Genome of Antarctic-endemic Copepod and Evolutionary Adaptation to Extreme Environments

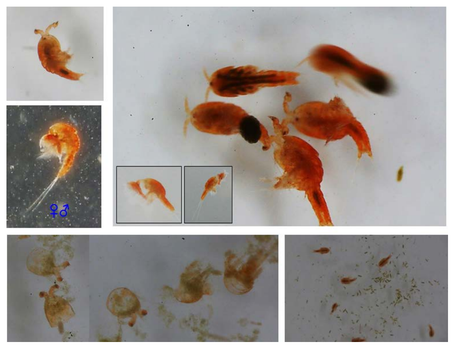
Seunghyun Kang[1], Do Hwan Ahn[2], Sanghee Kim[1]*, Hyun Park[1,2]*

[1]Division of Polar Life Sciences, Korea Polar Research Institute, Yeonsu-gu, Incheon 406-840, South Korea,
[2]Polar Sciences, University of Science & Technology, Yuseong-gu, Daejeon 305-333, Korea
*Co-corresponding author: sangheekim@kopri.re.kr, hpark@kopri.re.kr

Abstract

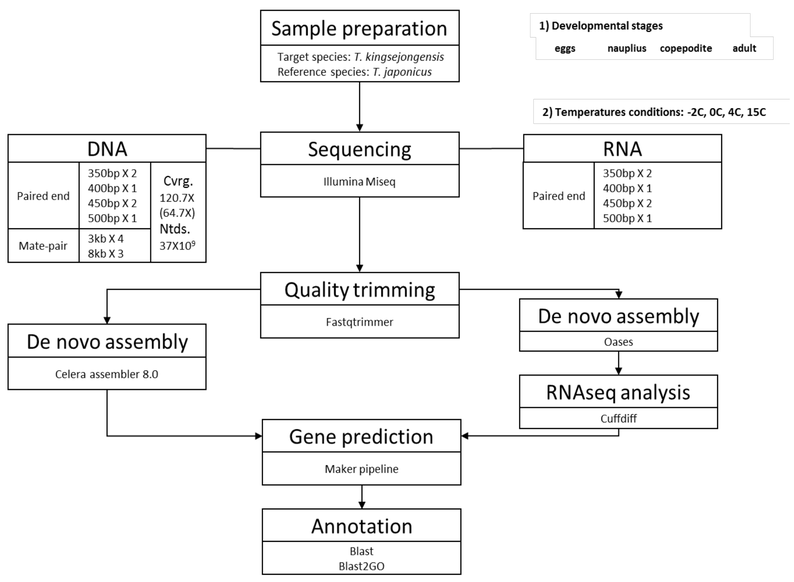
Antarctic marine invertebrates face extremely cold temperatures and many of decapod crustacean and fish groups became extinct because of extreme climate for over the last 30 million years. In that matter, species which have survived in Antarctic region may have evolutionary strategies and understanding their adaptation mechanisms in response to the extreme environment has received considerable attention. Here we present the first draft genome sequence and annotation for Antarctic copepoda *Tigriopus kingsejongensis*, the first Antarctic Crustacean to be sequenced using Illumina Miseq platform The final assembly consists of 48,368 contigs with an N50 contig length of 17.5 kilobases (kb) and 27,823 scaffolds with N50 contig length of 138.2 kb and a total of 39,717 coding genes were inferred using the MAKER annotation pipeline approach. The comparative genome analysis among 3,254 orthologs in 4 arthropod species (*T. kingsejongensis, Tigriopus japonicus, Daphnia pulex* and *Drosophila melanogaster*) revealed the *T. kingsejongensis* specific signals of molecular adaptation in genes associated with mitochondrial electron transport, deacetylase activity, proteasomal ubiquitin-dependent protein catabolic process, endoplasmic reticulum, and tryptophan metabolism. This suggest that *T. kingsejongensis* have changed adaptation mechanisms such as energy production and metabolism, proteolytic complex, and sterol biosynthesis. The results have important implications for understanding of Crustacean evolution and their adaptations to the Antarctic environment.

