Ice-binding proteins from the fungus *Antarctomyces psychrotrophicus* possibly originate from two different bacteria through horizontal gene transfer

Tatsuya Arai[1], Daichi Fukami[1], Tamotsu Hoshino[2], Hidemasa Kondo[1,2], and Sakae Tsuda[1–3 §]

[1] Graduate School of Life Science, Hokkaido University, Sapporo, 060-0810 Japan

[2] Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Sapporo, 062-8517 Japan

[3] OPERANDO Open Innovation Laboratory, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, 305-8563 Japan

[§] To whom correspondence should be addressed. E-mail: s.tsuda@aist.go.jp.

Key words: freeze resistance | ice-binding protein | ascomycete | horizontal gene transfer | protein structure

Databases: Nucleotide sequence data are available in the DDBJ database under the accession numbers LC378707, LC378707, LC378707 for AnpIBP1a, AnpIBP1b, AnpIBP2, respectively.

Running title: A fungus acquired IBP genes from bacteria.

Abbreviations: AnpIBP, *Antarctomyces psychrotrophicus* ice-binding protein; BI, Bayesian inference; DUF3494, domain of unknown function; HGT, horizontal gene transfer; IBP, ice-binding protein; IBS, ice-binding site; IR, ice recrystallization; IRI, ice-recrystallization inhibition; ML, maximum likelihood; M.w., molecular weight; RMSD, root mean square deviation; TH, thermal hysteresis.

Abstract

Various microbes, including fungi and bacteria, that live in cold environments produce ice-binding proteins (IBPs) that protect them from freezing. Ascomycota and Basidiomycota are two major phyla of fungi, and *Antarctomyces psychrotrophicus* is currently designated as the sole ascomycete that produces IBP (AnpIBP). However, its complete amino acid sequence, ice-binding property, and evolutionary history have not yet been clarified. Here, we determined the peptide sequences of three new AnpIBP isoforms by total cDNA analysis and compared them with those of other microbial IBPs. The AnpIBP isoforms and ascomycete-putative IBPs were found to be phylogenetically close to the bacterial ones but far from the basidiomycete ones, which is supported by the higher sequence identities to bacterial IBPs (40%–72%) than basidiomycete IBPs (29%–36%), although ascomycetes are phylogenetically distant from bacteria. In addition, two of the isoforms of AnpIBP share only 44% sequence identity and are not close in the phylogenetic tree. It is hence presumable that these two AnpIBP isoforms were independently acquired from different bacteria through horizontal gene transfer (HGT), which implies that ascomycetes and bacteria frequently exchange their IBP genes. The non-colligative freezing-point depression ability of a 300 μM solution of a recombinant AnpIBP isoform was not very high (0.7°C), whereas it exhibited significant abilities of ice-recrystallization inhibition, ice shaping, and cryo-protection against freeze-thaw cycles even at submicromolar concentrations. These results suggest that HGT is crucial for the cold-adaptive evolution of ascomycetes, and their IBPs offer freeze resistance to organisms to enable them to inhabit the icy environments of Antarctica.

Introduction

Microbes that adapt to low temperatures (or psychrophiles) can inhabit extremely cold environments, such as the Arctic and Antarctic regions[1,2]. These microbes have acquired several strategies against freezing that inevitably accompany ice crystal formation and overcome physical damages. An example of such a strategy is the production of cryoprotective substances, such as glycerol[3], trehalose[4], and ice-binding proteins (IBPs)[5], which inhibit ice crystal growth. Nevertheless, cold-survival mechanism and phylogenetic relationship of microbial IBPs, especially in fungi, are not well understood.

Biophysical studies have revealed that IBPs bind to specific planes of ice crystals to arrest further growth, which contributes to the survival of host organisms in icy environments[6]. One of the principle functions of IBPs is to inhibit ice recrystallization (IR)[7], a phenomenon of larger ice crystal formation at the expense of smaller ones that causes freezing damages to cells and tissues. IBPs can also modify the shape of ice crystals to mitigate their harsh texture around microbes[8]. In some animals that have naturally high (mg ml^{-1}) IBP concentrations, IBPs also cause the non-colligative freezing-point depression termed thermal hysteresis (TH)[9]. Although microbial IBPs are thought to exist at μg/ml concentrations in the extracellular space, they can exhibit TH when they are concentrated in the laboratory.

Various types of cold-adapted microbes, including bacteria[10], yeast[11], mushrooms[12], algae[8], and a copepod[13], are known to produce IBPs. Most of these microbial IBPs share a common sequence called the "domain of unknown function (DUF) 3494," which comprises approximately 190 amino acid residues exhibiting 30%–70% sequence identities. X-ray crystallographic studies have shown that these IBPs are commonly folded into an irregular beta-helical structure accompanying one long alpha-helix[14].

Ascomycetes and basidiomycetes are two major lineages of fungi and are closely related to each other. In the NCBI database (http://www.ncbi.nlm.nih.gov/), over 200 species of fungal DUF3494 sequences have been deposited, half of which is derived from ascomycetes and the other half from basidiomycetes. Since DUF3494s have been identified from microorganisms that live in both cold and non-cold regions, not all of them may be IBPs. So far none of the ascomycete-derived proteins containing a DUF3494 sequence has been examined for ice-binding activity, whereas the activity was observed for various IBPs from basidiomycetes[11,12,15].

Horizontal gene transfer (HGT) is the phenomenon of transmission and integration of genes between different organisms[16]. HGT is known to provide new adaptability in microorganisms and allow colonization in new environments[17]. Recently, the installation of IBP through HGT was suggested for several sea-ice algae[18,19]. The

evidence for this hypothesis include the (i) phylogenetic discrepancies of the species and genes, (ii) lack of IBP genes in closely-related species in warmer regions, and (iii) lack of intron in some of eukaryote-derived IBP genes. It is hence speculated that the cold adaptation of the microbes is also indebted to HGT. In addition, Raymond (2014)[20] reported that two putative IBPs from ascomycetes are less similar to those from basidiomycetes and suggested HGT in the ascomycetes. However, the primary-to-tertiary structures of ascomycete IBPs remain to be elucidated.

Antarctomyces psychrotrophicus is currently the sole ascomycete whose IBP (AnpIBP) production is confirmed[21]. Only a 20-residue peptide sequence was determined for the N-terminal region of AnpIBP, whereas the full sequence and homogeneity to known IBPs that contain the DUF3494 sequence remain unclear. Since *A. psychrotrophicus* is one of the dominant species isolated from several different substrates in Antarctica, such as soils[22], waters in lakes[23], and marine macroalgae[24], they are exposed to freeze-thaw cycles. IBP production may therefore be a key determinant of cold adaptation of *A. psychrotrophicus*. To clarify the evolutionary process and physiological role of ascomycete IBPs, we analyzed the cDNA and peptide sequences of AnpIBP and examined their phylogenetic relationship with other microbial IBPs. The model structure and ice-binding properties of AnpIBP were also examined and compared with other microbial IBPs. We discuss the cold-adaptation strategy of *A. psychrotrophicus* and the evolutionary relationship among microbial IBPs based on these results.

Results

Three AnpIBP isoforms were identified from total cDNA analysis

Total cDNA analyses of *A. psychrotrophicus* cultivated at cold (−1°C) and moderate (15°C) temperatures were performed to obtain the complete sequences of AnpIBP. We obtained a total of 27,265,280 clean reads (-1°C) and 27,764,816 clean reads (15°C) that are 94.21% and 94.09% of the raw reads, respectively. The de novo assembly identified a total of 16,659 unigenes of 1,162-nt average length for −1°C and a total of 13,831 unigenes of 1,249-nt average length for 15°C. Blast search identified three candidate sequences of AnpIBP from the unigenes. These sequences matched a lot of microbial IBP sequences, such as that of *Chloromonas brevispina* IBP, indicating that AnpIBPs are homologous to other microbial IBPs that contain the DUF3494 sequence. Two of the three cDNA sequences were identified only in the cells cultivated at −1°C, suggesting that their expression occurs only at low temperatures. Fig. 1A shows the previously identified 20-residue amino acid sequence[21] (native AnpIBP) and three newly translated amino acid sequences, designated as AnpIBP1a (236 residues), AnpIBP1b (245 residues), and AnpIBP2 (242 residues, identified as a partial sequence). The former two isoforms are perfectly identical except for their C-terminal regions (after the 203th residue), which were not observed in other microbial IBPs. In contrast, the former two isoforms (AnpIBP1a and 1b) and the latter (AnpIBP2) exhibited only 44% sequence identity. The C-terminal region of AnpIBP1a mainly comprises hydrophobic and positively charged residues, whereas that of AnpIBP1b contains one positive, two negative, and several hydrophobic residues. In both AnpIBP1a and 1b, the Ala1−Ala20 sequence almost matches the known 20-residue sequence of native AnpIBP secreted toward the extracellular space. Therefore, the 19-residue N-terminal sequence before Ala1 (Met18−Ala0: MVSAFMILCVLGSAFVSNA) was assigned to the signal peptide. The molecular weights (M.w.) excluding the signal peptide were hence estimated to be 21.4 and 23.0 kDa for AnpIBP1a and 1b, respectively. They are, however, not consistent with the SDS-PAGE result for purified native AnpIBP, indicating that it is 24 kDa (see below). This discrepancy can be attributed to the post-translational modification of the native protein, where Asn55 in the glycosylation sequence Asn-X-Ser/Thr (X is any amino acid residue, except proline) is presumably modified with N-linked glycan.

To obtain information on introns, the genes of AnpIBP1b and AnpIBP2 were amplified from the *A. psychrotrophicus* genome by PCR with primers based on the cDNA sequences. The AnpIBP2 gene amplified from the genome were consistent with the cDNA sequence, indicating the existence of a gene in the genome without an intron. In contrast,

the AnpIBP1b gene has two introns (Fig. 1B). Introns 1 and 2 were found to be of length 49 and 63 bp, respectively, and both followed the "GT-AG splicing rule." Intron 1 was located near the boundary between the signal peptide and mature IBP domain. In contrast, intron 2, located at the end of the IBP domain, contained the C-terminal sequence of AnpIBP1a, including the terminator codon. When intron 2 was spliced, the mRNA product matched perfectly with the cDNA sequence of AnpIBP1b. This finding indicates that intron 2 can be cleaved by an ambiguous splicing signal for the creation of two types of mRNA (AnpIBP1a and 1b) from one AnpIBP gene (Fig. 1C).

