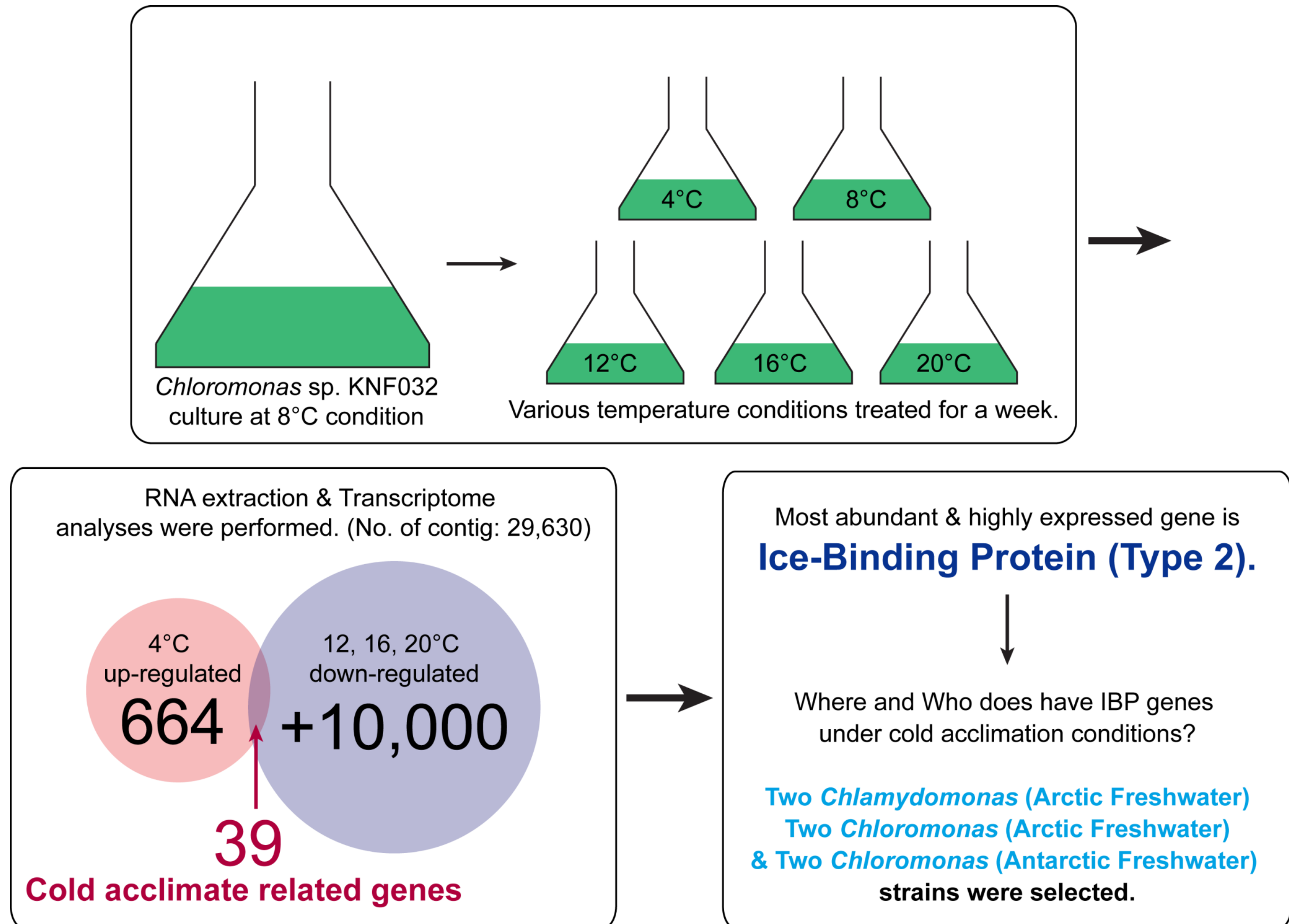


Identification of Ice-Binding Proteins of Arctic/Antarctic Chlorophyceae using Transcriptome Analyses

Sung-Mi Cho, Sanghee Kim, Han-Gu Choi, Hyun Park, and Jungeun Lee*
(jelee@kopri.re.kr)

