Manuscript Details

Manuscript number	PSL_2019_951_R1
Title	Influence of CNV on transcript levels of HvCBF genes at Fr-H2 locus revealed by resequencing in resistant barley cv. 'Nure' and expression analysis
Article type	Research Paper

Abstract

Resequencing in resistant cultivar 'Nure' and structural comparison with the same region of susceptible 'Morex' was performed in order to gain a better insight into barley Frost-resistance-H2 locus. Accurate annotation showed copy number variation (CNV) in the proximal part of the locus. In 'Nure', two exact copies of the HvCBF4-HvCBF2A region and one of the HvCBF4-HvCBF2B segment were observed, while in 'Morex' the corresponding region harboured a single HvCBF4-HvCBF2A (22 kb) segment. Abundance and diversity of repetitive element classes, gene function gain/losses, regulatory motifs and SNPs in gene sequences were identified. An expression study of key HvCBFs with/without CNV on selected genotypes contrasting for frost resistance and estimated HvCBF4-HvCBF2B copy number (2-10 copies) was also performed. Under light stimulus at warm temperature (23°C), CNV of HvCBF2A and HvCBF4 correlated with their expression levels and reported frost resistance of genotypes; moreover, expression levels of HvCBF2A and HvCBF14 were strongly correlated (r=0.908, p<0.01). On the other hand, frost resistance correlated to HvCBF14 expression (r=0.871, p<0.01) only after cold induction (6°C) in the dark. A complex interplay of HvCBFs expression levels under different light/temperature stimuli is discussed in light of CNV and presence/number of regulatory elements that integrate different signal transduction pathways.

Keywords	Barley; Frost resistance; Fr-H2 locus; CBF genes; CNV; RT-qPCR
Corresponding Author	Enrico Francia
Corresponding Author's Institution	University of Modena and Reggio Emilia
Order of Authors	Lavinia Mareri, Justyna Milc, Luca Laviano, Matteo Buti, Sonia Vautrin, Stéphane Cauet, Flavia Mascagni, Lucia Natali, Andrea Cavallini, Hélène Bergès, Nicola Pecchioni, Enrico Francia
Suggested reviewers	Tobias Würschum, Troy Frederiks, Shui-zhang Fei, Ana Casas

Submission Files Included in this PDF

File Name [File Type]

Cover Letter for revision.pdf [Cover Letter]

Answers to the reviewer.pdf [Response to Reviewers]

Highlights Mareri.docx [Highlights]

Mareri et al revisioned.doc [Manuscript File]

Fig 1.tif [Figure]

Fig 2 revised.tif [Figure]

Suppl Fig 1 revisioned.tif [Figure]

Conflict of Interest Declaration.pdf [Conflict of Interest]

To view all the submission files, including those not included in the PDF, click on the manuscript title on your EVISE Homepage, then click 'Download zip file'.

Research Data Related to this Submission

Data set   https://www.ncbi.nlm.nih.gov/nuccore/MN251600

The barley (Hordeum vulgare L.) cv. 'Nure' Frost resistance H2 (Fr-H2) locus sequence

Targeted Locus Study. A minimum tiling path (MTP) of 14 BAC clones spanning the entire 'Nure' Frost resistance-H2 genomic region was selected, sequenced using the PacBio RS II platform, and assembled. Sequence annotation was performed using Triannot pipeline, curated TREP datbase, and refined with manual annotation and curation.

UNIMORE
UNIVERSITÀ DEGLI STUDI DI
MODENA E REGGIO EMILIA

Dipartimento di Scienze della Vita

Sede di Modena
Via Giuseppe Campi, 287
41125 - Modena, Italia
T +39 059 2055014 · +39 059 2057143
F +39 059 2057146

Sede di Reggio Emilia
Viale Amendola, 2 (pad. Besta)
42122 - Reggio Emilia, Italia
T +39 0522 522054 · F +39 0522 522053

www.unimore.it
www.dsv.unimore.it

Reggio Emilia, 18/09/2019

Subject: Plant Science manuscript revision

Dear Editor,

Please consider the enclosed revised manuscript entitled: "Influence of CNV on transcript levels of *HvCBF* genes at *Fr-H2* locus revealed by resequencing in resistant barley cv. 'Nure' and expression analysis" for publication in Plant Science. The text has been modified as requested by the Reviewer and all revisions are kept visible in the text. Please note that, for the sake of clarity, the new references numbering had not been tracked. The following references were added to the manuscript according to Reviewer's comments/suggestions:

[11] E. Seo et al., Plant Cell. 21 (2009) 3185–3197. doi:10.1105/tpc.108.063883.

[12] M.M. Alonso-Peral et al., PLoS One. 6 (2011). doi:10.1371/journal.pone.0029456.

[28] R. Nitcher et al., Mol. Genet. Genomics. 288 (2013) 261–275. doi:10.1007/s00438-013-0746-8.

[29] J. Loscos et al., Front. Plant Sci. 5 (2014) 251. doi:10.3389/fpls.2014.00251.

[40] F. Rizza et al., Mol Breed. 36 (2016) 156. doi:10.1007/s11032-016-0571-y.

[49] R. Appels et al., Science. 361 (2018) eaar7191. doi:10.1126/science.aar7191

Yours Sincerely,

Enrico Francia, PhD
Department of Life Sciences
University of Modena and Reggio Emilia

Dipartimento di Scienze della Vita

Sede di Modena
Via Giuseppe Campi, 287
41125 - Modena, Italia
T +39 059 2055014 · +39 059 2057143
F +39 059 2057146

Sede di Reggio Emilia
Viale Amendola, 2 (pad. Besta)
42122 - Reggio Emilia, Italia
T +39 0522 522054 · F +39 0522 522053

www.unimore.it
www.dsv.unimore.it

Reggio Emilia, 18/09/2019

Subject: Plant Science manuscript revision

Dear Reviewer,

Thank You for reviewing our manuscript and for the positive evaluation. The text has been modified as requested and all revisions are kept visible in the text. Please note that, for the sake of clarity, the new references numbering had not been tracked. The following references were added to the manuscript according to Your comments/suggestions:

[11] E. Seo et al., Plant Cell. 21 (2009) 3185–3197. doi:10.1105/tpc.108.063883.
[12] M.M. Alonso-Peral et al., PLoS One. 6 (2011). doi:10.1371/journal.pone.0029456.
[28] R. Nitcher et al., Mol. Genet. Genomics. 288 (2013) 261–275. doi:10.1007/s00438-013-0746-8.
[29] J. Loscos et al., Front. Plant Sci. 5 (2014) 251. doi:10.3389/fpls.2014.00251.
[40] F. Rizza et al., Mol Breed. 36 (2016) 156. doi:10.1007/s11032-016-0571-y.
[49] R. Appels et al., Science. 361 (2018) eaar7191. doi:10.1126/science.aar7191

Following please find our answers (**in bold type**) to your comments listed one by one.

Reviewer Comments:

This is a nice work reporting the characterization of the frost resistance Fr-H2 locus in the barley genotype Nure. The study comprises the development of specific BAC libraries covering the Fr-H2 region, long-read sequencing, assembly of the reads, detailed annotation as well as comparison to the barley reference genotype Morex, which is frost susceptible. The study reveals abundant structural variation in the CBF genes, with CNV, diversity of repetitive elements, and SNPs in the region. The experimental approach is well designed and performed, and the results are clear.

The study includes an expression study of three CBF genes with CNV in a selected panel of genotypes. I have some concerns with this experiment and the conclusions reached in the manuscript. In Figure 2, the authors relate levels of CBF gene expression, according to the number of copies of the genes, with the growth habit of the genotypes and the Vrn-H1 allele they carry. Pamina and Lunet are facultative genotypes that carry 9-11 copies of HvCBF2A and HvCBF4. Nure is the only winter genotype, which has 4 copies of the genes. The rest of genotypes are spring-types, with different VrnH1 alleles, VrnH1-6 (Mellori and Ponente) which have 4-6 copies of those CBF genes, and two spring genotypes Tremois

(VrnH1-7) and Morex (VrnH1-1), both of them with 2 copies. It is well known that a winter genotype, requiring vernalization, contains a dominant winter allele of VrnH2 (ZCCT genes) and a winter allele in VrnH1. The authors do not mention VrnH2 anywhere in the manuscript. The information provided at the bottom of Figure 2, regarding growth habit of the different genotypes, is wrong. It does not agree with the results published by Rizza et al. 2016 (Mol Breeding 36:156), who characterized these seven genotypes, among others. In that paper (Rizza et al.) you can find data corresponding to the genetic and phenotypic characterization of the seven genotypes analyzed in the current work. Their data do not match yours

Thanks for the comments. The indication of the growth habit (Spring/Winter) has been corrected. In fact, the citation Rizza et al 2016 was not included by error in the submitted version of the manuscript and now has been included in the paper.

In my opinion, a conclusion of this study is that higher CBF CNV is observed in facultative genotypes that contain a winter VrnH1 allele, without VrnH2. Nure, Mellori and Ponente, all have VrnH2, somehow precluding the possibility of incorporating more copies of the CBF genes. The lower number of copies are found in the true spring genotypes, Tremois and Morex, null for VrnH2. What is apparent from this manuscript is that facultative genotypes, with a winter vrn1 and no VRN2 show higher CBF CNV. Had this relationship been reported previously? Loscos et al. 2014 showed increased HvFT1 CNV in barley genotypes without VrnH2.

We added the information on VrnH2 allele as suggested; however, as the relationship between this allele and CNV of CBFs had not been previously reported and as our data come from a very narrow genotype panel, we preferred not to stress too much this very putative correlation in the text.

Other comments or suggestions to modify the text:

The standard abbreviation for kilobases is kb, not kbp. It should be modified throughout the text, starting in the Abstract, line 47, then lines 283-288, and so on.

Done

Lines 50-51: "CNV of CBF2A and CBF4 correlated …" the correct name of the genes is HvCBF2A and HvCBF4.

Done

Line 52: expression levels of HvCBF2A and HvCBF14 were strongly correlated, verb is missing in that sentence.

Done

Line 78: What do you mean by 'ruled by'? The major locus responsible for vernalization requirement is VrnH1, a MADS box transcription factor orthologue of Arabidopsis AP1.

Done

Line 88: "…CBF expression is known to be downregulated by VRN-1 (AP1) itself or by a factor in the VRN-1 pathway". Deng et al [10] showed that the VRN1 protein binds to several CBFs (Hvcbf1, hvcbf2 and hvcbf9). The sentence is not clear enough.

Done. Please, also note that we implemented the fragment with more detailed information

Line 90: the citation by Shi et al. 2018 should be numbered [11].

Done

Line 103: replace resistant varieties by frost resistant varieties

Done

Line 106: IDELs should be INDELs

Done

Lines 109-110: "In particular in barley and other Triticeae crops, CNV was identified at Ppd-B1 and Vrn-A1 flower response loci [25], and Boron and Aluminum toxicity tolerance loci [26, 27]. Reference 25 corresponds to hexaploid, bread wheat; 26 is barley but 27 is maize. Is Zea mays a Triticeae crop? Nitcher et al. (2013, MGG 288:261-275) and Loscos et al. (2014, FPLS 5:251) showed CNV in barley VrnH3 (HvFT1), which could be mentioned here.

Done. The citation regarding maize was removed and substituted with the two citations suggested.

Line 116: Replace 'unpredencented' by 'unprecendent', it is not a verb

Done

Line 116-117: Replace 'investigation of' by 'investigating global polymorphisms'

Done

Lines 118-120: I cannot follow the two ways of investigating the role of CNV in frost resistance; 1) detailed resequencing and assembly of CBF elements; 2) assembly of Fr-1 and Fr-2? Something is missing in this sentence

The sentence has been rewritten.

Line 120: '…the genomic structure…'

Done

Line 122: 'the Fr-H2 locus…'

Done

Line 132: 'Young leaves of cv. Nure were harvested after 96 h of etiolation', to construct the BAC libraries. Are these seedlings germinated in the dark or young plants?

Done. The sentence has been rewritten to better clarify the developmental stage of the plants ("Young plants at first-true leaf stage that underwent 96h of etiolation")

Line 141: '…denaturated pool…' It should be '…denatured pools…'

Done

Line 144: Replace 'in' by 'with'. BAC clones were plated on medium with ….

Done

Line 148: Do you mean 'entire' instead of 'intire'?

Done

Line 153: Delate 'After sequencing' and start the sentence 'The excess of …'

Done

Line 169: Please provide a reference for BLASR, maybe Chaisson and Tesler 2012, BMC Bioinformatics

Done

Line 201: Replace 'sequence' by 'sequences'

Done

Line 204: Insert an article before Fr-H2, '…in the Fr-H2 locus…'

Done

Line 213: Replace 'basing' by 'based'

Done

Lines 215-222: Please correct, Mellori and Ponente are winter, not spring. The CNV information of HvCBF2A and HvCBF4 could be added to Figure 2, to facilitate interpretation of the expression results.