The beta-helical structure was modeled for AnpIBP

Based on the primary sequence information, the model structures of AnpIBP1a and AnpIBP2 were constructed using the MODELLER software with five microbial IBP structures (i.e., TisIBP8, LeIBP, FfIBP, ColIBP, and IBPv_a) as templates. The principle constituent of AnpIBP1a was apparently a seven-ladder β-helix with a triangular cross-section with an accompanying α-helix alongside (Fig. 2A), which is similar to the construction in other microbial IBPs[14]. The inward-pointing residues in the model constructed a hydrophobic core, which is generally indispensable for the stabilization of the β-helix. The three flat surfaces are called the A, B, and C faces. Among them, the B face is suggested to be an ice-binding site (IBS) in the known microbial IBP structure. This putative IBS comprises no regularly arrayed residues (Fig. 2B), different to the hyperactive IBPs identified from insects[25]. Significantly, most of the outward-pointing residues on the putative IBS were glycine, alanine, serine, and threonine, which contain a relatively small side-chain group. The "anchored clathrate water hypothesis[26]" predicts that the combination of hydroxyl (hydrophilic) and methyl (hydrophobic) groups tends to form a water network on IBS, which anchors the host protein to the ice crystal surfaces.

Although the overall structures of AnpIBP1a and AnpIBP2 were mostly identical (Cα RMSD = 0.482 Å), the amino acid compositions of their IBSs were different (Fig. 2B). The IBS of AnpIBP2 is more hydrophilic as it locates more Ser and Thr, but it has fewer Ala compared with AnpIBP1, suggesting that AnpIBP1 and AnpIBP2 possess different ice-binding properties.

In the model structure of AnpIBP1a, the N-glycosylation site (Asn55) is located at the N-terminal end of the long α-helix, which seems distant from the putative IBS. This assumption is supported by the fact that the removal of glycan did not affect the ice-binding property (see below). N-glycosylation may function for protein folding and stability rather than for ice binding, as has been suggested for known IBPs[27]. For both AnpIBP1a and 2, our modeling approach could not predict the structure of the C-terminal segment following

the DUF3494 sequence because this region has no structural template.

IBPs from ascomycetes and basidiomycetes have evolved independently

Protein Blast searches using the full amino acid sequences of AnpIBP1 and 2 showed that the related sequences were mostly DUF3494-containing proteins from ascomycetes or bacteria. Most of the latter were found to be actinobacteria, such as *Streptomyces*, *Arthrobactor*, and *Kitasatospora* species. The peptide sequence exhibiting the highest identity and similarity to AnpIBP1 was a putative IBP from an ascomycete *Zymoseptoria tritici* (58% and 66%) and that to AnpIBP2 was a putative IBP from a planctomycete *Paludisphaera borealis* (72% and 76%). Such high identities were also observed at the nucleotide level; identity of mRNA between AnpIBP1 and the putative IBP from *Z. tritici* was 59%, and that between AnpIBP2 and *P. borealis* was 61%. The mRNA-based identity was slightly lower for *P. borealis*, which might be due to a difference of codon usages (3rd codons); GC content of AnpIBP2 was 50%, while that of the IBP from *P. borealis* was 65%. In contrast, both AnpIBP1 and AnpIBP2 exhibited low sequence identity with basidiomycete-derived IBPs, such as TisIBP8 from *Typhula ishikariensis* and LeIBP from *Leucosporidium* sp. AY30 (29%–36%), although ascomycetes and basidiomycetes are phylogenetically close.

Another difference between ascomycete and basidiomycete IBPs was found at a specific portion of their sequences, which creates a capping structure at the end of their β-helical domains (Fig. 3). As can be seen in this figure, this portion exhibits a significant variation in the structures of microbial IBPs, and its construction in the AnpIBP model showed a significant similarity with that of *Flavobacteriaceae* bacterium isolate 3519-10 (IBPv_a). Do *et al*. (2014) proposed that microbial IBPs can be divided into two groups based on their capping structure[28]. Our sequence alignments, however, suggest that microbial IBPs can be divided into three groups. Group 1 and group 2 have large capping structures where a disulfide-bond is only contained in the latter, which is identical to the classification by Do *et al*. Our alignment further suggests the existence of group 3 that includes ascomycete AnpIBP, bacterial EfcIBP, StaIBP (putative IBP), and IBPv, all of which have a small capping structure. Similarity of DUF3494 sequences within each group are high, while those between the groups are low (Table S1). When all fungal putative IBP sequences in the database (~200) containing the capping loops were aligned, it was found that ascomycete- and basidiomycete-derived putative IBPs have small and large capping structure, respectively. In contrast, the capping structures of bacterial putative IBPs are variable depending on the species.

To investigate the evolutionary history of ascomycete IBPs, we created and

compared the phylogenetic trees of the microbes based on their DUF3494-peptide sequence (DUF3494 tree, Fig. 4A) and on their 16S/18S ribosomal RNA genes (rRNA tree, Fig. 4B). Comparison between these two trees is commonly used to identify the phylogeny of IBPs[19,29]. In Fig. 4B, ascomycetes, basidiomycetes, bacteria, and algae were separately categorized, and the former two shared a common node, which is consistent with the universal tree of their life. Significantly, the DUF3494 tree created by maximum likelihood (ML) method (Fig. 4A) showed that microbial DUF3494s are separated into three groups similarly to that in Fig. 3A. The sequences in group 1, 2, and 3 have large, disulfide-bonded, and small capping structures, respectively. Note that in Fig. 4A, group 3 was sub-divided into Groups 3a and 3b, as several bacterial sequences including IBPv are different from the other bacteria, although they have small capping structures. Thus, ascomycetes formed a different clade from basidiomycetes but shared a common node with several bacteria, which is inconsistent with the ribosomal RNA tree. The DUF3494 tree (Fig.4A) shows that AnpIBP1 and 2 (Antarctomyces psychrotropicus isoform 1 and 2) belong to different clades within Group 3A, and the latter shares a common node with a bacterium *P. borealis*. Such phylogenetic incongruence is generally explained by HGT. Note that almost the same phylogenetic relationships were observed in the Bayesian inference (BI) tree of DUF3494 sequences, whose bootstrap values were shown in Fig. 4A.

The DUF3494 phylogenetic tree is highly similar to that presented by Bayer-Giraldi, M. et al. (2010)[18] and also consistent with the other reports dealing with HGT of IBPs. For example, algal *Chloromonas brevispina* and *Chlamydomonas raudensis* IBPs are distant from the other algal IBPs but close to bacterial and/or ascomycete ones in the phylogenetic tree, which agrees with the previous reports hypothesizing probable acquisition of these IBPs by HGT[20,30]. Other examples are IBPs from bacterial symbiont of *Euplotes focardii*, which were separated into group 3a and 3b and shared a common node with *S. aurantiaca* or *Flavobacteriaceae* bacterium isolate 3519-10 in the phylogenetic tree. This also agree with the report presenting that EfcIBPs are possibly derived from close relatives of these species[31].

To examine whether a difference of the tree-topologies between Figs. 4A and 4B is ascribed to phylogenetic artefacts, statistical tests denoted “approximately unbiased (AU) test” and “Shimodaira-Hasegawa (SH) test” were performed. The upper trees in the red box of Fig. 5 shows the ML and BI trees of the DUF3494-peptide sequences from bacteria, ascomycetes, and basidiomycetes in group 2 and 3A. Again, topologies of the ML and BI trees are almost identical, and ascomycetes (reds) are phylogenetically close to bacteria (yellows) rather than basidiomycetes (greens). We then created the trees with a topological constraint (denoted Constrained tree 1) that intentionally separates

ascomycetes, basidiomycetes, and bacteria into different nodes (i.e., null hypothesis). The AU and SH tests evaluated small values (0.006–0.018) for ML and BI tress (Fig.5, Test1 table). These results support that a significant phylogenetic incongruence exists between the host species and DUF3494 sequences.

Additional statistical test was performed to confirm whether the AnpIBP isoforms belong to different clades or not. In this case, the evaluation was performed with using Constrained trees 2 (Fig. 5, blue boxes), where AnpIBP1 and 2 are monophyly and that they were placed next to each other. The AU and SH tests evaluated small values (0.000–0.002) for these trees (Fig.5, Test2 table), suggesting the independency of the evolutionary processes of these two AnpIBP isoforms.

AnpIBP1 is moderately active but creates a lemon-like ice crystal

The recombinant protein of AnpIBP1a was synthesized by employing the *Pichia pastoris* expression system and purified with the combined use of ion-exchange and gel-filtration chromatographies. The M.w. of purified recombinant AnpIBP1a was estimated to be 25 kDa by SDS-PAGE (Fig. 6A), inclusive of a 216-residue peptide (of approximately 21.4 kDa in size) and an N-linked glycan bound possibly to the Asn55 glycosylation site. A smeared band observed between the 25- and 35-kDa regions in SDS-PAGE may be due to multiple components of the glycan because *P. pastoris*-derived proteins generally do not migrate to this position. A slightly lower M.w. of native AnpIBP (24 kDa) suggests a difference in the compositions of the N-glycan attached to the two proteins.

Recombinant AnpIBP1a at a concentration of 150 μM was capable of stopping the growth of a single ice crystal and had a TH activity of 0.56°C (Fig. 6B, black). This value is identical to that of a purified native AnpIBP sample at the same concentration (Fig. 6B, green), which confirms the accuracy of the determined AnpIBP sequences. It has been well demonstrated that “moderately active IBP” shapes an ice crystal into a bipyramidal morphology and exhibits approximately 1°C of the TH activity at millimolar concentrations[32]. In contrast, a lemon-like ice crystal and approximately a 2°C–4°C temperature of high TH values is obtained for “hyperactive IBPs” at micromolar concentrations. Therefore, recombinant AnpIBP1a should be categorized into moderately active IBPs based on its TH value (0.7°C at 300 μM), although this protein simultaneously showed its ability to create lemon-like ice crystals (Fig. 6C) like a hyperactive IBP. This lemon-like ice crystal in the AnpIBP1a solution underwent a bursting growth along the *c*-axis to form a large hexagonal bipyramid (Fig. 6D) when the temperature fell below the depressed freezing point. These results are indicative of a unique ice-binding manner of AnpIBP1a.

AnpIBP1a contains an N-glycan possibly at Asn55 and an extra C-terminal segment after the 203th residue, which are not conserved in the other microbial IBPs. Hence, they may contribute to the ice-binding property of this protein. To examine such a possibility, a non-glycosylated mutant (AnpIBP1a_N55D) and a mutant containing no C-terminal segment (AnpIBP1_N55D) were recombinantly expressed by *P. pastoris* similarly to the wild type, and their TH activities were assessed. The substitution of Asn55 with Asp decreased M.w. from 25 to 21 kDa on SDS-PAGE (Fig. 6A), indicating the presence of an N-glycan at Asn55. The TH and ice-shaping activities of the two mutants were almost identical to those of the wild type (Fig. 6B). The His-tagged versions of these two mutants also demonstrated the same TH values as those without the tag. These results indicate that neither an N-glycan nor the C-terminal segment is involved in ice binding. The average yields of recombinant AnpIBP1 per culture were found to be approximately 10, 110, 90, 160, and 170 mg/L for AnpIBP1a, AnpIBP1a_N55D, AnpIBP1_N55D, His-AnpIBP1a_N55D, and His-AnpIBP1_N55D, respectively.