Unit of Polar Genomics, Korea Polar Research Institute, Incheon 21990 Korea

Workflow



PART I. Phylogeny of KOPRI strains

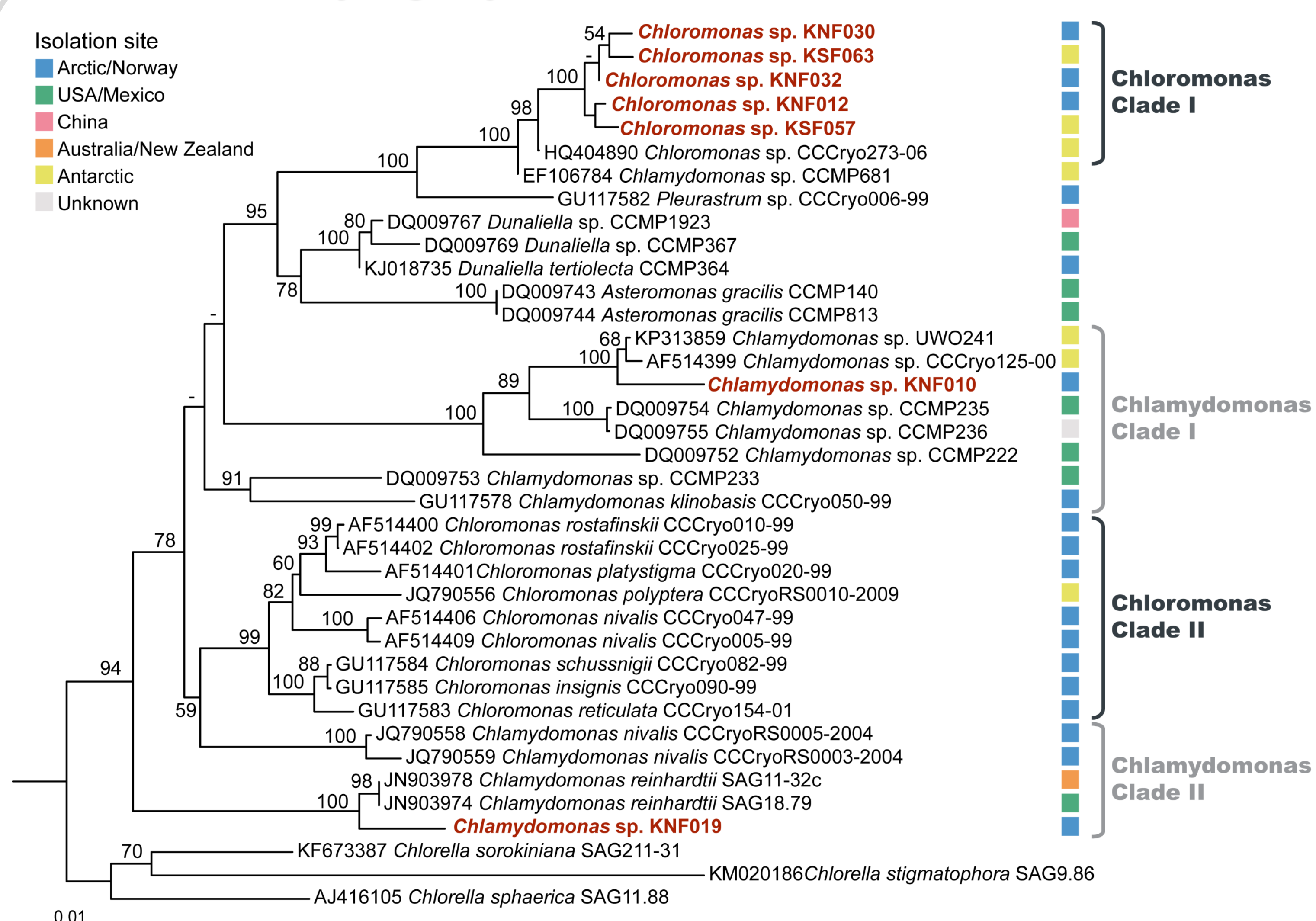


Figure 1. Neighbor-Joining tree of six KOPRI culture collection strains (two *Chlamydomonas* sp. and four *Chloromonas* sp.) inferred from 18S rRNA sequences (1190 nt, 38 taxa). Numbers above branches indicated NJ bootstrap value (>50%). Genus *Chlororella* (Trebouxiophyceae) served as outgroup taxa. Color boxes indicate isolation site of each strain.