Done. Figure 2 has been modified in the c. panel, and information on CNV has been added to the table. The caption of the figure was moreover implemented with details that were lacking

Line 241: '…gel The complementary…' Insert a full stop to separate the words gel and The

Done

Lines 250-251: '…a specific primer for detection of 2A and 2A/B forms (the latter harbored by 'Tremois') only was designed…' If I understood correctly, replace by '…a specific primer for detection of the 2A and 2A/B forms (the latter harbored by 'Tremois' only), was designed …'

Done

Line 262: Can you provide a reference for the GenStat software?

Done

Line 303: What is the correct name for that pseudogene - HvCBFIVΨB or HvCBF4ΨB, as it appears in Supplemental Figure 1?

HvCBFIVΨB is the correct nomenclature and it has been corrected in the Suppl Fig. 1

Line 328: Delete 'a' from the sentence '…'Nure' showed (and confirmed) copy…'

Done

Line 330: See previous comment regarding line 303

Done

Line 331: It should read 'had the HvCBF2B gene…'

Done

Line 335: See the previous comments for lines 303 and 330

Done

Line 337: '… two DNA transposons elements'. It should be two transposable elements

Done

Line 369: The interpretation of Figure 2 would be easier if the gene copy numbers appeared in it

Done

Line 372: The authors indicate that higher expression levels were observed for genotypes with high number of copies, especially for CBF2A. Looking at the Figure, I would say that high expression levels are detected for those same genotypes especially for CBF4, in the middle panel. Even though the correlation between CNV and expression seems higher for CBF2A

Done. The sentence has been rewritten accordingly

Line 387: Is there information from the new wheat genome published in 2018 in Science? It appears a number of 21 CBFs in that paper. The reference of 2005 [45] is accurate but there are more recent results

The detailed number of wheat CBFs (derived from the Science publication) and the suggested citation has been added

Line 394: Check the sentence. Winter growth lines accumulate transcripts at high levels when the plants are in the vegetative phase, or during the vegetative phase

Done. The sentence has been modified as: "…when plants are in the vegetative phase"

Line 426: Knox et al. highlighted the duplication 'of the HvCBF4-HvCBF2'

Done

Lines 437-438: Insert a verb. It should read '…it could be hypothesized…'

Done

Line 439: Instead of MY, do you mean million years? It is not provided as abbreviation

Done

Line 445: It should read 'More copies of the HvCBF4-HvCBF2A segment…'

Done

Line 447: CBF paralogs 'showed or have shown CNV'

Done

Lines 452-453: Something is missing in the following sentence 'Similar results were obtained by Babben et al. [67] in which DNA polymorphisms were significantly…'

The sentence has been rewritten

Line 462: 'copies for the HvCBF4-HvCF2A genomic segment'

Done

Lines 464-465: 'Results obtained showed…' The sentence should clarify that the current results agree with those already reported by Gierczik et al. [17].

Done

Line 466: Make it clear that this is a new contribution from the current study. Yet, Lunet has 10 copies of HvCBF2A but Pamina has 11, as mentioned in lines 215-222

We added Pamina as it was lacking due to a typing error

Lines 472-473: '…circadian pseudo-response regulators binding basic helix-loop-helix transcription factor, Phytochrome-interacting factor 4 – PIF4 [68].' Confusing name of the transcription factor. Is there any punctuation sign missing?

Added the brackets

Line 480: It should read 'CAMTA were identified in the promoters of CBF4…'

Done

Lines 483-485: The sentences should read 'While the majority of replacements were located in the C-terminal region of the proteins, no polymorphisms were identified in the entire sequence of genes coding for CBF2A, CBF4, CBF6B, CBF14 and CBF15A. Only CBF3 and CBF10A presented…'

Done

Line 497: CBF14 expression and temperature/light stimuli 'are relayed via' separate signaling routes. Are relayed by or relayed on?

Done. "relayed by"

Line 505: 'that copy number of the HvCBF4-HvCBF2A…'

Done

Line 512: 'retrieved in the promoters of CBF2C…'

Done

Line 514: '…modulated by the VRN-1 allelic state'

Done

Line 515: VRNH2 is present in vernalization-requiring winter genotypes

Information added

Line 516: VRN1-6 is recognized as a winter, recessive allele. It cannot be considered a spring allele, since it delays flowering without vernalization, as shown by Hemming et al. (2009, MGG, [7]) and Rizza et al. (2016).

Corrected

Line 519: '…in the promoter of HvCBF16,'

Done

Lines 532-533: The authors mention that a constitutive frost resistance mechanism would result in accumulation of protein pools of CBF2A and CBF4 at normal temperatures. Yet, the results of this work show higher expression of CBF genes. Does it translate in higher protein pools? Can the authors provide any reference for that?

Thanks for the interesting comment. However, after careful literature investigation, we did not find any proper reference describing the correspondence between higher transcript level and higher protein accumulation for the CBFs. As the sentence was just a speculation due to the lower binding activity reported in reference [56], we modified the text.

Figure 2: Modify the legend 'Expression patterns of …' Please detail what the different panels a, b and c are. The information in 2c is wrong. According to the results of Rizza et al. 2016, unless the authors have other new information, Pamina and Lunet are facultative, with winter, recessive vrn1 and no VRN2; Nure, Ponente and Mellori correspond to winter growth habit. Although they may have different vrn1

alleles, all of them have VRN2 and show vernalization response. Tremois and Morex are classified as spring growth habit, i.e. dominant Vrn1 and no Vrn2. The information in the Figure should be corrected and in the text.

All modifications have been done as for you suggestion

Yours Sincerely,

Enrico Francia, PhD
Department of Life Sciences
University of Modena and Reggio Emilia

Highlights:

- Differences in number, SNPs and regulatory motifs of *HvCBF* genes at 'Nure' vs 'Morex' *Fr-H2* locus

- CNV: two exact copies of *HvCBF4-HvCBF2A* region and one of *HvCBF4-HvCBF2B* in 'Nure'

- CNV of *HvCBF2A* and of *HvCBF4* correlated with their expression levels (23°C/ light)

- Expression levels of *HvCBF2A* and *HvCBF14* strongly correlated (23°C/ light)

- Frost resistance correlated to *HvCBF14* expression after cold induction (6°C/dark)

1 **Influence of CNV on transcript levels of *HvCBF* genes at *Fr-H2* locus revealed by**

2 **resequencing in resistant barley cv. 'Nure' and expression analysis**

3

4 Lavinia Mareri[1] [*], Justyna Milc[1], Luca Laviano[1], Matteo Buti[1] [§], Sonia Vautrin[2], Stéphane Cauet[2], Flavia Mascagni[3],

5 Lucia Natali[3], Andrea Cavallini[3], Hélène Bergès[2], Nicola Pecchioni[4,1], Enrico Francia[1] [$]

6

7 [1]: Department of Life Sciences, University of Modena and Reggio Emilia, via Amendola 2, Reggio Emilia, I-42122,

8 Italy

9 [2]: Centre National de Ressources Génomiques Végétales (CNRGV), Chemin de Borde Rouge 24-Auzeville CS 52627,

10 Castanet Tolosan Cedex, F-31326, France

11 [3]: Department of Agricultural, Food, and Environmental Sciences, University of Pisa, via del Borghetto 80, Pisa, I-

12 56124, Italy

13 [4]: Research Centre for Cereal and Industrial Crops (CREA-CI), S.S. 673, Km 25,200, Foggia, I-71122, Italy

14

15 [*]: Present address – Department of Life Sciences, University of Siena, via Mattioli 4, Siena, I-53100, Italy

16 [§]: Present address – Department of Agriculture, Food, Environment and Forestry (DAGRI), University of Florence, Via

17 Maragliano, 77, I-50144 Florence, Italy

18

19 [$]: Corresponding author: Enrico Francia, ORCID: 0000-0001-8709-1002

20 E-mail: enrico.francia@unimore.it Tel: +39-0522-522041

21

22 Lavinia Mareri, ORCID: 0000-0002-5729-7971

23 Justyna Milc, ORCID: 0000-0001-6145-5445

24 Luca Laviano, ORCID: 0000-0002-4202-4479

25 Matteo Buti, ORCID: 0000-0002-1430-7826

26 Sonia Vautrin, ORCID: 0000-0003-4446-1426

27 Stéphane Cauet, ORCID: 0000-0003-4985-8940

28 Flavia Mascagni, ORCID:

29 Lucia Natali, ORCID: 0000-0003-3179-5910

30 Andrea Cavallini, ORCID: 0000-0003-4977-2696

31 Hélène Bergès, ORCID: 0000-0002-5492-1062

32 Nicola Pecchioni, ORCID: 0000-0003-1704-2541

33

34 L. Mareri and J. Milc have equally contributed to this work

Highlights:

- Differences in number, SNPs and regulatory motifs of *HvCBF* genes at 'Nure' vs 'Morex' *Fr-H2* locus
- CNV: two exact copies of *HvCBF4-HvCBF2A* region and one of *HvCBF4-HvCBF2B* in 'Nure'
- CNV of *HvCBF2A* and of *HvCBF4* correlated with their expression levels (23°C/ light)
- Expression levels of *HvCBF2A* and *HvCBF14* strongly correlated (23°C/ light)
- Frost resistance correlated to *HvCBF14* expression after cold induction (6°C/dark)

Abstract

Resequencing in resistant cultivar 'Nure' and structural comparison with the same region of susceptible 'Morex' was performed in order to gain a better insight into barley *Frost-resistance-H2* locus. Accurate annotation showed copy number variation (CNV) in the proximal part of the locus. In 'Nure', two exact copies of the *HvCBF4-HvCBF2A* region and one of the *HvCBF4-HvCBF2B* segment were observed, while in 'Morex' the corresponding region harboured a single *HvCBF4-HvCBF2A* (22 ~~kbp~~kb) segment. Abundance and diversity of repetitive element classes, gene function gain/losses, regulatory motifs and SNPs in gene sequences were identified. An expression study of key *HvCBFs* with/without CNV on selected genotypes contrasting for frost resistance and estimated *HvCBF4-HvCBF2B* copy number (2-10 copies) was also performed. Under light stimulus at warm temperature (23°C), CNV of <u>*Hv*</u>*CBF2A* and <u>*Hv*</u>*CBF4* correlated with their expression levels and reported frost resistance of genotypes; moreover, expression levels of *HvCBF2A* and *HvCBF14* <u>were</u> strongly correlated ($r=0.908$, $p<0.01$). On the other hand, frost resistance correlated to *HvCBF14* expression ($r=0.871$, $p<0.01$) only after cold induction (6°C) in the dark. A complex interplay of *HvCBFs* expression levels under different light/temperature stimuli is discussed in light of CNV and presence/number of regulatory elements that integrate different signal transduction pathways.

Abbreviations

BAC, bacterial artificial chromosome; CAMTA, calmodulin-binding transcription activator; CBF genes, C-repeat binding factor genes; CCA1, circadian clock associated 1; CDS, coding sequence; SVs structural variants; CNRQ, copy-number relative quantity; CNV, copy number variation; COR genes, cold-regulated genes; ChIP-seq, chromatin immunoprecipitation sequencing; EF1α, elongation factor 1α; EREBP, ethylene-responsive element binding protein; ERF, ethylene response factor; FR, frost resistance; F_v/F_m, variable (F_v) to maximal (F_m) fluorescence ratio; HWM, high molecular weight; ICE1, inducer of CBF expression; INDELs, small insertions and deletions; LHY, late elongated hypocotyl; LINE, long interspersed nuclear element; LTR, long terminal repeat; MITE, miniature inverted-repeat transposable element; MTP, minimum tiling path; PPD, photoperiod sensitivity; QTL, quantitative trait loci; QV, quality value; RNA-seq, RNA sequencing; SNP, single nucleotide polymorphism; TE, transposable elements; VRN, vernalization requirement.

1. Introduction

Frost resistance (FR) refers to the ability to survive freezing temperatures by limiting plant injuries mainly at vegetative stage, and thus preventing yield losses. Comparative genetic studies in the *Triticeae* tribe of the *Poaceae* family, indicated that the genomic regions most frequently associated with FR and vernalization requirement are located on the long arm of chromosome group 5 [1–3]. Two separate QTL have been identified in barley and wheat [1,4]. The first one, referred to as *Frost resistance-1* (*Fr-1*), co-segregates with *VRN-1*, the major locus responsible for vernalization requirement, underlied~~and ruled~~ by an APETALA1 (AP1)-like MADS-box transcription factor [5]. Varieties carrying dominant *Vrn-1* spring alleles constitutively express VRN-1 to high levels, which confers an intrinsic reproductive competence to these genotypes [6]. In contrast, winter and facultative genotypes carrying recessive *vrn-1* alleles delay the VRN-1 accumulation, which results in a prolonged vegetative growth phase [7]. Moreover, VRN-1 binds to the promoter of its target genes to activate or repress their transcription in barley seedlings. Its potential targets include genes involved in hormone synthesis, reproductive development, and cold acclimation [8]. The second locus, referred to as *Frost-resistance-2* (*Fr-2*), co-localizes with QTL influencing transcript and protein accumulation of cold regulated (*COR*) genes, and segregates with a cluster of C-repeat binding factors (*CBF*) genes [3,9]. In barley, distal *Fr-H1/Vrn-H1*(*HvAP1*) and proximal *Fr-H2* (*HvCBFs*) loci are about 25 cM apart [1], corresponding to ca. 38.5 Mb of physical distance in the *Hordeum vulgare* r1 genome (Phytozome v12.1). Proteins of the CBF family are EREBP transcriptional activators and possess a plant-specific APETALA2 (AP2) protein domain that binds to the C-repeat elements (A/GCCCGAC) present in the promoter of *COR* genes [10]. On the other hand, CBF expression is known to be downregulated by VRN-1 (AP1) itself or by a factor in the VRN-1 pathway [8]. Transcript analysis has shown that VRN-1 directly regulates *CBF* genes and represses their expression, suggesting a negative feedback loop between cold acclimation and flowering transition [11,12]. Moreover, various transcription factors recognizing different *cis*-elements in *CBF* promoter regions and regulating their expression under cold stimulus were identified (see [13] for a review). The upregulated expression of *CBF* genes was observed in low temperature conditions in a wide range of species [14–17], and is dependent on a MYC-type basic helix-loop-helix transcription factor Inducer of CBF expression 1 (ICE1). The regulation also involves members of the calmodulin-binding transcription activator (CAMTA) family and the activation of phytochromes, which play an important role in light-mediated stimulus. Moreover, under warm temperatures, CBFs expression is regulated by the circadian clock through the action of a central oscillator (namely CCA1 and LHY), and day length (photoperiod) [18,19].

