

Comparative genome analysis of *Subtercola boreus*, an actinobacterium retrieved from Antarctic rocks, soil and freshwater

Yong-Joon Cho*, Ahnna Cho, Ok-Sun Kim, and Jong-Ik Lee

Unit of Antarctic K-route Expedition, Korea Polar Research Institute, Incheon, Republic of Korea

Introduction

The extreme environment is a great challenge for all living things. Especially most of the earth is a low temperature environment where general life forms can not live. However, many psychrophilic and psychrotolerant microorganisms survive through various strategies against low temperature to cope with such an environment.

The actinobacterium *Subtercola boreus* is psychrophilic species capable of growth at temperatures down to -2 °C with an optimum at 15-17 °C. This species was first discovered in permanently cold groundwater under glacial ridge (depth 18 m) in southern Finland but it was also isolated from several places at the Antarctica including rocks, terrestrial soil as well as freshwater. Generally microorganisms are restricted to limited environmental conditions. But how can this psychrophilic microorganism survive in very distant and various environments?

Here, we report the fully assembled genome sequence of *S. boreus* type strain to find out how this psychrophile can adapt to extreme environments. And the genomic structure and gene content of the type and 5 strains of this psychrophilic bacteria isolated from various Antarctic environments were analyzed and compared.

Results

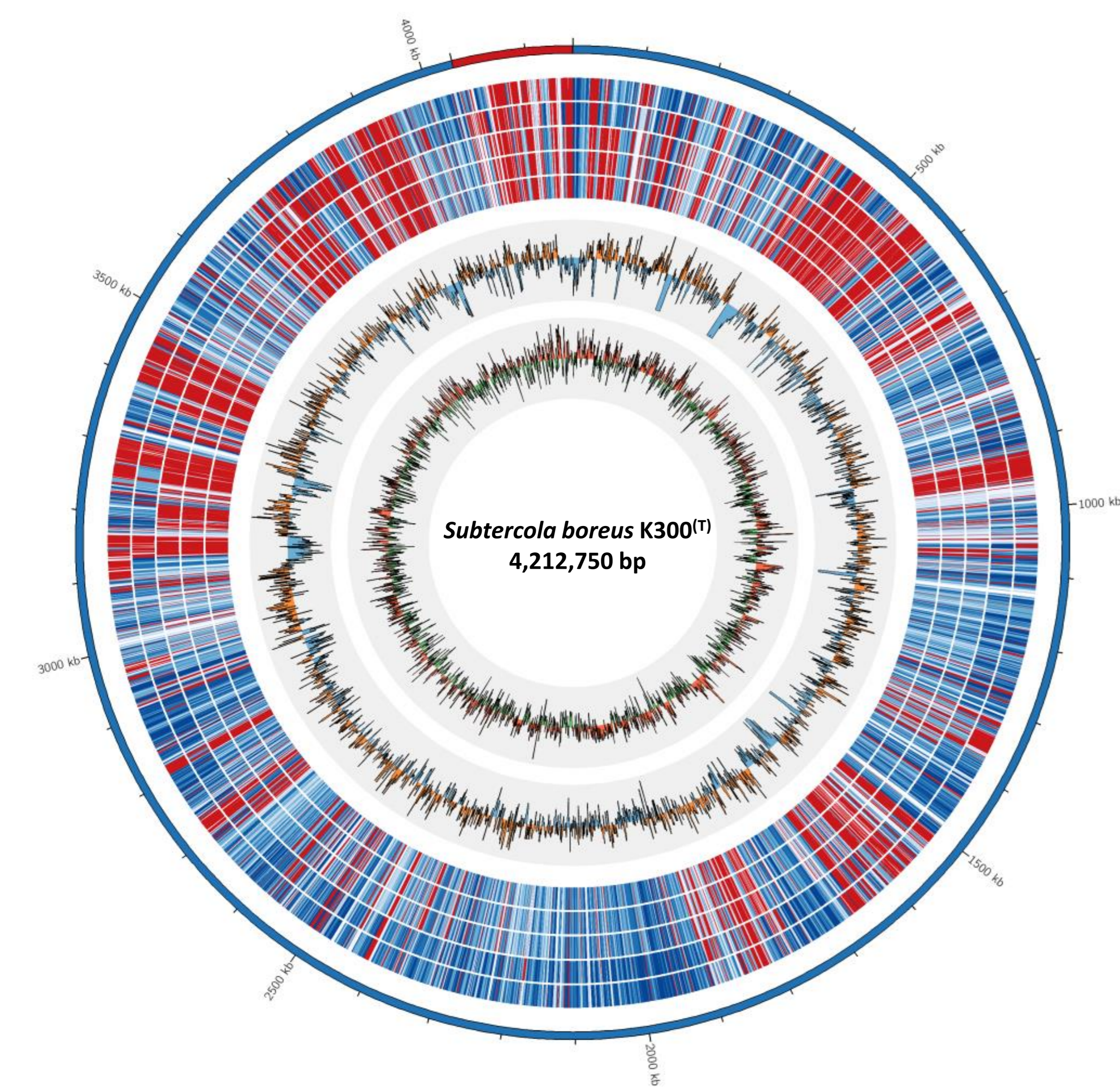


Figure 2. Comparative genome map of *S. boreus* K300^(T) and other 5 strains in various environments. Sequence similarity plots of PAMC 27444, PAMC 27479, PAMC 27924, PAMC 27979, PAMC 28004, and GC ratio, GC skew were depicted from outer circle.

<i>Subtercola boreus</i>	K300(T)	PAMC 27444	PAMC 27479	PAMC 27924	PAMC 27979	PAMC 28004
Source	Type strain	Water	Soil	Rocks 1	Rocks 2	Rocks 3
Genome status	Complete	Draft	Draft	Draft	Draft	Draft
Genome size (bp)	4,212,750	4,121,892	3,848,592	3,713,237	3,670,876	3,761,101
GC ratio (%)	68.0	67.1	66.9	67.1	67.2	66.9
Contigs	2	95	97	88	86	120
Total genes	3,898	3,915	3,716	3,463	3,424	3,521
Protein coding genes	3,843	3,861	3,662	3,412	3,373	3,458
Pseudo genes	128	152	208	196	188	206
RNA genes	54	53	53	50	50	62
Genes with function	2,193	2,123	1,969	1,944	1,928	1,972
Genes COG assigned	3,498	3,383	3,191	3,042	3,012	3,064

Table 1. Genome statistics of *S. boreus* type and other environmental strains.

