necessary step for subsequent YAP/TAZ

382 phosphorylation (99). Since *SAV1* is less expressed in MtrBTN2 cells the Hippo pathway may

383 be down- or inactivated.

384 The PI3K pathway also showed dysregulation. PIK3 pathway is involved in cell metabolism,

385 growth, proliferation, Cell-ECM interactions and survival. We observed a lower expression of

386 two genes, *PTEN* and *PIK3R1* acting as PIK3 antagonists (100). When the pathway is activated,

387 AKT1 is phosphorylated and inhibits the activity of TSC1/2. AKT-mediated phosphorylation of

388 TSC1/2 lifts its inhibition on RHEB activity, leading to activation of the complex mTORC1. Our

389 enrichment analysis (Chap. 2.3) has also highlighted the mTORC1 activation by the alternative
390 amino-acid sensing pathway. AKT1 plays also an indirect role in Wnt and Myc pathways by a
391 negative regulation of GSK3β through its phosphorylation (101). Thus, the inhibitory effect of
392 unphosphorylated GSK3β on MYC and its contribution in the destruction complex of β-
393 cathenin in Wnt pathway may be repressed in MtrBTN2.
394 MAD3, an antagonist of MYC for MAX binding (102), was less expressed in MtrBTN2 cells
395 suggesting that the MYC-MAX complex was favored and the pathway activated.
396 Among SFRPs, which act as Wnt pathway inhibitors by binding extracellular WNT ligands,
397 *SFRP1* was less and *SRFP3* was more expressed. Moreover, *APC,* which encodes a protein
398 constitutive of the Destruction Complex, was more expressed. APC is a multifunctional protein
399 and is involved, for instance, in the normal compaction of mitotic chromatin (103). We looked
400 for DEGs among the downstream genes of the Wnt pathway, which transcription is regulated
401 by unphosphorylated CTNNB1 mediation. We found that *CCND1*, *JAG1*, and *FN1* were more
402 expressed suggesting that the Wnt pathway was activated (104). CCND1 promotes cell
403 proliferation and JAG1 is a transmembrane ligand activating the Notch pathway. However,
404 FBXW7 that inhibits the NOTCHs cleavage and its consequent activation was more expressed
405 in MtrBTN2 cells. Moreover, *HES1* expression, which is promoted by activated NOTCH, was
406 lower in MtrBTN2 cells. HES1 negatively regulates expression of downstream target genes
407 such as tissue-specific transcription factors (105).
408 In order to identify potential relevant mutations in these signaling pathways, we investigated
409 for the presence of MtrBTN2-specific SNVs in the transcript sequences of the genes belonging
410 to the 10 oncogenic pathways by visualization of base variations and frequency with IGV (106).
411 Although the identification of these variations was complicated by the polyploidy of cancerous

412 cells and the short reads sequencing data, we managed to identify MtrBTN2-specific SNVs,
413 variants that were only observed in the three MtrBTN2 samples but not in the other 6 samples
414 (3 *M. edulis* and 3 *M. trossulus*). MtrBTN2-specific SNVs were found in 10 genes among those
415 involved in oncogenic pathways (Table 4).

416

			Substitution type			Missense substitution
Gene	**Transcript length**	**Pathway**	**Synonymous**	**Missense**	**Total number**	**in a domain**
PIK3CA	3144 bp	PI3K	26	1	27	-
PIK3CB	3156 bp	PI3K	3	0	3	-
PTEN	1398 bp	PI3K	2	0	2	-
AKT1	1458 bp	PI3K	3	1	4	1
KRAS	330 bp	RTK/RAS	3	0	3	-
MDM2	1707 bp	p53	26	22	48	9
p53	1795 bp	p53	2	1	3	-
MYC	1269 bp	Myc	0	1	1	1
FBXW7	2100 bp	Notch	1	-	1	-

417

418 **Table 4: Genes carrying MtrBTN2-specific SNVs among the most frequently altered genes in the oncogenic**
419 **signaling pathways**. *AKT1*, *MDM2* and *MYC* carried missense substitutions in a protein domain.

420 The highest number of MtrBTN2-specific SNVs were found in *PIK3CA* (PI3K pathway) and
421 *MDM2* (p53 pathway) genes, with 27 specific SNVs/3144 nucleotides and 48/1707
422 respectively. Most of these specific SNVs were synonymous substitutions but missense
423 substitutions were present in five genes: *PIK3CA*, *AKT1*, *MDM2*, *p53*, and *MYC*. In three of
424 these genes, the missense substitutions were located within protein domains. A missense
425 substitution in AKT1 sequence was located in the kinase domain, and a missense substitution
426 in *MYC* sequence, was located in the transcriptional activation domain. The real impact of
427 these missense substitutions found in *PIK3CA* and *AKT1* is still unknown. However, *MDM2* was
428 particularly affected with 9 missense mutations in functional protein domains: one

429 substitution in the p53-binding domain, four in the acidic domain, two in the Zn-finger domain,
430 and two in the RING-finger domain. Interestingly, in the *p53* gene which encodes both p53
431 and p63/73 proteins in mollusks (107) we found three MtrBTN2-specific SNVs. The first
432 (nucleotide 177, C in MtrBTN2 instead of T) and the second (nucleotide 816, T instead of C)
433 were synonymous substitutions previously described by Vassilenko et al. (2010) (108) in *M.*
434 *trossulus* affected by haemic neoplasia. The third was a missense substitution located in
435 nucleotide 1258 (T instead of A), newly described here. *M. trossulus p53* is characterized by a
436 deletion of 6 nucleotides if compared to *M. edulis p53* (109). We found this deletion in 100%
437 of reads coming from MtrBTN2 samples. Altogether, SNVs found in the MDM2-p53 sequences
438 of MtrBTN2 suggest that this pathway could contribute to MtrBTN2 oncogenesis.
439 To summarize these results, we represented a simplified view of the altered pathways and
440 their interactions in Fig. 6.

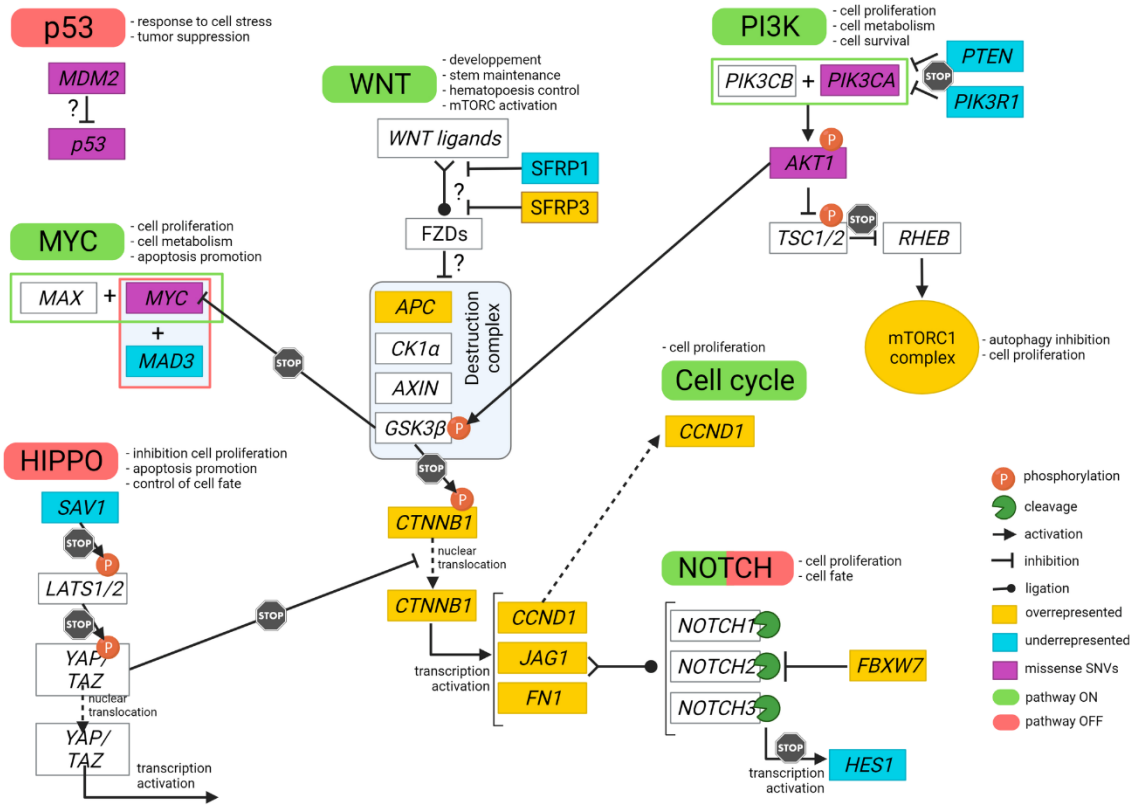


Figure 6: Seven oncogenic signaling pathways showed expression or sequence alterations in MtrBTN2 cells. Green pathways are expected to be activated; red, inactivated; grey, indeterminate. Purple genes brought missense substitutions; yellow were more expressed in MtrBTN2 cells than in hemocytes; blue were less expressed.