Take home messages,

- Three Arctic *Chloromonas* sp. (KNF012, KNF030 and KNF032) and two Antarctic *Chloromonas* sp. (KSF057 and KSF063) are closely related to *Chloromonas* sp. CCCryo273-66 strain isolated from Antarctic region.
- Temperature dependent transcriptome analysis of KNF032 results in 39 cold acclimate-related genes. Five among twelve annotated transcripts are Ice-Binding Protein (IBP). All of 17 IBP transcripts are denoted as type 2 IBPs which has not a DUF3494 domain.
- Type 2 IBP genes are found in Arctic/Antarctic freshwater *Chloromonas* strains. It is remarkable that type 2 IBP has not shared any domains and genetic structures of type 1 IBPs.
- Most of motifs are highly conserved as TFT. First Thr(T) of some motifs is often substituted to Lys(K), Asp(D) and Ser(S) less than second Phe(F) and third Thr(T).
- Predicted TF binding sites give a hint that these ibp genes are able to function as cold acclimation response like in plants.

PART II. Transcriptome analyses

Table 1. Description of 12 annotated transcripts among 39 cold acclimate-related genes. Five ice-binding proteins were remarkably expressed at low temperature condition.

| Contig ID | RPKM score | | | | | Description | E-Value |
|------------------|------------|--------|-------|-------|-------|-------------------------------------------|-----------|
| | 4°C | 8°C | 12°C | 16°C | 20°C | | |
| ArF32Contig690 | 358.41 | 136.89 | 54.62 | 17.87 | 15.03 | ice-binding protein-4 | 0 |
| ArF32Contig2971 | 269.44 | 88.97 | 45.55 | 13.52 | 4.57 | ice-binding protein-3 | 0 |
| ArF32Contig24913 | 147.25 | 41.66 | 24.42 | 32.45 | 31.50 | predicted protein | 1.08E-10 |
| ArF32Contig6739 | 138.63 | 50.19 | 40.04 | 23.75 | 24.78 | ice-binding protein-3 | 2.04E-158 |
| ArF32Contig14333 | 129.43 | 59.49 | 52.95 | 3.56 | 0.43 | hypothetical protein VOLCADRAFT_106879 | 1.45E-09 |
| ArF32Contig6901 | 89.86 | 38.28 | 18.18 | 9.38 | 8.99 | ice-binding protein-4 | 1.42E-81 |
| ArF32Contig11683 | 40.66 | 19.42 | 13.21 | 10.22 | 12.81 | acetyl-coa synthetase | 0 |
| ArF32Contig28905 | 30.54 | 14.58 | 2.52 | 1.19 | 2.48 | hypothetical protein VOLCADRAFT_121013 | 5.12E-48 |
| ArF32Contig21934 | 27.16 | 10.36 | 6.81 | 6.68 | 5.11 | flagellar associated protein | 1.02E-21 |
| ArF32Contig26817 | 20.21 | 6.54 | 1.82 | 0.64 | 0.38 | polysaccharide deacetylase family protein | 1.86E-18 |
| ArF32Contig26619 | 15.41 | 2.44 | 0.94 | 0.99 | 1.58 | ice-binding protein-4 | 6.89E-33 |
| ArF32Contig28814 | 6.35 | 2.83 | 1.45 | 2.55 | 1.22 | hypothetical protein VOLCADRAFT_92225 | 8.75E-23 |

