



OPEN

## An Antarctic lichen isolate (*Cladonia borealis*) genome reveals potential adaptation to extreme environments

Minjoo Cho<sup>1</sup>, Seung Jae Lee<sup>1</sup>, Eunkyung Choi<sup>1</sup>, Jinmu Kim<sup>1</sup>, Soyun Choi<sup>1</sup>, Jun Hyuck Lee<sup>2,3</sup>✉ & Hyun Park<sup>1</sup>✉

*Cladonia borealis* is a lichen that inhabits Antarctica's harsh environment. We sequenced the whole genome of a *C. borealis* culture isolated from a specimen collected in Antarctica using long-read sequencing technology to identify specific genetic elements related to its potential environmental adaptation. The final genome assembly produced 48 scaffolds, the longest being 2.2 Mbp, a 1.6 Mbp N50 contig length, and a 36 Mbp total length. A total of 10,749 protein-coding genes were annotated, containing 33 biosynthetic gene clusters and 102 carbohydrate-active enzymes. A comparative genomics analysis was conducted on six *Cladonia* species, and the genome of *C. borealis* exhibited 45 expanded and 50 contracted gene families. We identified that *C. borealis* has more *Copia* transposable elements and expanded transporters (ABC transporters and magnesium transporters) compared to other *Cladonia* species. Our results suggest that these differences contribute to *C. borealis*' remarkable adaptability in the Antarctic environment. This study also provides a useful resource for the genomic analysis of lichens and genetic insights into the survival of species isolated from Antarctica.

Lichens are composite organisms that exist in a symbiotic association between lichen-forming fungi (LFF, mycobiont) and partner algae or cyanobacteria (photobiont)<sup>1,2</sup>. This fungal lifestyle constitutes approximately 20% of all known fungal species<sup>2</sup>. The photobiont provides sustenance to the mycobiont via the fixation of carbohydrates and atmospheric nitrogen while sheltering under the protection of the LFF<sup>2</sup>. The lichen symbiosis is an interesting paradigm for the study of environmental adaptation. Lichens exhibit the capacity to endure in extreme environments, such as the cryogenic, arid, and high solar radiation conditions present in Antarctica<sup>3</sup>. However, lichen's astonishing environmental adaptation abilities are still not fully understood, and numerous studies are being conducted to study lichen's ability to adapt to the environment.

Lichens are one of the dominant autotrophs on the Antarctic Peninsula<sup>4</sup>. Antarctica and South Georgia Island are home to a diverse range of lichens, accommodating approximately 427 taxa<sup>5</sup>. Within the Antarctic vegetation, *Cladonia borealis* stands out as one of the dominant organisms<sup>1</sup>. *Cladonia borealis*, which belongs to the class Lecanoromycetes in the Ascomycota division of the Fungi kingdom, is found mainly in the soil, humus, and mosses of polar, subpolar, and alpine areas<sup>6</sup>. Although limited data is available for the growth rate of lichens in continental Antarctica, there is a general consensus that their growth is comparatively slower than that of lichens in other global regions<sup>7</sup>. Research efforts of *C. borealis* can contribute to advancing our understanding of the ecological impacts of lichen on the Antarctic region.

The ecological plasticity and evolutionary mechanisms that enable lichens to persist in various environments may be elucidated through the scrutiny of biological data, including genomic analyses. Consequently, advanced sequencing technologies have the potential to aid in unraveling these mysteries. For example, studies have shown that viruses can play an important role in symbioses, increasing the association's adaptability to the environment<sup>8</sup>. Diverse genetic elements are analyzed through genomic assembly and annotation process. Genetic elements have important roles in the survival of organisms. For instance, transposable elements (TEs) constitute a major part of any genome and play important roles in gene expression, genomic rearrangements, epigenetic variation,

<sup>1</sup>Department of Biotechnology, College of Life Sciences and Biotechnology, Korea University, Seoul 02841, South Korea. <sup>2</sup>Research Unit of Cryogenic Novel Material, Korea Polar Research Institute, Incheon 21990, South Korea. <sup>3</sup>Department of Polar Sciences, University of Science and Technology, Incheon 21990, South Korea. ✉email: junhyucklee@kopri.re.kr; hpark@korea.ac.kr

mutations, and phenotypic variation<sup>9</sup>. Most TEs remain quiescent and are activated by environmental cues. Lichens are an attractive source to study TE divergence, but not enough studies have been done.

Lichens are also known as putative natural product producers. Using bioinformatic data, diverse analyses to reveal lichen's secondary metabolites have been conducted. These metabolites have proven distinctive and bioactive, and over 1000 such compounds have been identified lichen exclusive to date<sup>10</sup>. Some exhibit pharmacological activities, including antibacterial, antifungal, and anti-inflammatory effects<sup>11</sup>. The lichen secondary metabolites atranorin and physciosporin have been shown to possess inhibitory properties against lung cancer cell activity<sup>12,13</sup>. Furthermore, some natural pigments derived from fungi are known to contribute to environmental adaptation and abiotic stress tolerance<sup>14</sup>. These examples suggest that lichen can produce compounds that can advance pharmaceutical science and ecology. Through genetic studies, we can discover biosynthetic gene clusters (BGCs) with high possibilities of producing these compounds, such as polyketide synthase (PKS), nonribosomal peptide synthetase (NRPS), terpene, and hybrid PKS-NRPS types of natural products<sup>10</sup>.

Fungal carbohydrate-active enzymes (CAZymes) are enzymes which can catalyze glycans and glycoconjugates. Fungi produce an array of CAZymes for utilization of substrates. Fungal carbohydrate-active enzymes can be classified into six main classes, glycoside hydrolases (GHs), Carbohydrate esterases (CEs), polysaccharide lyases (PLs), glycosyl transferases (GTs), auxiliary activities (AAs), and carbohydrate binding modules (CBMs). For processing carbohydrates, fungi mostly rely on GHs, GTs, and AAs<sup>15</sup>. In addition to the catalytic modules, CBMs are common in enzymes active in cell-wall hydrolysis<sup>16</sup>. CAZyme families play important roles in lignocellulose breakdown, simple sugar metabolism, biofilm formation, antimicrobial biosynthesis, and diverse nutrient cycling routes<sup>17</sup>. LFFs absorb and utilize nutrients from phototrophic organisms and are thus classified as composite autotrophs. CAZymes represent an intriguing target for investigation in the genome of LFF, as mycobionts do not directly absorb nutrients from organic matter. Resl et al. have substantiated that CAZyme loss did not occur in LFF<sup>18</sup>. Their result revealed intricate patterns of retention and loss, which are not consistently aligned with the notion of CAZyme erosion following phototrophic symbiotic acquisition.

Transporter proteins are essential types of transmembrane proteins, their function is important in cell nutrition, communication, stress resistance, and homeostasis. One of the largest transport protein superfamilies in plants and fungi, ABC transporters, which are membrane components, are vital for cell energy regulation, a function vital for cell survival<sup>19</sup>. In lichen genomes, a wide range of transporter proteins is present, similar to other organisms. Notably, lichen forms a symbiotic relationship with partner algae, underscoring the increased significance of the transport system in these organisms.