2.7 MtrBTN2 shares multiple oncogenic pathways with mammals

In their survey of bilaterian animals, Aktipis et al. (110) have revealed that all or nearly all bilaterians are susceptible to cancer. Mollusks (Lophotrochozoa) are no exception (19,21). Even if high levels of genotypic and phenotypic diversity have been described among cancers in mammals (111), molecular studies mainly in human and mice have identified common drivers of the cancerous process (24-25). Still, oncogenic processes have been poorly studied in invertebrate phyla, especially at the molecular level. Our present data show that most

453 common alterations related to oncogenesis found in mammals are present in MtrBTN2 cells,
454 with seven oncogenic signaling pathways altered. One of the most fundamental traits of
455 cancerous cells consists in their high proliferation rates. We earlier found a mean doubling
456 time of ~ 3 days in the case of MtrBTN2 (23). The present study provides information on the
457 molecular bases of cell proliferation and suggests apoptosis inhibition. MtrBTN2 proliferative
458 activity is supported by metabolic rewiring, as often seen in cancer cells (72). Indeed, we
459 highlight a central role for the Pi3K-AKT-mTORC1 pathway that is a master regulator in
460 coupling cell growth and amino-acid, lipid, and nucleotide metabolisms that are frequently
461 modified in many mammalian cancers of diverse origins (74,112-114). Dysregulation of
462 receptors and pathways involved in cell-ECM interactions (115) were also highlighted here,
463 which is consistent with our knowledge of MtrBTN2 biology, pathogenesis (within-host
464 disease progression) and its inter-host transmission. Indeed, MtrBTN2 cells are non-adherent
465 circulating cells, they infiltrate tissues and breach physiological barrier such as epithelium
466 during the transmission process. MtrBTN2 cells are also characterized by a dysregulated
467 differentiation state often seen in aggressive cancers (80), as reflected by the aberrant
468 activation of some meiosis-related genes that suggests a soma-to-germline transition. Such
469 reactivation of meiosis-related genes expression has been previously observed in a wide range
470 of non-germ cell cancers in humans (84). Finally, our transcriptomic data reveal a dampening
471 of host defenses in cancerous mussels likely to facilitate disease progression and pathological
472 outcomes. Evasion from immune destruction by disabling components of the immune system
473 is indeed another hallmark of human cancers (116).

474 **2.8 Future directions**

475 This transcriptomic analysis generated a large amount of information on gene expression in
476 MtrBTN2 cells and we identified genes and pathways linked to their cancerous state. These
477 results represent a significant step for understanding this disease at the cellular and molecular
478 level and set the ground for future research. A comparative oncology approach offers a unique
479 and strong opportunity to learn more about the evolutive mechanisms of cancers and
480 metastatic processes. Moreover, our model offers various advantages. Indeed, it is a naturally
481 occurring cancer and mollusks are not subject to ethical concerns. Yet, a huge amount of work
482 remains to do. The oncogenic processes are characterized by disturbances of mechanisms
483 essential for cell integrity and tissue homeostasis. However, to date, most of the molecular
484 dysfunctions have not been studied in mollusks and could diverge from vertebrates to some
485 extent. Priorities are i) to obtain well curated and annotated reference genomes, ii) the
486 characterization of gene products, and iii) the specific description of protein functions and
487 relationships within and between pathways. To reach these objectives gene manipulation
488 tools are deeply needed, which is a significant challenge in such non-model organisms.

489 **3. Conclusions**

490 We provide here the first transcriptomic profiling of a lophotrochozoan cancer, which in
491 addition has the characteristic to be transmissible. Our results raise a number of evolutionary
492 implications. First, we show that oncogenic processes across bilaterians are underpinned by
493 conserved molecular pathways. This suggests a remote origin of oncogenic processes in which
494 the central features of cooperation that characterize multicellularity are broken down by
495 cheating in proliferation inhibition, cell death, differentiation, resource allocation and
496 extracellular environment maintenance. We also show that the long-term evolution of
497 MtrBTN2, due to its transmissibility, has led to the selection of a large number of oncogenic

498 traits that makes them super-metastases. However, in the specific case of transmissible
499 cancers, an equilibrium between evolution and genome stability is probably necessary for the
500 survival of these cancerous lineages in mussel populations over time. Indeed, a spontaneous
501 question that arises is how these transmissible lineages deal with clonal degeneration for
502 hundreds or even thousands of years. Such evolutionary implications open the way to future
503 investigation.

504

505 **4. Materials and Methods**

506 **4.1 Mussel samplings and MtrBTN2 diagnostic**

507 We collected mussels at the end of 2019 in English Channel and Barents Sea.
508 Two hundred *Mytilus edulis* were sampled in a farm located in Agon-Coutainville
509 (49°0'44.784''N 1°35'55.643''O, Normandy, France) and immediately screened for cancer at
510 the LABEO laboratory (Caen, France). The presence of circulating cancerous cells and the
511 disease stage were first diagnosed by cytological examination on hemolymph samples (18).
512 We found 14 positive individuals among which three were at an advanced stage of the disease
513 (>95% of circulating cells were cancerous cells). After anesthesia (117), we drew a maximum
514 volume of hemolymph from the adductor muscle of these three cancerous mussels as well as
515 of three mussels diagnosed as MtrBTN2-free. Hemolymph was deposited individually in
516 RNase-free microtubes conserved on ice and centrifuged at 800 x g for 10 min at 4°C. The
517 pelleted cells were immediately resuspended in TRIzol® (Invitrogen) and conserved at -80°C
518 until RNA extraction. As non-transmissible circulating cancers also exist in mussels (19), we
519 confirmed the MtrBTN2 diagnosis by two qPCR, one specific to *M. trossulus* (and MtrBTN) and
520 targeting the nuclear marker EF1α (23) and one specific to the MtrBTN2 lineage targeting

521 mtCR (15). Briefly, a piece of mantle and gills were fixed in 96% ethanol and used for DNA
522 extraction done with the Nucleomag® 96 Tissue kit (MachereyNagel). We carried out both
523 qPCRs on cancerous and cancer-free *M. edulis* using the sensiFAST™SyBR® No-ROX Kit
524 (Bioline) and the LightCycler® 480 Real-Time PCR (Roche Life Science) system. We fixed a
525 transversal section of each individual in Davidson for 48h and included it paraffin (RHEM
526 facility). Sections of 3 µm-thin were realized and stained with hematoxylin and eosin to
527 confirm the health status of each mussel and to exclude the presence of other pathologies.
528 Three *M. trossulus* were collected in the wild in Mishukovo (69°02'39;34.3''N 33°01'39;45.9''E,
529 Kola Bay, Russia, Barents Sea). Hemolymph was drawn as for *M. edulis* individuals, the cells
530 were pelleted and resuspended in 100 µL of RNA*later*™ (Invitrogen) to be sent to our
531 laboratory where they were frozen at -80°C until RNA extraction. In addition, we received a
532 piece of mantle and gills from each individual fixed in 96% ethanol for genetic screening and
533 to exclude the presence of MtrBTN, and tissue sections in ethanol 70% after 48h of fixation in
534 Davidson's solution. *M. trossulus* individuals have been subjected to both qPCRs and
535 histological examination.