AnpIBP1 strongly inhibits IR

IR is a phenomenon of the increase in the ice crystal size within already frozen materials. One of the principle abilities of IBP is the inhibition of recrystallization through binding onto the ice crystal surfaces[6,7]. To evaluate this ability of IBP, termed IR inhibition (IRI), for recombinant AnpIBP1a wild type, we monitored the growth of ice crystals in a frozen AnpIBP1a solution supplemented with 30% sucrose after annealing for 60 min at subzero temperature (−3°C). In the absence of AnpIBP1a, an apparent growth of ice crystals was observed after 60 min of annealing, and the average size of the crystals was recorded to be approximately 393 μm^2 (Fig. 7). The ice crystal growth was significantly arrested with the addition of only a small amount of AnpIBP1a (5 μM), where the average ice crystal size was approximately 8 μm^2, indicating that this protein possesses strong IRI activity. On dilution to 0.5 μM, AnpIBP1a allowed crystal growth up to a size of 164 μm^2, but a further dilution to 0.05 μM almost failed to stop the growth (333 μm^2). These results are in good agreement with previous indications that native AnpIBP possesses IRI ability at approximately a 2-μM concentration[21].

Native AnpIBP affects the growth of ice crystals even at a very low concentration

We observed a trend of freezing of a potato dextrose broth (PDB) medium in the absence/presence of AnpIBP, for which a 3-μL droplet of each sample was frozen on the stage of a photomicroscope. For the pure PDB medium without AnpIBP, relatively large ice crystals with a rounded, leaf-like shape were observed (Fig.8A; right). This is the typical

observation for a solution without IBP. In contrast, numerous rock-like, fine ice crystals were generated in *A. psychrotrophicus*-containing medium cultivated at 4°C for 1 month, thereby including AnpIBP (Fig.8A; left). On a macroscale, the culture supernatant of *A. psychrotrophicus* turned white and opaque (Fig. 8B) when frozen at −20°C, indicating the formation of numerous fine ice crystals. When this culture medium was loaded onto SDS-PAGE, no significant band was detected by Coomassie brilliant blue (CBB) staining. These results again suggest that AnpIBP can create a fine ice texture at the micromolar level.

To assess the physiological functions of AnpIBP, we compared the growth of the mycelia of *A. psychrotrophicus* on potato dextrose agar (PDA) plates before and after 25 cycles of freeze and thawing. A snow mold fungus *T. ishikariensis* was also tested for comparison. These cell cultures were maintained at −1°C to induce IBP expression before starting the cycles. It was found that both fungi grew at 10°C without the cycles (Fig. 8C), where the growth rates were 0.5 cm/day for *A. psychrotrophicus* and 0.1 cm/day for *T. ishikariensis*. Significantly, *A. psychrotrophicus* grew normally on the plate even after the 25 freeze-thaw cycles, while *T. ishikariensis* did not show any growth, indicating that *A. psychrotrophicus* is extremely tolerate to freeze-thaw cycles.

Discussion

In cold environments, many psychrophilic microbes, including fungi, produce IBPs to control ice crystal growth, which has detrimental effects on their tissues. The fungi are categorized into two major phyla, Ascomycota and Basidiomycota. A sole ascomycete IBP from *A. psychrotrophicus* (AnpIBP) was initially characterized in our previous report. Here, we determined the cDNA and peptide sequences of three AnpIBP isoforms to examine their structures and ice-binding properties, which revealed the uniqueness of their phylogenetic background and functional abilities.

cDNA and peptide sequences reveal the structure of AnpIBP

Most of the microbial IBPs are known to contain the DUF3494 sequence, whereas the DUF3494-containing proteins detected in the ascomycete genomes or cDNA libraries have not been confirmed to exhibit the ice-binding property. The ascomycete DUF3494-proteins share a high homogeneity with the three AnpIBP isoforms, and some of them can be assigned to the microbial IBPs. For example, Kawahara *et al*.[33] reported that several Antarctic ascomycetes produce extracellular substances that modify the ice crystal shape, some of which could be IBPs.

The N-terminal sequences of AnpIBP1a and AnpIBP1b almost completely matched with the previously determined sequence of native AnpIBP (Fig. 1A), suggesting the secretions of both AnpIBP isoforms toward the extracellular space via the signal peptide. This finding is consistent with the present finding showing that recombinant AnpIBP1a possesses almost the same TH activity as that of the native AnpIBP (Fig. 6B). In contrast, AnpIBP2 was not found in the culture supernatant of *A. psychrotrophicus*. Thus, this isoform may exist in the intracellular space or on the cell surface. Recently, several reports have shown the existence of such non-secreted IBPs[34,35].

Our intron analysis revealed that AnpIBP1a and 1b are produced from one IBP-coding region through alternative splicing of the C-terminal non-ice-binding region (Fig. 1C). The existence of two isoforms that differ only at the C-terminal segment evoked a question about its relevance to the structure and function of this protein. However, the removal of the segment did not affect the ice-binding property of AnpIBP1a (Fig. 6B). The extension of the N- or C-terminal sequence has sometimes been observed in several IBPs, suggesting that these regions are mostly variable and not conserved. Although their functions have hardly been clarified, several reports have suggested that such additional sequences are responsible for dimerization[14] or anchoring of the protein to the cell surface[35,36]. The additional C-terminal sequences of AnpIBP1a and 1b have electrically opposite features (Fig. 1A). Thus, these segments may be involved in heterodimerization

or binding to different substrates other than ice.

AnpIBP isoforms may originate from different bacteria

The comparison of the AnpIBP sequences with that of other microbial IBPs revealed that AnpIBP1 and AnpIBP2 share high sequence identities with the DUF3434-containing proteins of both ascomycetes and bacteria (<72%) but poor identities with those of basidiomycetes (<36%). In the phylogenetic tree of DUF3494 domains (Fig. 4A), ascomycete DUF3494s, including AnpIBPs, were phylogenetically far from the basidiomycete ones but close to the bacterial ones in group 3, which is inconsistent with the phylogenetic relationship of their life (Fig. 4B). This phylogenetic incongruence suggests that ascomycete IBPs independently evolved from basidiomycete IBPs and that they were possibly acquired from bacteria through HGT. The statistical tests (Fig. 5, Test 1) further support this HGT hypothesis. These results are in good agreement with the proposition by Raymond[20] and support the structural similarity in the capping loop region of DUF3494 proteins (Fig. 3). All of the DUF3494 proteins in group 3 contain a small capping structure, whereas those from other groups have a different capping structure. This semi-conserved region may have a significant correlation with the evolutionary process of microbial DUF3494. Another evidence of the HGT hypothesis is the occurrence of differences in the number of introns. AnpIBP1 and 2 genes do not possess any introns in the DUF3494 sequence (Fig. 1B and C), which may be due to the occurrence of HGT from bacteria. It has been reported that basidiomycete IBPs from yeast[11] and mushrooms[12] have eight and four introns, respectively. Such a difference between IBPs from ascomycetes and basidiomycetes implies a difference in their evolutionary history and a relatively recent acquisition of ascomycete-derived IBP genes compared with those of basidiomycetes.

Such phylogenetic incongruence might also be explained by a different gene loss, in which DUF3494 genes were duplicated in an ancestor of fungi and bacteria, and they were lost differently during evolutionary process. If this is correct, some of the fungi should retain two or more kinds of DUF3494 that belong to different groups. All ascomycetes and basidiomycetes, however, possess only group 3 and group 1 DUF3494s, respectively. Furthermore, the DUF3494 gene has not been found in the other phyla of fungi, such as *Zygomycota* and *Chytridiomycota*. Since it is not likely that the DUF3494 genes were lost in all fungi other than ascomycetes and basidiomycetes, fungal IBP genes were possibly acquired by HGT.

The DUF3494 sequences of ascomycetes and bacteria in group 3 are scattered in the phylogenetic tree (Fig. 4A). Furthermore, several DUF3494s from bacteria such as proteobacteria and flavobacteria are separated into the different groups. These results

suggest that HGT of IBP occurs frequently.

Surprisingly, AnpIBP1 and AnpIBP2 in group 3 are apparently separated into different nodes. This finding is supported by the low sequence identity between AnpIBP1 and AnpIBP2 (44%), which is much lower than that between AnpIBP1 and ascomycete IBPs, such as *Z. tritici* IBP (58%), or that between AnpIBP2 and bacterial IBP, such as *P. borealis* IBP (72%). This significantly high sequence identity between AnpIBP2 and *P. borealis*-putative IBP indicates the installation of the IBP gene from this bacterium or its ancestor; thus, AnpIBP1 and AnpIBP2 presumably originate from different HGT events. This hypothesis was supported by another statistical test, which also negates monophyly of AnpIBP1 and 2 (Fig. 5, test 2).

In general, HGT hardly occurs between fungi (eukaryote) and bacteria (prokaryote) because of the difference in codon usage and incompatibility of promoters.[17]. However, it is known that HGT frequently occurs under specific stress from harsh environmental conditions and provides a selective advantage to the recipient organism[37]. Thus, extremely cold environments accompanied by ice crystal formation are likely to exert a strong selection pressure on microorganisms, and there is a strong demand for IBPs to adapt to this special environment. As a result of positive selection, several HGT events could occur and IBP-producing microorganisms could survive.

Most of the IBPs in group 3 were derived from ascomycetes or actinobacteria, both of which usually live in the soil, suggesting that the installation of the IBP gene through HGT progresses under cold soil environments. These microbes tend to form a community with the surrounding heterotrophic bacteria. In this dense community, physical contact between fungi and bacteria or among bacteria occurs frequently, thereby increasing the probability of HGT via a conjugation-like mechanism. Otherwise, the filamentous microbes may have acquired freely available IBP genes produced by lysis of neighboring microorganisms in the community. Another possibility is natural transformation due to cell membrane damage with high salt concentrations caused by ice crystal growth, which facilitate the incorporation of IBP genes into the cells. Kiko (2010)[13] proposed that sea-ice habitats create natural transformation conditions, enabling HGT. A similar scenario may be true in Antarctic soil because the continent is mostly covered with a large amount of snow and ice blocks.