During the last decade, several attempts have been undertaken in the *Triticeae* to reveal the underlying molecular mechanism involving the *CBF*s at *Fr-2* [20]. Physical and genetic high-resolution mapping have demonstrated that a common feature in grass genomes is that the *CBFs* are organized in a cluster of tandemly duplicated paralogs. In the economically most important cereals a total of 20, 13 and 37 elements characterize barley, *T. monococcum* and bread wheat genomes, respectively [9,21,22]. A differential expansion of the *CBF* gene cluster together with copy number variation (CNV), instead of mutations within individual gene(s), has been proposed as an explanation for the functional role played by the *CBF*s at *Fr-2* in genotypes showing different levels of frost resistance [18,23]. In particular, Knox et al. [23] and following Francia et al. [24] verified that frost resistant varieties of barley were characterized by a high number of copies for *HvCBF2* and *HvCBF4* genes, and maintained two distinct *HvCBF2* paralogs (*HvCBF2A* and *HvCBF2B*). On the contrary, susceptible genotypes had lower number of *HvCBF2* and *HvCBF4* copies.

Analysis of structural variations – such as CNV, INDELs and movement of transposable elements – was reported to be often related to environmental adaptation in crop plants. Alteration in copy number of genes/gene families caused by natural and man-driven (breeding) evolution may underlie the phenotypic variation of agronomically important traits

113   [25,26]. In particular, in barley and other *Triticeae* crops, CNV was identified at *Ppd-B1* and *Vrn-A1* flower response

114   loci [27]~~,~~ and *Vrn-H3* loci [28,29]~~Boron and Aluminum toxicity tolerance loci [26,27]~~. However, up to now, the number

115   of studies finely investigating the structural organization of *Fr-2* is limited, probably because of the difficulty in

116   sequence assembly caused by the high content of repetitive DNA. In barley, a first draft of *HvCBFs* organization within

117   *Fr-H2* locus has been reported by Knox et al. [23] who sequenced several CBF-harboring bacteriophage λ genomic

118   clones from two winter frost resistant cultivars ('Nure' and 'Dicktoo') and two spring frost susceptible varieties

119   ('Tremois' and 'Morex'). Subsequently, Pasquariello et al. [30] reported the first physical map and sequence of the

120   complete *Fr-H2* locus in the reference genome of cv. 'Morex'. Resequencing provides the unprecedent~~ed~~ opportunity

121   for investigati~~ng~~on ~~of~~ global polymorphism, and ultimately, bridging the gap of mapping all genetic variants to each

122   corresponding causative phenotypic variation in crop plants. Two ways to investigate the role of CNV in frost resistance

123   of *Triticeae* species is given by either skim resequencing of *CBF* elements in large germplasm collections, or detailed

124   resequencing and assembly of *Fr-1* and *Fr-2*~~skim or detailed resequencing and assembly of CBF elements in large~~

125   ~~germplasm collection and assembly of Fr-1 and Fr-2, respectively~~. In this regard, a comparison of genomic structure of

126   barley the *Fr-H2* locus in resistant vs susceptible cultivars has great importance for unravelling the molecular bases of

127   the QTL.

128   The present research aims to better investigate the structure and function of the *Fr-H2* locus in frost resistance, and the

129   putative role played by copy number variation of single *HvCBF*s. In the first part of the work, the complete structure of

130   *Fr-H2* in 'Nure', a winter-hardy frost resistant cultivar, has been described and compared to the previously sequenced

131   'Morex' [30]. In the second part of the work, the structural information was paralleled by a gene expression study on a

132   subset of barley *HvCBF* genes chosen based on their CNV and previous reports. Altogether, the obtained results

133   represent an important step forward in the understanding of the genetic basis of FR in *Triticeae*.

134

135

2. Materials and Methods

2.1. BAC libraries construction and screening

138   ~~Young l~~Leaves of barley cv. 'Nure' were harvested after exposing young plants (first-true leaf stage) to 96 h of

139   etiolation. About 20 g of frozen leaves were used to extract high molecular weight (HMW) DNA following the protocol

140   described by Peterson et al. [31]. HMW DNA was partially digested to obtain two genomic libraries: one with *EcoRI*

141   and one with *BamHI* (New England Biolabs). After two successive size selection steps, digested DNA from a 120-200

142   ~~kbp~~kb size range was eluted and ligated into pAGIBAC1 vectors prepared for high efficiency cloning with *EcoRI* – first

143   library, and *BamHI*– second library (Epicentre Biotechnologies, Madison, Wisconsin). Pulsed field electrophoresis,

144   DNA elutions, ligations and competent cell transformations (using DH10B strains from Invitrogen, Carlsbad,

145   California) were performed according to Chalhoub et al. [32]. The two resulting BAC libraries were named *Hvu-B-Nure*

146   and *Hvu-B-NureBI*. Unique BAC clones were distributed in pools including 1,500-2,000 clones and grown overnight at

147   37°C. A whole genome DNA amplification was performed on 1 µl of each one of these denatur~~ed~~ated pool (Genomiphi

148   v.2 kit, GE Healthcare, Chicago, Illinois). 1:200 diluted DNA pools were screened using previously developed specific

149   PCR markers ([30]; **Supplemental Table 1**) by real-time PCR LightCycler (Roche Diagnostics, Indianapolis, Indiana).

150   For each positive pool identified, around 4,000 BAC clones were plated on LB agar medium supplemented ~~in~~ with 12.5

151   µg/ml chloramphenicol and arranged in microplates. A secondary PCR screening step was done on this subset of clones

152   to identify individual coordinates of positive BAC clones. Additional PCR markers were developed on BAC sequences

153   in order to saturate and anchor the 'Nure' *Fr-H2* locus, and to obtain a minimum tiling path (MTP) of clones spanning

4

154 the ~~i~~entire genomic region. Primer pairs were designed on non-repetitive regions in order to avoid specificity issues.

155 PCR products obtained for each primer pair were sequenced to test the specificity of the new markers. Genomic DNA

156 amplification of the markers reported in **Supplemental Table 1** was carried out under specific validated conditions.

157 Purification of all amplicon products was performed by ethanol/sodium acetate precipitation and 15 ng were then used

158 for sequencing of both strands with the ABI BigDye® Terminator Reaction Ready Kit, Version 3.1 (PE Applied

159 Biosystems, California, USA) according to the manufacturer's instruction. ~~After sequencing, t~~The excess of labelled

160 dNTPs was removed by ethanol/sodium acetate precipitation and purified fragments were separated by capillary

161 electrophoresis using an ABI PRISM® 310 Genetic Analyzer (Applied Biosystems). Finally, the resulting sequences

162 were aligned on reference BAC sequences for validation using the bl2seq utility from the BLAST package (NCBI,

163 http://www.ncbi.nlm.nih.gov/).

164

165 **2.2. MTP sequencing and assembly**

166 Once the address of each BAC clone in the MTP was identified, single colonies were picked and plasmid DNA was

167 extracted using Nucleobond Xtra Midi Kit (Macherey Nagel). Plasmid DNA (2 μg) was pooled with additional BAC

168 clones and a single library was generated from 40 μg of pooled DNA using the standard PacBio library preparation

169 protocol (10 ~~kbp~~kb libraries). This library was sequenced in a PacBio RS II SMRT cell using the P6 polymerase in

170 combination with C4 chemistry at the Institute for Genomic Medicine (IGM) in San Diego, USA. The BAC assembly

171 was performed following the HGAP workflow (https://github.com/PacificBiosciences/Bioinformatics-

172 Training/wiki/HGAP). The work was conducted following the standard operating procedures of the manufacturer

173 (sequencing service provider using Pacific Biosciences PacBio RS II platform was GATC Biotech). The SMRT

174 analysis software v2.3 (http://www.pacb.com/products-and-services/analytical-software/smrt-analysis/), specifically

175 provided by PacBio, was then used to assemble the resulting reads. First, reads were aligned using BLASR [33] against

176 "*E. coli* str. K12 substr. DH10B, complete genome", then the identified *E. coli* reads together with low quality reads

177 (read quality < 0.80 and read length < 500 bp) were removed from data used for the BAC clone sequences assembly.

178 The resulting filtered reads were then preassembled to generate long and highly accurate sequences. To perform this

179 step, the smallest and longest reads (e.g. > 11 ~~kbp~~kb) were separated to correct read errors by mapping the first ones

180 onto the second ones. Obtained sequences were thus filtered against vector sequences, and the Celera assembler was

181 used to assemble the data and obtain a draft assembly. Finally, in the "polishing" step, the last of HGAP workflow, the

182 remaining INDELs and base substitution errors in the draft assembly were significantly reduced. A quality-aware

183 consensus algorithm, the Quiver algorithm, that uses the rich quality scores embedded in Pacific Biosciences bas.h5

184 files was applied (Quiver embedded in The SMRT® Analysis Versions 2.2.0; http://www.pacb.com/support/software-

185 downloads/). In order to calculate the error rate of PacBio RS II technology, the raw reads were aligned on the

186 assembled sequences with BLASR and alignment results were analyzed using Qualimap software. The assembly was

187 deposited in GenBank under accession number MN251600.

188

189 **2.3. Structural and functional annotation**

190 Identification of repetitive elements in the sequenced scaffolds of 'Nure' was performed in two steps using

191 RepeatMasker with the following parameters: –s –x –no_is –nolow. First, the scaffolds were masked against

192 TRansposable Elements Platform (TREP), a curated database of transposable elements (TEs) [34]. Secondly, a house-

193 made library of repetitive elements of barley (cv. 'Morex') was used to identify more elements and to improve the

194 annotation. Additionally, *de novo* identification of long terminal repeat retrotransposons (LTR-REs) was performed by

195   searching structural features with LTR-FINDER [35] and DOTTER [36]. All putative LTR-REs were subsequently
196   annotated using BLASTX and BLASTN against the non-redundant database of NCBI. The annotation of 'Morex'
197   repetitive elements was reviewed as follow. First, the scaffold sequences were analyzed by RepeatExplorer, a
198   computational pipeline accessible by Galaxy platform for the detection of transposable element protein coding domains
199   (https://repeatexplorer-elixir.ceritsc.cz/). The scaffolds were masked against TREP (using RepeatMasker adopting the
200   same parameters that were used for 'Nure') and LTR-FINDER was used for the identification of long terminal repeat
201   retrotransposons (LTR-REs).
202   TriAnnot Pipeline release 4.3.1 (http://wheaturgi.versailles.inra.fr/Tools/Triannot-Pipeline) was used for structural and
203   functional annotation of the sequenced scaffolds using default parameters [37]. Results obtained from TriAnnot were
204   further refined using BLASTN and BLASTX against the non-redundant nucleotide database and BLASTP against the
205   non-redundant protein database. Moreover, InterPro and Pfam web-based tools were used. Nomenclature of full-length
206   *HvCBFs* and truncated/mutated gene elements was inferred from alignments of their nucleotide and amino acids
207   sequence to known barley CBFs. 'Morex' sequences (previously annotated with TriAnnot Pipeline release 3.6) werewas
208   reannotated with the latest version of TriAnnot Pipeline (4.3.1) to make the annotations between the two genotypes
209   more comparable.
210   The cis-elements in the promoters of all *HvCBFs* in the *Fr-H2* locus of 'Nure' and 'Morex' were predicted using
211   ExactSearch, a web-based plant motif search tool for conserved cis-element sequences [38]. The promoter sequences
212   (arbitrarily assumed as 1,000 bp upstream the transcription starting site) were searched for DNA elements/motifs
213   recognized by transcription factors involved in light, circadian clock and cold regulation of *CBF* genes expression [13].
214   Nucleotide and protein similarity of CBF coding sequences were compared for all genes present in both varieties using
215   InterPro (v72.0) to identify polymorphisms in functional domains and regulatory regions [39].
216
217   **2.4. Gene expression studies: selection of genotypes and growth conditions**
218   Seven barley cultivars were used to assess the transcript profile of relevant *HvCBF*s. These genotypes were selected
219   from a panel of 41 accessions with different origin, growth habit and row-type basedbasing on their frost resistance
220   degree [40] and estimated copy number of *HvCBF4-HvCBF2A* genomic region after Francia et al. [24]. Two extreme
221   genotypes were selected: 'Pamina' (facultative, highly resistant) with ten copies of *HvCBF4* and eleven of *HvCBF2A*,
222   and 'Tremois' (spring, highly susceptible) with two copies of *HvCBF4* and the fused paralog *HvCBF2A/B* – 100 %
223   similar to *HvCBF2B* and *HvCBF2A* in its 5' and 3' sequence, respectively. In addition, five genotypes with different
224   CNV/FR levels were considered: reference genome cv. 'Morex' (spring, highly susceptible) with two copies of
225   *HvCBF4* and *HvCBF2A*; 'Nure' (winter, moderately resistant) with four copies of *HvCBF4* and *HvCBF2A*; 'Ponente'
226   (winterspring, moderately resistant) with six copies of *HvCBF4* and four of *HvCBF2A*; 'Mellori' (winterspring,
227   moderately resistant) with six copies of *HvCBF4* and *HvCBF2A*; 'Lunet' (facultative, highly resistant) with nine and ten
228   copies of *HvCBF4* and *HvCBF2A*, respectively. All copy numbers reported refer to diploid genome.
229   Six seeds of each genotype were sown into individual 5 cm diameter pots filled with peat composed of 70% organic
230   matter, 0.6% nitrogen pH 6.0 and with electric conductivity of 1.33 dS m^{-1}. Plants were placed into a controlled growth
231   chamber (Binder KBW 720, Tuttlingen, Germany) under an irradiance of 180 μmol m^{-2} s^{-1} (white fluorescent tubes
232   Fluora 18W/77, Osram, Munich, Germany), and a relative humidity of 60%. Plants were grown under 8 h/16 h (short
233   days, SD) light/dark cycle at 23°C/18°C for 10 days after coleoptile emergence from soil prior to temperature decrease.
234   The soil was kept moist during growth and cold treatment. At day ten, simultaneously with the switch from light to
235   dark, temperature was decreased from 23°C to 6°C and maintained constant for 10 h. Tissue samples were collected (i)