536 **4.2 Total RNA extraction, library preparation, and sequencing**

537 *M. trossulus* samples conserved in RNA*later*™ were centrifuged at 5000 x g for 10 min at 4°C,
538 supernatant was removed, and cells were resuspended in 500 µL of TRIzol®. *M. edulis*
539 conserved directly in TRIzol® were defrosted. Both sample types were incubated at room
540 temperature for 20 min under agitation to lyse the cells. RNA extraction was performed using
541 Direct-zol™ RNA MiniPrep according to manufacturer's instructions (Zymo Research).

542 The quantification and integrity of the total RNA were checked using a NanoDrop® ND-1000
543 spectrophotometer (Thermo Scientific) and by capillary electrophoresis on a 2100 Bioanalizer
544 (Agilent).

545 The polyadenylated RNA-seq library construction and the sequencing by Illumina® technology
546 were carried out by the GENEWIZ® Company (Germany). Four hundred ng of total RNA at a
547 concentration of 10 ng/µL, and OD260/280 comprised between 1.85 and 2.21 were used for
548 library preparation. The NEBnext® Ultra™ II Directional RNA kit was used for the cDNA library
549 preparation and 9 cycles of enrichment PCR were run. Sequencing was performed on Illumina®
550 NovaSeq™, with a 150bp paired-end configuration, and a sequencing depth of 100M raw
551 paired-end reads per sample.

552 **4.3 De novo transcriptome assembly and functional annotation**

553 Raw reads were processed with RCorrector v1.04
554 (https://anaconda.org/bioconda/rcorrector/files?version=1.0.4) with default settings and -rf
555 configuration to correct sequencing errors (118). Then, we removed uncorrectable reads using
556 FilterUncorrectabledPEfastq tool
557 (https://github.com/harvardinformatics/TranscriptomeAssemblyTools/blob/master/FilterUn
558 correctabledPEfastq.py). The output reads were further processed for adapter removal and
559 trimming with TrimGalore! v0.6.4 (https://github.com/FelixKrueger/TrimGalore) with default
560 parameters and -q 28, --length 100. Ribosomal RNAs potentially still present after polyA
561 capture were removed through alignment against the SILVA Ribosomal database with Bowtie2
562 v2.4.1 (119). Read quality was assessed before and after read trimming with FastQC v0.11.9
563 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

564 In France, MtrBTN2 infects *M. edulis* hosts but originated in a *M. trossulus* founder host. These
565 two species are closely related, hybridize when they come into contact, either naturally or via
566 human-induced introductions (120-122) and introgression between them is observed.
567 However, *M. trossulus* are not present in France and individuals of this species have been
568 sampled in a different environment (Barent Sea *vs* Channel Sea for MtrBTN2). Thus, in order
569 to make possible the differential gene expression analysis of the cancerous cells all the
570 retained reads (from MtrBTN2 cells, *M. edulis* hemocytes and *M. trossulus* hemocytes) were
571 assembled in a pantranscriptome with Trinity v2.8.5 (123) using the default options.
572 TransDecoder v3.0.1 (https://github.com/TransDecoder/TransDecoder/wiki) was run on
573 these contigs to identify CDS, with a minimum length of 90 amino acids. Finally, CDS sequence
574 redundancy was reduced with CD-HIT-EST v4.8.1 (https://github.com/weizhongli/cdhit/wiki)
575 with the following options: -n 6, -c 0.86, -G 0, -aL 1, -aS 1. We assessed the quality of the
576 assembly using several tests. The assembly was first analyzed with TrinityStats and the final
577 pantranscriptome completeness was estimated with BUSCO v5.1.1 (124) against the
578 conserved single-copy metazoan genes database (n = 978). Finally, filtered reads were
579 mapped back on the filtered pantranscriptome to evaluate individual mapping rate with
580 Bowtie2 v2.4.1 (119). To infer that our pantranscriptome strategy was reliable, we performed
581 a Redundancy Analysis (RDA) on normalized read counts to analyze the impact of explanatory
582 variables ("cell type": hemocyte/MtrBTN2 cells, species: trossulus/edulis, "environment":
583 Barents Sea/Channel Sea) on response variables (gene expression) followed by ANOVA like
584 permutation test (nperm=999, model="full") (https://cran.r-
585 project.org/web/packages/vegan/vegan.pdf) (Supplementary Figure S1). Both "cell type" and
586 "environment" were retained as significant explanatory variables ($p<0.05$). This confirms that

587 the inclusion of healthy *M. edulis* mussels in the analysis is necessary to subtract the
588 environment effect (Channel Sea *vs* Barent Sea).

589 For functional annotation, the transcripts were searched against the Uniprot (Swissprot and
590 TrEMBL) protein reference database (125) using PLASTX v2.3.3 algorithm with an e-value
591 cutoff of 1.0E-3 (126). Domain prediction against the InterPro database (127) was carried out
592 with InterProScan v5.48-83.0. Both results were combined and we used the OmicsBox
593 program v2.0 (51,52) to assign GO terms to the annotated sequences with an e-value hit filter
594 of 1.10−3, an annotation cutoff of 55, a HSP-Hit coverage of 60%, and an evidence code of 0.8.
595 We search for supplementary correspondences with EggNog v5.0 (128) by an orthology
596 analysis.

597 **4.4 Biological interpretation of gene expression profiles**

598 In the context of a non-model species, we used three different approaches to interpret the
599 biological relevance of DEGS.

600 ***Enrichment analysis***

601 We performed a GO term enrichment analysis focusing on biological processes with the
602 GO_MWU tool (https://github.com/z0on/GO_MWU) using adaptive clustering and a rank-
603 based statistical test (Mann–Whitney U-test). We used the following parameters for adaptive
604 clustering: largest = 0.5; smallest = 10; clusterCutHeight = 0.5. We took into consideration
605 both the level of expression and the significance of the differential expression: we attributed
606 the log2 fold change value to genes that were significantly differentially expressed
607 (adjusted $p < 0.05$), while we attributed a zero to the others. We considered as enriched a
608 biological process with a FDR < 1%. To represent the results synthetically, we used the

609 Enrichment Map v3.3.3 tool (Merico et al., 2010) in Cytoscape v3.9.1 (129). The intensity of
610 the enrichment was evidenced in the network and was calculated as follows: i) for the
611 processes enriched with over-expressed genes, "number of genes significantly over-expressed
612 in the process/total number of genes in the process"; (ii) for the processes enriched with
613 under-expressed genes, "−1 × (number of genes significantly downregulated in the
614 process/total number of genes in the process)".

615 ***Hub and top expressed gene identification***

616 The top 20 genes in connectivity ranking in the PPI network were selected as Hub genes. We
617 used the Search Tool for the Retrieval of Interacting Genes (STRING) which is a biological
618 database designed for predicting PPI networks (130). The results were visualized in Cytoscape
619 v3.9.1 using NetworkAnalyzer v4.4.8 visualization software (131) that can construct
620 comprehensive models of biologic interactions. Isolated and partially connected nodes were
621 not included.

622 To identify marker genes of the cancerous condition we took into account both differential
623 expression (Log2FC) and expression level (RLE) values to build a plot graph. We defined
624 arbitrarily thresholds (log2FC values > |5|and RLE > 500) to highlight the most discriminating
625 genes between the two conditions (cancerous/healthy circulating cells).

626 ***Targeted analyses***

627 We carried out a targeted analysis focusing on genes and pathways that have been identified
628 as altered at high frequencies across many different human cancer types by Sanchez-Vega et
629 al. (25). Based on these results, we search for the presence of MtrBTN2 DEGs or genes carrying
630 MtrBTN2-specific alleles among these frequently altered genes. Allele visualization was

631 carried out with IGV tool v2.12.3 (106) on BAM files obtained after read alignment on the
632 transcript sequence with Bowtie2. v2.4.1 (119).