Sampling site and Methods

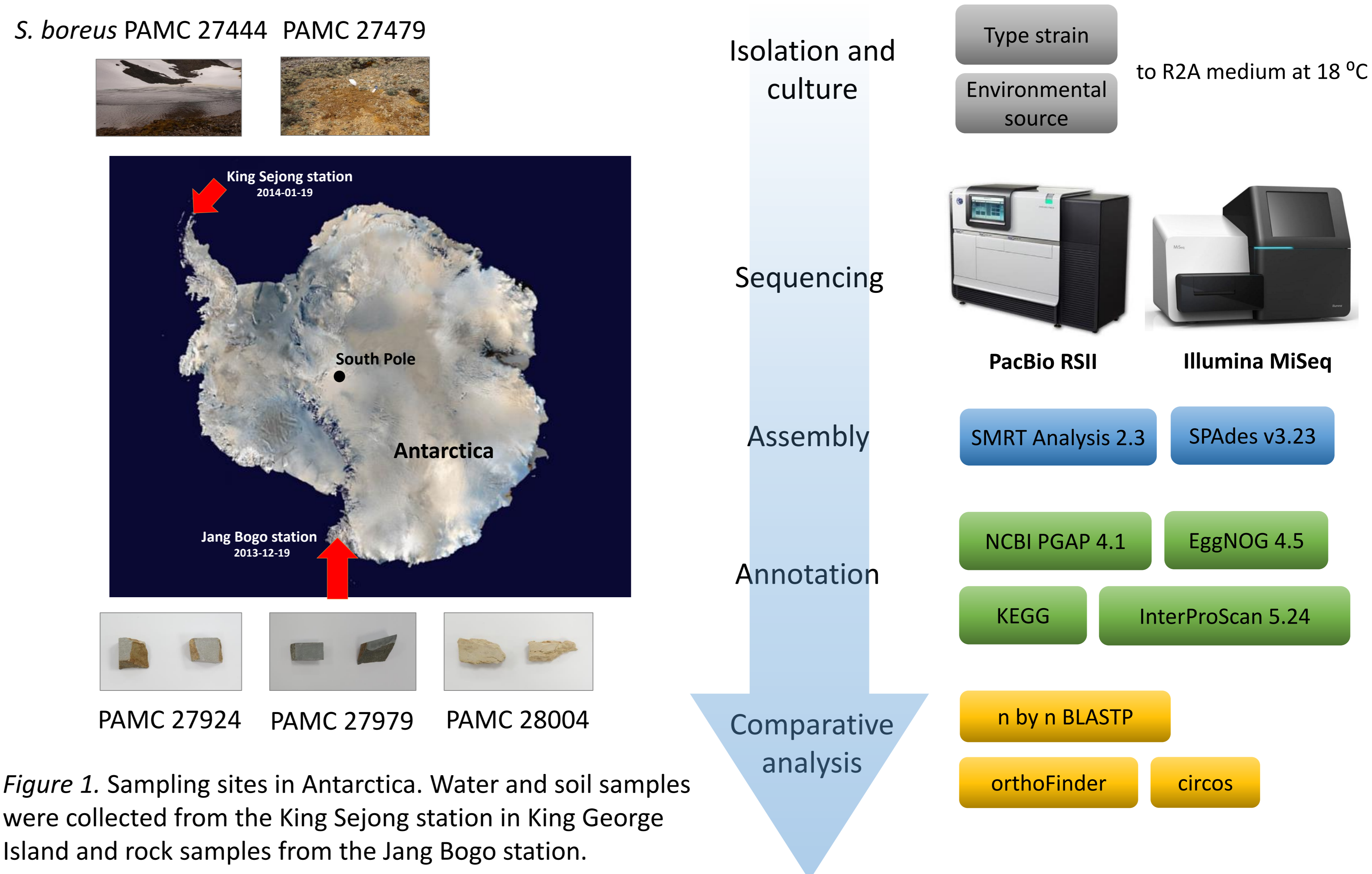


Figure 1. Sampling sites in Antarctica. Water and soil samples were collected from the King Sejong station in King George Island and rock samples from the Jang Bogo station.