just before switching off the light (23°C, 8 h after dawn), and (ii) after exposure to 6°C for 10 h in the dark. The chosen conditions allowed us to separate the effect of two external stimuli: 10 h of cold treatment in the dark and 8 h after dawn when a peak in *CBF* transcript accumulation was reported by Gierczik et al. [19]. For each sample, three biological replicates made up from the bulked tissue harvested from six seedlings were collected. Once collected, plant tissues were snap–frozen and stored at −80°C until use.

2.5. RNA isolation, quality control, and cDNA synthesis

Total RNA was extracted from 60 mg of crown tissue (1–1.5 cm segment of the white, non-photosynthetic tissue between the upper photosynthetic green shoot and the primary root; [16]) using Trizol™ method (Invitrogen, Carlsbad, CA, USA) according to manufacturer instructions. Genomic DNA was removed by DNase treatment (Invitrogen, Carlsbad, CA, USA). RNA was quantified using Nanodrop ND-1000 spectrophotometer (NanoDrop Technologies, Inc., USA) and its quality was assessed by non-denaturing 1.2% agarose gel. The complementary DNA (cDNA) was synthetized from 500 ng of total RNA using the SuperScript II Reverse Transcriptase (Invitrogen Life Technologies, CA) according to manufacturer protocol.

2.6. RT-qPCR of candidate *HvCBFs*

HvCBF4 and *HvCBF2A* – included in genomic region subjected to CNV, and *HvCBF14* – known to exist in single copy in barley, were chosen for analysis according to previous sequence and functional evidences [23,41]. Owing to Francia et al. [24], both *HvCBF4* and *HvCBF14* exist in a single form, while *HvCBF2* has two forms: *HvCBF2A* and *HvCBF2B*. Only *HvCBF2A* is comprised in the repeated region and subjected to CNV, and thus was considered in the present study for primer design. Therefore, a specific primer for detection of the *2A* and *2A/B* forms (the latter harbored by 'Tremois') only) was designed according to Liu et al. [42] with the 3' end of the forward primer coincident with a SNP [A/G] at position 189 of the CDS, and an additional mismatch introduced at -3 bp from the 3' end of primer to increase the discrimination capacity [43,44]. Primer sequences are reported in **Supplemental Table 2**.

Quantitative real-time PCR was performed in the 7300 Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) in 25 μl reactions containing 12.5 μl of SYBR Green PCR Master Mix with ROX (Applied Biosystems), 2.5 μl of each primer (0.1-0.8 μM), 5 μl of cDNA (2 ng/μl) and 2.5 μl of water. PCR conditions were: 95°C for 10 min, then 40 cycles of 95°C for 15 s and 60°C for 1 min. Three biological and two technical replicates for each sample, along with two negative controls per plate were performed. The relative expression levels were calculated by the ΔCt method [45] using Elongation Factor 1 α (*EF1α*) as reference for normalization [46].

For investigating either simple association or cause-and effect relationships between *HvCBF* genes relative expression obtained in the present study, and CNV values and FR levels (F_v/F_m) from previous studies [24], correlation analysis was performed for any pair of quantitative variables using GenStat for Windows 17th Edition (VSN International Ltd., Hemel Hempstead, UK, 2014). Charts and graphs were prepared using GIMP-2.8 version.

3. Results

3.1 'Nure' *Fr-H2* sequencing and assembly

In order to cover and physically delimit the *Fr-H2* region in 'Nure', two independent non-gridded BAC libraries were constructed: *Hvu-B-Nure* including 814,460 clones representing 17.6X genome equivalents and *Hvu-B-NureBI* including 550,398 clones representing 12.3X coverage. Details of each library (e.g. number of clones, insert size and

277 coverage) are reported in **Supplemental Table 3**. Clones spanning gene-specific markers known to reside at the locus
278 were selected by PCR screening with a set of selected amplicons. An additional set of markers was developed along
279 with the availability of BAC sequences to facilitate the screening process. Specific primer pairs targeting XP-G/RAD2
280 DNA repair endonuclease family (*XPG-I*) gene and a multidrug efflux pump domain (*MatE*) gene allowed identification
281 of the proximal and distal bordering elements of the *HvCBF* cluster, and were used to delimit the *Fr-H2* physical map
282 according to Pasquariello et al. [30]. A minimum tiling path (MTP) of 14 BAC clones spanning the entire 'Nure' *Fr-H2*
283 genomic region was selected for PacBio RS II sequencing. Data generated by the SMRT cell of the PacBio RS II
284 platform (P6C4 chemistry) were processed by the SMRT analysis software v2.3 and assembly was performed following
285 the workflow HGAP. After raw data cleaning and vector reads suppression, quality value (QV) of assembly was
286 calculated. The mean QV range – between 48.25 and 48.87 – indicated a remote possibility of error in the basecall
287 process (**Table 1**). The final coverage, estimated by mapping the raw reads of the assembly contigs with the BLASR
288 software, ranged from 139X for 21H6 to 502X for 186A1. As for the contig total size, the results confirmed the mean
289 size estimated by gel electrophoresis ranging from 107 ~~kbp~~kb for clone 21H6 to 181 ~~kbp~~kb for clone 13J22. Notably, a
290 single contig was obtained for each sequenced clone. The individual sequences of the positive clones were then
291 assembled in two sets of overlapping BACs corresponding to the proximal and distal part of the locus: Contig-I and
292 Contig-II, respectively (**Fig. 1**). Order and orientation of clones were also verified via alignment with the 'Morex' *Fr-*
293 *H2* physical map. The final physical size estimation of the two contigs was 225 ~~kbp~~kb and 730 ~~kbp~~kb, respectively. A
294 single non overlapping clone (32K2, ca. 110 ~~kbp~~kb in size) positive to the putative BRCT-like gene specific probe was
295 positioned in the central part of the locus, nearly closing the sequence gap between the two contigs.
296

297 **3.2 Structural and functional annotation of 'Nure' *Fr-H2* locus in comparison to 'Morex'**
298 The gene composition of 'Nure' *Fr-H2* sequence was annotated with Triannot pipeline [37]. A total of 53 genes were
299 predicted (**Table 2** and **Supplemental Fig. 1**); 19 were full-length *HvCBF* genes, five were *pseudoCBF* genes, two
300 were classified as *Related to AP2 Triticeae*- RAPT genes. All *HvCBF* annotations were retrieved from NCBI GenBank
301 database: full correspondence to 'Nure' *CBF* sequence*s* (Evalue=0; identity=100%; query coverage=100%) for
302 *HvCBF2A*, *HvCBF2B*, *HvCBF4*, *HvCBF14*, *HvCBF15A*, *HvCBF12*, *HvCBF16*, *HvCBF3*, *HvCBF10A*, *HvCBF10B* and
303 *HvCBF6* were obtained. *HvCBF9* was annotated due to its 100% correspondence to 'Dicktoo' ortholog; *HvCBF2C*,
304 *HvCBF12C* and *HvCBF6B* due to their high correspondence to the 'Morex' elements (Evalue=0; coverage 100%;
305 identity=99%). Moreover, *HvCBF13* gene was identified due to its homology to 'Morex' pseudogene *HvCBF13ψ;*
306 however, a functional *HvCBF13* gene was found in 'Nure'. GenBank searches also allowed to identify five *AP2/CBF*
307 sequences that were classified as pseudogenes based on typical degenerative features such as premature stops,
308 frameshift mutations, and truncations. These pseudogenes showed a high similarity (Evalue=0; coverage 100%;
309 identity=97-99%) to 'Morex' *HvCBF13ψB*, *HvCBF14ψ* and *HvCBFIVψB*; therefore, the same nomenclature was used.
310 Finally, twelve putative protein-coding sequences were annotated based on their high similarity to sequences present in
311 databases. In particular, a DNA repair endonuclease XPG-I and a putative glutaredoxin (GRX) protein-coding gene
312 were identified in Contig-I. A gene coding for BRCT (C-Terminal BReast Cancer susceptibility) protein domain and
313 two genes coding for Aspartyl protease were annotated in the central 32K2 BAC clone. Moreover, a phosphoinositide-
314 binding structural domain (PX domain), a DNA-Binding Domain (DBINO) coding sequences a Pre-SET zinc binding
315 motif, a Phosphatidyl-Serine Decarboxylase (PSD) and a Multi Antimicrobial Extrusion (MatE) protein-coding
316 sequences were identified in Conting-II. Finally, 15 hypothetical protein encoding genes were retrieved in the *Fr-H2*
317 locus. In summary, while total gene count was similar between 'Nure' and 'Morex' (53 vs 48), there were 19 and 13

HvCBF genes in 'Nure' and 'Morex', respectively (**Table 2** and **Supplemental Fig. 1**). As expected, no major differences in average CBF gene size were observed between the two genotypes.

Two different libraries of repetitive elements were used for the annotation of 'Nure' repetitive sequences; the first one was based on reviewed 'Morex' annotation [30], while the second one was based on curated TREP database of transposable elements (TEs) [34]. As expected, low-complexity sequences represented the majority of the 'Nure' Fr-H2. Class I retroelements and Class II DNA transposons accounted together for around 65% of the locus sequence (**Table 2**). Among Class I retroelements, Copia elements were the most represented, followed by Gypsy and LINE. The most abundant elements among Class II DNA transposons were CACTA and Harbinger. Similarly, in 'Morex' repetitive elements accounted together for the major part (ca. 73%) of the sequence. Comparative analysis of the composition of each TE class showed interesting differences. While for Class I retroelements, Gypsy and LINE were numerically similar in the sequence of both varieties (135 vs 136 and 10 vs 9, for 'Nure' and 'Morex' respectively), total sequence length of Gypsy elements was double in 'Morex' (39.4 % of Fr-H2). As for Class II DNA transposons, CACTA elements were the most represented type of DNA transposons accounting for 8.90% and 4.58% of 'Morex' and 'Nure' sequence, respectively. On the other hand, Harbinger and MITE elements were more abundant in 'Nure' than in 'Morex' representing 1.90% and 0.21% vs 0.79% and 0.07 % of the sequence, respectively (**Table 2**).