633 **References**

634 1. Ostrander EA, Davis BW, Ostrander GK. Transmissible tumors: breaking the
635 cancer paradigm. Trends Genet. 2016;32(1): 1-15. doi: 10.1016/j.tig.2015.10.001.
636 2. Dujon AM, Gatenby RA, Bramwell G, MacDonald N, Dohrmann E, Raven N, et al.
637 Transmissible cancers in an evolutionary perspective. iScience. 2020;23(7):
638 101269. doi: 10.1016/j.isci.2020.101269.
639 3. Nowinsky M. Zurfrageueber die Impfung der krebsigengeschwuelste. Zentralbl
640 Med Wissensch. 1876;14: 790-791.
641 4. Murgia C, Pritchard JK, Kim SY, Fassati A, Weiss RA. Clonal origin and evolution of
642 a transmissible cancer. Cell. 2006;126(3): 477-87. doi: 10.1016/j.cell.2006.05.051.
643 5. Baez-Ortega A, Gori K, Strakova A, Allen JL, Allum KM, Bansse-Issa L, Bhutia TN,
644 Bisson JL, Briceño C, Castillo Domracheva A, Corrigan AM, Cran HR, Crawford JT,
645 Davis E, de Castro KF, B de Nardi A, de Vos AP, Delgadillo Keenan L, Donelan EM,
646 Espinoza Huerta AR, Faramade IA, Fazil M, Fotopoulou E, Fruean SN, Gallardo-
647 Arrieta F, Glebova O, Gouletsou PG, Häfelin Manrique RF, Henriques JJGP, Horta
648 RS, Ignatenko N, Kane Y, King C, Koenig D, Krupa A, Kruzeniski SJ, Kwon YM, Lanza-
649 Perea M, Lazyan M, Lopez Quintana AM, Losfelt T, Marino G, Martínez Castañeda
650 S, Martínez-López MF, Meyer M, Migneco EJ, Nakanwagi B, Neal KB, Neunzig W,
651 Ní Leathlobhair M, Nixon SJ, Ortega-Pacheco A, Pedraza-Ordoñez F, Peleteiro MC,
652 Polak K, Pye RJ, Reece JF, Rojas Gutierrez J, Sadia H, Schmeling SK, Shamanova O,

653 Sherlock AG, Stammnitz M, Steenland-Smit AE, Svitich A, Tapia Martínez LJ, Thoya
654 Ngoka I, Torres CG, Tudor EM, van der Wel MG, Viţălaru BA, Vural SA, Walkinton
655 O, Wang J, Wehrle-Martinez AS, Widdowson SAE, Stratton MR, Alexandrov LB,
656 Martincorena I, Murchison EP. Somatic evolution and global expansion of an
657 ancient transmissible cancer lineage. Science. 2019;365(6452):eaau9923. doi:
658 10.1126/science.aau9923.

659 6. Siddle HV, Kreiss A, Eldridge MD, Noonan E, Clarke CJ, Pyecroft S, et al.
660 Transmission of a fatal clonal tumor by biting occurs due to depleted MHC
661 diversity in a threatened carnivorous marsupial. Proc Natl Acad Sci USA.
662 2007;104(41): 16221-16226. doi: 10.1073/pnas.0704580104.

663 7. Hamede RK, Bashford J, McCallum H, Jones M. Contact networks in a wild
664 Tasmanian devil (*Sarcophilus harrisii*) population: using social network analysis to
665 reveal seasonal variability in social behaviour and its implications for transmission
666 of devil facial tumour disease. Ecol Lett. 2009;12(11): 1147-1157. doi:
667 10.1111/j.1461-0248.2009.01370.x.

668 8. Hamede RK, McCallum H, Jones M. Biting injuries and transmission of Tasmanian
669 devil facial tumour disease. J Anim Ecol. 2013;82(1): 182-190. doi:
670 10.1111/j.1365-2656.2012.02025.x.

671 9. Hawkins CE, Baars C, Hesterman H, Hocking GJ, Jones ME, Lazenby B, et al.
672 Emerging disease and population decline of an island endemic, the Tasmanian
673 devil *Sarcophilus harrisii*. Biol Conserv. 2006;131(2): 307-324. doi:
674 10.1016/j.biocon.2006.04.010.

675 10. Lachish S, Jones M, McCallum H. The impact of disease on the survival and
676 population growth rate of the Tasmanian devil. J Anim Ecol. 2007;76(5): 926-936.
677 doi: 10.1111/j.1365-2656.2007.01272.x.
678 11. McCallum H, Jones M, Hawkins C, Hamede R, Lachish S, Sinn DL, Beeton N,
679 Lazenby B. Transmission dynamics of Tasmanian devil facial tumor disease may
680 lead to disease-induced extinction. Ecology. 2009;90(12): 3379-3392. doi:
681 10.1890/08-1763.1.
682 12. Lazenby BT, Tobler MW, Brown WE, Hawkins CE, Hocking GJ, Hume F, et al.
683 Density trends and demographic signals uncover the long-term impact of
684 transmissible cancer in Tasmanian devils. J Appl Ecol. 2018;55(3): 1368-1379. doi:
685 10.1111/1365-2664.13088.
686 13. Metzger MJ, Reinisch C, Sherry J, Goff SP. Horizontal transmission of clonal cancer
687 cells causes leukemia in soft-shell clams. Cell. 2015;161(2): 255-263. doi:
688 10.1016/j.cell.2015.02.042.
689 14. Metzger MJ, Villalba A, Carballal MJ, Iglesias D, Sherry J, Reinisch C, et al.
690 Widespread transmission of independent cancer lineages within multiple bivalve
691 species. Nature. 2016;534(7609): 705-709. doi: 10.1038/nature18599.
692 15. Yonemitsu MA, Giersch RM, Polo-Prieto M, Hammel M, Simon A, Cremonte F, et
693 al. A single clonal lineage of transmissible cancer identified in two marine mussel
694 species in South America and Europe. elife. 2019;8: e47788. doi:
695 10.7554/eLife.47788.
696 16. Garcia-Souto D, Bruzos AL, Diaz S, Rocha S, Pequeño-Valtierra A, Roman-Lewis CF,
697 et al. Mitochondrial genome sequencing of marine leukaemias reveals cancer

contagion between clam species in the Seas of Southern Europe. eLife, 2022;11: e66946.doi: 10.7554/eLife.66946

17. Riquet F, Simon A, Bierne N. Weird genotypes? Don't discard them, transmissible cancer could be an explanation. Evol Appl. 2017;10(2): 140-145. doi: 10.1111/eva.12439.
18. Burioli EAV, Trancart S, Simon A, Bernard I, Charles M, Oden E, Bierne N, Houssin M. Implementation of various approaches to study the prevalence, incidence and progression of disseminated neoplasia in mussel stocks. J Invertebr Pathol. 2019;168: 107271. doi: 10.1016/j.jip.2019.107271.
19. Hammel M, Simon A, Arbiol C, Villalba A, Burioli EAV, Pépin JF, et al. Prevalence and polymorphism of a mussel transmissible cancer in Europe. Mol Ecol. 2022;31(3): 736-751. doi: 10.1111/mec.16052.
20. Skazina M, Odintsova N, Maiorova M, Ivanova A, Väinölä R, Strelkov P. First description of a widespread *Mytilus trossulus*-derived bivalve transmissible cancer lineage in M. trossulus itself. Sci Rep. 2021;11(1): 5809. doi: 10.1038/s41598-021-85098-5.
21. Carballal MJ, Barber BJ, Iglesias D, Villalba A. Neoplastic diseases of marine bivalves. J Invertebr Pathol. 2015;131: 83-106. doi: 10.1016/J.JIP.2015.06.004.
22. Benadelmouna A, Saunier A, Ledu C, Travers MA, Morga B. Genomic abnormalities affecting mussels (*Mytilus edulis-galloprovincialis*) in France are related to ongoing neoplastic processes, evidenced by dual flow cytometry and cell monolayer analyses. J Invertebr Pathol. 2018;157: 45-52. doi: 10.1016/j.jip.2018.08.003.