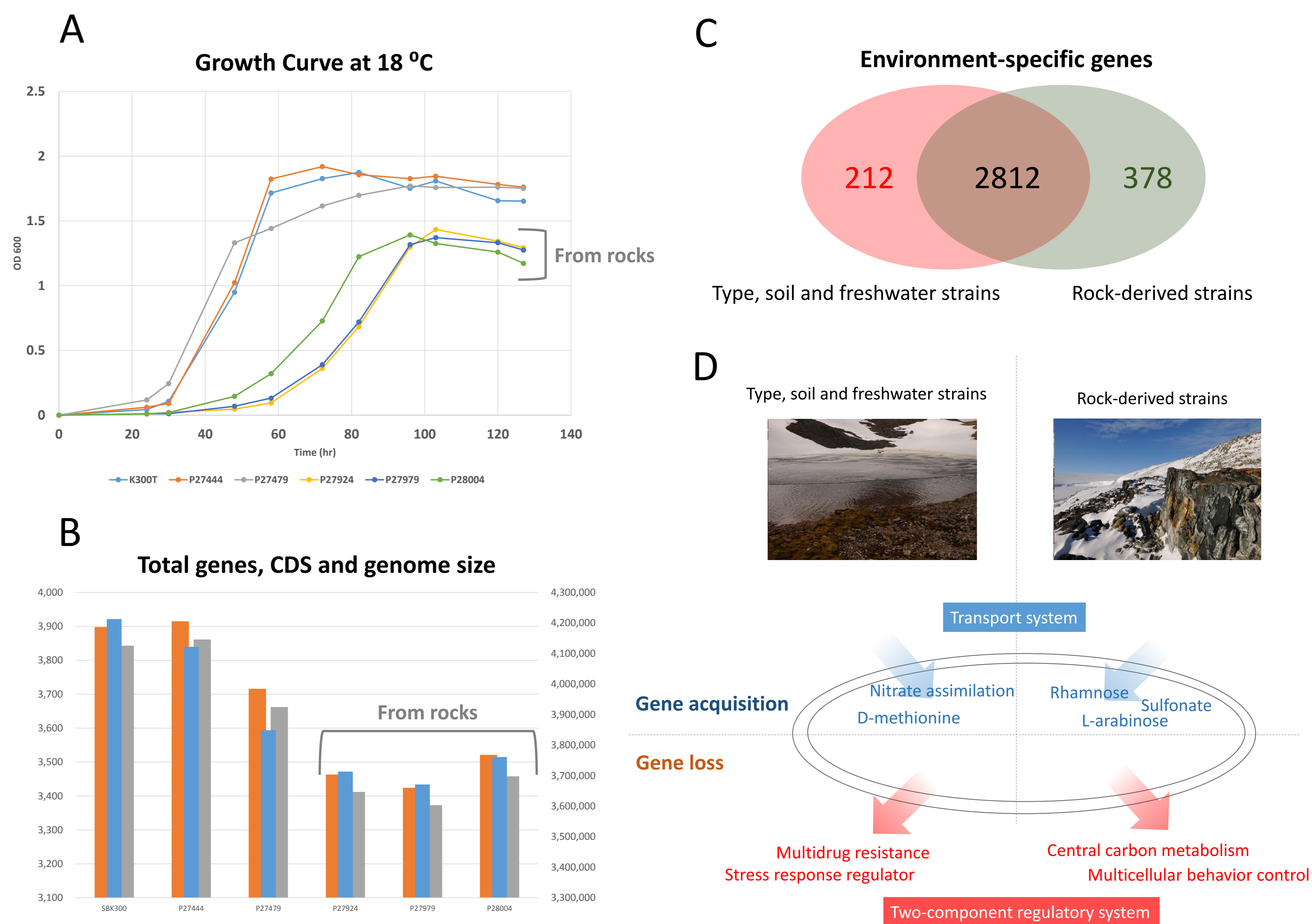


Figure 3. Comparative genome map of *S. boreus* K300^(T) and other 5 strains in various environments. Rock-derived strains clearly decreased their growth rate compared with other strains, and have large deletions in genome contents.

Conclusions

- Even within the same species, the content of the genome and growth rate is decreasing toward the extreme environment in order to deal with the limited condition.
- Various transport system related genes were acquired and many redundant functional genes were lost in rock-derived strains for adaptation to limiting nutrient in the environment.
- Genome contents show that many genes came from close biological source by phage and active horizontal gene transfer in spite of harsh environment.

References

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455-477.
- Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* 16:157.
- Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende DR, Sunagawa S, Kuhn M, Jensen LJ, von Mering C, Bork P. 2016. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res* 44:D286-293.

Acknowledgement 
This study was supported by the research project Pioneering inland route to the Antarctic inland and deep glacier ice bottom lake drilling (PE17110) of Korea Polar Research Institute (KOPRI)