Introduction

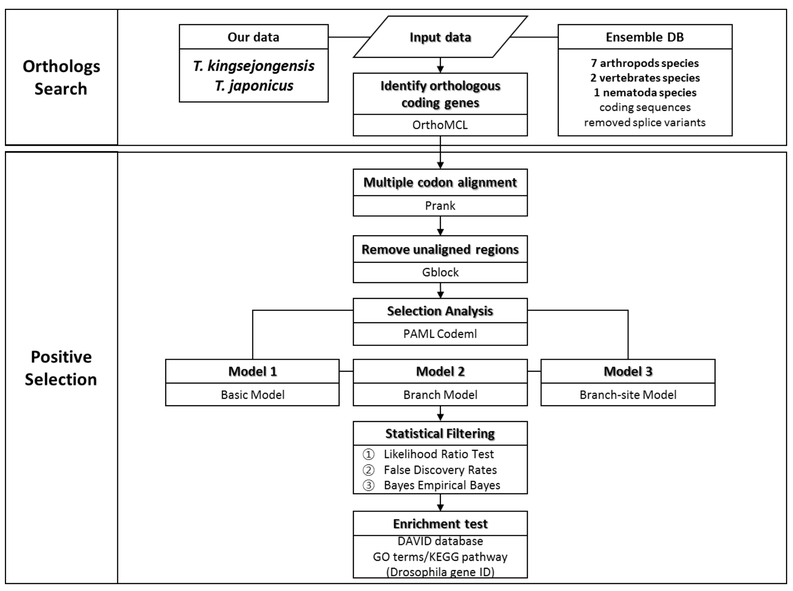


- Copepods
  - Highly diverged with long evolutionary history
  - Can survive temperatures below 0°C in polar regions and 45°C in hot springs
  - Survive under a wide range of environments are also advantageous in genetic comparative analysis for physical and biological studies
  - Potential good model organisms
- Antarctic *Tigriopus kingsejongensis*
- Temperate *Tigriopus japonicus*
- Report the first draft genome sequence of *T. kingsejongensis*
- Exhibit genome-wide signatures of adaptation to Antarctic environment

Genome Sequencing, Assembly and Annotation Pipeline



Gene Families and Positive Selection Analysis Workflow



Results

1. Comparative Genome Analyses

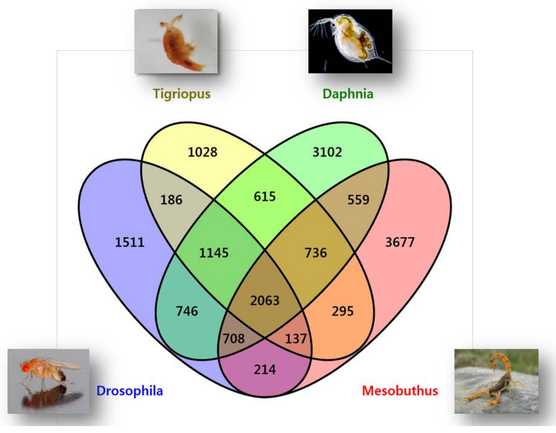


Figure 1. Venn diagram of the shared gene families. Comparison among the four arthropod species (copepod, mosquito, fruit fly and water flea) identified that 2,063 gene families were shared by the four species. *T. kingsejongenesis* shares 4,562 (73.5%) gene families with *Daphnia pulex* the most which belong to the same Crustacean lineage. The *T. kingsejongensis* specific gene families were 1,028 and significantly enriched in ATPase activity (8 genes; P < 0.01) and active transmembrane transporter activity (12 genes; P < 0.01).

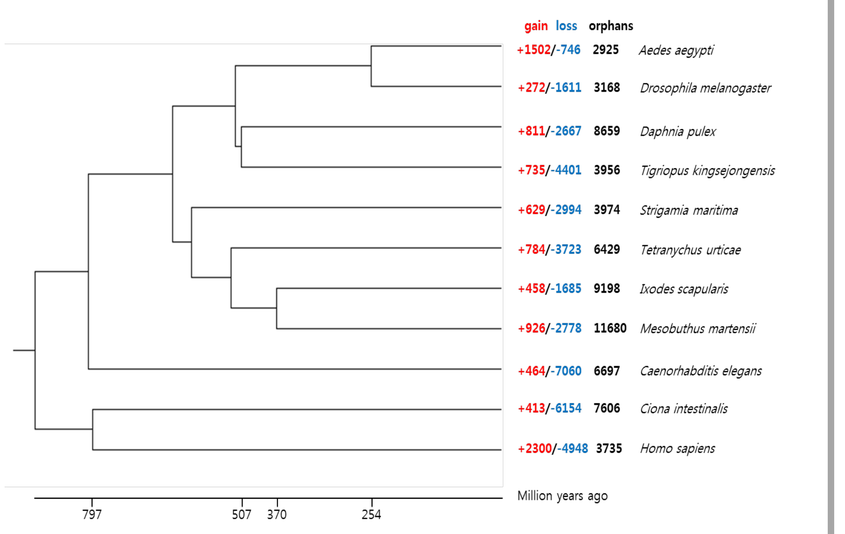


Figure 2. Gene gain and loss in the *T. kingsejongensis* genome. The gene gain-and-loss analysis were performed on 11 representative species and *T. kingsejongensis* gained 735 gene families and lost 4,401 gene families and showed the greatest gene family turnover of 5,136 which is the largest value among the eight arthropods.