721 23. Burioli EAV, Hammel M, Bierne N, Thomas F, Houssin M, Destoumieux-Garzón D,
722 Charrière GM. Traits of a mussel transmissible cancer are reminiscent of a
723 parasitic life style. Sci Rep. 2021;11(1): 24110. doi: 10.1038/s41598-021-03598-
724 w.
725 24. Hanahan D. Hallmarks of Cancer: New Dimensions. Cancer Discov. 2022;12(1): 31-
726 46. doi: 10.1158/2159-8290.CD-21-1059.
727 25. Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, et al. Oncogenic
728 signaling pathways in the cancer genome atlas. Cell. 2018;173(2): 321-337.e10.
729 doi: 10.1016/j.cell.2018.03.035.
730 26. Giersch RM, Hart SFM, Reddy SG, Yonemitsu MA, Orellana Rosales MJ, Korn M,
731 Geleta BM, Countway PD, Fernández Robledo JA, Metzger MJ. Survival and
732 Detection of Bivalve Transmissible Neoplasia from the Soft-Shell Clam Mya
733 arenaria (MarBTN) in Seawater. Pathogens. 2022;11(3):283. doi:
734 10.3390/pathogens11030283.
735 27. Sotiriou C, Piccart MJ. Taking gene-expression profiling to the clinic: when will
736 molecular signatures become relevant to patient care? Nat Rev Cancer.
737 2007;7(7): 545-553. doi: 10.1038/nrc2173.
738 28. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics.
739 Nat Rev Genet. 2009;10(1): 57-63. doi: 10.1038/nrg2484.
740 29. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et
741 al. A survey of best practices for RNA-seq data analysis. Genome Biol. 2016;17:
742 13. doi: 10.1186/s13059-016-0881-8. Erratum in: Genome Biol. 2016;17(1): 181.

743 30. Martin JA, Wang Z. Next-generation transcriptome assembly. Nat Rev Genet.
744 2011;12(10): 671-682. doi: 10.1038/nrg3068.
745 31. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a
746 universal tool for annotation, visualization and analysis in functional genomics
747 research. Bioinformatics. 2005;21: 3674-3676. doi:
748 10.1093/bioinformatics/bti610.
749 32. Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, et al. High-
750 throughput functional annotation and data mining with the Blast2GO
751 suite. Nucleic acids res. 2008;36(10): 3420-3435. doi: 10.1093/nar/gkn176.
752 33. Hu J, Zhang Z, Shen WJ, Azhar S. Cellular cholesterol delivery, intracellular
753 processing and utilization for biosynthesis of steroid hormones. Nutr Metab
754 (Lond). 2010;7: 47. doi: 10.1186/1743-7075-7-47.
755 34. Murphy-Ullrich JE, Poczatek M. Activation of latent TGF-β by thrombospondin-1:
756 mechanisms and physiology. Cytokine Growth Factor Rev. 2000;11(1-2): 59-69.
757 doi: 10.1016/S1359-6101(99)00029-5.
758 35. Letterio JJ, Roberts AB. Regulation of immune responses by TGF-β. Annu Rev
759 Immunol. 1998;16(1): 137-161. doi: 10.1146/annurev.immunol.16.1.137.
760 36. Ashcroft GS. Bidirectional regulation of macrophage function by TGF-β. Microbes
761 Infect. 1999;1(15): 1275-1282. doi: 10.1016/S1286-4579(99)00257-9.
762 37. Läubli H, Borsig L. Selectins promote tumor metastasis. Semin Cancer Biol.
763 2010;20(3): 169-177. doi: 10.1016/j.semcancer.2010.04.005.

764 38. Fabricius HÅ, Starzonek S, Lange T. The role of platelet cell surface P-Selectin for
765 the direct platelet-tumor cell contact during metastasis formation in human
766 tumors. Front Oncol. 2021;11: 642761. doi: 10.3389/fonc.2021.642761.
767 39. Anderson KM, Seed T, Vos M, Mulshine J, Meng J, Alrefai W, Ou D, Harris JE. 5-
768 Lipoxygenase inhibitors reduce PC-3 cell proliferation and initiate nonnecrotic cell
769 death. Prostate. 1998;37(3): 161-173. doi: 10.1002/(sici)1097-
770 0045(19981101)37:3<161::aid-pros5>3.0.co;2-d.
771 40. Bishayee K, Khuda-Bukhsh AR. 5-Lipoxygenase antagonist therapy: a new
772 approach towards targeted cancer chemotherapy. Acta Biochim Biophys Sin.
773 2013;45(9): 709-719. doi: 10.1093/abbs/gmt064.
774 41. Lee JH, Lee GT, Woo SH, Ha YS, Kwon SJ, Kim WJ, Kim IY. BMP-6 in renal cell
775 carcinoma promotes tumor proliferation through IL-10-dependent M2
776 polarization of tumor-associated macrophages. Cancer Res. 2013;73(12): 3604-
777 3614. doi: 10.1158/0008-5472.CAN-12-4563.
778 42. Redondo-García S, Peris-Torres C, Caracuel-Peramos R, Rodríguez-Manzaneque
779 JC. ADAMTS proteases and the tumor immune microenvironment: Lessons from
780 substrates and pathologies. Matrix Biol Plus. 2020;9: 100054. doi:
781 10.1016/j.mbplus.2020.100054.
782 43. Rucci N, Sanità P, Angelucci A. Roles of metalloproteases in metastatic niche. Curr
783 Mol Med. 2011;11: 609-622. doi: 10.2174/156652411797536705.
784 44. Huang T, Sun L, Yuan X, Qiu H. Thrombospondin-1 is a multifaceted player in
785 tumor progression. Oncotarget. 2017;8(48): 84546-84558. doi:
786 10.18632/oncotarget.19165.

787 45. Bach DH, Park HJ, Lee SK. The dual role of bone morphogenetic proteins in cancer.
788 Mol Ther Oncolytics. 2017;8: 1-13. doi: 10.1016/j.omto.2017.10.002.
789 46. Yee KO, Connolly CM, Duquette M, Kazerounian S, Washington R, Lawler J. The
790 effect of thrombospondin-1 on breast cancer metastasis. Breast Cancer Res Treat.
791 2009;114(1): 85-96. doi: 10.1007/s10549-008-9992-6.
792 47. Nucera C, Porrello A, Antonello ZA, Mekel M, Nehs MA, Giordano TJ, et al. B-
793 Raf(V600E) and thrombospondin-1 promote thyroid cancer progression. Proc
794 Natl Acad Sci USA. 2010;107(23): 10649-10654. doi: 10.1073/pnas.1004934107.
795 48. Borsotti P, Ghilardi C, Ostano P, Silini A, Dossi R, Pinessi D., et al. Thrombospondin-
796 1 is part of a Slug-independent motility and metastatic program in cutaneous
797 melanoma, in association with VEGFR-1 and FGF-2. Pigment Cell Melanoma Res.
798 2015;28: 73-81. doi: 10.1111/pcmr.12319.
799 49. Huang T, Wang L, Liu D, Li P, Xiong H, Zhuang L, Sun L, Yuan X, Qiu H. FGF7/FGFR2
800 signal promotes invasion and migration in human gastric cancer through
801 upregulation of thrombospondin-1. Int J Oncol. 2017;50(5): 1501-1512. doi:
802 10.3892/ijo.2017.3927.
803 50. Iwamura T, Caffrey TC, Kitamura N, Yamanari H, Setoguchi T, Hollingsworth MA.
804 P-selectin expression in a metastatic pancreatic tumor cell line (SUIT-2). Cancer
805 Res. 1997;57(6): 1206-1212.
806 51. Ding XZ, Kuszynski CA, El-Metwally TH, Adrian TE. Lipoxygenase inhibition induced
807 apoptosis, morphological changes, and carbonic anhydrase expression in human
808 pancreatic cancer cells. Biochem Biophys Res Commun. 1999;266(2): 392-399.
809 doi: 10.1006/bbrc.1999.1824.