AnpIBP facilitate the cold adaptation of A. psychrotrophicus

Because *A. psychrotrophicus* is isolated from various environments of Antarctica[22-24], AnpIBP could play an important role in its cold adaptation. Our results suggest that AnpIBP1 is secreted into the extracellular space and is believed to be

responsible for controlling ice growth around the cells. Nevertheless, the TH activity of AnpIBP1 (0.7°C at 300 μM) was not significant, although the environmental temperature is extremely low. For example, the minimum temperature at King George Island, where *A. psychrotrophicus* was first isolated, is −19.9°C[22]. AnpIBP1 showed both IRI (Fig. 7) and ice-shaping activities (Fig. 8A and B), even at extremely low concentrations (5 μM). Thus, the main function of this IBP will be the formation of fine ice crystals during freezing and the inhibition of IR during melting rather than freezing-point depression, which may lead to the prevention of freezing damages to the host organism. Indeed, *A. psychrotrophicus* can tolerate 25 freeze–thaw cycles, whereas the snow mold *T. ishikariensis* cannot (Fig. 8C). HGT between ascomycetes and bacteria may have significant importance for their cold adaptation and evolution, leading to the wide distribution of *A. psychrotrophicus* in Antarctica.

Materials and methods

Preparation of cDNA library, and total cDNA analysis of A. psychrotrophicus

A. psychrotrophicus strain Syw-1 was isolated from soil collected from Kizahashi-hama, Skarvsnes on the Soya coast, Antarctica, in 2006 and cultured in our laboratory in PDB medium. The cultures were kept at either −1°C or 15°C for 1 month to compare the efficiency of IBP expression between the two temperatures. Total RNA of *A. psychrotrophicus* was extracted and purified using the RNeasy Plant Mini Kit (QIAGEN, CA, USA), according to the suggested protocol. Poly (A+) RNA was isolated by oligo (dT) cellulose chromatography (New England Biolabs, MA, USA). cDNA was synthesized using the Marathon cDNA Amplification Kit (BD Biosciences Clontech, CA, USA). The constructed library was sequenced with the Illumina HiSeq™ 2000. The raw reads were cleaned by removing adaptor-only reads, repeated reads, and low-quality reads by Illumina Sequencing Analysis Pipeline ver. 1.6. Transcriptome de novo assembly was carried out with short reads assembling program Trinity[38]. IBP sequences were identified using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

To confirm the accuracy of the determined cDNA sequences and to obtain intron information, the AnpIBP1b gene was amplified from the genome of *A. psychrotrophicus* by PCR with forward primer 5′-ATGGTTTCCGCCTTCATGATCC-3′ and reverse primer 5′-TTAGACCTTGAAGAACTTGGCAGA-3′, which are the 5′- and 3′-terminal cDNA sequences of AnpIBP1b, respectively. AnpIBP2 gene was also amplified with a forward primer 5′-GCATCGTTCGCTGTTCTAGGAG-3′ and a reverse primer 5′-GGAGGTAGTAGTAGTGGTTGTGGT-3′. The genome of *A. psychrotrophicus* was purified, as described in [39]. The PCR reaction was performed with TaKaRa Ex Taq (TaKaRa, Shiga, Japan). The PCR product was cloned into a pMD20-T vector (Mighty TA Cloning Kit; TaKaRa, Shiga, Japan), transformed into *Escherichia coli* JM109, and then sequenced.

Sequence alignment and structural modeling of AnpIBP

Multiple sequence alignment was performed with the Clustal omega tool (http://www.ebi.ac.uk/Tools/msa/clustalo/) using default parameters. The same tool was used to examine the sequence identities.

The structural models of AnpIBP1a and AnpIBP2 were prepared with MODELLER (http://salilab.org/modeller/) and visualized with Chimera (http://www.cgl.ucsf.edu/chimera/). TisIBP8 (pdb code = 5B5H), LeIBP (3UYU), FfIBP (4NU2), ColIBP (3WP9), and IBPv_a (5UYT) were used as templates. A total of 10 models were created for each isoform, and an energy-minimized structure was selected on the basis

of the MODELLER score (DOPE).

Phylogenetic analysis of microbial DUF3494 domains

The MEGA7 software (http://www.megasoftware.net/) was used to create the phylogenetic trees based on the microbial DUF3494s and 16/18S ribosomal RNA sequences. The former (DUF3494 tree) was made with maximum likelihood method based on the WAG+F+G model. The bootstrap values were obtained with 500-resamplings. The latter tree (rRNA tree in Fig. 4B) was prepared by employing the neighbor-joining method based on the TN93+G model. Only Bayesian inference tree was made by using Mrbayes software ver. 3.2.6 (http://www.mrbayes.sourceforge.net/download.php) with using 400,000 generations (four chains), sampling once every 100 rounds, random starting tree, WAG amino acid substitution model, and discarding a burn-in of 1,000. The AU test[40] and SH test[41] were performed with using Treefinder software (http://www.treefinder.de/).

Expression and purification of recombinant and native AnpIBPs

Codon-optimized AnpIBP1a gene (Ala^{1}–Val^{216}) with 5′-terminal *Xho*I and Kex2 signal cleavage sites (ctcgagaaaaaga) and 3′-terminal *Not*I (gcggccgc) sites was synthesized and digested with the restriction enzymes. The resultant DNA fragment was ligated into pPICZα (Thermo Fisher Scientific, MA, USA) digested with the two restriction enzymes. The expression vector containing the AnpIBP1a gene was linearized using the *Pme*I restriction enzyme. Transformation of the *P. pastoris* X33 strain was performed using the *Pichia* EasyComp Transformation Kit (Thermo Fisher Scientific, MA, USA). The transformant was selected by Zeocin resistance and the IBP expression level. The selected transformant was cultured in 80 mL of buffered glycerol complex medium (1% yeast extract, 2% peptone, 100 mM potassium phosphate at pH 6.0, 1.34% YNB, 4×10^{-5}% biotin, and 1% glycerol) at 28°C with agitation. When OD_{600} reached 10, the cells were harvested by centrifugation at 7500 rpm for 10 min at room temperature and were then resuspended in 450 mL of 2× buffered methanol complex medium (2% yeast extract, 4% peptone, 200 mM potassium phosphate at pH 6.0, 2.68% YNB, 8×10^{-5}% biotin, and 0.5% methanol). The cells were cultivated in a 1-L BMJ-01P fermenter (ABLE, Tokyo, Japan) at 20°C. Methanol was continuously added to the fermenter through a peristaltic pump at a manually controlled flow rate of 0.5–3.5 mL/h. After 96–120 h of cultivation, the medium containing IBP was collected by centrifugation and was dialyzed against 20 mM Gly-HCl buffer (pH 3.0). The dialysate was loaded into a High-S (Bio-Rad, CA, USA) column and eluted using a 0–300-mM NaCl linear gradient over 10-column bed volumes. The antifreeze active fractions were recovered and dialyzed against 20 mM Tris-HCl (pH 8.0). The dialysate was

applied into a High-Q (Bio-Rad, CA, USA) column and eluted using a 0–300-mM NaCl linear gradient over 10-column bed volumes. The recovered fraction was concentrated to a 5-mL volume and further purified by gel-filtration chromatography using Superdex 200 (GE-Healthcare, Amersham, UK) equilibrated with 20 mM Tris-HCl buffer containing 500 mM NaCl. The antifreeze active fractions were collected and dialyzed against water. The purity of the sample was checked using 16% SDS-PAGE stained with CBB.

All mutants were prepared using the KOD-plus Mutagenesis Kit (Toyobo, Osaka, Japan) and confirmed by DNA sequencing. Then, they were expressed and purified by the same procedure as that used for the wild types. In the case of His-tagged mutants, the first purification on the High-S column was substituted with the Ni-NTA column (QIAGEN, CA, USA) equilibrated with 20 mM Tris-HCl (pH 8.0) containing 500 mM NaCl. His-tagged AnpIBPs were eluted with the same buffer containing 250 mM imidazole.

Native AnpIBP was purified from the culture supernatant of *A. psychrotrophicus*, as described elsewhere[21].

Measurement of TH activity and IRI

The TH activity was measured with a photomicroscope system equipped with a temperature-controlling system, as described in[42], with slight modifications. Briefly, IBP was dissolved in 20 mM Tris-HCl (pH 8.0). The sample solution (1 μL) in a glass capillary was placed on the stage and cooled rapidly to approximately −25°C until frozen. The frozen sample was melted slowly until a single ice crystal remained. The temperature at which the ice crystals melted was recorded as *Tm*. The formed ice-crystal was kept at *Tm* − 0.1°C for 5 min. The temperature was slowly lowered (0.1°C/min). The freezing point (T_f) of the sample was determined as the temperature at which rapid ice growth was observed. The TH values were evaluated from the absolute value of the difference between T_m and T_f (i.e., $TH = |T_f - T_m|$).

IRI of IBPs was measured by the observation of the change in the ice crystal morphology under the photomicroscope. Recombinant AnpIBP1 dissolved in 30% sucrose solution (1.2 μL) was sandwiched between two glass plates. The sample was cooled at 55°C/min until the entire solution was frozen. Then, it was warmed at 20°C/min up to −3°C and incubated for 1 h. A snapshot was taken every 15 min. The size of the ice crystals was measured using ImageJ software (http://imagej.nih.gov/ij/).

Observation of ice crystal morphology in the AnpIBP solution

The *A. psychrotrophicus* culture (3 μL) cultivated at 4°C or pure PDB medium was placed on a glass plate and observed under a photomicroscope equipped with a

temperature controller. The sample solution was frozen entirely by decreasing the temperature to approximately −25°C, which created a tight assembly of numerous single ice crystals. These crystals were then melted by increasing the temperature to −1°C in order to decrease the number of ice crystals formed. The ice crystals were then allowed to grow by decreasing the temperature at the rate of 1°C/min, and their morphologies were observed.

To observe the macroscale ice texture, 25 mL of *A. psychrotrophicus* culture or pure PDB medium in a 50-mL tube was frozen at −20°C for 24 h.

Freezing–thawing cycle experiments of A. psychrotrophicus

The mycelia of *A. psychrotrophicus* and *T. ishikariensis* were cultured in 50 mL of PDB medium at 10°C. After growth of the mycelia, the cultures were transferred to an incubator at −1°C to induce IBP expression. After 1 month, the cultures were made to undergo 25 cycles of freezing at −20°C and thawing at room temperature. The cells were grown on a PDA plate at 10°C, and the growth rates of the mycelia were recorded.

Acknowledgements

This work was supported by Grant-in-aid 15K13760 for Scientific Research from the Japan Society for the Promotion of Science (JSPS).

Author contributions

TA and ST designed the research. TA, DF, TH, and HK performed the research and analyzed the data. TA and ST wrote the paper.