Discussion

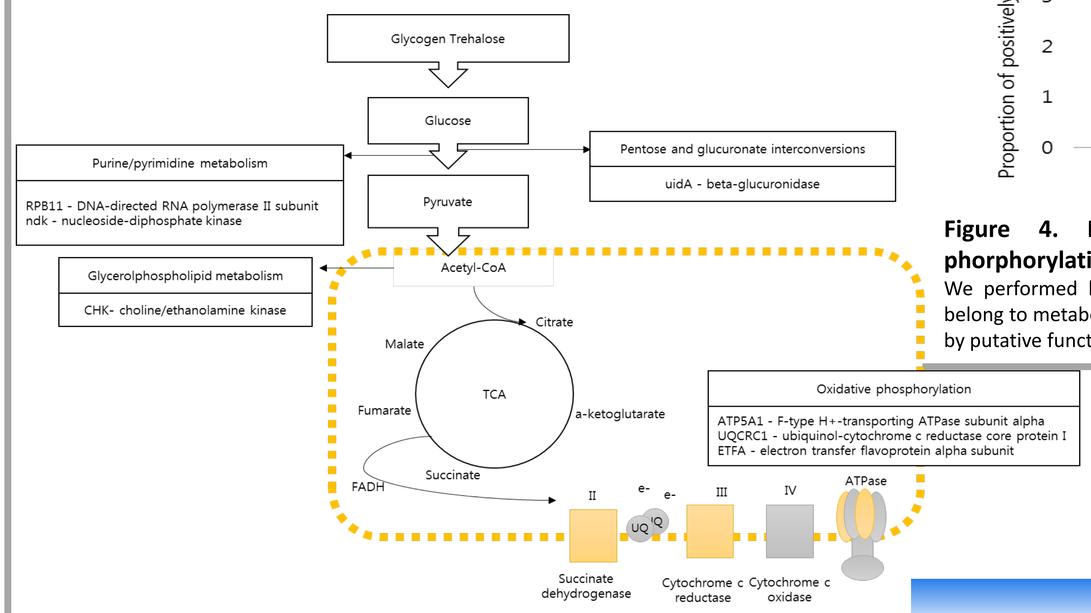


Figure 5. Total enzyme coding genes were PGS involved in metabolic pathways.

2. Adaptive Evolution

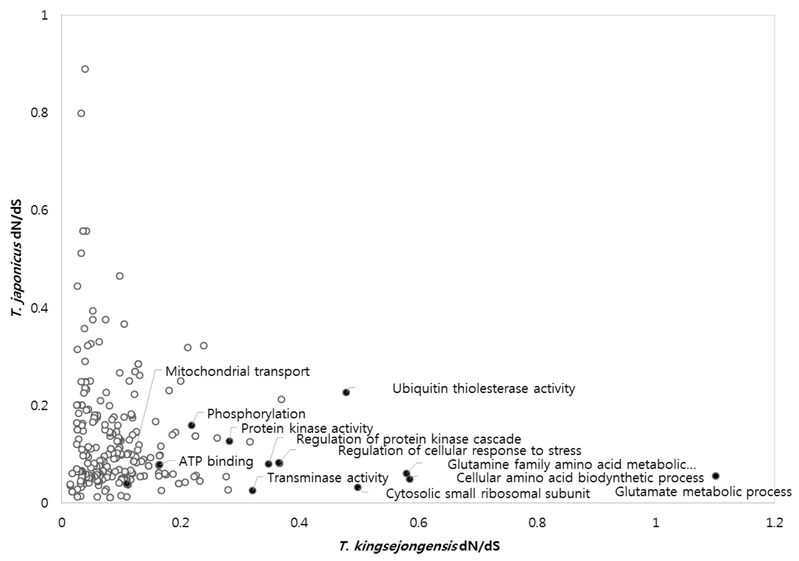


Figure 3. Global mean *dN/dS* distribution by GO categories of *T. kingsejongensis* and *T. japonicus*. The gene ontology (GO) categories showing accelerated evolution in *T. kingsejongensis* were 'regulation of cellular response to stress', carbohydrate metabolism, including 'glutamate metabolic process, 'glutamine family amino acid metabolic pathway' and 'cellular amino acid biosynthetic process' and energy metabolism such as 'ATP binding' and 'mitochondrial transport'.

3. Positively Selected Genes

Table 1. Enriched GO categories identified by PSGs from the *T. kingsejongensis*.