810 52. Gupta S, Srivastava M, Ahmad N, Sakamoto K, Bostwick DG, Mukhtar H.
811 Lipoxygenase-5 is overexpressed in prostate adenocarcinoma. Cancer.
812 2001;91(4): 737-743. doi: 10.1002/1097-0142(20010215)91:4<737::aid-
813 cncr1059>3.0.co;2-f.
814 53. Ohd JF, Nielsen CK, Campbell J, Landberg G, Löfberg H, Sjölander A. Expression of
815 the leukotriene D4 receptor CysLT1, COX-2, and other cell survival factors in
816 colorectal adenocarcinomas. Gastroenterology. 2003;124(1): 57-70. doi:
817 10.1053/gast.2003.50011.
818 54. Skovrlj B, Koehler SM, Anderson PA, Qureshi SA, Hecht AC, Iatridis JC, Cho SK.
819 Association Between BMP-2 and Carcinogenicity. Spine (Phila Pa 1976).
820 2015;40(23): 1862-1871. doi: 10.1097/BRS.0000000000001126.
821 55. Pfeiffer V, Lingner J. Replication of telomeres and the regulation of telomerase.
822 Cold Spring Harbor Perspect Biol. 2013;5(5): a010405. doi:
823 10.1101/cshperspect.a010405.
824 56. Bell RJ, Rube HT, Xavier-Magalhães A, Costa BM, Mancini A, Song JS, Costello JF.
825 Understanding TERT Promoter Mutations: A Common Path to Immortality. Mol
826 Cancer Res. 2016;14(4): 315-323. doi: 10.1158/1541-7786.MCR-16-0003.
827 57. Mills GB, Schmandt R, McGill M, Amendola A, Hill M, Jacobs K, et al. Expression
828 of TTK, a novel human protein kinase, is associated with cell proliferation. J Biol
829 Chem. 1992;267(22): 16000-16006. doi: 10.1016/S0021-9258(19)49633-6.
830 58. Tucker L, Zhang Q, Sheppard G, Lou P, Jiang F, McKeegan E, et al. Ectopic
831 expression of methionine aminopeptidase-2 causes cell transformation and

832 stimulates proliferation. Oncogene. 2008;27: 3967-3976. doi:
833 10.1038/onc.2008.14.
834 59. Nguyen JT, Ray C, Fox AL, Mendonça DB, Kim JK, Krebsbach PH. Mammalian EAK-
835 7 activates alternative mTOR signaling to regulate cell proliferation and migration.
836 Sci Adv. 2018;4(5):eaao5838. doi: 10.1126/sciadv.aao5838.
837 60. Sancak Y, Bar-Peled L, Zoncu R, Markhard AL, Nada S, Sabatini DM. Ragulator-Rag
838 complex targets mTORC1 to the lysosomal surface and is necessary for its
839 activation by amino acids. Cell. 2010;141(2): 290-303. doi:
840 10.1016/j.cell.2010.02.024.
841 61. Pastuszak-Lewandoska D, Czarnecka KH, Migdalska-Sęk M, Nawrot E, Domańska
842 D, Kiszałkiewicz J, et al. Decreased FAM107A expression in patients with non-
843 small cell lung cancer. Adv Exp Med Biol. 2015;852: 39-48. doi:
844 10.1007/5584_2014_109.
845 62. Saharinen J, Hyytiäinen M, Taipale J, Keski-Oja J. Latent transforming growth
846 factor-beta binding proteins (LTBPs)-structural extracellular matrix proteins for
847 targeting TGF-beta action. Cytokine Growth Factor Rev. 1999;10(2): 99-117. doi:
848 10.1016/s1359-6101(99)00010-6.
849 63. Wu BX, Zeidan YH, Hannun YA. Downregulation of neutral ceramidase by
850 gemcitabine: Implications for cell cycle regulation. Biochim Biophys Acta.
851 2009;1791(8): 730-739. doi: 10.1016/j.bbalip.2009.03.012.
852 64. Loveday C, Tatton-Brown K, Clarke M, Westwood I, Renwick A, Ramsay E, et al.
853 Mutations in the PP2A regulatory subunit B family genes PPP2R5B, PPP2R5C and

854 PPP2R5D cause human overgrowth. Hum Mol Genet. 2015;24(17): 4775-4779.
855 doi: 10.1093/hmg/ddv182.
856 65. Fu M, Wang C, Li Z, Sakamaki T, Pestell RG. Cyclin D1: normal and abnormal
857 functions. Endocrinology. 2004;145(12): 5439-5447. doi: 10.1210/en.2004-0959.
858 66. Shimobayashi M, Hall M. Multiple amino acid sensing inputs to mTORC1. Cell Res.
859 2016;26: 7-20. doi: 10.1038/cr.2015.146.
860 67. Cargnello M, Tcherkezian J, Roux PP. The expanding role of mTOR in cancer cell
861 growth and proliferation. Mutagenesis. 2015;30(2):169-76. doi:
862 10.1093/mutage/geu045.
863 68. Ricoult SJ, Yecies JL, Ben-Sahra I, Manning BD. Oncogenic PI3K and K-Ras stimulate
864 de novo lipid synthesis through mTORC1 and SREBP. Oncogene.
865 2016;35(10):1250-60. doi: 10.1038/onc.2015.179.
866 69. Valvezan AJ, Turner M, Belaid A, Lam HC, Miller SK, McNamara MC, Baglini C,
867 Housden BE, Perrimon N, Kwiatkowski DJ, Asara JM, Henske EP, Manning BD.
868 mTORC1 Couples Nucleotide Synthesis to Nucleotide Demand Resulting in a
869 Targetable Metabolic Vulnerability. Cancer Cell. 2017;32(5):624-638.e5. doi:
870 10.1016/j.ccell.2017.09.013.
871 70. Ben-Sahra I, Hoxhaj G, Ricoult SJH, Asara JM, Manning BD. mTORC1 induces
872 purine synthesis through control of the mitochondrial tetrahydrofolate cycle.
873 Science. 2016;351(6274):728-733. doi: 10.1126/science.aad0489.
874 71. Martínez-Reyes I, Chandel NS. Mitochondrial TCA cycle metabolites control
875 physiology and disease. Nat Commun. 2020;11: 102. doi: 10.1038/s41467-019-
876 13668-3.

877 72. DeBerardinis RJ, Lum JJ, Hatzivassiliou G, Thompson CB. The biology of cancer:
878 metabolic reprogramming fuels cell growth and proliferation. Cell Metab.
879 2008;7(1):11-20. doi: 10.1016/j.cmet.2007.10.002.
880 73. Nishimura K, Ishiai M, Horikawa K, Fukagawa T, Takata M, Takisawa H, Kanemaki
881 MT. Mcm8 and Mcm9 form a complex that functions in homologous
882 recombination repair induced by DNA interstrand crosslinks. Mol Cell. 2012;47(4):
883 511-522. doi: 10.1016/j.molcel.2012.05.047.
884 74. Hanada K, Budzowska M, Modesti M, Maas A, Wyman C, Essers J, Kanaar R. The
885 structure-specific endonuclease Mus81-Eme1 promotes conversion of
886 interstrand DNA crosslinks into double-strands breaks. EMBO J. 2006;25(20):
887 4921-49. doi: 10.1038/sj.emboj.7601344.
888 75. Chatterjee N, Walker GC. Mechanisms of DNA damage, repair, and mutagenesis.
889 Environ Mol Mutagen. 2017;58(5): 235-263. doi: 10.1002/em.22087.
890 76. Di Micco R, Fumagalli M, Cicalese A, Piccinin S, Gasparini P, Luise C, et al.
891 Oncogene-induced senescence is a DNA damage response triggered by DNA
892 hyper-replication. Nature. 2006;444(7119): 638-642. doi: 10.1038/nature05327.
893 77. Haber JE. Partners and pathways repairing a double-strand break. Trends Genet.
894 2000;16: 259-264. doi: 10.1016/s0168-9525(00)02022-9.
895 78. Kozakiewicz CP, Fraik AK, Patton AH, Ruiz-Aravena M, Hamilton DG, Hamede R, et
896 al. Spatial variation in gene expression of Tasmanian devil facial tumors despite
897 minimal host transcriptomic response to infection. BMC Genomics. 2021;22(1):
898 698. doi: 10.1186/s12864-021-07994-4.