A comparison of the general organization of Fr-H2 sequence in the two cultivars revealed several macroscopic structural variations (**Supplemental Fig. 1**). In the proximal part of the locus, 'Nure' showed (and confirmed) ~~a~~ copy number variation of the HvCBF4-HvCBF2A region. The segment was around 22 ~~kbp~~kb long and included four genes in the following order: HvCBF4-HvCBFIVψ-PutativeGRX-HvCBF2A and was repeated twice, while the third repetition had the HvCBF2B gene form instead of HvCBF2A. Interestingly, insertion of repetitive elements in the promoter regions (< 1,000 bp) of HvCBF2A and HvCBF4 characterized 'Nure' Fr-H2. For example, a full-length, 5,236 bp-long Copia retrotransposon was identified in the promoter region of the proximal copy of the HvCBF4 in 'Nure', 662 bp before the beginning of the coding sequence, while it was absent in 'Morex'. In the other two copies of HvCBF4, this Copia retroelement was instead present in a truncated form. The 3'UTR region of the HvCBFIVψ pseudogene contained a MITE element of 275 bp, while its 5'UTR harbored two Harbinger Rong elements of 234 bp and 4,503 bp. PutativeGRX had in its 3'UTR two DNA transposo~~sablens~~ elements, Mariner and Mutator (210 and 273 bp long, respectively). The corresponding region in 'Morex' was 16 ~~kbp~~kb long, and composed of the same genes, in the same order, with no repetitions. Moreover, 'Morex' PutativeGRX gene did not include any Mariner and Mutator elements in its UTR regions. Another kind of SV found in Contig-II compared to the reference genotype 'Morex' was represented by sequence inversions (**Supplemental Fig. 1**). A 20 ~~kbp~~kb segment with gene order HvCBF6B-HvCBF13-HvCBF3 was found in 'Nure', while in 'Morex' this region was shorter (ca. 15 ~~kbp~~kb) and harbored HvCBF3-Hypothetical protein-HvCBF13ψA-HvCBF6B. A second inversion of around 20 ~~kbp~~kb characterized the genomic region containing Pre-Set and MatE genes in the most distal part of the locus. Moreover, presence/absence variation of single (pseudo)genes between the two genotypes was revealed: 'Nure' lacked the HvCBF6ψB pseudogene and harbored both HvCBF10A and HvCBF10B, while 'Morex' had only HvCBF10A. Finally, in 'Nure' only 2 RAPT genes were detected, while in 'Morex' 3 RAPT genes dispersed along the Fr-H2 region were reported.

The coding sequence of all HvCBFs in common between 'Nure' and 'Morex' were compared and identified amino acid substitutions/polymorphisms were summarized in **Supplemental Table 4**. Most of replacement features could be classified as conservative replacements in the C-terminal region of the protein. While genes coding for CBF3 and CBF10A presented amino acid replacements within the AP2/ERF DNA binding domain, no polymorphisms were found in the entire sequence of genes encoding CBF2A, CBF4, CBF6B, CBF14, CBF15A.

359   Aiming at identification of putative regulatory elements within the promoter region (i.e. 1,000 bp upstream the

360   transcription starting site) of all *HvCBFs*, conserved cis-elements were predicted in both 'Nure' and 'Morex' using the

361   ExactSearch tool [38]; the identified motifs are presented in **Table 3**. A relatively high number of MYC elements for

362   the ICE1 transcription factor was revealed in all *HvCBFs* of both varieties, while in most cases other motifs were

363   present once/twice. Interestingly, genes in the central part of the locus harbored the conserved CCGAC CRT/DRE

364   regulatory cis-element that could be recognized by the CBF proteins themselves in a putative auto-regulatory

365   mechanism.

366

367   **3.3. Gene expression analysis of *HvCBF* candidate genes**

368   With the aim of evaluating the putative effect of CNV and variability in the promoter regions on *HvCBFs* expression, a

369   quantitative real time approach was applied to determine the relative gene expression levels of selected candidate genes.

370   Two genes with CNV (*HvCBF4* and *HvCBF2A*) and one with no CNV reported so far *(HvCBF14)* were chosen for

371   analysis. The timings were accurately chosen to consider one external stimulus known to induce *CBF* expression at a

372   time in order to investigate the putative additive contribution of CNV to induction. Eight hours after dawn (light, no

373   cold stimulus) in short-day grown plants was thus selected as the time point where the maximum daily expression for

374   *CBF* coding genes was already reported [19]. Timing at ten hours at 6°C without light stimulus was chosen to evaluate

375   the relative expression under the cold induction. As shown in **Fig. 2** and in **Table 4**, a correlation between the CNV of

376   *CBF2A* and *CBF4* genes and their relative expression could be observed 8 h after dawn. In presence of light, with no

377   cold stimulus, higher expression levels were observed for genotypes with higher number of copies (e.g. 'Pamina' and

378   'Lunet') ~~especially for *CBF2A*~~. In general, a higher accumulation of *CBF2A* and *CBF4* transcripts was observed under

379   light condition without cold stress application respect to cold stress imposed under dark in all seven genotypes. On the

380   other hand, the cold treatment (in the dark) caused high accumulation of *HvCBF14* transcript especially in varieties

381   characterized by high FR values. Thus, the expression level of *HvCBF14* during cold treatment is not correlated to its

382   copy number, but driven by other factors.

383

384

385   **4. Discussion**

386   The predicted climate change, where more frequent weather extremes will affect cereal productivity, will have to be

387   faced not only by increasing yield potential of current germplasm, but also by improving yield stability through

388   enhanced tolerance to abiotic stresses, among which, low temperature is one of the most harmful. Frost affects all

389   aspects of molecular functions thus compromising plant physiology and reducing crop yield [47]. In barley, similarly to

390   other temperate cereals (e.g. *T. monococcum* and *T. aestivum*), a central role in the acquisition of FR is played by the

391   two major loci *Frost resistance-H1* (*Fr-H1*) and *Frost resistance-H2* (*Fr-H2*).

392   Cereal species have a large and complex *CBF* family with up to 25 different *CBF* genes [21]. The large number of

393   *CBFs* harbored by wheat and barley genomes are similar [48], in wheat the recent manual curation annotated 17 *CBF*

394   genes from 5A, 19 *CBF*s from 5B, and 18 *CBF*s from 5D chromosomes. It also included genes from 6A (3), 6B (2), 6D

395   (3) [49]. On the other hand, ~~while~~ there are only ten and six CBF/DREBs in rice and Arabidopsis, respectively. It is not

396   known why freezing tolerant/resistant cereals have evolved and maintained so many *CBFs* [21]. More than 13 distinct

397   *CBF* CDSs reside at barley *Fr-H2*. Many are present in two or more copies of identical paralogs indistinguishable in

398   their coding sequence [23,24]. Interestingly, transcripts for eight of the 13 *CBF* elements were consistently detectable

399   by RNA blot hybridization [16,48]. In fact, in contrast to Arabidopsis and tomato, in which high transcript levels rise

only after plants are exposed to cold, relatively high levels are detected at normal growth temperatures in cereals [16,21,50–56]. Moreover, winter growth habit lines accumulate transcripts at relatively high levels when the plants are in the vegetative phase but they do not sustain those levels after the plant has transitioned to the reproductive phase [16,57].

Resequencing provides an opportunity for the investigation of global polymorphism, population structure, and ultimately, bridging the gap of mapping all genetic variants to each corresponding causative phenotypic variations in crop plants. In this regard, a comparison of genomic structure of *Fr-H2* locus in resistant *vs* susceptible cultivars has great importance for unravelling the molecular bases of the QTL. In the present work, we aimed to generate a high quality sequence of *Fr-H2* locus in 'Nure' a winter-hardy frost resistant genotype. The resulting sequence was thus compared with the corresponding one available for the frost susceptible reference genotype 'Morex' [30] in order to identify the structural variants that may explain the determinants of the QTL. The two genotypes belong to different breeding material (North American vs European) and differ for many characteristics. 'Morex' ('Cree' x 'Bonanza'; released by the Minnesota Agriculture Experiment Station, USA in 1978) is a six-row malting variety selected as the reference genotype for the International Barley Genome Sequencing Consortium. It has spring (*Vrn-H1/vrn-H2*) growth habit and is highly susceptible to frost. 'Nure' ('Baraka' x ('Fior40' x 'Alpha'²); released by the Istituto Sperimentale per la Cerealicoltura, Italy in 1998) is a two-row feeding cultivar. It has winter (*vrn-H1/Vrn-H2*) growth habit and is moderately resistant to frost.

To obtain the sequence of 'Nure' *Fr-H2* locus, a clone-by-clone (or hierarchical shotgun) sequencing approach was used. Sequencing of selected clones was performed by PacBio RS II technology, a third-generation sequencing technology. Such method of sequencing has several advantages over the other available sequencing techniques. Compared to the pyrosequencing technologies whose read length is 200-400 bp, PacBio produces much longer reads (average > 10,000 bp, some reads 60,000 bp) without the need of any amplification step [58]. These two characteristics greatly facilitate the assembly of *de novo* genomes and identification of SVs. The advantage of using PacBio technology becomes evident by comparing the assembly metrics of *Fr-H2* locus of 'Nure' and 'Morex', whose final sequence was obtained with the Roche/454 GS FLX platform. For 'Morex', the average read length was 358 bp and, for each clone, a variable number and size of sequence contigs were obtained (ranging from 3 to 29 contigs per clone and 122.4 ~~kbp~~kb to 502 bp per contig). Conversely, for 'Nure' the average read length was much longer 7.5 ~~kbp~~kb and for all sequenced clones (except one) a single contig was obtained. In addition to this, contig size was uniform ranging from 107 ~~kbp~~kb to 181 ~~kbp~~kb. Compared to the final assembly of 'Morex' *Fr-H2* (42 scaffolds, 26 oriented), 'Nure' sequence is covered by two oriented scaffolds and a single clone in the central part of the locus. The general composition of *Fr-H2* sequences was similar in both the cultivars with the repetitive DNA elements being the most represented. The structural and functional annotation of genes showed that in 'Nure' a total of 19 *HvCBFs* was predicted against 13 *HvCBFs* for 'Morex'.

Variable copy number of specific genes involved in adaptation may be advantageous to face the threats posed by environmental changes. Knox et al. [23] already highlighted the duplication of the *HvCBF4-HvCBF2* genomic region in 'Dicktoo' and 'Nure' resistant genotypes by sequencing genomic lambda clones. Afterwards, Francia et al. [24] estimated the copy number relative quantity of all *HvCBF* elements residing at *Fr-H2*. However, in the present study, a single sequence covering the proximal region of the locus was obtained showing the presence of two identical copies of the *HvCBF4-HvCBF2A* region combined with a third *HvCBF4-HvCBF2B* segment. Insertion of TEs close to genes can have a strong influence on their expression (e.g. by providing novel *cis*-acting regulatory enhancer sites or by modifying chromatin state; [59]), affecting plant phenotype by a variety of genetic mechanisms [60–62]. A full-length, 5,236 bp-

11

long *Copia* retrotransposon was identified in the promoter region of the proximal copy of the *HvCBF4* gene in 'Nure' *Fr-H2* locus, 662 bp before the transcription starting site, as already reported by Knox et al. [23]. According to the substitution rate calculated for barley genes [63,64], and to length and number of mutations between 5'- and 3'- long terminal repeats, the mentioned *Copia* element insertion could be dated at approximately one million years ago. Since in the present study only a part of the retrotransposon sequence was found in the two distal copies of *HvCBF4*, it could be hypothesized that the *HvCBF4-HvCBF2* segment replication in 'Nure' started from the proximal copy, and occurred in a time following the retroelement insertion, and thus during the last ~~MY~~million years. It would be interesting to verify the presence of this element in a large germplasm collection including both winter and spring genotypes. Moreover, additional sub-microscopic SVs were also found as inversions in the distal *HvCBF* cluster (**Supplemental Fig. 1**). The expected availability of a high-quality gene repertoire derived from the sequence assembly of the barley pan-genome [65] will provide a catalogue of structural variation between diverse genotypes.

In barley, Muñoz-Amatriaín et al. [66] found that winter cultivars/accessions harbored less copies of *HvCBF3* than spring ones. More copies of the *HvCBF4–HvCBF2A* segment in the proximal subcluster and less copies of *HvCBF3* in the distal subcluster of the *Fr-H2* locus were harbored by the most resistant genotypes [24]. On the other hand, different subgroups of *CBF* paralogs show~~n~~ed CNV at the *Fr-A2* locus in wheat (diploid and polyploid) species [67–69], while no reports are available on copy number variation in rye (*Secale cereale* L.). Würschum et al. [70] examined a panel of 407 diverse European winter wheat varieties genotyped by GBS method and performed association mapping of winter hardiness. Although copy number variation at *CBF-A14* explained most part of phenotypic variance, two other causative polymorphisms – within *Fr-A2* yet far from the gene – were identified, suggesting that the causal SNPs might be in one of the other *CBFs* or in their regulatory regions. Similar results were obtained by Babben et al. [71] who identified DNA polymorphisms in *CBF-A3*, *CBF-A13* and *CBF-A15* along with *CBF-A14* that were significantly associated with frost resistance.~~in which DNA polymorphisms were significantly associated with frost resistance in CBF-A3, CBF-A13 and CBF-A15 along with CBF-A14.~~ Moreover, SNPs/INDELs in *CBF-A13* were shown to disrupt the protein structure. What is interesting, as observed by Knox et al. [23] and confirmed in the present study, a functional copy of *HvCBF13* was found in 'Nure' while a pseudo-*CBF13* was present in 'Morex'. Thus, even if several evidences confirmed the involvement of *CBF14* in frost resistance of *Triticeae*, it appears clear that also other *CBF* elements are involved in a complex interplay based on copy number variation of different segments/genes, presence-absence variants of *CBF* elements/pseudogenes.