In this study, we performed whole genome sequencing of *C. borealis* using PacBio single-molecule real-time (SMRT) sequencing, which has been widely used for de novo assembly<sup>20</sup>. A comparative analysis using genomes of the genus *Cladonia* was conducted. Secondary metabolite BGCs, CAZymes, and putative environmental adaptation related genes were also analyzed. This study provides a basic resource for the whole-genome sequencing of lichen and genetic insights into Antarctic lichen isolate.

## Results

### Genome assembly of *C. borealis*

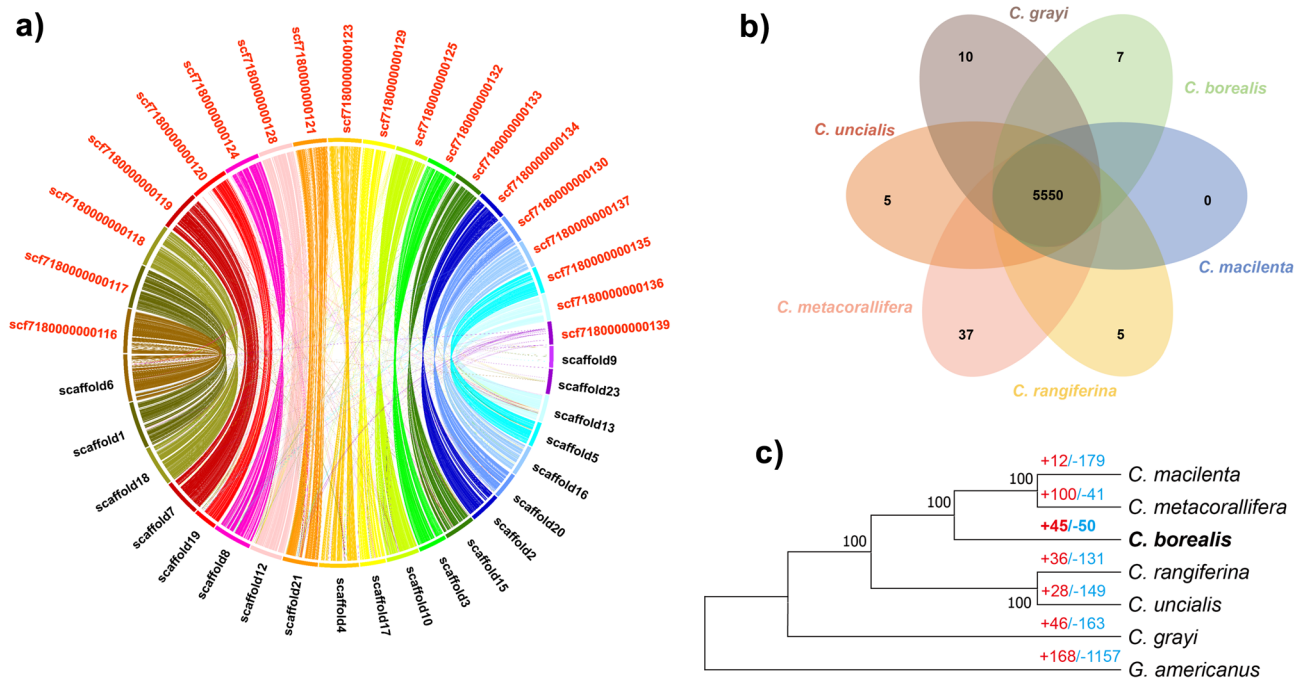
PacBio SMRT sequencing produced 635,943 long reads and 4.9 Gbp data. Reads over 10 Kbp accounted for more than 58% of the total reads (Table S1). Through de novo assembly, we produced 48 scaffolds, the longest being 2.2 Mbp, with an N50 scaffold length of 1.6 Mbp and a 36 Mbp total scaffolds length after the polishing process (Table 1). Assembly completeness was confirmed with Benchmarking Universal Single-Copy Orthologs (BUSCO), among 1706 Ascomycota orthologous genes, 1601 (93.9%) complete genes were identified. Of these complete BUSCOs, 1595 (93.5%) and six (0.4%) genes were single-copy and duplicated BUSCOs, respectively. As fragmented and missing BUSCOs, 18 (1.1%) and 87 (5.0%) genes were confirmed (Table 2). The genome synteny between *C. borealis* and *C. metacorallifera* showed a high level of conservation (Fig. 1a).

Assembly	
Number	48
Total size (bp)	36,013,978
Longest scaffolds (bp)	2,231,325
N50 scaffold length (bp)	1,657,173
Number of scaffolds > 1 Mb	19
G + C content (%)	45.1
Annotation database	Annotated no.
No. of genes	10,749
Pfam	3930 (36.6%)
SignalP	9703 (90.3%)
TmHMM	1966 (18.3%)
Swissprot blastx	6280 (58.4%)
Swissprot blastp	6437 (59.9%)

**Table 1.** Assembly and annotation information for the *C. borealis* genome.

Ascomycota_odb10	No.
Complete BUSCOs (C)	1601 (93.9%)
Complete and single-copy BUSCOs (S)	1595 (93.5%)
Complete and duplicated BUSCOs (D)	6 (0.4%)
Fragmented BUSCOs (F)	18 (1.1%)
Missing BUSCOs (M)	87 (5.0%)
Total BUSCO groups searched	1706

**Table 2.** Assembly completeness of the *C. borealis* genome, estimated using BUSCO v4.1.2.



**Figure 1.** Genome synteny and comparisons between six *Cladonia* species. (a) Circos plot showing conservation of synteny between *C. borealis* and *C. metacorallifera*. The synteny map represents 20 *C. metacorallifera* scaffolds (black) and 19 *C. borealis* scaffolds (red). (b) Venn diagram, created using the Orthovenn 2 web platform, showing the numbers of gene families shared among *Cladonia* species and unique to each species. (c) Maximum-likelihood hood tree, created using MegaX, based on the sequences of single-copy orthologous proteins. The numbers of gene families that have expanded (red, +) and contracted (blue, -), as inferred using CAFÉ, are represented.

### Genome annotation

A total of 11,677 repeat elements were identified, which accounted for 10.34% of the total *C. borealis* genome. Repeat sequences included long terminal repeats (LTRs, 5.91%), long interspersed nuclear elements (LINEs, 0.01%), and DNA transposons (0.43%) (Table S2).

Through EvidenceModeler annotation, 10,749 protein-coding genes were annotated. The total length of exons was 17 Mbp, with an average number of 3.11 exons per gene. For the functional annotation, 3930 (36.6%), 9703 (90.3%), 1966 (18.3%), 6280 (58.4%), and 6437 (59.9%) genes were annotated in Pfam, SignalP, TmHMM, Swissprot blastx, and Swissprot blastp, respectively (Table 1). The distribution of the top 20 gene ontology (GO) terms in level 5 is shown in Fig. S1. Among them, 1280 genes were annotated as intracellular membrane-bounded organelle (GO:0043231) related genes, making this the most abundant GO term in the *C. borealis* genome.