References

1. Wang M, Tian J, Xiang M, & Liu X (2017) Living strategy of cold-adapted fungi with the reference to several representative species. *Mycology* **8**, 178–188.
2. De Maayer P, Anderson D, Cary C & Cowan DA (2014) Some like it cold: understanding the survival strategies of psychrophiles. *EMBO Rep* **15**, 508–517.
3. Su Y, Jiang X, Wu W, Wang M, Hamid MI, Xiang M & Liu X (2016) Genomic, transcriptomic, and proteomic analysis provide insights into the cold adaptation mechanism of the obligate psychrophilic fungus *Mrakia psychrophila*. *G3 Bethesda* **6**, 3603–3613.
4. Phadtare S (2004) Recent developments in bacterial cold-shock response. *Curr Issues Mol Biol* **6**, 125-136.
5. Duman JG & Olsen TM (1993) Thermal hysteresis protein activity in bacteria, fungi, and phylogenetically diverse plants. *Cryobiology* **30**, 322–328.

6. Davies PL (2014) Ice-binding proteins: a remarkable diversity of structures for stopping and starting ice growth. *Trends Biochem Sci* **39**, 548–555.
7. Knight CA, Hallett J & DeVries AL (1988) Solute effects on ice recrystallization: an assessment technique. *Cryobiology* **25**, 55–60.
8. Janech MG, Krell A, Mock T, Kang JS & Raymond JA (2006) Ice-binding proteins from sea ice diatoms (Bacillariophyceae). *J Phycol* **42**, 410–416.
9. Raymond JA & DeVries AL (1977) Adsorption inhibition as a mechanism of freezing resistance in polar fishes. *Proc Natl Acad Sci USA* **74**, 2589–2593.
10. Sun X, Griffith M, Pasternak JJ & Glick BR (1995) Low temperature growth, freezing survival, and production of antifreeze protein by the plant growth promoting rhizobacterium *Pseudomonas putida* GR12-2. *Can J Microbiol* **41**, 776–784.
11. Lee JK, Park KS, Park S, Park H, Song YH, Kang SH & Kim HJ (2010) An extracellular ice-binding glycoprotein from an Arctic psychrophilic yeast. *Cryobiology* **60**, 222–228.
12. Raymond JA & Janech MG (2009) Ice-binding proteins from enoki and shiitake mushrooms. *Cryobiology* **58**, 151–156.
13. Kiko R (2010) Acquisition of freeze protection in a sea-ice crustacean through horizontal gene transfer? *Polar Biol* **33**, 543–556.
14. Lee JH, Park AK, Do H, Park KS, Moh SH, Chi YM & Kim HJ (2012) Structural basis for antifreeze activity of ice-binding protein from arctic yeast. *J Biol Chem* **287**, 11460−11468.
15. Hoshino T, Kiriaki M, Ohgiya S, Fujiwara M, Kondo H, Nishimiya Y, Yumoto I & Tsuda S (2003) Antifreeze proteins from snow mold fungi. *Canad J Bot* **81**, 1175−1181.
16. Jones D & Sneath PH (1970). Genetic transfer and bacterial taxonomy. *Bacteriol Rev* **34**, 40−81.
17. Fitzpatrick DA (2012) Horizontal gene transfer in fungi. *FEMS Microbiol Lett* **329**, 1–8.
18. Bayer-Giraldi M, Uhlig C, John U, Mock T & Valentin K (2010) Antifreeze proteins in polar sea ice diatoms: diversity and gene expression in the genus *Fragilariopsis*. *Environ Microbiol* **12**, 1041−1052.
19. Raymond JA & Kim HJ (2012) Possible role of horizontal gene transfer in the colonization of sea ice by algae. *PloS one* **7**, e35968.
20. Raymond JA (2014) The ice-binding proteins of a snow alga, *Chloromonas brevispina*: probable acquisition by horizontal gene transfer. *Extremophiles* **18**, 987–994.

21. Xiao N, Suzuki K, Nishimiya Y, Kondo H, Miura A, Tsuda S & Hoshino T (2010) Comparison of functional properties of two fungal antifreeze proteins from *Antarctomyces psychrotrophicus* and *Typhula ishikariensis*. *FEBS J* **277**, 394–403.
22. Stchigel AM, Josep CANO, Mac Cormack W & Guarro J (2001) *Antarctomyces psychrotrophicus* gen. et sp. nov., a new ascomycete from Antarctica. *Mycol Res* **105**, 377-382.
23. Gonçalves VN, Vaz AB, Rosa CA & Rosa LH (2012) Diversity and distribution of fungal communities in lakes of Antarctica. *FEMS Microbiol Ecol* **82,** 459–471.
24. Loque CP, Medeiros AO, Pellizzari FM, Oliveira EC, Rosa CA & Rosa LH (2010) Fungal community associated with marine macroalgae from Antarctica. *Polar Biol* **33**, 641–648.
25. Graether SP, Kuiper MJ, Gagné SM, Walker VK, Jia Z, Sykes BD & Davies PL (2000). β-Helix structure and ice-binding properties of a hyperactive antifreeze protein from an insect. *Nature* **406**, 325–328.
26. Garnham CP, Campbell RL & Davies PL (2011) Anchored clathrate waters bind antifreeze proteins to ice. *Proc Natl Acad Sci USA* **108**, 7363–7367.
27. Park KS, Do H, Lee JH, Park SI, jung Kim E, Kim SJ, Kang SH & Kim HJ (2012) Characterization of the ice-binding protein from Arctic yeast *Leucosporidium* sp. AY30. *Cryobiology* **64**, 286–296.
28. Do H, Kim SJ, Kim HJ & Lee JH (2014) Structure-based characterization and antifreeze properties of a hyperactive ice-binding protein from the Antarctic bacterium *Flavobacterium frigoris* PS1. *Acta Crystallogr D Biol Crystallogr* **70**, 1061–1073.
29. Sorhannus U (2011) Evolution of antifreeze protein genes in the diatom genus Fragilariopsis: evidence for horizontal gene transfer, gene duplication and episodic diversifying selection. *Evol Bioinform* **7**, 279–289.
30. Raymond JA & Morgan-Kiss R (2013) Separate origins of ice-binding proteins in Antarctic *Chlamydomonas* species. *PLoS One* **8**, e59186.
31. Pucciarelli S, Chiappori F, Devaraj RR, Yang G, Yu T, Ballarini P & Miceli C (2014). Identification and analysis of two sequences encoding ice-binding proteins obtained from a putative bacterial symbiont of the psychrophilic Antarctic ciliate *Euplotes focardii*. *Antarct Sci* **26**, 491-501.
32. Scotter AJ, Marshall CB, Graham LA, Gilbert JA, Garnham CP & Davies PL (2006) The basis for hyperactivity of antifreeze proteins. *Cryobiology* **53**, 229–239.
33. Kawahara H, Takemura T & Obata H (2006) Function analysis and screening of antifreeze material from fungi. *Cryobio Cryotech* **52**, 151–155.

34. Gwak Y, Jung W, Lee Y, Kim JS, Kim CG, Ju JH & Jin E (2014). An intracellular antifreeze protein from an Antarctic microalga that responds to various environmental stresses. *FASEB J* **28**, 4924–4935.
35. Vance TD, Graham LA & Davies PL (2018) An ice-binding and tandem beta-sandwich domain-containing protein in *Shewanella frigidimarina* is a potential new type of ice adhesin. *FEBS J* **285**, 1511-1527.
36. Guo S, Garnham CP, Whitney JC, Graham LA & Davies PL (2012) Re-evaluation of a bacterial antifreeze protein as an adhesin with ice-binding activity. *PloS one* **7**, e48805.
37. Martínez-Rosales C, Fullana N, Musto H & Castro-Sowinski S (2012) Antarctic DNA moving forward: genomic plasticity and biotechnological potential. *FEMS Microbiol Lett* **331**, 1–9.
38. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N & Regev A (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* **29**, 644-652
39. Al-Samarrai TH & Schmid J (2000) A simple method for extraction of fungal genomic DNA. *Lett Appl Microbiol* **30**, 53–56.
40. Shimodaira H (2002) An approximately unbiased test of phylogenetic tree selection. *Syst Biol* **51**, 492-508.
41. Shimodaira H & Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol* **16**, 1114-1116.
42. Takamichi M, Nishimiya Y, Miura A & Tsuda S (2007) Effect of annealing time of an ice crystal on the activity of type III antifreeze protein. *FEBS J* **274**, 6469–6476.

Figure legends

Figure 1. Sequence alignment of AnpIBP isoforms and genome structure of the AnpIBP1 gene. (A) The sequence alignment of translated amino acid sequences of AnpIBP isoforms identified by total cDNA analysis. Putative N-glycosylation (Asn55) sites of AnpIBP1 are underlined. Cylinders and arrows indicate alpha-helices and beta-sheets of the AnpIBP1 model, respectively. Black arrows particularly indicate beta-sheets that comprise the putative IBS of AnpIBP1. (B) Full-length AnpIBP1 gene in the genome. Introns are underlined. Terminator codons are shown in italic. (C) Schematic diagram of

the genome structure and mature mRNA of AnpIBP. Ca and Cb indicate the C-terminal segment of AnpIBP1a and AnpIBP1b, respectively. Dotted lines indicate the splice positions. The AnpIBP1 gene in the genome comprises two introns (I) and three exons (E). The two types of mRNA are produced by alternative splicing of Exon 2. In all the figures, the signal peptide, untranslated region, DUF3494 region, and C-terminal segment are colored by blue, black, red, and yellow, respectively.

Figure 2. Structural model of AnpIBP. (A) The overall structure of AnpIBP1a is represented with a ribbon model in spectral color gradation from blue (N-terminus) to red (C-terminus). (B) Amino acid residues of the putative ice-binding sites (B face) of AnpIBP1a and AnpIBP2. Large and small letters indicate the outward- and inward-pointing residues, respectively. Serine and threonine residues are indicated in orange and positively charged residues in red.

Figure 3. Sequence alignment of the capping structure between microbial IBPs. (A) Multiple sequence alignment among IBPs from *A. psychrotrophicus* (AnpIBP1, 2), bacterial symbiont of *Euplotes focardii* (EfcIBP), *Leucosporidium* sp. AY30 (LeIBP), *S. aurantiaca* (StaIBP), *T. ishikariensis* (TisIBP8), *Colwellia* sp. SLW05 (ColIBP), *Flavobacterium frigoris* (FfIBP), *Navicula glaciei* (NagIBP8), and *Flavobacteriaceae* isolate 3519-10 (IBPv). Sequences corresponding to the capping structure are in blue box. Cysteine residues involved in the disulfide bond are boxed with yellow. (B) Microbial IBP structures from each group. The crystal structures of ColIBP (pdb code: 3WP9), TisIBP8 (5B5H), IBPv_a (5UYT), and EfcIBP (6EIO) and the structural model of AnpIBP1a are shown as ribbon models, in which the capping structures are colored in blue.