Category	GO IDs	GO descriptions	Genes	%
Cellular process	GO:0005811	lipid particle	7	9.9
Biological Process	GO:0006091	generation of precursor metabolites and energy	6	8.5
Biological Process	GO:0006091	generation of precursor metabolites and energy	6	8.5
Molecular Function	GO:0022890	inorganic cation transmembrane transporter activity	5	7.0
Cellular process	GO:0044455	mitochondrial membrane part	5	7.0
Biological Process	GO:0032268	regulation of cellular protein metabolic process	5	7.0
Biological Process	GO:0006119	oxidative phosphorylation	5	7.0
Molecular Function	GO:0051540	metal cluster binding	4	5.6
Molecular Function	GO:0051536	iron-sulfur cluster binding	4	5.6
Molecular Function	GO:0015078	hydrogen ion transmembrane transporter activity	4	5.6
Molecular Function	GO:0015077	monovalent inorganic cation transmembrane transporter activity	4	5.6
Molecular Function	GO:0008135	translation factor activity, nucleic acid binding	4	5.6
Molecular Function	GO:0004386	helicase activity	4	5.6
Biological Process	GO:0010608	posttranscriptional regulation of gene expression	4	5.6
Cellular process	GO:0005753	mitochondrial proton-transporting ATP synthase complex	3	4.2
Cellular process	GO:0045259	proton-transporting ATP synthase complex	3	4.2
Biological Process	GO:0015986	ATP synthesis coupled proton transport	3	4.2
Biological Process	GO:0015985	energy coupled proton transport, down electrochemical gradient	3	4.2
Biological Process	GO:0007277	pole cell development	3	4.2

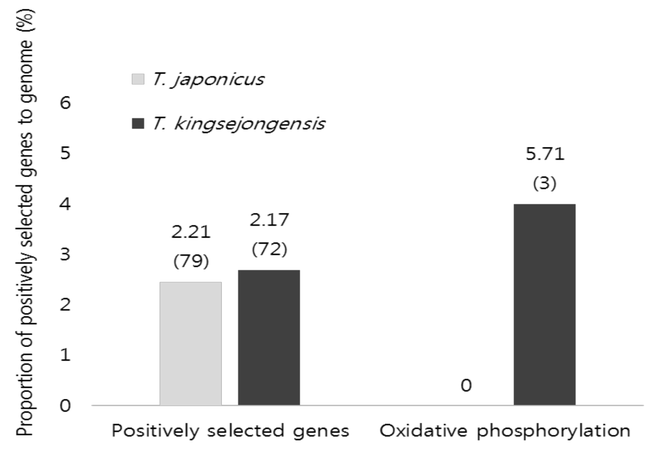


Figure 4. Proportion of PSGs to genome and oxidative phorphorylation pathway in the *T. kingsejongensis* and *T. japonicus*. We performed branch-site model analysis to test the hypothesis that the genes belong to metabolic pathways were significantly undergone positive selection process by putative functional divergence in the *T. kingsejongensis*.

Positively Selected Genes Involved in Metabolic Pathways

ATP synthase have a catalytic core composed of two catalytic domains alpha (ATP5A) and beta (ATP5B) subunits and these undergo a subsequent conformational changes and form ATP from ADP. Generally, mtDNA- and nuclear- encoded mitochondrial genes especially enzymes from the OxPhos pathway are highly conserved even between the distantly related species [1,2]. In that respect, mutations in mitochondrial genes are known to be lead to variety of negative effects including increased oxidative stress, reduction in body mass and survival, metabolic disorders [3]. However, polymorphism of ATP5A reported in ovenbirds is result in higher individual fitness leading by increased body mass [4] implying possible role in environmental adaptations. Additionally, UQCRC1 polymorphism in human is significantly associated with body lipid accumulation [5] and ETF polymorphism also discovered in human altered resistance to thermal stability of enzyme activity [6]. As *T. kingsejongenesis* face extreme conditions, especially constant low temperatures and seasonal food scarcity similar with other Antarctic invertebrates, efficient energy production, adequate energy sources to save and cold tolerance mechanisms is important for organisms living in high altitude. Additionally, we found four PSGs belong in metabolic pathways: nucleotide, lipid, and carbohydrate metabolic pathways.

References

1 Gray, M. W., Burger, G. & Lang, B. F. Mitochondrial evolution. *Science* **283**, 1476-1481 (1999).
2 Boore, J. L. Animal mitochondrial genomes. *Nucleic Acids Res.* **27**, 1767-1780 (1999).
3 Johannsen, D. L. & Ravussin, E. The role of mitochondria in health and disease. *Curr. Opin. Pharm.* **9**, 780-786 (2009).
4 Toms, J. D., Eggert, L. S., Arendt, W. J. & Faaborg, J. A genetic polymorphism in the sex-linked ATP5A1 gene is associated with individual fitness in Ovenbirds (Seiurus aurocapilla). *Ecology and evolution* **2**, 1312-1318 (2012).
5 Kunej, T. *et al.* Functional UQCRC1 polymorphisms affect promoter activity and body lipid accumulation. *Obesity* **15**, 2896-2901 (2007).
6 Henriques, B. J., Fisher, M. T., Bross, P. & Gomes, C. M. A polymorphic position in electron transfer flavoprotein modulates kinetic stability as evidenced by thermal stress. *FEBS Lett.* **585**, 505-510 (2011).