The structural data reporting the presence of CNV at *HvCBF4-HvCBF2A* was thus further investigated in order to study how it influences transcript responsiveness of the corresponding genes. To answer this question, we assessed the expression of *HvCBF4* and *HvCBF2A* in seven genotypes with a different degree of FR and a different number of copies for the *HvCBF4-HvCBF2A* genomic segment ~~[22]~~ (**Fig. 2**). The expression analysis also included *HvCBF14* gene for which no CNV has been reported in barley. The information was implemented with the *in silico* analysis of putative regulatory motifs in the promoter regions of all *HvCBFs* annotated in 'Nure' and 'Morex'. The current ~~R~~results ~~obtained~~ showed that the expression of *HvCBF2A* and *HvCBF4* genes is clearly influenced by light, in accordance with Gierczik et al. [19]. ~~For HvCBF2A, particularly~~ Noteworthy, for both genes, elevated transcript abundance was observed for the facultative (*vrn-H1/vrn-H2*) genotypes 'Pamina' and 'Lunet' that harbor a notably high number of copies respect to the winter (*vrn-H1/Vrn-H2*) cultivars 'Nure', 'Mellori' and 'Ponente'. The lower CNV of the true spring (*Vrn-H1/vrn-H2*) genotypes, 'Tremois' and 'Morex', was accompanied by the lower expression. Loscos et al. [29] revealed a similar situation in which *HvFT1*(*Vrn-H3*) copy number was higher in the barley genotypes deleted for *Vrn-H2*. However, given the limited number of cultivars tested in the present study, further investigation is required to

prove the putative relationship between a null (deleted) *Vrn-H2* and CNV at *Fr-H2.* Although *HvCBF2*, *HvCBF4* and *HvCBF14* are known to follow circadian rhythm with expression peak observed right before the simulated sunset in both long- and short-day conditions [19,21], apparently no circadian motifs (binding motifs for LHY and CCA1) were found in their promoter regions neither in 'Nure' nor in 'Morex'. Moreover, no G-box elements were found in *HvCBF2A* while they were present in promotors of *HvCBF2C, HvCBF4, HvCBF9, HvCBF12C, HvCBF14* of both varieties. In Arabidopsis, G-Box-Like motifs are necessary for transcriptional regulation by circadian pseudo-response regulators binding basic helix-loop-helix transcription factor, Phytochrome-interacting factor 4 —(PIF4) [72].

Plants efficiently control the acquisition of freezing resistance using two different signaling pathways in response to a gradual temperature decrease during seasonal changes and a sudden temperature drop during the night [73]. Membrane rigidificaton-activated mechano-sensitive or ligand activated Ca^{2+} channels lead to calcium influx into cytosol and its binding by Ca-sensors such as calmodulins [74]. As an example, Calmodulin binding transcription activator 3 (CAMTA3) and CAMTA5 respond to a rapid decrease in temperature, but these proteins do not respond to a gradual one. Moreover, they are active during day and night, in contrast to key circadian components. Putative binding sites for CAMTA were identified in the promoters of *CBF4*, *CBF12*, *CBF15* and *CBF16* in both genotypes (and thus those *HvCBFs* might be involved in a common, rapid response to night temperature decrease).

Amino acid substitutions between 'Morex' and 'Nure' identified in the present work were compared to CBF sequence alignment reported by Pasquariello et al. [30]. While the majority of replacements were located in the C-terminal region of the proteins, and no polymorphisms were identified in the entire sequence of genes coding for CBF2A, CBF4, CBF6B, CBF14, CBF15A., o Only CBF3 and CBF10A presented amino acid replacements within the AP2/ERF DNA binding domain (**Supplemental Table 4**).

In the present study, no evident correlation between CNV and expression level for *HvCBF4* and *HvCBF2A* genes was observed after the cold treatment in dark conditions. Conversely, Dhillon et al. [18] reported the association between *HvCBF4*–*HvCBF2A* copy numbers and *HvCBF2* transcript levels; however, in their experiment plants were subjected to light and cold temperature (6°C) stimuli at the same time. Consequently, this implies that the observed expression was the sum/interaction of both environmental factors, each influencing *HvCBF* expression. On the other hand, in the present study only *HvCBF14* seemed to be induced just by low temperature under dark conditions (**Fig. 2**). The importance of *HvCBF14* in FR has been demonstrated via association and expression studies leading to the conclusions that transcript accumulation is induced by temperature shifts independently from light conditions, and that the effects of temperature and light treatments are additive [41,75,76]. Rye *ScCBF14* showed similar expression patterns in response to the treatment when studied using RTq-PCR [77]. Based on these data, it can be assumed that the level of *CBF14* expression and temperature/light stimuli are relayed by via separate signaling routes [41]. In the present study, numerous putative MYC motifs recognized by ICE1 transcription factors were identified in the promoters of all *CBFs* in both genotypes (**Table 3**), confirming the crucial role of this transcription factor in their regulation. An interesting difference in number of MYC motifs in the promoter of *CBF14* was detected between the resistant 'Nure' and susceptible 'Morex': nine and five, respectively, that might underlie different levels of expression in response to cold.

Besides their primary role in regulating *COR* gene expression, individual members of the *CBF* family are also involved in transcriptional auto-regulation [78] and modulation of distinct sets of target genes [79]. Dhillon et al. [18] reported that barley *CBF14* expression correlated with transcription levels of *CBF2*. Those observations supported the scenario that copy number of the *HvCBF4-HvCBF2A* genomic segment affects the expression of other *CBFs*. Moreover, chromatin immunoprecipitation assay evidenced the presence of *CBF2* protein at *CBF12* and *CBF16* promoters at normal growth temperatures [18]. Overexpressing *HvCBF2A* in susceptible 'Golden Promise' increased transcript levels

13

523 of *HvCBF12*, *HvCBF15* and *HvCBF16* and freezing resistance of transformed lines [80]. These data indicate that

524 *HvCBF2A* – under copy number variation – activates target genes at warm temperatures and that transcript

525 accumulation for some of these targets is greatly enhanced by cold temperatures [80]. Confirming these observations,

526 the promoter regions of all *HvCBFs* were searched for putative motifs recognized by other AP2-CBF factors in the

527 present study. Such motifs were retrieved in promoter of *CBF2C*, *CBF12*, *CBF12C*, *CBF14*, *CBF15* and *CBF16* in both

528 varieties (**Table 3**), suggesting an extensive interplay of *CBF* gene family in response to external stimuli.

529 Expression of some *CBF* elements at *Fr-2* is also known to be negatively modulated by the *VRN-1* allelic state [16]. In

530 vernalization-requiring winter genotypes, harboring dominant *Vrn-2* allele, *VRN-1* is transcriptionally activated by

531 prolonged cold to trigger flowering. Spring genotypes that flower without vernalization typically carry dominant alleles

532 at *VRN-H1* (e.g. *VRN1-1*, *VRN1-6*, and *VRN1-7*; [7]) that are actively transcribed without cold, reducing or eliminating

533 the requirement for vernalization [81]. VRN-1 is a MADS-box transcription factor that binds to CargG motifs in

534 promoter of the target genes. Such DNA elements were here identified in *HvCBF2A* and *HvCBF15A* of both 'Nure' and

535 'Morex' and in the promoter of *HvCBF16*, *HvCBF6B* and *HvCBF2B* of 'Nure' only (**Table 3**). A certain interplay

536 between the allelic state at *VRN-1* and CNV at *CBF4* and *CBF2A* genes could be hypothesized also on the basis of the

537 results obtained in this study. In genotypes harboring dominant spring alleles at *VRN-1*, a high number of copies of the

538 target genes could – at least in part – counterbalance its dampening effect on the expression of *CBF* genes.

539 Correspondingly, in winter genotypes – where the expression of the recessive *VRN-1* alleles is practically null at the

540 beginning of acclimation – the number of copies of *CBF*s can greatly influence the pool of key transcripts that lead to

541 acquisition of resistance. Finally, while *CBF2A* and *CBF4* binding activity to CRT motifs in *COR* genes promoter was

542 low-temperature dependent [56] and null at warm temperatures [48], low gene expression of *COR14B* and *DHN5* using

543 real time PCR in lines overexpressing *CBF2A* was observed also at warm temperatures [80]; however, they were much

544 higher at cold temperatures. Noteworthy, in the present study we could observe much higher transcript accumulation

545 during the warm temperature (and under light condition). One hypothesis could be that at "steady state" (neither

546 chilling, nor acclimation, nor frost) the frost resistant genotypes accumulate more transcripts as a result or their higher

547 number of *CBF* gene copies. Although such a mechanism would be energy consuming, it mightcould provide a kind of

548 "constitutive" FR consisting in accumulation of protein pools of CBF2A and CBF4 at normal temperatures with no (or

549 low) binding activity, and such activity is triggered once plants face cold. Further research is however needed to verify

550 if the higher transcript level of *CBF*s observed in the present study indeed corresponds to higher protein accumulation.

551 Deeper structural and functional characterization – that go beyond the scope of this work – is also requiredneeded to

552 definitely clarify the biological role of the CNV at *CBF4-CBF2A* segment and its effect in increasing FR in the

553 *Triticeae*. Rapid generation advancements (i.e. speed breeding; [82]), reciprocal near isogenic lines or large germplasm

554 collections harboring contrasting copy-number alleles would provide *ad hoc* genetic resources. This combined with

555 shotgun sequencing of transcriptome (RNA-seq) or chromatin immunoprecipitated-DNA fragments (ChIP-seq), could

556 be applied for studying the modulation of all *CBF*s residing at *Fr-2* in response to rapid/prolonged exposure to

557 low/freezing temperature, and to the effects of different light and cold stimuli combinations.

558

559

5. Conclusions

561 In the present work, a re-sequencing and an accurate annotation of the *Fr-H2* genomic locus in the frost resistant cv.

562 'Nure' is presented for the first time along with its structural comparison with the same region of susceptible cv.

563 'Morex'. The CNV reported in this region for the *HvCBF4-HvCBF2A* segment was confirmed and several sequence

564 differences as far as abundance and diversity of repetitive elements (i.e. Class I retroelements and Class II DNA
565 transposons) were identified. Either presence/absence or function gain/loss of specific genes (i.e. *HvCBF2B*, *HvCFB13*,
566 *HvCBF10B*), differences in regulatory motif elements and SNPs in coding sequences were also identified suggesting
567 possible functional explanation of the locus. Moreover, an interesting and complex interplay of gene expression levels
568 of some *HvCBF*s (with and without CNV) and putatively involved regulation pathways were identified. The number of
569 copies of *HvCBF2A* and *HvCBF4* genes resulted correlated to their expression under light stimulus at room
570 temperature. On the other hand, FR correlated with *HvCBF14* expression but only after cold induction in the dark.
571 Noteworthy, under light stimulus at room temperature, expression levels of *HvCBF4* and *HvCBF14* were correlated
572 with *HvCBF2A*, while FR was shown to correlate with the copy number of *HvCBF2A* and *HvCBF4*. Our results suggest
573 that *Fr-H2* contribution to FR is given by a fine-tuned network of CNV and differences in expression levels of single-
574 copy *HvCBF* genes.

Conflict of Interest

593 The authors declare that they have no conflict of interest.

References

[1]   E. Francia, F. Rizza, L. Cattivelli, A.M. Stanca, G. Galiba, B. Tóth, P.M. Hayes, J.S. Skinner, N. Pecchioni,
       Two loci on chromosome 5H determine low-temperature tolerance in a "Nure" (winter) x Tremois' (spring)
       barley map, Theor. Appl. Genet. 108 (2004) 670–680. doi:10.1007/s00122-003-1468-9.

[2]   P.M. Hayes, T. Blake, T.H.H. Chen, S. Tragoonrung, F. Chen, A. Pan, B. Liu, Quantitative trait loci on barley
       (Hordeum vulgare L.) chromosome 7 associated with components of winterhardiness, Genome. 36 (1993) 66–
       71. doi:10.1139/g93-009.

[3]   A. Vágújfalvi, G. Galiba, L. Cattivelli, J. Dubcovsky, The cold-regulated transcriptional activator Cbf3 is linked
       to the frost-tolerance locus Fr-A2 on wheat chromosome 5A, Mol. Genet. Genomics. 269 (2003) 60–67.
       doi:10.1007/s00438-003-0806-6.

15

[4]   G. Galiba, S.A. Quarrie, J. Sutka, A. Morgounov, J.W. Snape, RFLP mapping of the vernalization (Vrn1) and frost resistance (Fr1) genes on chromosome 5A of wheat, Theor. Appl. Genet. 90 (1995) 1174–1179. doi:10.1007/BF00222940.

[5]   L. Yan, A. Loukoianov, G. Tranquilli, M. Helguera, T. Fahima, J. Dubcovsky, Positional cloning of the wheat vernalization gene VRN1, PNAS. 100 (2003) 6263–6268. doi:10.1073/pnas.0937399100.

[6]   B. Trevaskis, M.N. Hemming, E.S. Dennis, W.J. Peacock, The molecular basis of vernalization-induced flowering in cereals, Trends Plant Sci. 12 (2007) 352–357. doi:10.1016/j.tplants.2007.06.010.

[7]   M.N. Hemming, S. Fieg, W.J. Peacock, E.S. Dennis, B. Trevaskis, Regions associated with repression of the barley (Hordeum vulgare) VERNALIZATION1 gene are not required for cold induction, Mol. Genet. Genomics. 282 (2009) 107–117. doi:10.1007/s00438-009-0449-3.