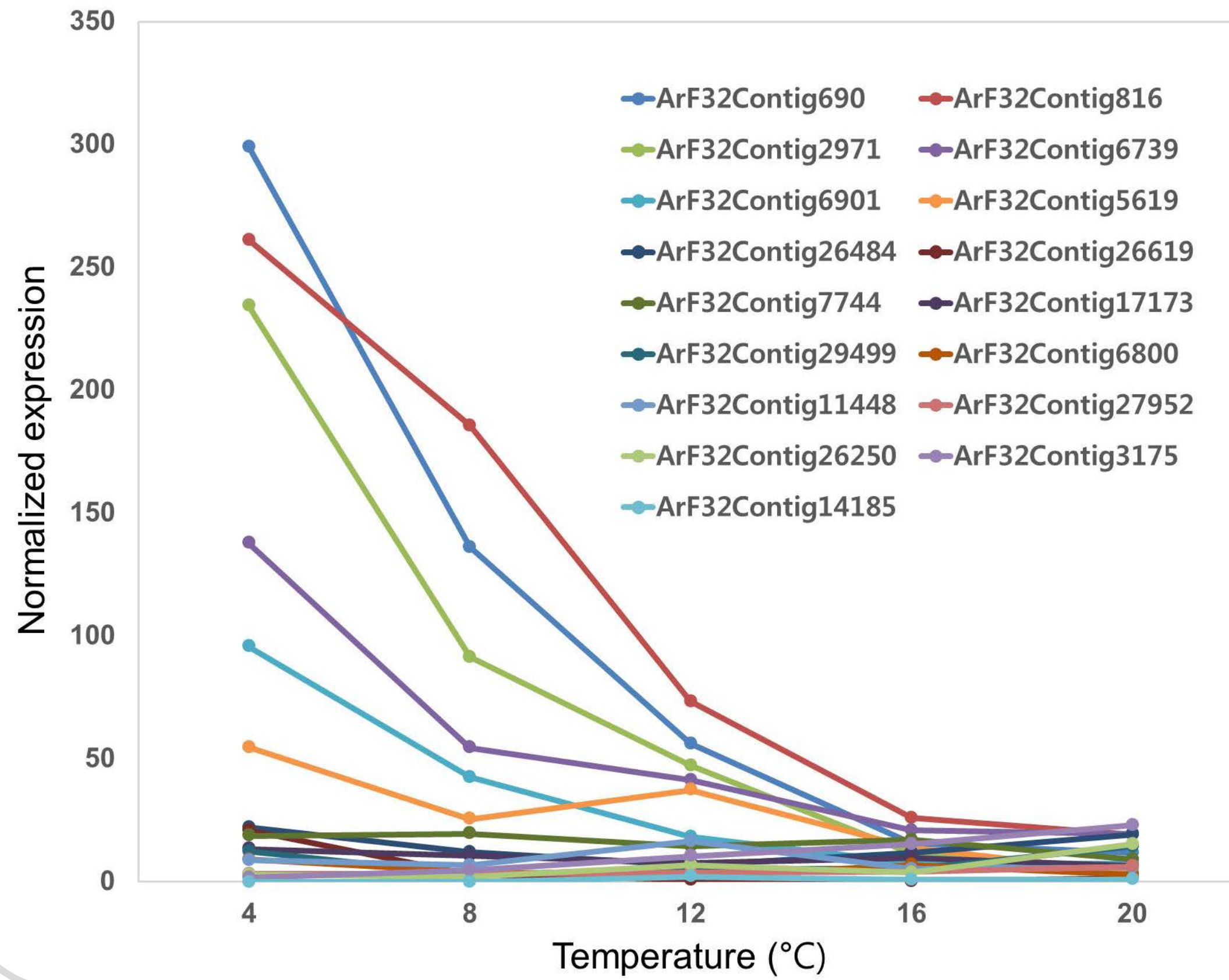


Figure 2. Normalized expression level of all 17 ice-binding protein (IBP) transcripts according to temperature changes. All transcripts were highly expressed at 4°C condition and then decreased according to temperature increasing. Most strongly expressed contig690, 816 and 2971 were selected to further study.