Genome of Antarctic-endemic Copepod and Evolutionary Adaptation to Extreme Environments

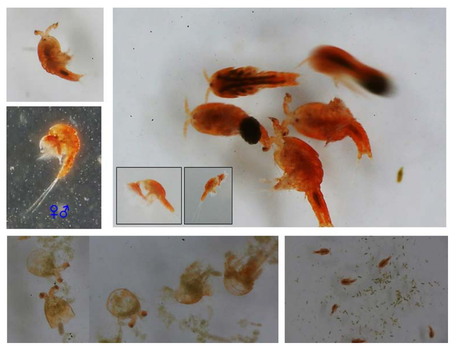
Seunghyun Kang[1], Do Hwan Ahn[2], Sanghee Kim[1]*, Hyun Park[1,2]*

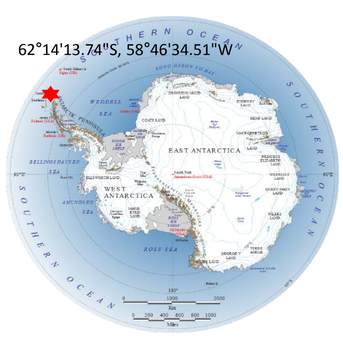
[1]Division of Polar Life Sciences, Korea Polar Research Institute, Yeonsu-gu, Incheon 406-840, South Korea,
[2]Polar Sciences, University of Science & Technology, Yuseong-gu, Daejeon 305-333, Korea
*Co-corresponding author: sangheekim@kopri.re.kr, hpark@kopri.re.kr

Abstract

Antarctic marine invertebrates face extremely cold temperatures and many of decapod crustacean and fish groups became extinct because of extreme climate for over the last 30 million years. In that matter, species which have survived in Antarctic region may have evolutionary strategies and understanding their adaptation mechanisms in response to the extreme environment has received considerable attention. Here we present the first draft genome sequence and annotation for Antarctic copepoda *Tigriopus kingsejongensis*, the first Antarctic Crustacean to be sequenced using Illumina Miseq platform The final assembly consists of 48,368 contigs with an N50 contig length of 17.5 kilobases (kb) and 27,823 scaffolds with N50 contig length of 138.2 kb and a total of 39,717 coding genes were inferred using the MAKER annotation pipeline approach. The comparative genome analysis among 3,254 orthologs in 4 arthropod species (*T. kingsejongensis, Tigriopus japonicus, Daphnia pulex* and *Drosophila melanogaster*) revealed the *T. kingsejongensis* specific signals of molecular adaptation in genes associated with mitochondrial electron transport, deacetylase activity, proteasomal ubiquitin-dependent protein catabolic process, endoplasmic reticulum, and tryptophan metabolism. This suggest that *T. kingsejongensis* have changed adaptation mechanisms such as energy production and metabolism, proteolytic complex, and sterol biosynthesis. The results have important implications for understanding of Crustacean evolution and their adaptations to the Antarctic environment.