[8]   W. Deng, M.C. Casao, P. Wang, K. Sato, P.M. Hayes, E.J. Finnegan, B. Trevaskis, Direct links between the vernalization response and other key traits of cereal crops, Nat. Commun. 6 (2015) 5882. doi:10.1038/ncomms6882.

[9]   E. Francia, D. Barabaschi, A. Tondelli, G. Laidò, F. Rizza, A.M. Stanca, M. Busconi, C. Fogher, E.J. Stockinger, N. Pecchioni, Fine mapping of a HvCBF gene cluster at the frost resistance locus Fr-H2 in barley, Theor. Appl. Genet. 115 (2007) 1083–1091. doi:10.1007/s00122-007-0634-x.

[10]   A. Tondelli, E. Francia, D. Barabaschi, M. Pasquariello, N. Pecchioni, Inside the CBF locus in Poaceae, Plant Sci. 180 (2011) 39–45. doi:10.1016/j.plantsci.2010.08.012.

[11]   E. Seo, H. Lee, J. Jeon, H. Park, J. Kim, Y.S. Noh, I. Lee, Crosstalk between cold response and flowering in Arabidopsis is mediated through the flowering-time gene SOC1 and its upstream negative regulator FLC, Plant Cell. 21 (2009) 3185–3197. doi:10.1105/tpc.108.063883.

[12]   M.M. Alonso-Peral, S.N. Oliver, M.C. Casao, A.A. Greenup, B. Trevaskis, The promoter of the cereal VERNALIZATION1 gene is sufficient for transcriptional induction by prolonged cold, PLoS One. 6 (2011). doi:10.1371/journal.pone.0029456.

[13]   Y. Shi, Y. Ding, S. Yang, Molecular regulation of CBF signaling in cold acclimation, Trends Plant Sci. 23 (2018) 623–637. doi:10.1016/j.tplants.2018.04.002.

[14]   T.S. Artlip, M.E. Wisniewski, J.L. Norelli, Field evaluation of apple overexpressing a peach CBF gene confirms its effect on cold hardiness, dormancy, and growth, Environ. Exp. Bot. 106 (2014) 79–86. doi:10.1016/j.envexpbot.2013.12.008.

[15]   M. Buti, M. Pasquariello, D. Ronga, J.A. Milc, N. Pecchioni, V.T. Ho, C. Pucciariello, P. Perata, E. Francia, Transcriptome profiling of short-term response to chilling stress in tolerant and sensitive Oryza sativa ssp. Japonica seedlings, Funct. Integr. Genomics. 18 (2018). doi:10.1007/s10142-018-0615-y.

[16]   E.J. Stockinger, J.S. Skinner, K.G. Gardner, E. Francia, N. Pecchioni, Expression levels of barley Cbf genes at the Frost resistance-H2 locus are dependent upon alleles at Fr-H1 and Fr-H2, Plant J. 51 (2007). doi:10.1111/j.1365-313X.2007.0141.x.

[17]   X. Zhang, S.G. Fowler, H. Cheng, Y. Lou, S.Y. Rhee, E.J. Stockinger, M.F. Thomashow, Freezing-sensitive tomato has a functional CBF cold response pathway, but a CBF regulon that differs from that of freezing-tolerant Arabidopsis, Plant J. 39 (2004) 905–919. doi:10.1111/j.1365-313X.2004.02176.x.

[18]   T. Dhillon, K. Morohashi, E.J. Stockinger, CBF2A-CBF4B genomic region copy numbers alongside the circadian clock play key regulatory mechanisms driving expression of FR-H2 CBFs, Plant Mol. Biol. 94 (2017) 333–347. doi:10.1007/s11103-017-0610-z.

[19]   K. Gierczik, A. Novák, M. Ahres, A. Székely, A. Soltész, Á. Boldizsár, Z. Gulyás, B. Kalapos, I. Monostori, L. Kozma-Bognár, G. Galiba, A. Vágújfalvi, Circadian and light regulated expression of CBFs and their upstream signalling genes in barley, Int J Mol Sci. 18 (2017). doi:10.3390/ijms18081828.

[20]   G. Galiba, E.J. Stockinger, E. Francia, J. Milc, G. Kocsy, N. Pecchioni, Freezing tolerance in the Triticeae, in: R.K. Varshney, R. Tuberosa (Eds.), Transl. {Genomics} {Crop} {Breeding}, John Wiley & Sons Ltd, 2013: pp. 99–124. doi:10.1002/9781118728482.ch7.

[21] M. Badawi, J. Danyluk, B. Boucho, M. Houde, F. Sarhan, The CBF gene family in hexaploid wheat and its relationship to the phylogenetic complexity of cereal CBFs, Mol Genet Genomics. 277 (2007) 533–554. doi:10.1007/s00438-006-0206-9.

[22] A.K. Miller, G. Galiba, J. Dubcovsky, A cluster of 11 CBF transcription factors is located at the frost tolerance locus Fr-Am2 in Triticum monococcum, Mol. Genet. Genomics. 275 (2006) 193–203. doi:10.1007/s00438-005-0076-6.

[23] A.K. Knox, T. Dhillon, H. Cheng, A. Tondelli, N. Pecchioni, E.J. Stockinger, CBF gene copy number variation at Frost Resistance-2 is associated with levels of freezing tolerance in temperate-climate cereals, Theor Appl Genet. 121 (2010) 21–35. doi:10.1007/s00122-010-1288-7.

[24] E. Francia, C. Morcia, M. Pasquariello, V. Mazzamurro, J.A. Milc, F. Rizza, V. Terzi, N. Pecchioni, Copy number variation at the HvCBF4–HvCBF2 genomic segment is a major component of frost resistance in barley, Plant Mol. Biol. 92 (2016) 161–175. doi:10.1007/s11103-016-0505-4.

[25] G. Corrado, R. Rao, Towards the genomic basis of local adaptation in landraces, Diversity. 9 (2017). doi:10.3390/d9040051.

[26] A. Dolatabadian, D.A. Patel, D. Edwards, J. Batley, Copy number variation and disease resistance in plants, Theor. Appl. Genet. (2017). doi:10.1007/s00122-017-2993-2.

[27] A. Díaz, M. Zikhali, A.S. Turner, P. Isaac, D.A. Laurie, Copy number variation affecting the photoperiod-B1 and vernalization-A1 genes is associated with altered flowering time in wheat (Triticum aestivum), PLoS One. (2012). doi:10.1371/journal.pone.0033234.

[28] R. Nitcher, A. Distelfeld, C. Tan, L. Yan, J. Dubcovsky, Increased copy number at the HvFT1 locus is associated with accelerated flowering time in barley, Mol. Genet. Genomics. 288 (2013) 261–275. doi:10.1007/s00438-013-0746-8.

[29] J. Loscos, E. Igartua, B. Contreras-Moreir, M. Pilar Gracia, A.M. Casas, HvFT1 polymorphism and effect— survey of barley germplasm and expression analysis, Front. Plant Sci. 5 (2014) 251. doi:10.3389/fpls.2014.00251.

[30] M. Pasquariello, D. Barabaschi, A. Himmelbach, B. Steuernagel, R. Ariyadasa, N. Stein, F. Gandolfi, E. Tenedini, I. Bernardis, E. Tagliafico, N. Pecchioni, E. Francia, The barley Frost resistance-H2 locus, Funct. Integr. Genomics. 14 (2014) 85–100. doi:10.1007/s10142-014-0360-9.

[31] D.G. Peterson, J.P. Tomkins, D.A. Frisch, R.A. Wing, A.H. Paterson, Construction of plant bacterial artificial chromosome (BAC) libraries: an illustrated guide., J. Agric. Genomics. 5 (2000) 1–3. https://www.cabdirect.org/cabdirect/abstract/20023037607 (accessed August 25, 2017).

[32] B. Chalhoub, H. Belcram, M. Caboche, Efficient cloning of plant genomes into bacterial artificial chromosome (BAC) libraries with larger and more uniform insert size, Plant Biotechnol. J. 2 (2004) 181–188. doi:10.1111/j.1467-7652.2004.00065.x.

[33] M.J. Chaisson, G. Tesler, Mapping single molecule sequencing reads using basic local alignment with successive refinement (BLASR): application and theory, BMC Bioinformatics. 13 (2012) 1–18. doi:10.1186/1471-2105-13-238.

[34] T. Wicker, F. Sabot, A. Hua-Van, J.L. Bennetzen, P. Capy, B. Chalhoub, A. Flavell, P. Leroy, M. Morgante, O. Panaud, E. Paux, P. SanMiguel, A.H. Schulman, A unified classification system for eukaryotic transposable elements, Nat. Rev. Genet. 8 (2007) 973–982. doi:10.1038/nrg2165.

[35] Z. Xu, H. Wang, LTR-FINDER: an efficient tool for the prediction of full-length LTR retrotransposons, Nucleic Acids Res. 35 (2007) W265--268. doi:10.1093/nar/gkm286.

[36] E.L.L. Sonnhammer, R. Durbin, A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis, Gene. 167 (1995) GC1--GC10. doi:10.1016/0378-1119(95)00714-8.

[37] P. Leroy, N. Guilhot, H. Sakai, A. Bernard, F. Choulet, S. Theil, S. Reboux, N. Amano, T. Flutre, C. Pelegrin, H. Ohyanagi, M. Seidel, F. Giacomoni, M. Reichstadt, M. Alaux, E. Gicquello, F. Legeai, L. Cerutti, H. Numa, T. Tanaka, K. Mayer, T. Itoh, H. Quesneville, C. Feuillet, TriAnnot: A versatile and high performance pipeline

for the automated annotation of plant genomes, Front Plant Sci. 3 (2012). doi:10.3389/fpls.2012.00005.

[38]   C. Gunasekara, A. Subramanian, J.V.R.K. Avvari, B. Li, S. Chen, H. Wei, ExactSearch: A web-based plant motif search tool, Plant Methods. (2016). doi:10.1186/s13007-016-0126-6.

[39]   D. Piovesan, F. Tabaro, L. Paladin, M. Necci, I. Mieti, C. Camilloni, N. Davey, Z. Dosztányi, B. Mészáros, A.M. Monzon, G. Parisi, E. Schad, P. Sormanni, P. Tompa, M. Vendruscolo, W.F. Vranken, S.C.E. Tosatto, MobiDB 3.0: More annotations for intrinsic disorder, conformational diversity and interactions in proteins, Nucleic Acids Res. (2018). doi:10.1093/nar/gkx1071.

[40]   F. Rizza, I. Karsai, C. Morcia, F.-W. Badeck, V. Terzi, D. Pagani, T. Kiss, A.M. Stanca, Association between the allele compositions of major plant developmental genes and frost tolerance in barley (Hordeum vulgare L.) germplasm of different origin, Mol Breed. 36 (2016) 156. doi:10.1007/s11032-016-0571-y.

[41]   A. Novák, Á. Boldizsár, K. Gierczik, A. Vágújfalvi, É. Ádám, L. Kozma-Bognár, G. Galiba, Light and temperature signalling at the level of CBF14 gene expression in wheat and barley., Plant Mol Biol Rep. 35 (2017) 399–408. doi:10.1007/s11105-017-1035-1.

[42]   J. Liu, S. Huang, M. Sun, S. Liu, Y. Liu, W. Wang, X. Zhang, H. Wang, W. Hua, An improved allele-specific PCR primer design method for SNP marker analysis and its application, Plant Methods. 8 (2012) 34. doi:10.1186/1746-4811-8-34.

[43]   S. Kwok, S.Y. Chang, J.J. Sninsky, A. Wang, A guide to the design and use of mismatched and degenerate primers, PCR Methods Appl. 3 (1994) S39--47.

[44]   R.S. Cha, H. Zarbl, P. Keohavong, W.G. Thilly, Mismatch amplification mutation assay (MAMA): application to the c-H-ras gene, PCR Methods Appl. 2 (1992) 14–20.

[45]   K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method, Methods. 25 (2001) 402–408. doi:10.1006/meth.2001.1262.

[46]   A. Janská, J. Hodek, P. Svoboda, J. Zámečník, I.T. Prášil, E. Vlasáková, L. Milella, J. Ovesná, The choice of reference gene set for assessing gene expression in barley (Hordeum vulgare L.) under low temperature and drought stress, Mol. Genet. Genomics. 288 (2013) 639–649. doi:10.1007/s00438-013-0774-4.

[47]   V. Chinnusamy, J.-K. Zhu, R. Sunkar, Gene regulation during cold stress acclimation in plants, Methods Mol. Biol. 639 (2010) 39–55. doi:10.1007/978-1-60761-702-0_3.

[48]   J.S. Skinner, J. von Zitzewitz, P. Szűcs, L. Marquez-Cedillo, T. Filichkin, K. Amundsen, E.J. Stockinger, M.F. Thomashow, T.H.H. Chen, P.M. Hayes, Structural, functional, and phylogenetic characterization of a large CBF gene family in barley, Plant Mol Biol. 59 (2005) 533–551. doi:10.1007/s11103-005-2498-2.