899 79. Mizejewski GJ. Role of integrins in cancer: survey of expression patterns. Proc Soc
900 Exp Biol Med. 1999; 222(2):124-38. doi: 10.1046/j.1525-1373.1999.d01-122.x.
901 80. Jilkine A, Gutenkunst RN. Effect of dedifferentiation on time to mutation
902 acquisition in stem cell-driven cancers. PLoS Comput Biol. 2014;10(3):e1003481.
903 doi: 10.1371/journal.pcbi.1003481.
904 81. Bocker T, Barusevicius A, Snowden T, Rasio D, Guerrette S, Robbins D, et al.
905 hMSH5: A human MutS homologue that forms a novel heterodimer with hMSH4
906 and is expressed during spermatogenesis. Cancer res. 1999;59(4): 816-822.
907 82. Syrjänen JL, Pellegrini L, Davies OR. A molecular model for the role of SYCP3 in
908 meiotic chromosome organisation. eLife. 2014;3: e02963. doi:
909 10.7554/eLife.02963.
910 83. Adelman CA, Petrini JH. ZIP4H (TEX11) deficiency in the mouse impairs meiotic
911 double strand break repair and the regulation of crossing over. PLoS Genet.
912 2008;4(3):e1000042. doi: 10.1371/journal.pgen.1000042.
913 84. McFarlane RJ, Wakeman JA. Meiosis-like Functions in Oncogenesis: A New View
914 of Cancer. Cancer Res. 2017;77(21):5712-5716. doi:10.1158/0008-5472.CAN-17-
915 1535.
916 85. Simpson AJ, Caballero OL, Jungbluth A, Chen YT, Old LJ. Cancer/testis antigens,
917 gametogenesis and cancer. Nat Rev Cancer. 2005;5(8):615-
918 625.doi:10.1038/nrc1669.
919 86. Hubert F, Noel T, Roch P. A member of the arthropod defensing family from edible
920 Mediterranean mussels (*Mytilus galloprovincialis*). Eur J Biochem. 1996;240(1):
921 302-306. doi: 10.1111/j.1432-1033.1996.0302h.x.

922 87. Mitta G, Hubert F, Noël T, Roch P. Myticin, a novel cysteine-rich antimicrobial
923 peptide isolated from haemocytes and plasma of the mussel *Mytilus*
924 *galloprovincialis*. Eur J Biochem. 1999;265(1): 71-78. doi: 10.1046/j.1432-
925 1327.1999.00654.x.
926 88. Krasity BC, Troll JV, Weiss JP, McFall-Ngai MJ. LBP/BPI proteins and their relatives:
927 conservation over evolution and roles in mutualism. Biochem Soc T. 2011;39(4):
928 1039-1044. doi: 10.1042/BST0391039.
929 89. Zhang L, Li L, Zhang G. A *Crassostrea gigas* Toll-like receptor and comparative
930 analysis of TLR pathway in invertebrates. Fish Shellfish Immunol. 2011;30(2): 653-
931 660. doi: 10.1016/j.fsi.2010.12.023.
932 90. Garlatti V, Belloy N, Martin L, Lacroix M, Matsushita M, Endo Y, et al. Structural
933 insights into the innate immune recognition specificities of L- and H-ficolins.
934 EMBO j. 2007;26(2): 623-633. doi: 10.1038/sj.emboj.7601500.
935 91. Schlosser A, Thomsen T, Moeller JB, Nielsen O, Tornøe I, Mollenhauer J, Moestrup
936 SK, Holmskov U. Characterization of FIBCD1 as an acetyl group-binding receptor
937 that binds chitin. J Immunol. 2009 ;183(6): 3800-3809. doi:
938 10.4049/jimmunol.0901526.
939 92. Dheilly NM, Duval D, Mouahid G, Emans R, Allienne JF, Galinier R, et al. A family
940 of variable immunoglobulin and lectin domain containing molecules in the snail
941 *Biomphalaria glabrata*. Dev Comp Immunol. 2015;48(1): 234-243. doi:
942 10.1016/j.dci.2014.10.009.
943 93. Trino S, Laurenzana I, Lamorte D, Calice G, De Stradis A, Santodirocco M,
944 Sgambato A, Caivano A, De Luca L. Acute myeloid leukemia cells functionally

945 compromise hematopoietic stem/progenitor cells inhibiting normal
946 hematopoiesis through the release of extracellular vesicles. Front Oncol. 2022;12:
947 824562. doi: 10.3389/fonc.2022.824562.
948 94. Kent ML, Elston RA, Wilkinson MT, Drum AS. Impaired defense mechanisms in bay
949 mussels, *Mytilus edulis*, with hemic neoplasia. J Invertebr Pathol. 1989;53(3): 378-
950 386. doi: 10.1016/0022-2011(89)90103-1.
951 95. Assenov Y, Ramírez F, Schelhorn SE, Lengauer T, Albrecht M. Computing
952 topological parameters of biological networks. Bioinformatics. 2008;24(2): 282-
953 284. doi: 10.1093/bioinformatics/btm554.
954 96. Pankov R, Yamada KM. Fibronectin at a glance. J Cell Sci. 2002;115(20): 3861-
955 3863. doi: 10.1242/jcs.00059.
956 97. Cotteret S, Chernoff J. The evolutionary history of effectors downstream of Cdc42
957 and Rac. Genome Biol. 2002;3(2):REVIEWS0002. doi: 10.1186/gb-2002-3-2-
958 reviews0002
959 98. Porter AG, Jänicke RU. Emerging roles of caspase-3 in apoptosis. Cell Death Differ.
960 1999;6(2): 99-104. doi: 10.1038/sj.cdd.4400476.
961 99. Yu FX, Guan KL. The Hippo pathway: regulators and regulations. Genes Dev.
962 2013;27(4): 355-371. doi: 10.1101/gad.210773.112.
963 100. Carracedo A, Pandolfi P. The PTEN–PI3K pathway: of feedbacks and cross-talks.
964 Oncogene. 2008;27: 5527-5541. doi: 10.1038/onc.2008.247.
965 101. Cohen P, Frame S. The renaissance of GSK3. Nat Rev Mol Cell Biol. 2001;2: 769-
966 776. doi: 10.1038/35096075.

967 102. Amati B, Land H. Myc-Max-Mad: a transcription factor network controlling cell
968 cycle progression, differentiation and death. Curr Opin Genet Dev. 1994;4(1):
969 102-108. doi: 10.1016/0959-437x(94)90098-1.
970 103. Dikovskaya D, Khoudoli G, Newton IP, Chadha GS, Klotz D, Visvanathan A, et al.
971 The Adenomatous Polyposis Coli protein contributes to normal compaction of
972 mitotic chromatin. PLOS One. 2012;7(6): e38102. doi:
973 10.1371/journal.pone.0038102.
974 104. Lecarpentier Y, Schussler O, Hébert JL, Vallée A. Multiple targets of the
975 canonical WNT/β-Catenin signaling in cancers. Front Oncol. 2019;9: 1248. doi:
976 10.3389/fonc.2019.01248.
977 105. Iso T, Kedes L, Hamamori Y. HES and HERP families: multiple effectors of the
978 Notch signaling pathway. J Cell Physiol. 2003;194(3): 237-255. doi:
979 10.1002/jcp.10208.
980 106. Robinson JT, Thorvaldsdóttir H, Wenger AM, Zehir A, Mesirov JP. Variant review
981 with the Integrative Genomics Viewer (IGV). Cancer Res. 2017;77(21): 31-34. doi:
982 10.1158/0008-5472.CAN-17-0337.
983 107. Kelley ML, Winge P, Heaney JD, Stephens RE, Farell JH, Beneden RJV, et al.
984 Expression of homologues for p53 and p73 in the softshell clam (*Mya arenaria*),
985 a naturally-occuring model for human cancer. Oncogene. 2001;20: 748-758. doi:
986 10.1038/sj.onc.1204144.
987 108. Vassilenko EI, Muttray AF, Schulte PM, Baldwin SA. Variations in p53-like cDNA
988 sequence are correlated with mussel haemic neoplasia: A potential molecular-