### Comparative genomics

The genome completeness of all *Cladonia* species (*C. macilenta*, *C. metacorallifera*, *C. rangiferina*, *C. uncialis*, and *C. grayi*) and *Gomphillus americanus* were identified to be above 90%, and the number of annotated genes ranged from 8200 to 10,825 (Table 3, Table S3). We classified repeat elements based on superfamily and found that the *C. borealis* genome contained the greatest total lengths of Kolobok-H and PIF-Harbinger types of DNA transposons and simple repeats. Also, regarding types of LTR transposable elements (TEs), *Copia*, ERVK, and Pao types of LTR TE were the most common types in the *C. borealis* genome (Table S2).

Organisms	Abbr.	Accession No.	Project no.	Sampling location
<i>Cladonia borealis</i>	Cbo	GCA_018257855.1	PRJNA693578	King George Island, Antarctica
<i>Cladonia grayi</i>	Cgr	Cgr/DA2myc/ss v2.0	PRJNA731936	North Carolina, USA
<i>Cladonia macilenta</i>	Cma	GCA_000444155.1	PRJNA210603	Mt. Cangshan, China
<i>Cladonia metacorallifera</i>	Cme	GCA_000482085.2	PRJNA219240	Mt. Seorak, South Korea
<i>Cladonia rangiferina</i>	Crg	GCA_006146055.1	PRJNA418168	Deciduous woodland, USA
<i>Cladonia uncialis</i>	Cuc	GCA_002927785.1	PRJNA348097	Manitoba, Canada
<i>Gomphillus americanus</i>	Gam	GCA_905337335.1	PRJEB42325	Granitic flatrock border woodland, USA

**Table 3.** Species and information for the genomes used in this study. *C. grayi* was downloaded from JGI and other species were obtained from NCBI.

### Prediction of BGCs and CAZymes

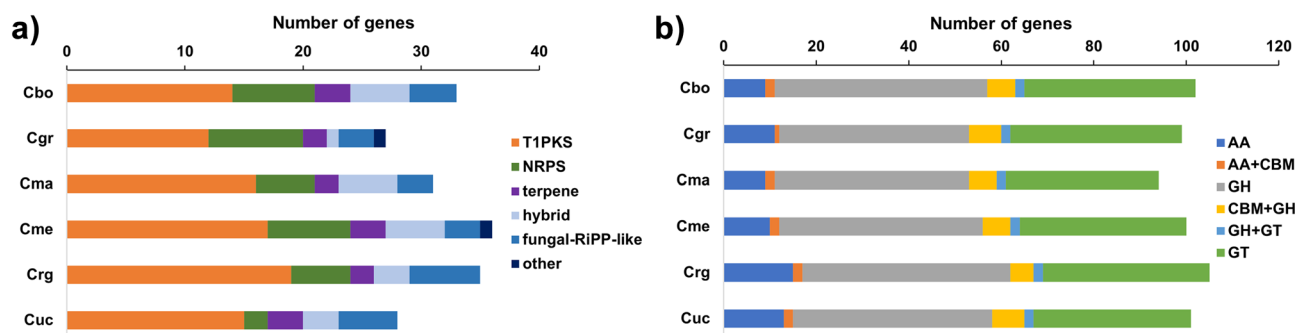
In total, 33, 27, 31, 36, 35, and 28 BGCs were identified in *C. borealis*, *C. grayi*, *C. macilenta*, *C. metacorallifera*, *C. rangiferina*, and *C. uncialis* genomes, respectively. Fourteen type I PKS (T1PKS), seven NRPS, three terpenes, five hybrid clusters, and four fungal ribosomally synthesized and posttranslationally modified peptide product (RiPP)-like clusters were involved in the secondary metabolism of *C. borealis* (Fig. 2a). Among the 33 gene clusters of *C. borealis*, 14 were matched with similar known clusters. Six predicted natural compounds were found only in *C. borealis* genome: CJ-16,173/CJ-15,696/citridone B/citridone A/citridone B', pseurotin/azaspirene, secalonic acids, xenoacremone A, 6-hydroxymellein, and wortmanamide A/wortmanamide B with similarity score 12% to 83%.

We identified 102, 99, 94, 100, 105, and 101 CAZymes in *C. borealis*, *C. grayi*, *C. macilenta*, *C. metacorallifera*, *C. rangiferina*, and *C. uncialis* genome, respectively. In *C. borealis*, CAZymes included 11 auxiliary activities (AA) distributed in two families, AA1 and AA3; 45 glycoside hydrolases (GH) in 22 families; and 81 glycosyl transferases (GT) in 19 families. Additionally, enzymes containing two domains were detected, two AA/carbohydrate binding modules (CBM), six GH/CBM, and two GH/GT (Fig. 2b, Table S6).

### Expanded genes in *C. borealis*

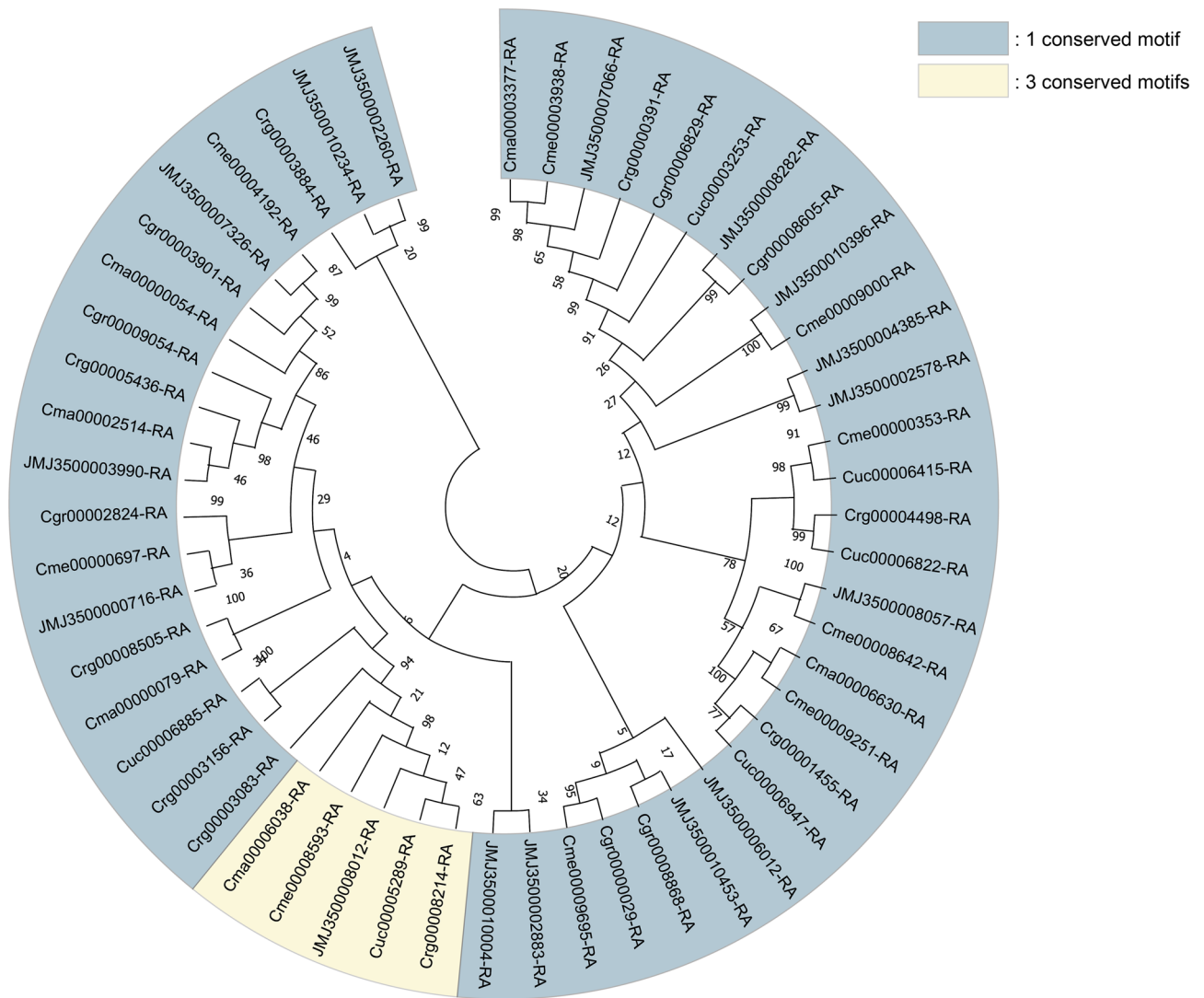
In the six *Cladonia* mycobionts, a total of 5550 gene families were shared and seven gene families were identified as *C. borealis* specific. Seven gene families were annotated with GO terms (GO:0055085, transmembrane transport; GO:0004674, protein serine/threonine kinase activity; GO:0016705, oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen; and 4 not represented in the GO database) (Fig. 1b).