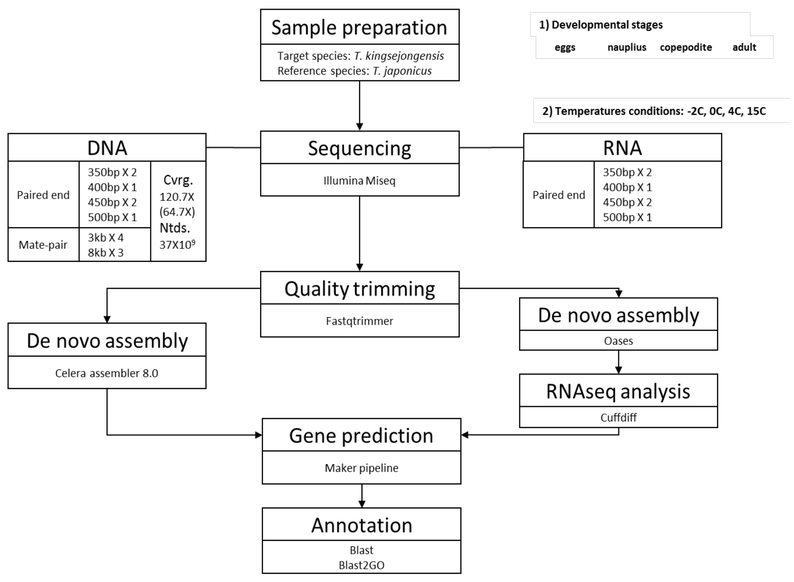
Introduction



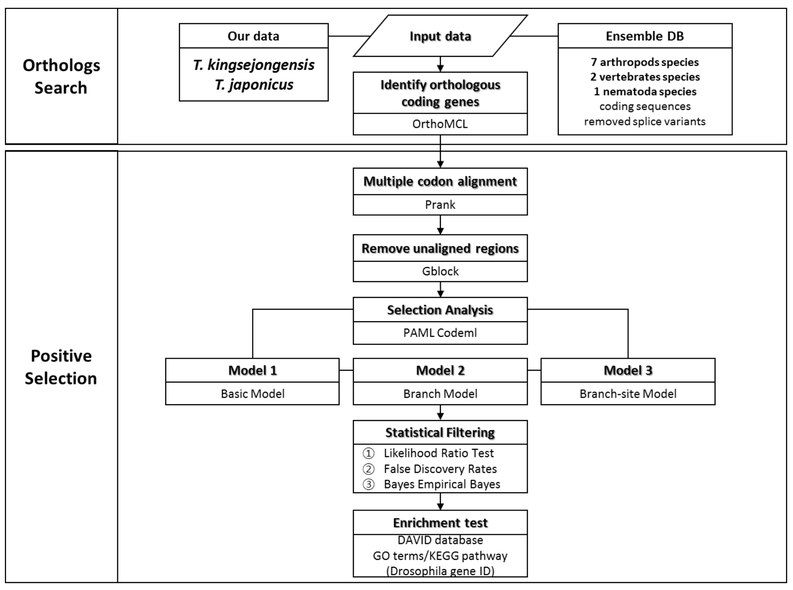


- Copepods
  - Highly diverged with long evolutionary history
  - Can survive temperatures below 0°C in polar regions and 45°C in hot springs
  - Survive under a wide range of environments are also advantageous in genetic comparative analysis for physical and biological studies
  - Potential good model organisms
- Antarctic *Tigriopus kingsejongensis*
- Temperate *Tigriopus japonicus*
- Report the first draft genome sequence of *T. kingsejongensis*
- Exhibit genome-wide signatures of adaptation to Antarctic environment

Genome Sequencing, Assembly and Annotation Pipeline



Gene Families and Positive Selection Analysis Workflow



Results

1. Comparative Genome Analyses

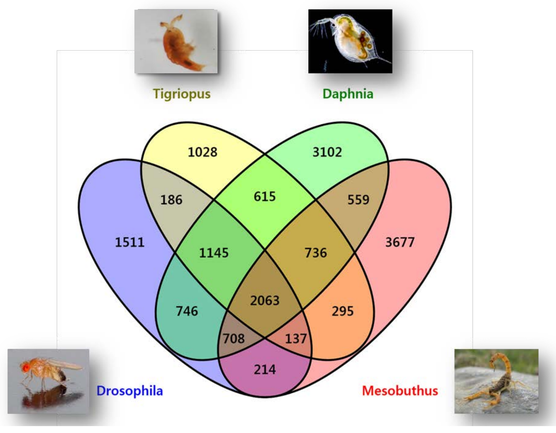


Figure 1. Venn diagram of the shared gene families. Comparison among the four arthropod species (copepod, mosquito, fruit fly and water flea) identified that 2,063 gene families were shared by the four species. *T. kingsejongenesis* shares 4,562 (73.5%) gene families with *Daphnia pulex* the most which belong to the same Crustacean lineage. The *T. kingsejongensis* specific gene families were 1,028 and significantly enriched in ATPase activity (8 genes; $P < 0.01$) and active transmembrane transporter activity (12 genes; $P < 0.01$).

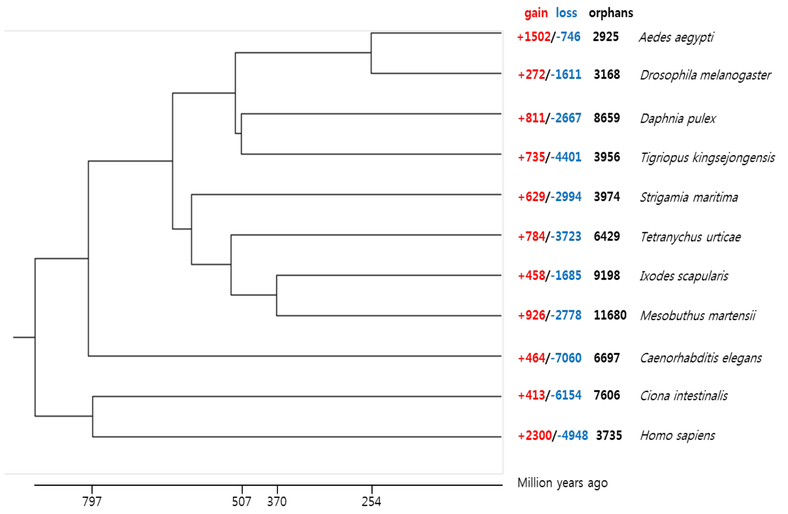


Figure 2. Gene gain and loss in the *T. kingsejongensis* genome. The gene gain-and-loss analysis were performed on 11 representative species and *T. kingsejongensis* gained 735 gene families and lost 4,401 gene families and showed the greatest gene family turnover of 5,136 which is the largest value among the eight arthropods.