989 level tool for biomonitoring. Mutat Res. 2010;701(2): 145-152. doi:
990 10.1016/j.mrgentox.2010.06.001.
991 109. Muttray AF, Cox RL, Reinisch CL, Baldwin SA. Identification of DeltaN isoform
992 and polyadenylation site choice variants in molluscan p63/p73-like homologues.
993 Mar Biotechnol (NY). 2007;9(2): 217-230. doi: 10.1007/s10126-006-6045-1.
994 110. Aktipis CA, Boddy AM, Jansen G, et al. Cancer across the tree of life:
995 cooperation and cheating in multicellularity. Philos Trans R Soc Lond B Biol Sci.
996 2015;370(1673): 20140219. doi:10.1098/rstb.2014.0219.
997 111. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al.
998 Mutational heterogeneity in cancer and the search for new cancer-associated
999 genes. Nature. 2013;499(7457): 214-218. doi: 10.1038/nature12213.
1000 112. Hoxhaj G, Manning BD. The PI3K-AKT network at the interface of oncogenic
1001 signalling and cancer metabolism. Nat Rev Cancer. 2020;20(2):74-88. doi:
1002 10.1038/s41568-019-0216-7.
1003 113. Magaway C, Kim E, Jacinto E. Targeting mTOR and Metabolism in Cancer:
1004 Lessons and Innovations. Cells. 2019;8(12):1584. doi: 10.3390/cells8121584.
1005 114. Zou Z, Tao T, Li H, Zhu X. mTOR signaling pathway and mTOR inhibitors in
1006 cancer: progress and challenges. Cell Biosci. 2020;10:31. doi: 10.1186/s13578-
1007 020-00396-1.
1008 115. Hamidi H, Ivaska J. Every step of the way: integrins in cancer progression and
1009 metastasis. Nat Rev Cancer. 2018;18(9):533-548. doi: 10.1038/s41568-018-0038-
1010 z.

1011 116. Hanahan D. Hallmarks of Cancer: New Dimensions. Cancer Discov. 2022;12(1):
1012 31-46. doi: 10.1158/2159-8290.
1013 117. Suquet M, de Kermoysan G, Gonzalez Araya R, Queau I, Lebrun L, Le Souchu
1014 P, Mingant C. Anesthesia in Pacific oyster *Crassostrea gigas*. Aquat Living Resour.
1015 2009;22: 29-34. doi: 10.1051/alr/2009006.
1016 118. Song L, Florea L. Rcorrector: efficient and accurate error correction for Illumina
1017 RNA-seq reads. GigaScience. 2015;4(1): 48. doi: 10.1186/s13742-015-0089-y.
1018 119. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat
1019 Methods. 2012;9(4): 357-359. doi:10.1038/nmeth.1923.
1020 120. Fraïsse C, Belkhir K, Welch JJ, Bierne N. Local interspecies introgression is the
1021 main cause of extreme levels of intraspecific differentiation in mussels. Mol Ecol.
1022 2016;25(1): 269-286. doi: 10.1111/mec.13299.
1023 121. Popovic I, Matias AMA, Bierne N, Riginos C. Twin introductions by independent
1024 invader mussel lineages are both associated with recent admixture with a native
1025 congener in Australia. Evol Appl. 2020;13(3): 515-532. doi: 10.1111/eva.12857.
1026 122. Simon A, Arbiol C, Nielsen EE, Couteau J, Sussarellu R, Burgeot T, et al.
1027 Replicated anthropogenic hybridisations reveal parallel patterns of admixture in
1028 marine mussels. Evol Appl. 2020;13(3): 575-599. doi: 10.1111/eva.12879.
1029 123. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-
1030 length transcriptome assembly from RNA-Seq data without a reference genome.
1031 Nat Biotechnol. 2011;29(7): 644-652. doi: 10.1038/nbt.1883.
1032 124. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO:
1033 assessing genome assembly and annotation completeness with single-copy

1034 orthologs. Bioinformatics. 2015;31: 3210-3212. doi:
1035 10.1093/bioinformatics/btv351.
1036 125. Bairoch A, Apweiler R. The SWISS-PROT protein sequence database and its
1037 supplement TrEMBL in 2000. Nucleic Acids Res. 2000;28: 45-48. doi:
1038 10.1093/nar/28.1.45.
1039 126. Nguyen VH, Lavenier D. PLAST: parallel local alignment search tool for database
1040 comparison. BMC Bioinf. 2009;10: 329. doi: 10.1186/1471-2105-10-329.
1041 127. Blum M, Chang H, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, et al. The
1042 InterPro protein families and domains database: 20 years on. Nucleic Acids Res.
1043 2021;49(D1): 344-354. doi: 10.1093/nar/gkaa977.. doi: 10.1093/nar/gkaa977.
1044 128. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H,
1045 et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated
1046 orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res.
1047 2019; 47(D1): 309-314. doi: 10.1093/nar/gky1085.
1048 129. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape:
1049 a software environment for integrated models of biomolecular interaction
1050 networks. Genome Res. 2003;13(11): 2498-2504. doi: 10.1101/gr.1239303.
1051 130. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING
1052 database in 2021: customizable protein–protein networks, and functional
1053 characterization of user-uploaded gene/measurement sets. Nucleic Acids Res.
1054 2021;49(D1): 605-612. doi: 10.1093/nar/gkaa1074.

1055 131. Assenov Y, Ramírez F, Schelhorn SE, Lengauer T, Albrecht M. Computing
1056 topological parameters of biological networks. Bioinformatics. 2008;24(2):282-4.
1057 doi: 10.1093/bioinformatics/btm554.

1058 **Acknowledgments**

1059 We acknowledge the Animal Health-Clinical Biology unit of the LABÉO laboratory (Caen,
1060 France), especially Maryline Houssin and Ludovic Petit, for their help during cytological
1061 screening and Professor Petr Strelkov from the Saint Petersburg State University for the
1062 sampling of *M. trossulus* individuals. We thank the "Réseau d'Histologie Expérimentale de
1063 Montpellier" - RHEM facility - supported by SIRIC Montpellier Cancer (Grant
1064 INCa_Inserm_DGOS_12553), the European regional development foundation and the
1065 Occitanian region (FEDER-FSE 2014-2020 Languedoc Roussillon), for processing our animal
1066 tissues, histology technics and expertise. This study is set within the framework of the
1067 'Laboratoires d'Excellence (LABEX)' Tulip (ANR- 10-LABX-41). **Funding:** this work was
1068 supported by the French National Agency for Research, TRANCAN project (ANR-18-CE35-
1069 0009) and the Montpellier Université d'Excellence (MUSE), BLUECANCER project. FT is
1070 supported by the MAVA Foundation. **Author contributions:** Conceptualization: EAVB, DDG,
1071 GM, GMC, EV, JVD; Investigation, Formal analysis: EAVB, MH; Methodology: EAVB, EV, JVD;
1072 Funding Acquisition: DDG, GMC, NB, FT; Original Draft preparation: EAVB; Reviewing&Editing:
1073 DDG, GMC, MH, EV, JVD, NB, FT, GM. **Competing interests:** The authors declare that they have
1074 no competing interests. **Data and materials availability:** Raw read sequences were deposited
1075 with links to BioProject accession number PRJNA749800 in the NCBI BioProject database
1076 (https://www.ncbi.nlm.nih.gov/bioproject/).

1077