The *C. borealis* genome showed expansion in 45 and contraction in 50 gene families (Fig. 1c). For the expanded gene families, Fisher's exact tests were used to identify enriched GO terms using Omicsbox. The 20 most significantly enriched GO terms of *C. borealis* are shown in Fig. S2. Transporter-related GO terms were the most enriched annotation in expanded gene families (Fig. S2). We manually annotated transporter-related proteins, and genes were classified according to their function. *Cladonia borealis* had the highest number of ATP-binding cassette (ABC) transporters and magnesium transporters (MGTs) among the studied species with 19 ABC transporter proteins and 16 MGTs (Table 4). The ABC transporters were located in 12 scaffolds, while MGTs were located in 10 (Tables S4, S5). Most of the ABC transporter proteins included a signature



**Figure 2.** Secondary metabolite synthase gene clusters and CAZyme genes found in the six *Cladonia* species genomes. (a) Secondary metabolite clusters, including type I polyketide synthases (T1PKS), non-ribosomal peptide synthetases (NRPS), terpenes, hybrid, fungal-RiPP-like, and other clusters, in the *Cladonia* genomes. (b) Number of CAZyme genes, including auxiliary activities (AA), glycoside hydrolases (GH), and glycosyl transferases (GT) and combined two-gene domains comprising GH and GT (GH+GT) or carbohydrate-binding modules (CBM) and AA (AA+CBM) or GH (CBM+GH), in the *Cladonia* genomes. See Table 3 for species abbreviations.





**Figure 4.** Phylogenetic tree depicting the relationships among the magnesium transporter proteins of the six *Cladonia* species. Except for five, all genes contained one conserved domain. These five exhibited three domains, with *C. grayi* being the only species not containing any genes with the three conserved motifs. Each protein identifier contains the species abbreviation (see Table 3); “JM” represents *C. borealis*.

potential adaptation behind this Antarctic native lichen. We performed comparative genomics analyses among different *Cladonia* species using a variety of bioinformatic tools to identify genes specific to *C. borealis*. The *Cladonia* species utilized in this study were selected for the purpose of comparing lichens isolated from Antarctic regions with those isolated from temperate climate environments.

Repeat elements represented 5.73 to 12.76% of the total genomes of the six *Cladonia* species, and about one-third of the repeat elements in each genome were uncharacterized. Like typical fungal genomes, LTR *Gypsy* transposons were more common than *Copia* elements in every species<sup>23</sup>. While repeat elements occupy a large amount in the genomes of animals and plants, these results indicate that repeat elements represent a relatively small proportion in lichen, but so many repeat elements are still unknown<sup>24</sup>. The *C. borealis* genome contained the most diverse set of repeat element types among the six species. We observed that the LTR *Copia* superfamily was expanded in the *C. borealis* genome. The activation of retrotransposons in response to stressors and extreme change is common in eukaryotes<sup>25–27</sup>. In grasses, it has been observed that *Copia* transposable elements are more frequently inserted near genes associated with stress response<sup>28</sup>. Similarly, in pitaya plants and *Dendrobium officinale*, *Copia* retrotransposons exhibit transcriptional activation, and among environmental stressors, transcriptional activation was highest in response to cold stress in particular<sup>29,30</sup>. Previous research, along with the findings of this study, suggests that transposable elements play a role in lichen’s remarkable environmental adaptability, especially in the harsh Antarctic environment.

In many organisms, including lichen-forming fungi, secondary metabolites have putative roles in responses to environmental pressures<sup>14</sup>. Also, some terpenoid compounds are derived from plant metabolites and are known to protect plants from abiotic stressors or have diverse functions in the interactions between plants and their environment<sup>31</sup>. But our results showed that *C. borealis*, *C. metacorrallifera*, and *C. uncialis* genome revealed

same number of terpene gene clusters. We have verified the presence of six distinct natural compounds within the *C. borealis* genome. However, an assessment of their potential influence on Antarctic adaptation remains pending. While our research has been conducted to explore the genetic mechanisms underlying adaptation in an Antarctic lichen, thus far, no specific BGCs have been identified that are specifically associated with survival in Antarctica. Uncovering the potential genetic pathways involved in adaptation to this unique environment will require further exploration. Also, we compared six *Cladonia* species to characterize CAZyme divergence in each genome. The *C. borealis* genome possesses the largest numbers of GHs and GTs and diverse CAZyme families, but the differences between the *Cladonia* species were insignificant.

In symbiotic partner relationships, transporter proteins play a key role in nutrient and signal exchanges. We found transmembrane transporter protein types that are expanded in *C. borealis* through Computational Analysis of Gene Family Evolution (CAFÉ), OrthoFinder, and Orthovenn 2 analyses, and with functional classification, we identified the specificity of ABC transporters and MGTs in the genome of *C. borealis*.

Previous studies have found that genomes of cold-tolerant strains of bacteria harbor increased numbers of transport proteins, including ABC transporters, and that ABC transporters play an important role in abiotic stress response including cold adaptation<sup>32–35</sup>. Fungal ABC transporters, in particular, have been shown to work as natural toxin or natural product efflux pumps<sup>36</sup>. Under extreme cold stress, such as is found in the Antarctic environment, organisms have difficulty in accessing nutrients and maintain cell homeostasis. Under such conditions, this class of transporters might perform key functions in maintaining the consistency of nutrient uptake and energy availability, important processes for lichen survival.