PART III. IBPs of Arctic/Antarctic Chlorophyceae

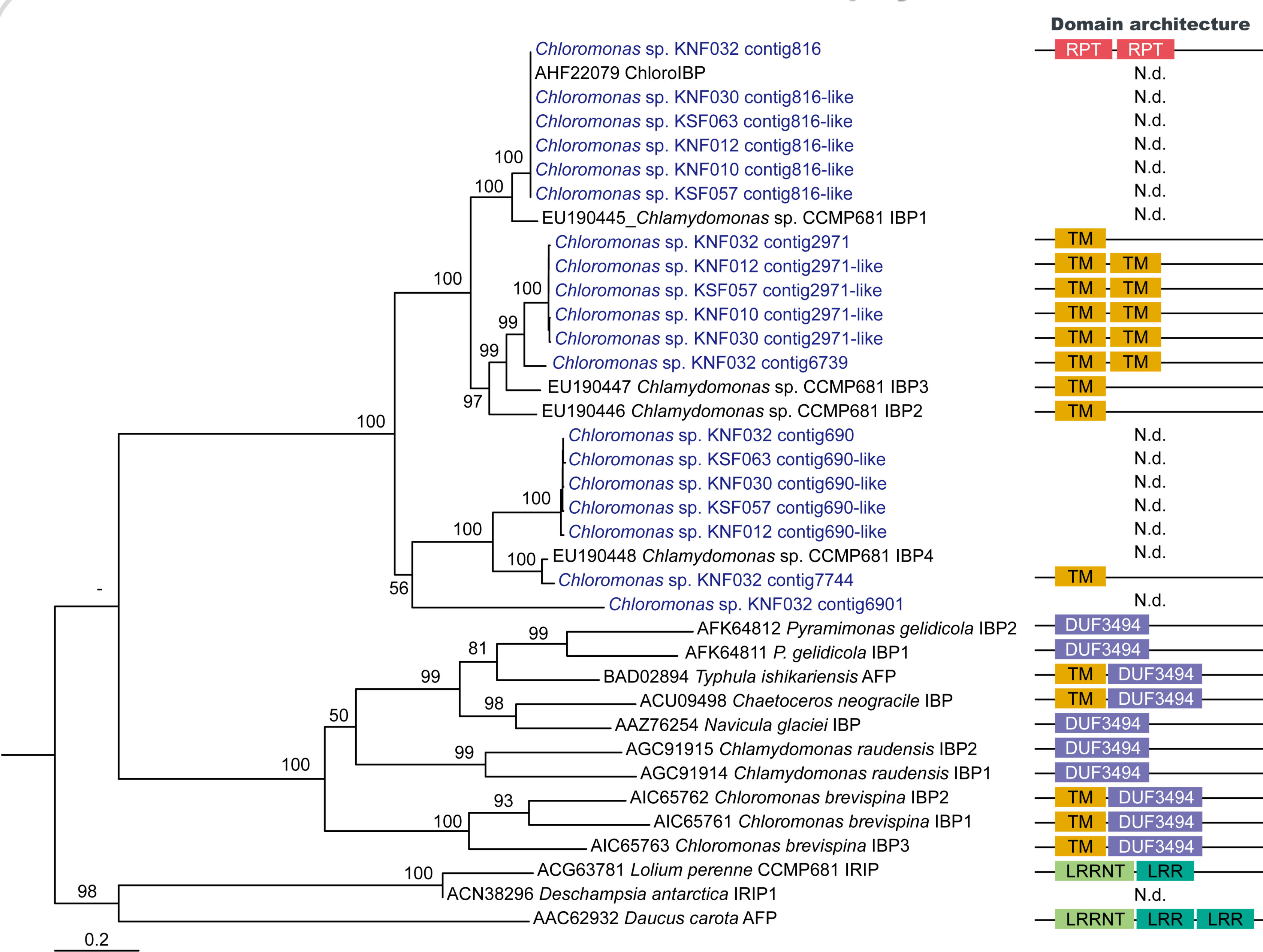


Figure 3. Neighbor-Joining tree of IBP amino acid sequences including 13 newly analyzed IBPs from six KOPRI strains. Numbers above branches indicated NJ bootstrap value (>50%). Anti-freezing protein (AFP) and ice-recrystallization inhibiting protein (IRIP) from Planta served as outgroup. Domain architecture was predicted using SMART homepage with PFAM database. Type 2 IBP was clearly distinguished from type 1 IBPs containing DUF3494 domain. Some proteins have not been detected any domains. Domain name was abbreviated as follow: RPT, repeated domain in UCH-protine; TM, transmembrane; DUF3494, unknown domain; LRRNT, leucine rich repeat N-terminal domain; LRR, leucine rich repeat.

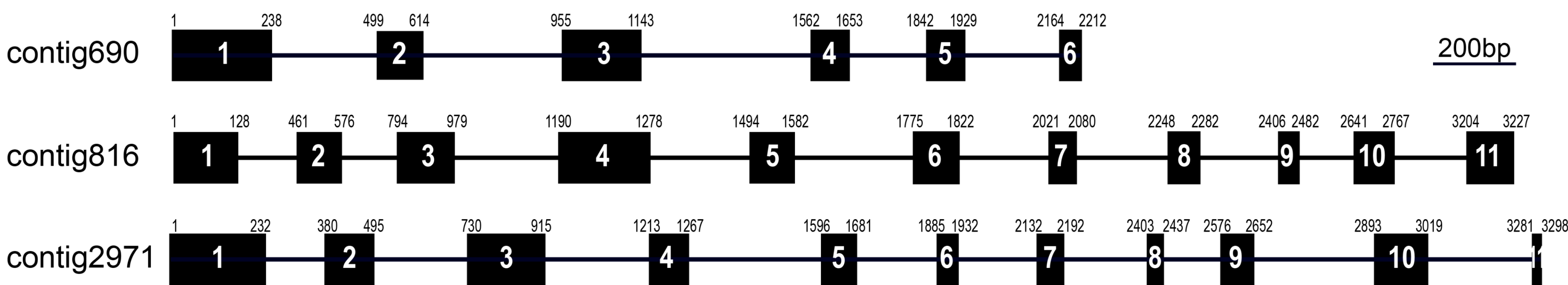


Figure 5. Gene structures of contig690, 816 and 2971 CDS region. A total of 11 exons were found in contig816 and 2971.