Discussion

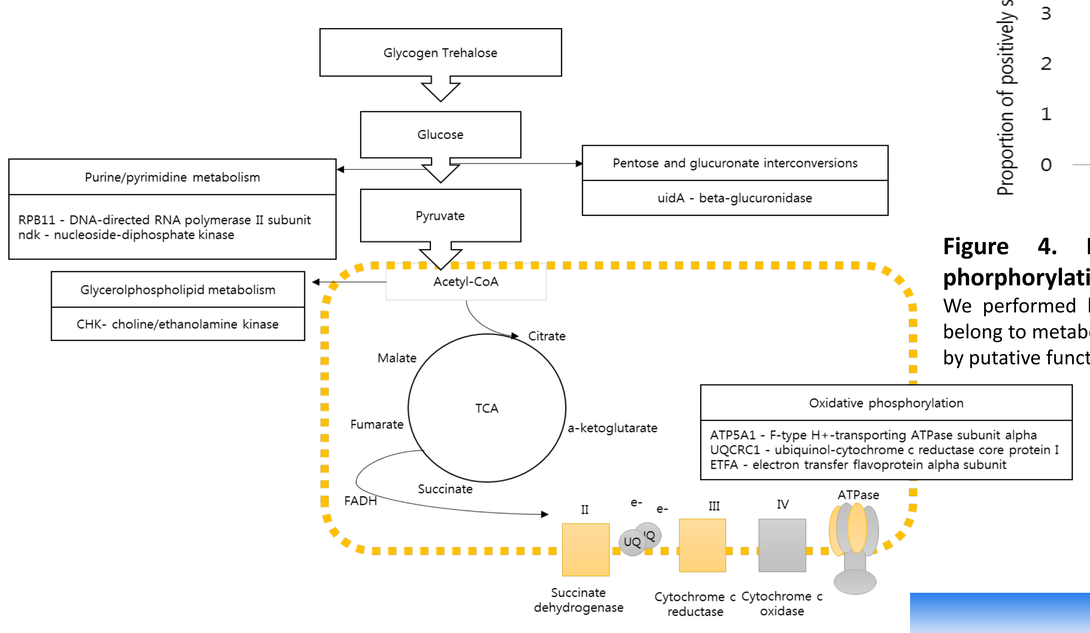


Figure 5. Total enzyme coding genes were PGS involved in metabolic pathways.

2. Adaptive Evolution

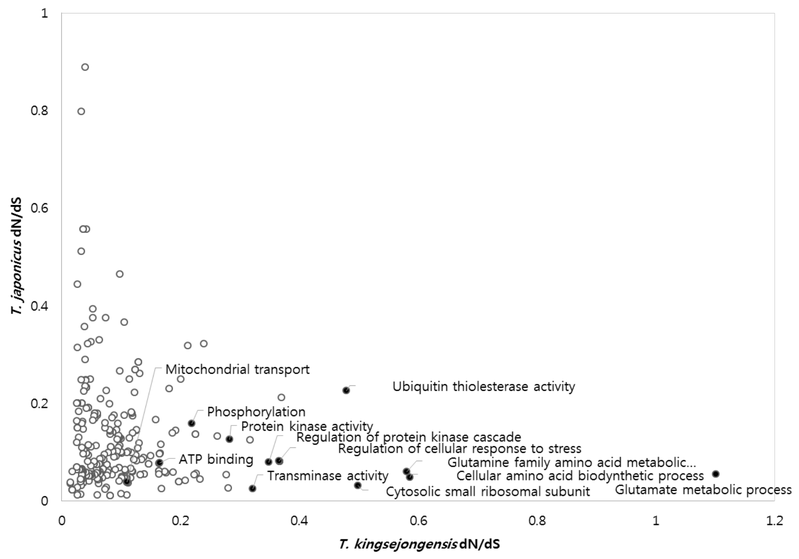


Figure 3. Global mean *dN/dS* distribution by GO categories of *T. kingsejongensis* and *T. japonicus*. The gene ontology (GO) categories showing accelerated evolution in *T. kingsejongensis* were 'regulation of cellular response to stress', carbohydrate metabolism, including 'glutamate metabolic process, 'glutamine family amino acid metabolic pathway' and 'cellular amino acid biosynthetic process' and energy metabolism such as 'ATP binding' and 'mitochondrial transport'.

3. Positively Selected Genes

Table 1. Enriched GO categories identified by PSGs from the *T. kingsejongensis*.