Our analysis has confirmed the presence of numerous MGTs in the *C. borealis* genome, which has directed our attention towards investigating the role of magnesium in the Antarctic environment. Lichens do not depend directly on the soil but rather utilize nutrients from the air<sup>37</sup>. On King George Island, Antarctica, insoluble airborne particulates in ice are composed of aluminum, potassium, magnesium, iron, and calcium. Furthermore, in Antarctic ice-rich sediments, extensive salt accumulation with major cations (Na<sup>+</sup>, Ca<sup>2+</sup>, K<sup>+</sup>, and Mg<sup>2+</sup>) occurs<sup>38</sup>. These findings suggest that the Antarctic atmosphere contains a significant amount of magnesium, and these aerosols may contribute to the accumulation of Mg<sup>2+</sup> in the soil. These environmental conditions expose living organisms to extreme osmotic and abiotic stress. Osmotic homeostasis is significant for cell survival: under osmotic stress, cells remain viable by importing compatible solutes via membrane transport proteins to regulate osmotic pressure and maintain turgor. Magnesium efflux and import systems maintain magnesium homeostasis, which is critical for organisms to adapt to high-magnesium environments<sup>39</sup>. Thus, the function of the magnesium transporter becomes crucial in maintaining cellular homeostasis during osmotic stress and response to abiotic stress<sup>40,41</sup>. *Cladonia borealis* can live in the Mg<sup>2+</sup>-rich Antarctic environment, which suggests a requirement for more MGTs. Also, previous study has revealed that MGTs are upregulated under cold stress in *Yersinia pseudotuberculosis*<sup>42</sup>. *Cladonia borealis*'s significantly higher number of MGTs among the *Cladonia* species might represent potential adaptation to the environment of Antarctica. Our results provide evidence of putative cold environmental adaptation related genes in Antarctic dominant *C. borealis* through a comparative analysis with other *Cladonia* species isolated from temperate climate.

## Conclusion

Here, we report the genome of *C. borealis*, based on PacBio RSII long-read DNA sequencing data. The genes specific to the Antarctic-derived lichens can be identified through comparison between Antarctic and temperate species isolates. In the genome of *C. borealis*, we have identified *Copia* transposable elements and an expanded set of ABC transporters and MGTs that potentially play helpful roles in enduring the extreme Antarctic environment. Transposable elements are known to be involved in significant genetic phenomena, and the expanded *Copia* TEs could potentially provide support for the survival of living organisms under extreme environmental stress. Additionally, the expanded ABC transporters and magnesium transporters may play roles in maintaining cellular homeostasis in *C. borealis* compared to species from temperate climates. This study should serve as a useful resource for the genomic analysis of lichens, contributing to the advancement of understanding and harnessing the genetic resources within this group of organisms. The genome sequencing data of *C. borealis* has revealed information on genes potentially related to environmental adaptation in the Antarctic and has contributed to advancing understanding of an Antarctic dominant lichen mycobiont.

## Methods

### Sample preparation and genome sequencing

The *Cladonia borealis* specimen was collected from King George Island, Antarctica (S 61° 59' 17", W 58° 1' 10"). We removed contaminants and isolated the *C. borealis* mycobiont. The isolated specimen was pure cultured in the laboratory for PacBio sequencing. The fungal isolate was grown under dark conditions at 15 °C on malt-yeast (MY) agar medium (15 g malt extract, 15 g/L agar). The lichen voucher material is stored in the 'Korean Lichen & Allied Bioresource Center' (<https://cc.aris.re.kr/kolabic/app/main/mainView.do>) with resource number 022262. Mycobionts' genomic DNA was extracted using a method previously described by Varela-Alvarez et al.<sup>43</sup>, with some modifications. The 400 mg of tissue from *C. borealis* was homogenized with liquid nitrogen in the mortar. Then, 8 mL of lysis buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 20 mM EDTA, 2% SDS, Proteinase K 20 mg/mL) was added and the sample was mixed. Next 2.4 mL of pre-heated CTAB buffer was added and the sample was heated at 65 °C for 1 h. One volume of phenol-chloroform-isoamyl alcohol (25:24:1) was added to the sample and mixed by inversion 10 times and then centrifuged for 10 min at 13,200 rpm. The aqueous phase was collected into a clean microcentrifuge tube and the rest was discarded. Two volumes of absolute ethanol were added and mixed gently. The sample was left for at least 30 min at – 20 °C and afterwards centrifuged for 10 min at 13,200 rpm. The supernatant was discarded, and the pellet was washed in 70% ethanol and dried at room temperature.

The pellet was dissolved in 200  $\mu$ L of TE (1X) buffer (1 mM Tris HCl pH 8.0, 0.1 mM EDTA pH 8.0). The eluted DNA was purified with PowerClean DNA Clean-up kit following manufacturer's instructions (MoBio Laboratories, USA). After isolation, the quality and the size of the DNA was checked by electrophoresis using a 0.6% agarose gel and a 1Kb DNA marker (Takara, Japan), and the DNA concentration was quantified with a Qubit 2.0 Fluorometer (Invitrogen, Merelbeke, Belgium). To confirm that the specimen was *C. borealis*, we checked the ITS sequence with ITS1F/ITS4 primer<sup>44</sup>. The *C. borealis* genome was sequenced by DNALink Inc. using the PacBio RS II platform (Pacific Biosciences, USA). The PacBio SMRT bell library was long-read sequenced in five SMRT cells (Pacific Biosciences) using C4 chemistry (DNA Sequencing Reagent 4.0), and 1  $\times$  240-min movies were captured for each SMRT cell. Illumina sequencing was used for correction during assembly: a paired-end (PE) library with an insert size of 350 bp was constructed in accordance with the manufacturer's protocol and sequenced using the Illumina HiSeq platform (Illumina, Inc., San Diego, CA, USA).

Total RNA was extracted from the samples using the mirVana™ miRNA Isolation Kit (Ambion) following the manufacturer's recommended procedures. The purity was assessed on a NanoDrop8000 spectrophotometer with 1  $\mu$ L of the total RNA extract. The integrity of the total RNA was evaluated using an Agilent Technologies 2100 Bioanalyzer, and the RNA Integrity Number (RIN) value was determined. The RNA-seq data was de novo assembled with Trinity (v2.15.0), and the Program to Assemble Spliced Alignments (PASA) annotation tool was used for gene predictions<sup>45</sup>.

### Genome assembly

De novo genome assembly was conducted using the PacBio long-read sequencing data with the FALCON-Unzip tool (v0.4)<sup>46</sup>. To enhance the assembly accuracy and correct errors in the draft *C. borealis* genome, Pilon (v1.22) was employed with the default parameters utilizing RNA-seq data generated from Illumina HiSeq<sup>47</sup>. To determine the completeness of the assembly, BUSCO (v4.1.2) was used with the Ascomycota\_odb10 dataset<sup>48</sup>.

### Genome annotation

RepeatModeler (v2.0.1, RepeatModeler), with default parameters, was utilized to reveal transposable elements<sup>49</sup>. We used LTR\_Retriever (v2.9.0) to identify long terminal repeat retrotransposons (LTR-RTs)<sup>50</sup>. The repeat elements were identified by constructing a de novo repeat library using RepeatMasker (v4.0.9)<sup>51</sup>.