Figure 4. Phylogenetic discrepancy between microbial 16S or 18S ribosomal RNA and DUF3494s. (A) Maximum likelihood tree based on the amino acid sequences of microbial DUF3494. The sequences confirmed to show ice-binding activity are underlined. (B)

Neighbor phylogenetic tree of 16S or 18S ribosomal RNA of microorganisms. In both the figures, the numbers at the nodes indicate the bootstrap values for 500 replications (<30 are not shown), and the grouping (as in Fig. 3) is shown on the right side. In the DUF3494 tree, bootstrap values by ML and BI methods were shown as black and red, respectively. The edge colors of the bacterial and the algae markers indicate bacterial phyla as Proteobacteria (black), Planctomycetes (green), Firmicutes (cyan), Actinobacteria (orange), Bacteroidetes (purple), and unidentified (yellow) for bacteria, and *Bacillariophyta* (Black), *Haptophyta* (pink), and *Chlorophyta* (green) for algae.

Figure 5. Statistical tests performed on the phylogenetic trees prepared from the DUF3494 gene and the host species. Statistical tests of HGT tree against constrained tree in which ascomycetes, basidiomycetes, and bacteria form different clades (Red box). Statistical tests of HGT tree against constrained tree in which AnpIBP isoforms are monophyly (Blue boxes). AU and SH test p-values are shown in right side tables. Bootstrap values for 500 replications are shown at the nodes. Markers are same as in Fig. 4.

Figure 6. SDS-PAGE, TH activities, and ice crystal morphologies of recombinant AnpIBP1. (A) SDS-PAGE of purified AnpIBP1a and its mutants. M: Molecular weight marker; Lane 1, AnpIBP1a; Lane 2, AnpIBP1a N55D; Lane 3, AnpIBP1 N55D; Lane 4, His-tagged AnpIBP1 N55D; Lane 5, native AnpIBP. (B) The TH activity of AnpIBP and its mutants. Error bars indicate standard deviations. (C) The comparison of ice crystal shapes in the IBP solution. (D) Ice crystal morphology of AnpIBP1a. The black arrow indicates the c-axis direction. Scale bar = 50 μm.

Figure 7. Concentration-dependence of IRI activity of AnpIBP1a. (A) IBP concentration-dependent ice crystal morphologies before and after −3°C incubation. Ice crystals were observed under a photomicroscope. Scale bar = 50 μm.

Figure 8. Effect of AnpIBP on ice crystal growth. (A) Microscopic view of ice crystal growth in PDB medium with or without native AnpIBP. In IBP (+) medium, *A. psychrotrophicus* was cultured at 4°C for 1 month. In the *A. psychrotrophicus* culture, the

size of the ice crystals was smaller than that in pure PDB medium. (B) Comparison of frozen PDB medium with or without AnpIBP. Formed ice was opaque in the presence of AnpIBP, whereas pure PDB medium was frozen as semi-transparent ice. (C) Mycelium growth on PDA plates of *A. psychrotrophicus* and *T. ishikariensis* with and without 25 cycles of freezing–thawing. *A. psychrotrophicus* is on the left side, and *T. ishikariensis* is on the right side.

Supporting information

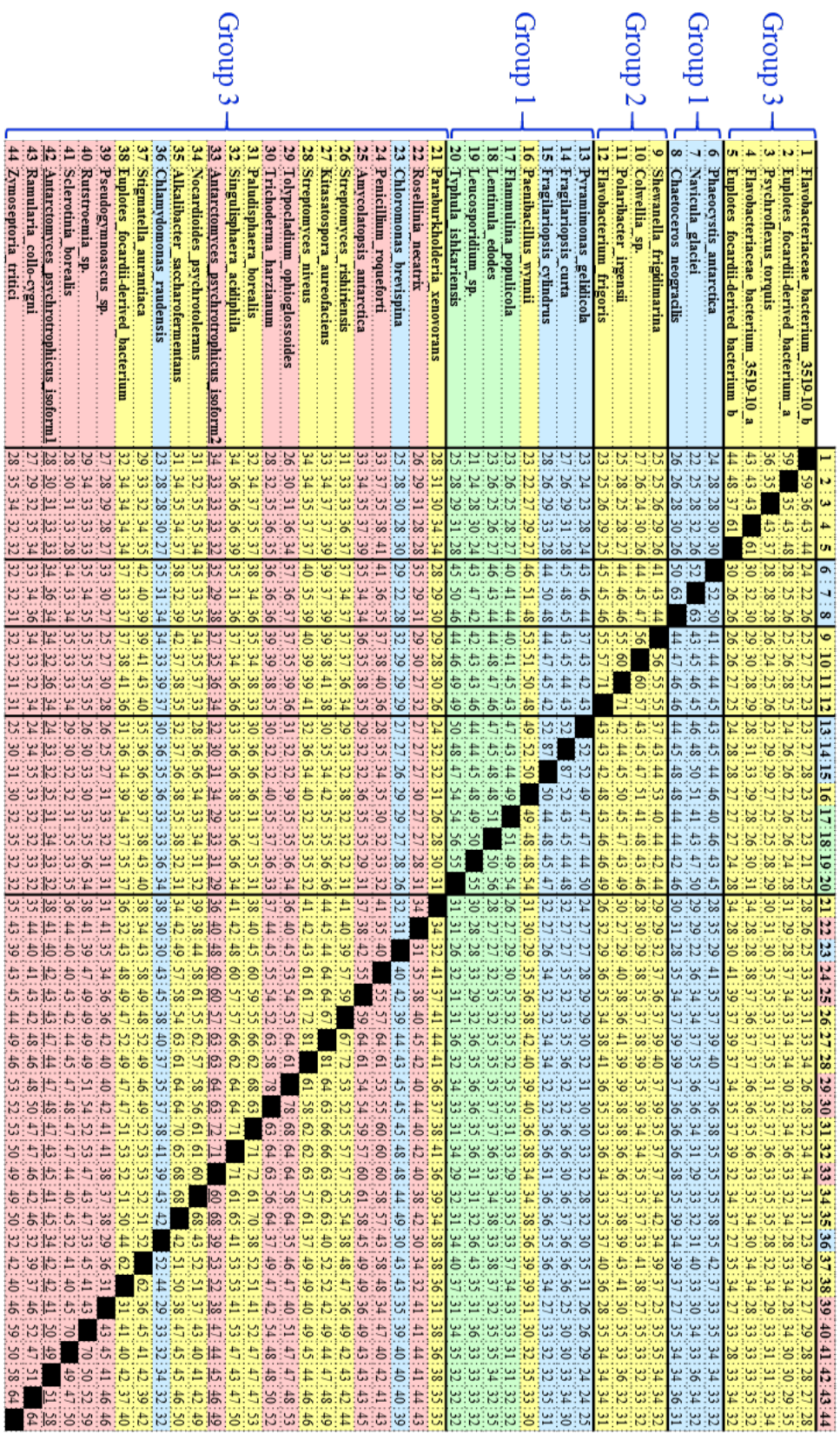


Table S1. Sequence identities (%) between selected microbial DUF3494s. The sequences used here are the same as those used for the phylogenetic analysis (mentioned in Fig. 4A). Rows of bacteria, algae, basidiomycetes, and ascomycetes are indicated in yellow, blue, green, and red, respectively. The grouping is the same as in Fig. 3.

Figure 1

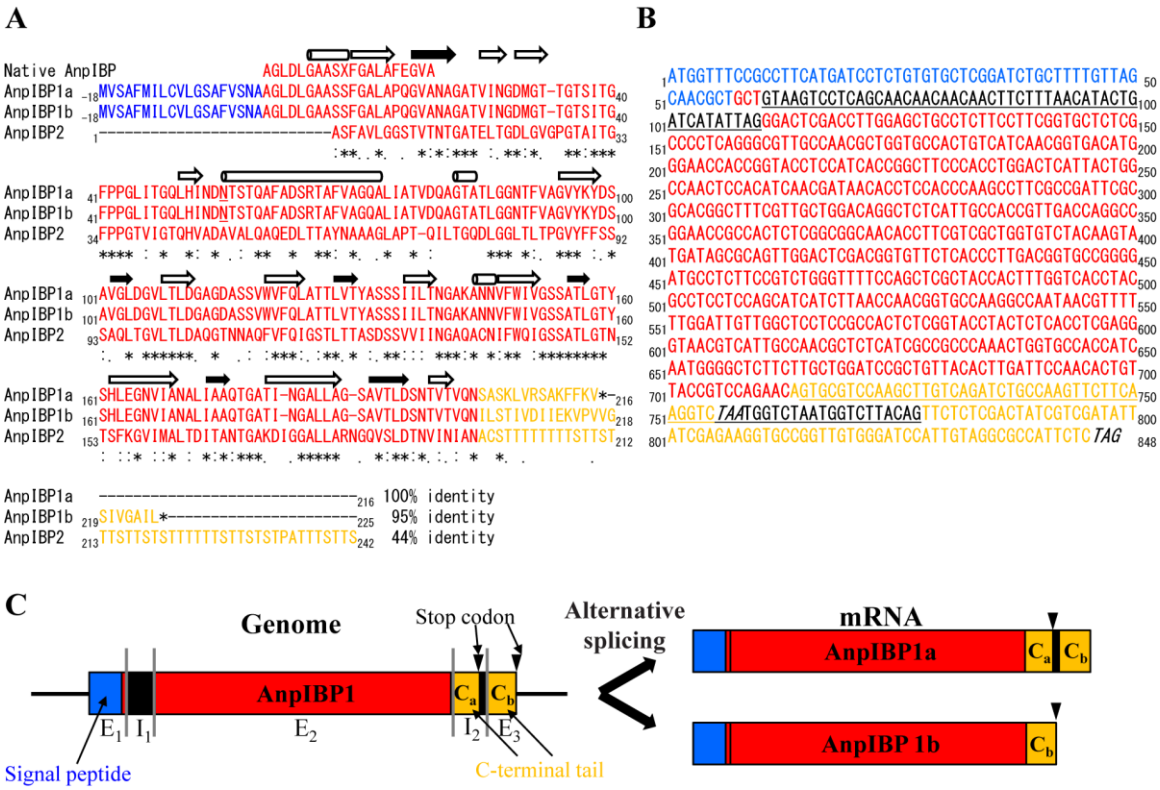

A
B
C
Native AnpIBP
AnpIBP1a
AnpIBP1b
AnpIBP2
AGLDLGAASXFGALAFEGVA
100% identity
95% identity
44% identity
Genome
Stop codon
Alternative splicing
mRNA
AnpIBP1
AnpIBP1a
AnpIBP 1b
Signal peptide
C-terminal tail