Category	GO IDs	GO descriptions	Genes	%
Cellular process	GO:0005811	lipid particle	7	9.9
Biological Process	GO:0006091	generation of precursor metabolites and energy	6	8.5
Biological Process	GO:0006091	generation of precursor metabolites and energy	6	8.5
Molecular Function	GO:0022890	inorganic cation transmembrane transporter activity	5	7.0
Cellular process	GO:0044455	mitochondrial membrane part	5	7.0
Biological Process	GO:0032268	regulation of cellular protein metabolic process	5	7.0
Biological Process	GO:0006119	oxidative phosphorylation	5	7.0
Molecular Function	GO:0051540	metal cluster binding	4	5.6
Molecular Function	GO:0051536	iron-sulfur cluster binding	4	5.6
Molecular Function	GO:0015078	hydrogen ion transmembrane transporter activity	4	5.6
Molecular Function	GO:0015077	monovalent inorganic cation transmembrane transporter activity	4	5.6
Molecular Function	GO:0008135	translation factor activity, nucleic acid binding	4	5.6
Molecular Function	GO:0004386	helicase activity	4	5.6
Biological Process	GO:0010608	posttranscriptional regulation of gene expression	4	5.6
Cellular process	GO:0005753	mitochondrial proton-transporting ATP synthase complex	3	4.2
Cellular process	GO:0045259	proton-transporting ATP synthase complex	3	4.2
Biological Process	GO:0015986	ATP synthesis coupled proton transport	3	4.2
Biological Process	GO:0015985	energy coupled proton transport, down electrochemical gradient	3	4.2
Biological Process	GO:0007277	pole cell development	3	4.2

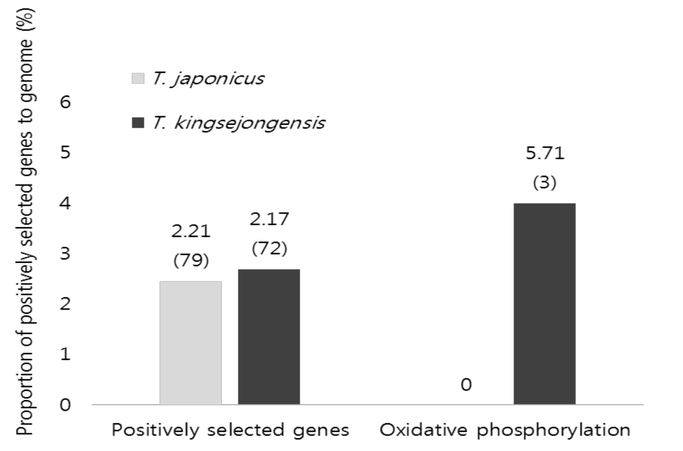


Figure 4. Proportion of PSGs to genome and oxidative phorphorylation pathway in the *T. kingsejongensis* and *T. japonicus*. We performed branch-site model analysis to test the hypothesis that the genes belong to metabolic pathways were significantly undergone positive selection process by putative functional divergence in the *T. kingsejongensis*.

Positively Selected Genes Involved in Metabolic Pathways

ATP synthase have a catalytic core composed of two catalytic domains alpha (ATP5A) and beta (ATP5B) subunits and these undergo a subsequent conformational changes and form ATP from ADP. Generally, mtDNA- and nuclear- encoded mitochondrial genes especially enzymes from the OxPhos pathway are highly conserved even between the distantly related species [1,2]. In that respect, mutations in mitochondrial genes are known to be lead to variety of negative effects including increased oxidative stress, reduction in body mass and survival, metabolic disorders [3]. However, polymorphism of ATP5A reported in ovenbirds is result in higher individual fitness leading by increased body mass [4] implying possible role in environmental adaptations. Additionally, UQCRC1 polymorphism in human is significantly associated with body lipid accumulation [5] and ETF polymorphism also discovered in human altered resistance to thermal stability of enzyme activity [6]. As *T. kingsejongenesis* face extreme conditions, especially constant low temperatures and seasonal food scarcity similar with other Antarctic invertebrates, efficient energy production, adequate energy sources to save and cold tolerance mechanisms is important for organisms living in high altitude. Additionally, we found four PSGs belong in metabolic pathways: nucleotide, lipid, and carbohydrate metabolic pathways.

References

1 Gray, M. W., Burger, G. & Lang, B. F. Mitochondrial evolution. *Science* **283**, 1476-1481 (1999).
2 Boore, J. L. Animal mitochondrial genomes. *Nucleic Acids Res.* **27**, 1767-1780 (1999).
3 Johannsen, D. L. & Ravussin, E. The role of mitochondria in health and disease. *Curr. Opin. Pharm.* **9**, 780-786 (2009).
4 Toms, J. D., Eggert, L. S., Arendt, W. J. & Faaborg, J. A genetic polymorphism in the sex-linked ATP5A1 gene is associated with individual fitness in Ovenbirds (Seiurus aurocapilla). *Ecology and evolution* **2**, 1312-1318 (2012).
5 Kunej, T. *et al.* Functional UQCRC1 polymorphisms affect promoter activity and body lipid accumulation. *Obesity* **15**, 2896-2901 (2007).
6 Henriques, B. J., Fisher, M. T., Bross, P. & Gomes, C. M. A polymorphic position in electron transfer flavoprotein modulates kinetic stability as evidenced by thermal stress. *FEBS Lett.* **585**, 505-510 (2011).