Gene prediction was executed using PASA (v2.5.1)<sup>52</sup>. Followed by ab initio gene prediction using GeneMark (v4.2.9) and Augustus (v3.4.0) based on the repeat masked genome data<sup>53,54</sup>. Specifically, in GeneMark gene prediction, GeneMark-ES was utilized as the ab initio prediction in the first step. Through the ProtHint pipeline, hints were generated using fungi protein information from the ortho database, which we constructed for lichen annotation. Subsequently, protein equality alignment was performed in GeneMark-EP+<sup>55</sup>. Next, RNA-Seq data was used for PASA pipeline to provide evidence for the transcriptome data, which leads to more accurate genome prediction. All gene prediction data—including ab initio, transcriptome, and reference protein data—were integrated using EVIDENCEModeler (v1.1.1, EVM), which has been widely used for eukaryote gene prediction<sup>52</sup>. We reused PASA for the modification of the final gene models. Genome Annotation Generator (v2.0.1) was utilized to incorporate start and stop codons and generate a gff3 file for the genome annotation<sup>56</sup>.

### Functional annotation

Functional annotation was performed using BlastP (v2.2.29) based on a custom fungal database (E-value < 10<sup>-5</sup>)<sup>57</sup>. The custom fungal database was constructed by downloading the Fungi RefSeq database from the National Center for Biotechnology Information (NCBI). Protein domains were analyzed using InterProScan 5<sup>58</sup>. For the annotation of transcriptome assembly data, Trinotate (v3.2.0) was employed to provide comprehensive annotation information<sup>59</sup>. To decode the amino acid sequences, TransDecoder (v5.5.0) was utilized. Protein signal peptides were predicted using SignalP (v6.0)<sup>60</sup>. Additionally, protein family predictions were conducted using Pfam<sup>61</sup>. A GO analysis was conducted using Omicsbox software (v2.1.2)<sup>62</sup>.

### Comparative analysis within the genus *Cladonia*

For the comparative genomics analysis, we obtained the sampling locations and genomic sequences of five *Cladonia* species (*C. macilenta*<sup>63</sup>, *C. metacorallifera*<sup>64</sup>, *C. rangiferina*, *C. uncialis*, and *C. grayi*<sup>65</sup>) from NCBI and JGI which were isolated and sequenced from temperate climate (Table 3)<sup>66</sup>. The completeness of all genome assemblies was confirmed with BUSCO. To compare repeat element divergence between *Cladonia* species, we used LTR\_retriever and RepeatMasker. We annotated four species (*C. grayi*, *C. macilenta*, *C. rangiferina*, and *C. uncialis*) with MAKER (v2.28) and *C. metacorallifera* was annotated on the GenSAS (v5.1, GenSAS, <https://www.gensas.org>) online platform<sup>67</sup>. For the *C. borealis* and *C. metacorallifera* genome alignment, MUMmer 4 was used<sup>68</sup>. Genome synteny over 1 Mbp in length were visualized using Circos (v0.69-8)<sup>69,70</sup>.

To compare the six *Cladonia* species, OrthoFinder 2 was used with *Gomphillus americanus* protein sequences, downloaded from NCBI, as an outgroup<sup>71</sup>. *G. americanus* was used to ensure a meaningful comparison at the genus level, we chose species from different family that were registered in the annotation completed NCBI database. The orthologous gene families were gathered between the seven species based on their protein sequence similarities. Using single-copy ortholog data, evolutionary relationships were inferred by maximum-likelihood (ML) analysis using MEGA X (v10.2.4) with branch support assessed using 1000 bootstraps<sup>72</sup>.

### Prediction of BGCs and CAZymes

To explore secondary metabolite related BGCs, we searched our lichen genome sequences using antiSMASH fungal v7.0 with the default parameters<sup>73</sup>. CAZymes and CBM domains were annotated using run\_dbcan 3 software with the default parameters by combining the search results of four databases: the EC number, HMMER,



eCAMI, and DIAMOND databases<sup>74</sup>. To classify CAZyme families, we utilized the outcomes obtained from two or more databases.

### Expanded genes in *C. borealis*

Gene gain and loss were analyzed using CAFÉ 4.0 with gene family expansions or contractions considered significant at  $P < 0.05$ <sup>75</sup>. A Venn diagram was drawn using the Orthovenn 2 web server to visualize orthologous gene families among *Cladonia* species<sup>76</sup>. Based on the CAFÉ analysis result, *C. borealis*-specific genes were identified. We manually annotated expanded genes in *C. borealis* and the five other *Cladonia* species using four databases—including NCBI reference sequences, Uniprot reviewed sequences, GO terms, and Pfam—based on a BlastP E-value cut-off of  $1e-12$  and a homology of 50% or above<sup>77</sup>. Among transporter proteins, we sorted transmembrane transporters and classified genes by function. Annotated sequences were once again confirmed with the NCBI Conserved Domain Database (CDD) database. The evolutionary relationships between *Cladonia* species, based on the expanded or contracted protein families were analyzed through ML inference using MEGA X with branch support based on 1000 bootstrap sample trees<sup>72</sup>.

### Data availability

The dataset utilized in this research can be accessed through an online repository. The name of the repository and accession number are BioProject ID; PRJNA693578. This whole genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number JAFEKC000000000. The version described in this paper is version JAFEKC020000000.