Figure 2

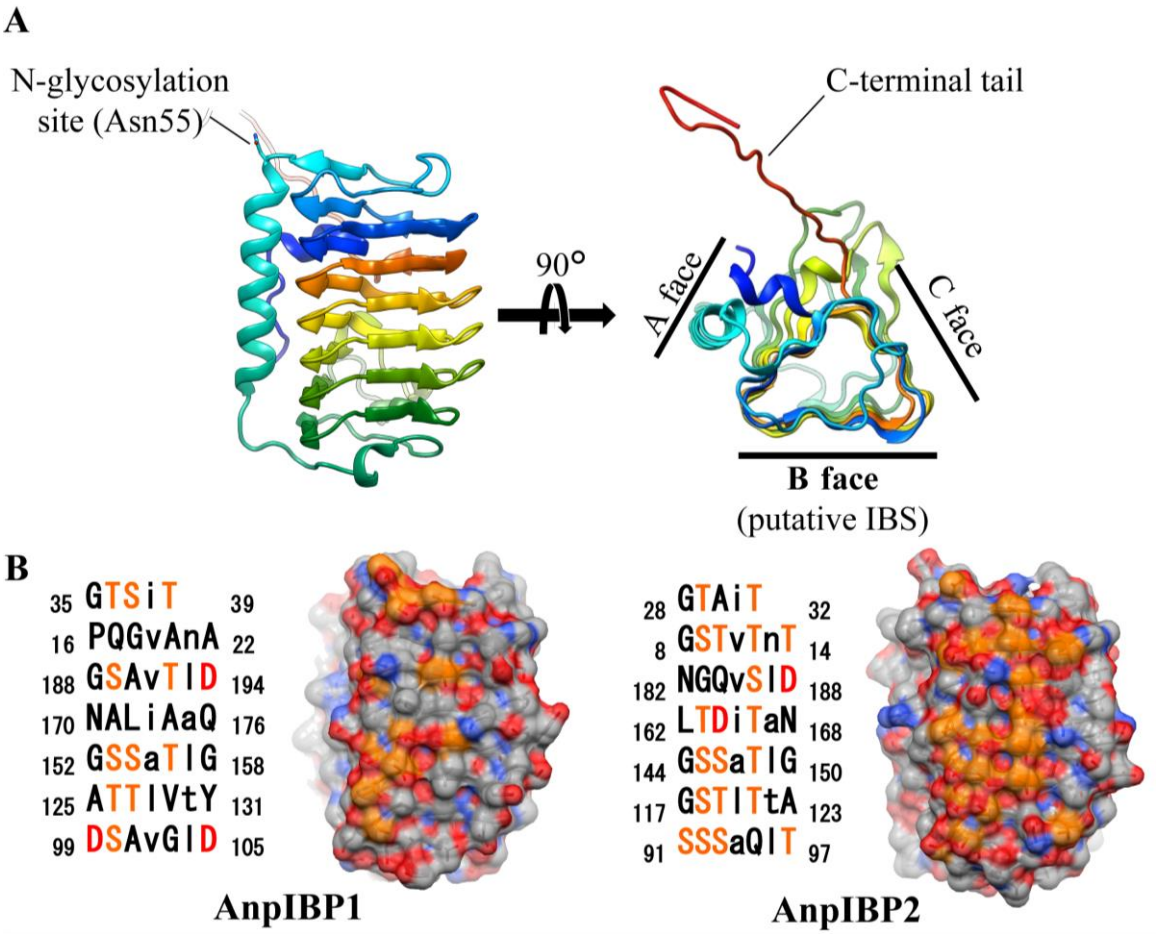


Figure 3

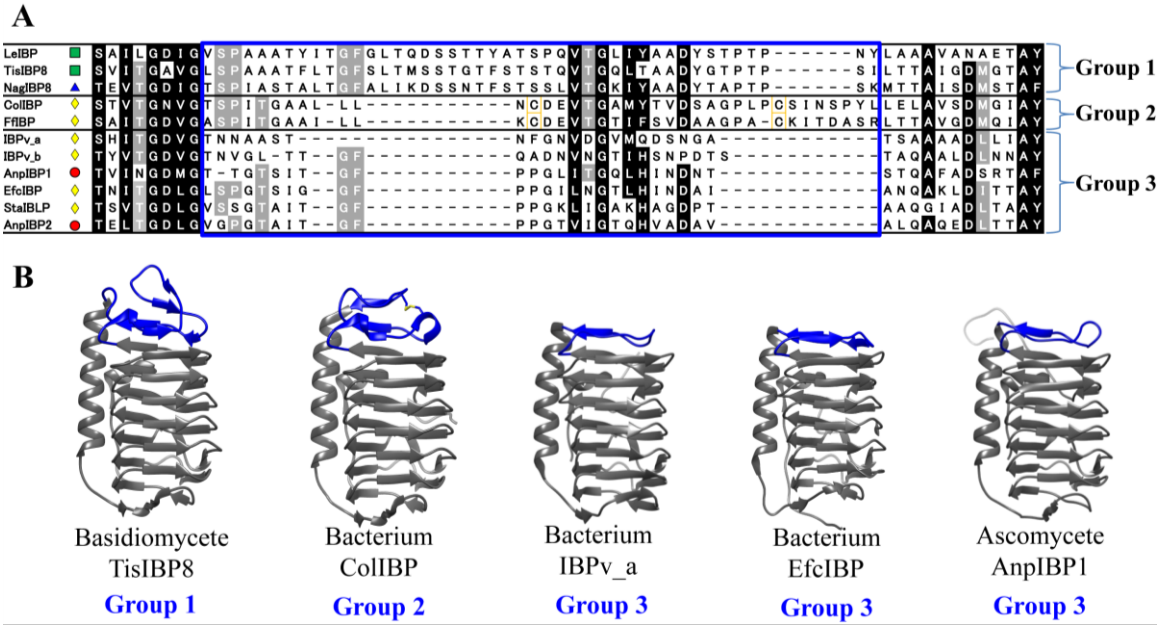


Figure 4

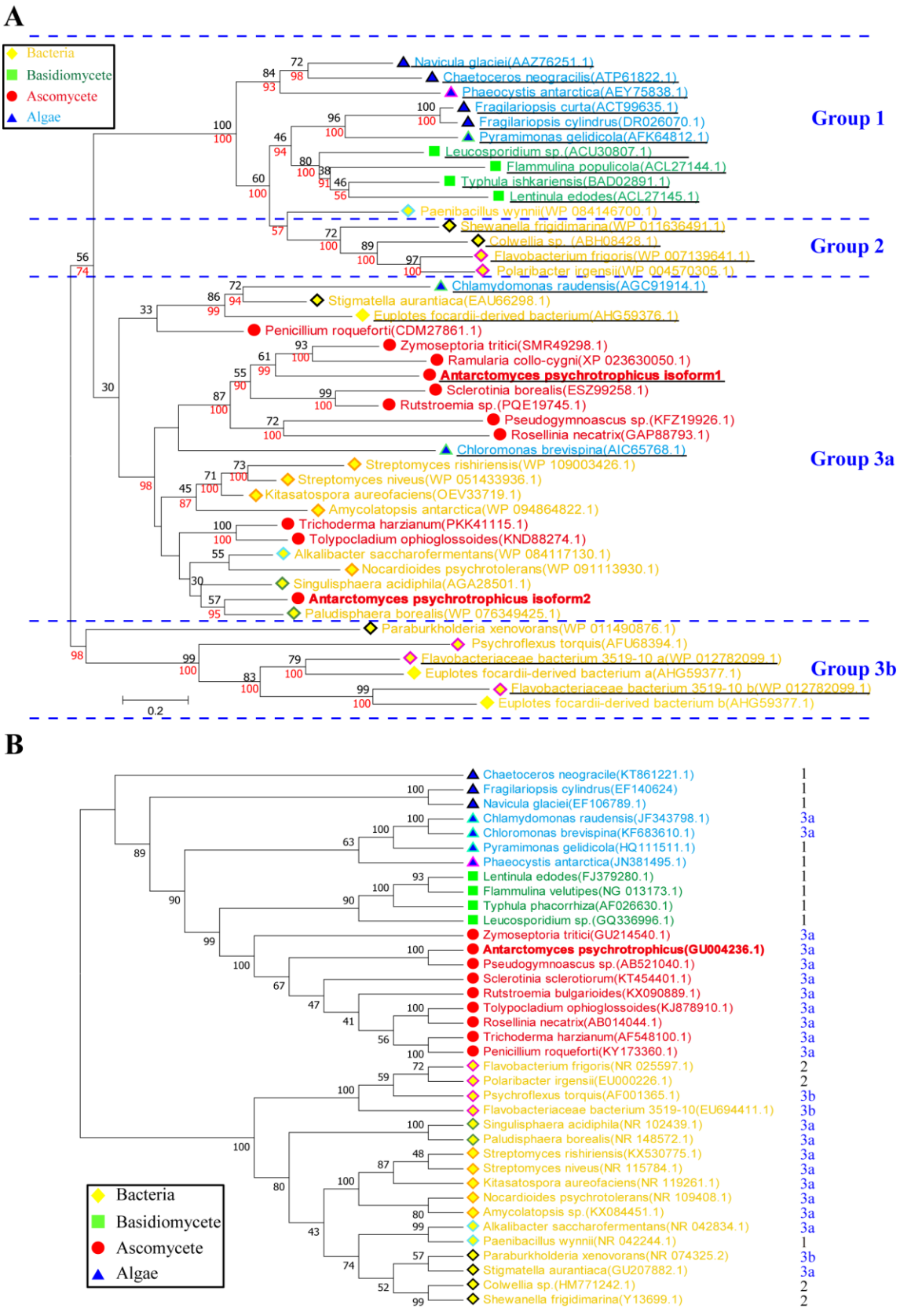
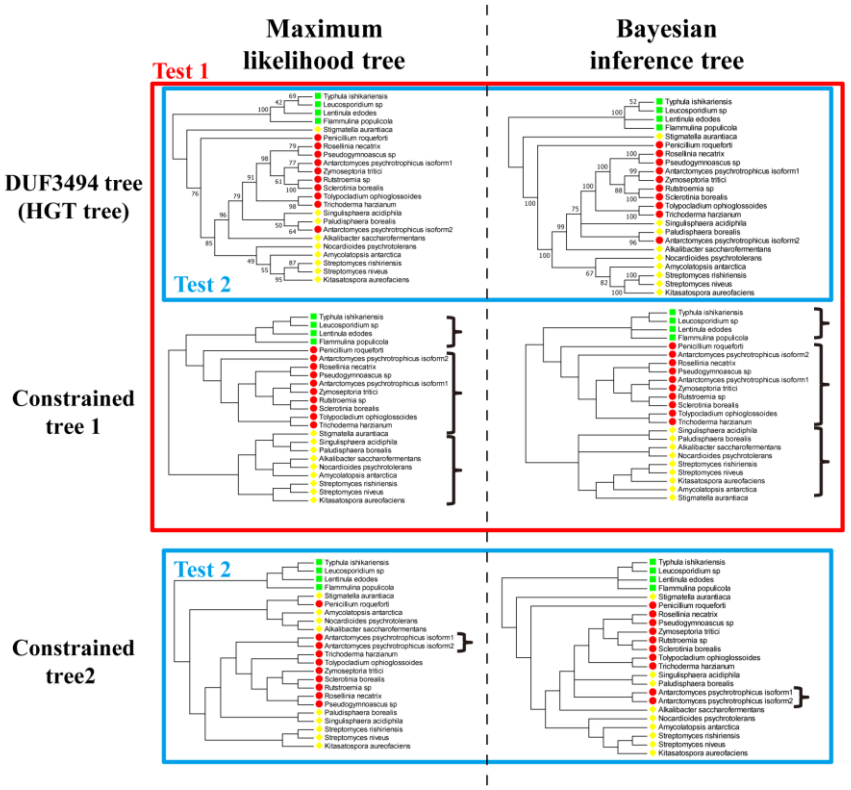