| | Domain architecture | | | | | | |
|------------------------|---------------------|-----|-----|-----|-----|-----|-----|
| | I | II | III | IV | V | VI | VII |
| KNF032 contig6739 | TFT | TFT | TFT | TFT | TFT | TFT | TWT |
| CCMP681 IBP3 | TFT | TFT | TFT | TFT | TFT | TFT | TWT |
| KNF032 contig2971 | TLT | TFT | TFT | TFT | TFT | TFT | TWT |
| KNF010 contig2971-like | TLT | TFT | TFT | TFT | TFT | TFT | TWT |
| KNF012 contig2971-like | TLT | TFT | TFT | TFT | TFT | TFT | TWT |
| KNF030 contig2971-like | TLT | TFT | TFT | TFT | TFT | TFT | TWT |
| KSF057 contig2971-like | TLT | TFT | TFT | TFT | TFT | TFT | TWT |
| KNF032 contig6901 | TWT | TFT | TFT | TFT | TFT | TFT | TFT |
| CCMP681 IBP2 | TFT | TFT | TFT | VFT | TFT | TFT | TWT |
| KNF032 contig816 | TFT | TFT | TFT | TFT | TFT | TFT | TWT |
| KNF010 contig816-like | TFT | TFT | TFT | TFT | TFT | TFT | TWT |
| KNF012 contig816-like | TFT | TFT | TFT | TFT | TFT | TFT | TWT |
| KNF030 contig816-like | TFT | TFT | TFT | TFT | TFT | TFT | TWT |
| KSF057 contig816-like | TFT | TFT | TFT | TFT | TFT | TFT | TWT |
| KSF063 contig816-like | TFT | TFT | TFT | TFT | TFT | TFT | TWT |
| CCMP681 IBP1 | TFT | TFT | TFT | TFT | TFT | TFT | TWT |
| ChloroIBP | TFT | TFT | TFT | TFT | TFT | TFT | TWT |
| KNF032 contig690 | TFT | TFT | SFT | TFT | TFT | DFT | KWT |
| KNF012 contig690-like | TFT | TFT | SFT | TFT | TFT | DFT | KWT |
| KNF030 contig690-like | TFT | TFT | SFT | TFT | TFT | DFT | KWT |
| KSF057 contig690-like | TFT | TFT | SFT | TFT | TFT | DFT | KWT |
| KSF063 contig690-like | TFT | TFT | SFT | TFT | TFT | DFT | KWT |
| KNF032 contig7744 | KFT | TFT | SFT | TFT | TFT | DFT | KWT |
| CCMP681 IBP4 | KFT | TFT | SFT | TFT | TFT | DFT | KWT |

Figure 4. Highly conserved TXT motifs discovered in all of type 2 IBP proteins. Multiple alignments were performed by ClustalW program. According to previously studied IBP proteins of *Chlamydomonas* sp. CCMP681 was used to select TXT motifs. Most motifs are expressed as TFT but some motifs at I, III, VI and VII positions were changeable. First Thr(T) is often substituted to Lys(K), Asp(D) and Ser(S) than second Phe(F) and third Thr(T).

Table 2. Predicted transcription factor binding at N-terminus of IBP transcript. Number of transcription factor binding site was reported in parenthesis

| Contig | B3 | bZIP | bHLH | ERF | MYB | Trihelix | WRKY |
|--------|-----------------|--------------------------|---------|--------|----------|----------|-------------------------|
| 690 | ABI3(1), VP1(1) | DPBF(2), OBF4(1), RSG(1) | N.d | CBF(2) | GAMYB(1) | N.d | WRKY18(1) |
| 816 | ABI3(1), VP1(2) | OBF4(4) | N.d | CBF(2) | N.d | GT-1(1) | WRKY1,2,3(3) |
| 2971 | ABI3(1) | N.D | PIF3(3) | CBF(1) | N.d | GT-2(2) | WRKY1,2,3(2), WRKY18(1) |