Received: 15 September 2023; Accepted: 10 January 2024

Published online: 16 January 2024

### References

1. Cho, S. M., Lee, H., Hong, S. G. & Lee, J. Study of ecophysiological responses of the antarctic fruticose lichen *Cladonia borealis* using the PAM fluorescence system under natural and laboratory conditions. *Plants* **9**, 85. <https://doi.org/10.3390/plants9010085> (2020).
2. Song, H. *et al.* A comparative genomic analysis of lichen-forming fungi reveals new insights into fungal lifestyles. *Sci. Rep.* **12**, 10724. <https://doi.org/10.1038/s41598-022-14340-5> (2022).
3. Wang, Y., Zhang, X., Zhou, Q., Zhang, X. & Wei, J. Comparative transcriptome analysis of the lichen-forming fungus *Endocarpon pusillum* elucidates its drought adaptation mechanisms. *Sci. China Life Sci.* **58**, 89–100. <https://doi.org/10.1007/s11427-014-4760-9> (2015).
4. Colesie, C., Büdel, B., Hurry, V. & Green, T. G. A. Can Antarctic lichens acclimatize to changes in temperature? *Glob. Change Biol.* **24**, 1123–1135 (2018).
5. Øvstedal, D. O. & Smith, R. I. L. *Lichens of Antarctica and South Georgia: A Guide to Their Identification and Ecology* (Cambridge University Press, 2001).
6. Park, C. H., Jeong, G. & Hong, S. G. Possible multiple introductions of *Cladonia borealis* to King George Island. *Antarct. Sci.* **24**, 359–366. <https://doi.org/10.1017/s0954102012000223> (2012).
7. Sancho, L. G. & Pintado, A. Evidence of high annual growth rate for lichens in the maritime Antarctic. *Polar Biol.* **27**, 312–319. <https://doi.org/10.1007/s00300-004-0594-4> (2004).
8. Merges, D., Dal Grande, F., Greve, C., Otte, J. & Schmitt, I. Virus diversity in metagenomes of a lichen symbiosis (*Umbilicaria phaea*): Complete viral genomes, putative hosts and elevational distributions. *Environ. Microbiol.* **23**, 6637–6650 (2021).
9. Su, Y., Huang, Q., Wang, Z. & Wang, T. High genetic and epigenetic variation of transposable elements: Potential drivers to rapid adaptive evolution for the noxious invasive weed *Mikania micrantha*. *Ecol. Evol.* **11**, 13501–13517 (2021).
10. Calcott, M. J., Ackerley, D. F., Knight, A., Keyzers, R. A. & Owen, J. G. Secondary metabolism in the lichen symbiosis. *Chem. Soc. Rev.* **47**, 1730–1760 (2018).
11. Shishido, T. K. *et al.* Microbial communities of *Cladonia* lichens and their biosynthetic gene clusters potentially encoding natural products. *Microorganisms* **9**, 347. <https://doi.org/10.3390/microorganisms9071347> (2021).
12. Zhou, R. *et al.* The lichen secondary metabolite atranorin suppresses lung cancer cell motility and tumorigenesis. *Sci. Rep.* **7**, 8136. <https://doi.org/10.1038/s41598-017-08225-1> (2017).
13. Yang, Y. *et al.* Lichen secondary metabolite, physciosporin, inhibits lung cancer cell motility. *PLoS ONE* **10**, e0137889. <https://doi.org/10.1371/journal.pone.0137889> (2015).
14. Singh, G. *et al.* Climate-specific biosynthetic gene clusters in populations of a lichen-forming fungus. *Environ. Microbiol.* **23**, 4260–4275 (2021).
15. Baltar, F., Zhao, Z. & Herndl, G. J. Potential and expression of carbohydrate utilization by marine fungi in the global ocean. *Microbiome* **9**, 106. <https://doi.org/10.1186/s40168-021-01063-4> (2021).
16. Hervé, C. *et al.* Carbohydrate-binding modules promote the enzymatic deconstruction of intact plant cell walls by targeting and proximity effects. *Proc. Natl. Acad. Sci.* **107**, 15293–15298. <https://doi.org/10.1073/pnas.1005732107> (2010).
17. Barcoto, M. O. *et al.* Fungus-growing insects host a distinctive microbiota apparently adapted to the fungiculture environment. *Sci. Rep.* **10**, 12384 (2020).
18. Resl, P. *et al.* Large differences in carbohydrate degradation and transport potential among lichen fungal symbionts. *Nat. Commun.* **13**, 2634. <https://doi.org/10.1038/s41467-022-30218-6> (2022).
19. Ye, Z., Lu, Y. & Wu, T. The impact of ATP-binding cassette transporters on metabolic diseases. *Nutr. Metab.* **17**, 1–14 (2020).
20. Zhou, Z. *et al.* The sequence and *de novo* assembly of *Takifugu bimaculatus* genome using PacBio and Hi-C technologies. *Sci. Data* **6**, 187. <https://doi.org/10.1038/s41597-019-0195-2> (2019).
21. Fetsch, E. E. & Davidson, A. L. Vanadate-catalyzed photocleavage of the signature motif of an ATP-binding cassette (ABC) transporter. *Proc. Natl. Acad. Sci.* **99**, 9685–9690 (2002).
22. Hollenstien, K., Dawson, R. J. P. & Locher, K. P. Structure and mechanism of ABC transporter proteins. *Curr. Opin. Struct. Biol.* **17**, 412–418. <https://doi.org/10.1016/j.sbi.2007.07.003> (2007).
23. Daboussi, M. J. & Capy, P. Transposable elements in filamentous fungi. *Annu. Rev. Microbiol.* **57**, 275–299. <https://doi.org/10.1146/annurev.micro.57.030502.091029> (2003).
24. McKenzie, S. K., Walston, R. F. & Allen, J. L. Complete, high-quality genomes from long-read metagenomic sequencing of two wolf lichen thalli reveals enigmatic genome architecture. *Genomics* **112**, 3150–3156. <https://doi.org/10.1016/j.ygeno.2020.06.006> (2020).