A
Bacteria
Basidiomycete
Ascomycete
Algae
Navicula glaciei(AAZ76251.1)
Chaetoceros neogracilis(ATP61822.1)
Phaeocystis antarctica(AEY75838.1)
Fragilariopsis curta(ACT99635.1)
Fragilariopsis cylindrus(DR026070.1)
Pyramimonas gelidicola(AFK64812.1)
Leucosporidium sp.(ACU30807.1)
Flammulina populicola(ACL27144.1)
Typhula ishkariensis(BAD02891.1)
Lentinula edodes(ACL27145.1)
Paenibacillus wynnii(WP 084146700.1)
Shewanella frigidimarina(WP 011636491.1)
Colwellia sp. (ABH08428.1)
Flavobacterium frigoris(WP 007139641.1)
Polaribacter irgensii(WP 004570305.1)
Chlamydomonas raudensis(AGC91914.1)
Stigmatella aurantiaca(EAU66298.1)
Euplotes focardii-derived bacterium(AHG59376.1)
Penicillium roqueforti(CDM27861.1)
Zymoseptoria tritici(SMR49298.1)
Ramularia collo-cygni(XP 023630050.1)
Antarctomyces psychrotrophicus isoform1
Sclerotinia borealis(ESZ99258.1)
Rutstroemia sp.(PQE19745.1)
Pseudogymnoascus sp.(KFZ19926.1)
Rosellinia necatrix(GAP88793.1)
Chloromonas brevispina(AIC65768.1)
Streptomyces rishiriensis(WP 109003426.1)
Streptomyces niveus(WP 051433936.1)
Kitasatospora aureofaciens(OEV33719.1)
Amycolatopsis antarctica(WP 094864822.1)
Trichoderma harzianum(PKK41115.1)
Tolypocladium ophioglossoides(KND88274.1)
Alkalibacter saccharofermentans(WP 084117130.1)
Nocardioides psychrotolerans(WP 091113930.1)
Singulisphaera acidiphila(AGA28501.1)
Antarctomyces psychrotrophicus isoform2
Paludisphaera borealis(WP 076349425.1)
Paraburkholderia xenovorans(WP 011490876.1)
Psychroflexus torquis(AFU68394.1)
Flavobacteriaceae bacterium 3519-10 a(WP 012782099.1)
Euplotes focardii-derived bacterium a(AHG59377.1)
Flavobacteriaceae bacterium 3519-10 b(WP 012782099.1)
Euplotes focardii-derived bacterium b(AHG59377.1)
0.2
Group 1
Group 2
Group 3a
Group 3b
B
Chaetoceros neogracile(KT861221.1) 1
Fragilariopsis cylindrus(EF140624) 1
Navicula glaciei(EF106789.1) 1
Chlamydomonas raudensis(JF343798.1) 3a
Chloromonas brevispina(KF683610.1) 3a
Pyramimonas gelidicola(HQ111511.1) 1
Phaeocystis antarctica(JN381495.1) 1
Lentinula edodes(FJ379280.1) 1
Flammulina velutipes(NG 013173.1) 1
Typhula phacorrhiza(AF026630.1) 1
Leucosporidium sp.(GQ336996.1) 1
Zymoseptoria tritici(GU214540.1) 3a
Antarctomyces psychrotrophicus(GU004236.1) 3a
Pseudogymnoascus sp.(AB521040.1) 3a
Sclerotinia sclerotiorum(KT454401.1) 3a
Rutstroemia bulgarioides(KX090889.1) 3a
Tolypocladium ophioglossoides(KJ878910.1) 3a
Rosellinia necatrix(AB014044.1) 3a
Trichoderma harzianum(AF548100.1) 3a
Penicillium roqueforti(KY173360.1) 3a
Flavobacterium frigoris(NR 025597.1) 2
Polaribacter irgensii(EU000226.1) 2
Psychroflexus torquis(AF001365.1) 3b
Flavobacteriaceae bacterium 3519-10(EU694411.1) 3b
Singulisphaera acidiphila(NR 102439.1) 3a
Paludisphaera borealis(NR 148572.1) 3a
Streptomyces rishiriensis(KX530775.1) 3a
Streptomyces niveus(NR 115784.1) 3a
Kitasatospora aureofaciens(NR 119261.1) 3a
Nocardioides psychrotolerans(NR 109408.1) 3a
Amycolatopsis sp.(KX084451.1) 3a
Alkalibacter saccharofermentans(NR 042834.1) 3a
Paenibacillus wynnii(NR 042244.1) 1
Paraburkholderia xenovorans(NR 074325.2) 3b
Stigmatella aurantiaca(GU207882.1) 3a
Colwellia sp.(HM771242.1) 2
Shewanella frigidimarina(Y13699.1) 2
Bacteria
Basidiomycete
Ascomycete
Algae

Figure 5



Test 1	AU	SH
HGT ML	0.985	1.000
Con ML	0.016	0.018
HGT Bay	0.994	1.000
Con Bay	0.006	0.008

Test 2	AU	SH
HGT ML	0.988	1.000
Con ML	0.000	0.002
HGT Bay	0.889	1.0000
Con Bay	0.000	0.000

Figure 6

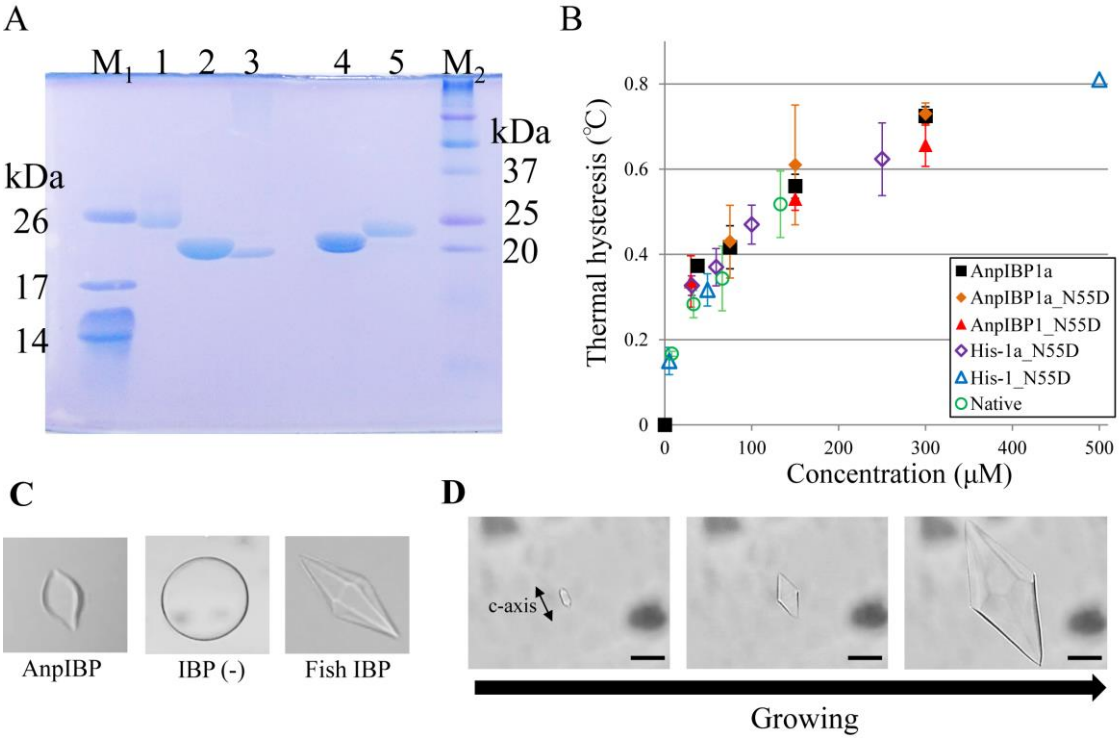


Figure 7

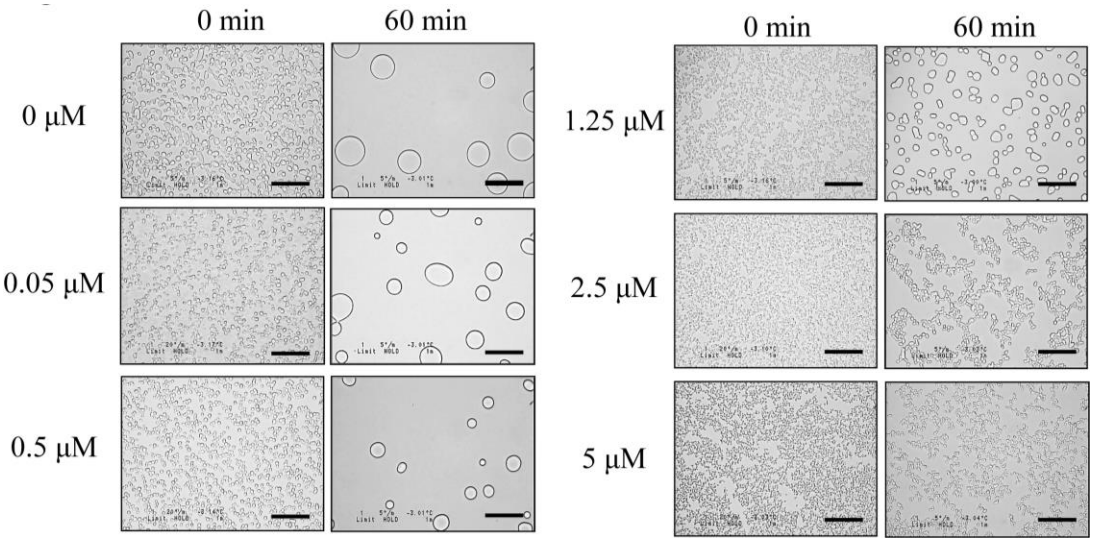


Figure 8