25. Grandbastien, M.-A. Activation of plant retrotransposons under stress conditions. *Trends Plant Sci.* **3**, 181–187. [https://doi.org/10.1016/S1360-1385\(98\)01232-1](https://doi.org/10.1016/S1360-1385(98)01232-1) (1998).
26. Almer, J. *et al.* Symbiont-specific responses to environmental cues in a threesome lichen symbiosis. *Mol. Ecol.* **32**, 1045–1061 (2022).
27. Deneweth, J., Van de Peer, Y. & Vermeirssen, V. Nearby transposable elements impact plant stress gene regulatory networks: A meta-analysis in *A. thaliana* and *S. lycopersicum*. *BMC Genom.* **23**(18), 2022. <https://doi.org/10.1186/s12864-021-08215-8> (2022).
28. Ramachandran, D., McKain, M. R., Kellogg, E. A. & Hawkins, J. S. Evolutionary dynamics of transposable elements following a shared polyploidization event in the tribe andropogoneae. *G3 Genes Genomes Genet.* **10**, 4387–4398. <https://doi.org/10.1534/g3.120.401596> (2020).
29. Nie, Q., Qiao, G., Peng, L. & Wen, X. Transcriptional activation of long terminal repeat retrotransposon sequences in the genome of pitaya under abiotic stress. *Plant Physiol. Biochem.* **135**, 460–468. <https://doi.org/10.1016/j.plaphy.2018.11.014> (2019).
30. Gao, Y. Activation of Ty1-*copia* group retrotransposons of *Dendrobium officinale* under abiotic stress conditions. *Hered. Genet.* **5**, 2161 (2016).
31. Trapp, S. C. & Croteau, R. B. Genomic organization of plant terpene synthases and molecular evolutionary implications. *Genetics* **158**, 811–832. <https://doi.org/10.1093/genetics/158.2.811> (2001).
32. Buechel, E. R. & Pinkett, H. W. Transcription factors and ABC transporters: From pleiotropic drug resistance to cellular signaling in yeast. *FEBS Lett.* **594**, 3943–3964 (2020).
33. Williams, T. J. *et al.* Cold adaptation of the Antarctic haloarchaea *Halohasta litchfieldiae* and *Halorubrum lacusprofundi*. *Environ. Microbiol.* **19**, 2210–2227. <https://doi.org/10.1111/1462-2920.13705> (2017).
34. Mykytczuk, N. C. *et al.* Proteomic insights into cold adaptation of psychrotrophic and mesophilic *Acidithiobacillus ferrooxidans* strains. *Antonie Van Leeuw.* **100**, 259–277 (2011).
35. Wang, X.-Y., Yan, J. & Xie, J. Co-Culture of *Acinetobacter johnsonii* and *Shewanella putrefaciens* contributes to ABC transporter that impacts cold adaption in aquatic food storage environment (2023).
36. Batool, W. *et al.* AGC/AKT Protein kinase SCH9 is critical to pathogenic development and overwintering survival in *Magnaporthe oryzae*. *J. Fungi* **8**, 810 (2022).
37. Kuldeep, S. & Prodyut, B. Lichen as a bio-indicator tool for assessment of climate and air pollution vulnerability. *Int. Res. J. Environ. Sci.* **4**, 107–117 (2015).
38. Cuozzo, N. *et al.* Silicate weathering in antarctic ice-rich permafrost: Insights using magnesium isotopes. *Geochim. Cosmochim. Acta* **278**, 244–260. <https://doi.org/10.1016/j.gca.2019.07.031> (2020).
39. Tang, Y. *et al.* Uncovering the role of wheat magnesium transporter family genes in abiotic responses. *Front. Plant Sci.* **14**, 299 (2023).
40. Wendel, B. M. *et al.* A central role for magnesium homeostasis during adaptation to osmotic stress. *MBio* **13**, e00092 (2022).
41. Heidari, P., Abdullah, F. S. & Poczai, P. Magnesium transporter gene family: Genome-wide identification and characterization in *Theobroma cacao*, *Corchorus capsularis*, and *Gossypium hirsutum* of family *Malvaceae*. *Agronomy* **11**, 1651 (2021).
42. Virtanen, J.-P., Keto-Timonen, R., Jaakkola, K., Salin, N. & Korkeala, H. Changes in transcriptome of *Yersinia pseudotuberculosis* IP32953 grown at 3 and 28 C detected by RNA sequencing shed light on cold adaptation. *Front. Cell. Infect. Microbiol.* **8**, 416 (2018).
43. Varela-Álvarez, E. *et al.* Genomic DNA isolation from green and brown algae (*Caulerpa* and *Fucales*) for microsatellite library construction 1. *J. Phycol.* **42**, 741–745 (2006).
44. White, T. J., Bruns, T., Lee, S. & Taylor, J. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR Protoc.* **18**, 315–322 (1990).
45. Haas, B. J. *et al.* De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512 (2013).
46. Kingan, S. B. *et al.* A high-quality de novo genome assembly from a single mosquito using PacBio sequencing. *Genes* **10**, 62. <https://doi.org/10.3390/genes10010062> (2019).
47. Walker, B. J. *et al.* Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* **9**, e112963 (2014).
48. Simao, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212. <https://doi.org/10.1093/bioinformatics/btv351> (2015).
49. Flynn, J. M. *et al.* RepeatModeler2 for automated genomic discovery of transposable element families. *Proc. Natl. Acad. Sci.* **117**, 9451–9457 (2020).
50. Ou, S. & Jiang, N. LTR\_retriever: A highly accurate and sensitive program for identification of long terminal repeat retrotransposons. *Plant Physiol.* **176**, 1410–1422. <https://doi.org/10.1104/pp.17.01310> (2018).
51. Chen, N. Using repeat masker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinform.* **5**, 11–14 (2004).
52. Haas, B. J. *et al.* Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced alignments. *Genome Biol.* **9**, R7. <https://doi.org/10.1186/gb-2008-9-1-r7> (2008).
53. Lukashin, A. V. & Borodovsky, M. GeneMark. hmm: New solutions for gene finding. *Nucleic Acids Res.* **26**, 1107–1115 (1998).
54. Stanke, M. *et al.* AUGUSTUS: Ab initio prediction of alternative transcripts. *Nucleic Acids Res.* **34**, W435–W439 (2006).
55. Brůna, T., Lomsadze, A. & Borodovsky, M. GeneMark-EP+: Eukaryotic gene prediction with self-training in the space of genes and proteins. *NAR Genom. Bioinform.* **2**, 026 (2020).
56. Geib, S. M. *et al.* Genome annotation generator: A simple tool for generating and correcting WGS annotation tables for NCBI submission. *Gigascience* **7**, 018 (2018).
57. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
58. Paysan-Lafosse, T. *et al.* InterPro in 2022. *Nucleic Acids Res.* **51**, D418–D427 (2023).
59. Bryant, D. M. *et al.* A tissue-mapped axolotl de novo transcriptome enables identification of limb regeneration factors. *Cell Rep.* **18**, 762–776 (2017).
60. Teufel, F. *et al.* SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat. Biotechnol.* **40**, 1023–1025 (2022).
61. Sonnhammer, E. L., Eddy, S. R. & Durbin, R. Pfam: A comprehensive database of protein domain families based on seed alignments. *Proteins Struct. Funct. Bioinform.* **28**, 405–420 (1997).
62. Conesa, A. *et al.* Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674–3676 (2005).
63. Park, S.-Y. *et al.* Draft genome sequence of *Cladonia macilenta* KoLRI003786, a lichen-forming fungus producing biruloquinone. *Genome Announc.* **1**, 613. <https://doi.org/10.1128/genomea.00695-00613> (2013).
64. Park, S.-Y. *et al.* Draft genome sequence of lichen-forming fungus *Cladonia metacoralifera* strain KoLRI002260. *Genome Announc.* **2**, 13. <https://doi.org/10.1128/genomea.01065-01013> (2014).
65. Armaleo, D. *et al.* The lichen symbiosis re-viewed through the genomes of *Cladonia grayi* and its algal partner *Asterochloris glomerata*. *BMC Genom.* **20**, 1–33 (2019).
66. Armaleo, D. & May, S. Sizing the fungal and algal genomes of the lichen *Cladonia grayi* through quantitative PCR. *Symbiosis* **49**, 43–51. <https://doi.org/10.1007/s13199-009-0012-3> (2009).

67. Humann, J. L., Lee, T., Ficklin, S. & Main, D. Structural and functional annotation of eukaryotic genomes with GenSAS. In *Gene Prediction: Methods and Protocols* (ed. Kollmar, M.) 29–51 (Springer, 2019).
68. Kurtz, S. *et al.* Versatile and open software for comparing large genomes. *Genome Biol.* **5**, 1–9 (2004).
69. Krzywinski, M. *et al.* Circos: An information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009).
70. Cantarel, B. L. *et al.* MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* **18**, 188–196 (2008).
71. Emms, D. M. & Kelly, S. OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 1–14 (2019).
72. Tamura, K., Stecher, G., Peterson, D., Filipiński, A. & Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729 (2013).
73. Blin, K. *et al.* antiSMASH 6.0: Improving cluster detection and comparison capabilities. *Nucleic Acids Res.* **49**, W29–W35. <https://doi.org/10.1093/nar/gkab335> (2021).
74. Yin, Y. *et al.* dbCAN: A web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **40**, W445–W451 (2012).
75. Kim, B.-M. *et al.* Antarctic blackfin icefish genome reveals adaptations to extreme environments. *Nat. Ecol. Evol.* **3**, 469–478. <https://doi.org/10.1038/s41559-019-0812-7> (2019).
76. Xu, L. *et al.* OrthoVenn2: A web server for whole-genome comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res.* **47**, W52–W58 (2019).
77. Klosterman, S. J. *et al.* Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens. *PLoS Pathog.* **7**, e1002137 (2011).

### Author contributions

H.P. and J.H.L. conceived the idea of the study. H.P. supervised the study. J.H.L. collected the samples. S.J.L. performed the genome sequencing and data curation. E.C., J.K. and S.C. performed genome assembly. M.C. analyzed after genome sequencing and wrote the manuscript. All authors reviewed the manuscript.

### Funding

This research was a part of the project titled “Development of potential antibiotic compounds using polar organism resources (KIMST 20200610, KOPRI Grant Number PM23030)”, funded by Ministry of Oceans and Fisheries, Korea and a grant from Korea University.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-51895-x>.

**Correspondence** and requests for materials should be addressed to J.H.L. or H.P.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024