[49]   R. Appels, K. Eversole, C. Feuillet, B. Keller, J. Rogers, N. Stein, C.J. Pozniak, F. Choulet, A. Distelfeld, J. Poland, G. Ronen, O. Barad, K. Baruch, G. Keeble-Gagnère, M. Mascher, G. Ben-Zvi, A.A. Josselin, A. Himmelbach, F. Balfourier, J. Gutierrez-Gonzalez, M. Hayden, C.S. Koh, G. Muehlbauer, R.K. Pasam, E. Paux, P. Rigault, J. Tibbits, V. Tiwari, M. Spannagl, D. Lang, H. Gundlach, G. Haberer, K.F.X. Mayer, D. Ormanbekova, V. Prade, T. Wicker, D. Swarbreck, H. Rimbert, M. Felder, N. Guilhot, G. Kaithakottil, J. Keilwagen, P. Leroy, T. Lux, S. Twardziok, L. Venturini, A. Juhasz, M. Abrouk, I. Fischer, C. Uauy, P. Borrill, R.H. Ramirez-Gonzalez, D. Arnaud, S. Chalabi, B. Chalhoub, A. Cory, R. Datla, M.W. Davey, J. Jacobs, S.J. Robinson, B. Steuernagel, F. Van Ex, B.B.H. Wulff, M. Benhamed, A. Bendahmane, L. Concia, D. Latrasse, M. Alaux, J. Bartoš, A. Bellec, H. Berges, J. Doležel, Z. Frenkel, B. Gill, A. Korol, T. Letellier, O.A. Olsen, H. Šimková, K. Singh, M. Valárik, E. Van Der Vossen, S. Vautrin, S. Weining, T. Fahima, V. Glikson, D. Raats, H. Toegelová, J. Vrána, P. Sourdille, B. Darrier, D. Barabaschi, L. Cattivelli, P. Hernandez, S. Galvez, H. Budak, J.D.G. Jones, K. Witek, G. Yu, I. Small, J. Melonek, R. Zhou, T. Belova, K. Kanyuka, R. King, K. Nilsen, S. Walkowiak, R. Cuthbert, R. Knox, K. Wiebe, D. Xiang, A. Rohde, T. Golds, J. Čížkova, B.A. Akpinar, S. Biyiklioglu, L. Gao, A. N'Daiye, J. Číhalíková, M. Kubaláková, J. Šafář, F. Alfama, A.F. Adam-Blondon, R. Flores, C. Guerche, M. Loaec, H. Quesneville, A.G. Sharpe, J. Condie, J. Ens, R. Maclachlan, Y. Tan, A. Alberti, J.M. Aury, V. Barbe, A. Couloux, C. Cruaud, K. Labadie, S. Mangenot, P. Wincker, G. Kaur, M. Luo, S. Sehgal, P. Chhuneja, O.P. Gupta, S. Jindal, P. Kaur, P. Malik, P. Sharma, B. Yadav, N.K. Singh, J.P. Khurana, C. Chaudhary, P. Khurana, V. Kumar, A. Mahato, S. Mathur, A. Sevanthi, N. Sharma, R.S. Tomar, K. Holušová, O. Plíhal, M.D. Clark, D. Heavens, G. Kettleborough, J. Wright, B. Balcárková, Y. Hu, N. Ravin, K. Skryabin, A. Beletsky, V. Kadnikov, A. Mardanov, M. Nesterov, A. Rakitin, E. Sergeeva, H. Kanamori, S. Katagiri, F. Kobayashi, S. Nasuda, T. Tanaka, J. Wu, F. Cattonaro, M. Jiumeng, K. Kugler, M. Pfeifer, S.

Sandve, X. Xun, B. Zhan, J. Batley, P.E. Bayer, D. Edwards, S. Hayashi, Z. Tulpová, P. Visendi, L. Cui, X. Du, K. Feng, X. Nie, W. Tong, L. Wang, Shifting the limits in wheat research and breeding using a fully annotated reference genome, 361 (2018) eaar7191. (2018). doi:10.1126/science.aar7191.

[50]   C. Campoli, M.A. Matus-Cádiz, C.J. Pozniak, L. Cattivelli, D.B. Fowler, Comparative expression of Cbf genes in the Triticeae under different acclimation induction temperatures, Mol Genet Genomics. 282 (2009) 141–152. doi:10.1007/s00438-009-0451-9.

[51]   S.G. Fowler, D. Cook, M.F. Thomashow, Low temperature induction of Arabidopsis CBF1, 2, and 3 is gated by the circadian clock, Plant Physiol. 137 (2005) 961–968. doi:10.1104/pp.104.058354.

[52]   F. Kobayashi, S. Takumi, S. Kume, M. Ishibashi, R. Ohno, K. Murai, C. Nakamura, Regulation by Vrn-1/Fr-1 chromosomal intervals of CBF-mediated Cor/Lea gene expression and freezing tolerance in common wheat, J. Exp. Bot. 56 (2005) 887–895. doi:10.1093/jxb/eri081.

[53]   A.E. Limin, D.B. Fowler, Low-temperature tolerance and genetic potential in wheat (Triticum aestivum L.): response to photoperiod, vernalization, and plant development, Planta. 224 (2006) 360–366. doi:10.1007/s00425-006-0219-y.

[54]   J.C. Pennycooke, H. Cheng, S.M. Roberts, Q. Yang, S.Y. Rhee, E.J. Stockinger, The low temperature-responsive, Solanum CBF1 genes maintain high identity in their upstream regions in a genomic environment undergoing gene duplications, deletions, and rearrangements, Plant Mol. Biol. 67 (2008) 483–497. doi:10.1007/s11103-008-9333-5.

[55]   A. Vágújfalvi, A. Aprile, A. Miller, J. Dubcovsky, G. Delugu, G. Galiba, L. Cattivelli, The expression of several Cbf genes at the Fr-A2 locus is linked to frost resistance in wheat, Mol. Genet. Genomics. 274 (2005) 506–514. doi:10.1007/s00438-005-0047-y.

[56]   G.-P. Xue, The DNA-binding activity of an AP2 transcriptional activator HvCBF2 involved in regulation of low-temperature responsive genes in barley is modulated by temperature, Plant J. 33 (2003) 373–383.

[57]   T. Dhillon, S.P. Pearce, E.J. Stockinger, A. Distelfeld, C. Li, A.K. Knox, I. Vashegyi, A. Vágújfalvi, G. Galiba, J. Dubcovsky, Regulation of freezing tolerance and flowering in temperate cereals: the VRN-1 connection, Plant Physiol. 153 (2010) 1846–1858. doi:10.1104/pp.110.159079.

[58]   A. Rhoads, K.F. Au, PacBio sequencing and its applications, Genomics. Proteomics Bioinformatics. 13 (2015) 278–289. doi:10.1016/j.gpb.2015.08.002.

[59]   C.D. Hirsch, N.M. Springer, Transposable element influences on gene expression in plants, Biochim. Biophys. Acta - Gene Regul. Mech. 1860 (2017) 157–165. doi:10.1016/j.bbagrm.2016.05.010.

[60]   M. Buti, T. Giordani, M. Vukich, C. Pugliesi, L. Natali, A. Cavallini, Retrotransposon-related genetic distance and hybrid performance in sunflower (Helianthus annuus L.), Euphytica. 192 (2013) 289–303. doi:10.1007/s10681-013-0883-5.

[61]   A. Kanazawa, B. Liu, F. Kong, S. Arase, J. Abe, Adaptive Evolution Involving Gene Duplication and Insertion of a Novel Ty1/copia-Like Retrotransposon in Soybean, J. Mol. Evol. 69 (2009) 164–175. doi:10.1007/s00239-009-9262-1.

[62]   L. Wei, X. Cao, The effect of transposable elements on phenotypic variation: insights from plants to humans, Sci. China Life Sci. 59 (2016) 24–37. doi:10.1007/s11427-015-4993-2.

[63]   B.S. Gaut, B.R. Morton, B.C. McCaig, M.T. Clegg, Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene Adh parallel rate differences at the plastid gene rbcL., Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 10274–9.

[64]   K.F.X. Mayer, M. Martis, P.E. Hedley, H. Simková, H. Liu, J.A. Morris, B. Steuernagel, S. Taudien, S. Roessner, H. Gundlach, M. Kubaláková, P. Suchánková, F. Murat, M. Felder, T. Nussbaumer, A. Graner, J. Salse, T. Endo, H. Sakai, T. Tanaka, T. Itoh, K. Sato, M. Platzer, T. Matsumoto, U. Scholz, J. Dolezel, R. Waugh, N. Stein, Unlocking the barley genome by chromosomal and comparative genomics., Plant Cell. 23 (2011) 1249–63. doi:10.1105/tpc.110.082537.

[65]   C. Monat, M. Schreiber, N. Stein, M. Mascher, Prospects of pan-genomics in barley, Theor. Appl. Genet.

(2018). doi:10.1007/s00122-018-3234-z.

[66]   M. Muñoz-Amatriaín, S.R. Eichten, T. Wicker, T.A. Richmond, M. Mascher, B. Steuernagel, U. Scholz, R. Ariyadasa, M. Spannagl, T. Nussbaumer, K.F.X. Mayer, S. Taudien, M. Platzer, J.A. Jeddeloh, N.M. Springer, G.J. Muehlbauer, N. Stein, Distribution, functional impact, and origin mechanisms of copy number variation in the barley genome, Genome Biol. 14 (2013) R58. doi:10.1186/gb-2013-14-6-r58.

[67]   T. Dhillon, E.J. Stockinger, Cbf14 copy number variation in the A, B, and D genomes of diploid and polyploid wheat, Theor. Appl. Genet. 126 (2013) 2777–2789. doi:10.1007/s00122-013-2171-0.

[68]   A.-N. Sieber, C.F.H. Longin, W.L. Leiser, T. Würschum, Copy number variation of CBF-A14 at the Fr-A2 locus determines frost tolerance in winter durum wheat, Theor. Appl. Genet. 129 (2016) 1087–1097. doi:10.1007/s00122-016-2685-3.

[69]   J. Zhu, S. Pearce, A. Burke, D.R. See, D.Z. Skinner, J. Dubcovsky, K. Garland-Campbell, K.G. Campbell, Copy number and haplotype variation at the VRN-A1 and central FR-A2 loci are associated with frost tolerance in hexaploid wheat, Theor. Appl. Genet. 127 (2014) 1183–1197. doi:10.1007/s00122-014-2290-2.

[70]   T. Würschum, C.F.H. Longin, V. Hahn, M.R. Tucker, W.L. Leiser, Copy number variations of CBF genes at the Fr-A2 locus are essential components of winter hardiness in wheat, Plant J. 89 (2017) 764–773. doi:10.1111/tpj.13424.

[71]   S. Babben, E. Schliephake, P. Janitza, T. Berner, J. Keilwagen, M. Koch, F.A. Arana-Ceballos, S.E. Templer, Y. Chesnokov, T. Pshenichnikova, J. Schondelmaier, A. Börner, K. Pillen, F. Ordon, D. Perovic, Association genetics studies on frost tolerance in wheat (Triticum aestivum L.) reveal new highly conserved amino acid substitutions in CBF-A3, CBF-A15, VRN3 and PPD1 genes, BMC Genomics. 19 (2018). doi:10.1186/s12864-018-4795-6.

[72]   T.L. Liu, L. Newton, M.-J. Liu, S.-H. Shiu, E.M. Farré, A G-Box-like motif is necessary for transcriptional regulation by circadian Pseudo-Response Regulators in Arabidopsis., Plant Physiol. 170 (2016) 528–539. doi:10.1104/pp.15.01562.

[73]   S. Kidokoro, K. Yoneda, H. Takasaki, F. Takahashi, K. Shinozaki, K. Yamaguchi-Shinozaki, Different cold-signaling pathways function in the responses to rapid and gradual decreases in temperature, Plant Cell. 29 (2017) 760–774. doi:10.1105/tpc.16.00669.

[74]   X. Wang, D. Wu, Q. Yang, J. Zeng, G. Jin, Z.-H. Chen, G. Zhang, F. Dai, Identification of mild freezing shock response pathways in barley based on transcriptome profiling, Front Plant Sci. 7 (2016). doi:10.3389/fpls.2016.00106.

[75]   A. Fricano, F. Rizza, P. Faccioli, D. Pagani, P. Pavan, A. Stella, L. Rossini, P. Piffanelli, L. Cattivelli, Genetic variants of HvCbf14 are statistically associated with frost tolerance in a European germplasm collection of Hordeum vulgare, Theor Appl Genet. 119 (2009) 1335–1348. doi:10.1007/s00122-009-1138-7.

[76]   A. Novák, Á. Boldizsár, É. Ádám, L. Kozma-Bognár, I. Majláth, M. Båga, B. Tóth, R. Chibbar, G. Galiba, Light-quality and temperature-dependent CBF14 gene expression modulates freezing tolerance in cereals, J. Exp. Bot. 67 (2016) 1285–1295. doi:10.1093/jxb/erv526.

[77]   W.J. Jung, Y.W. Seo, Identification of novel C-repeat binding factor (CBF) genes in rye (Secale cereale L.) and expression studies, Gene. 684 (2019) 82–94. doi:10.1016/j.gene.2018.10.055.

[78]   F. Novillo, J.M. Alonso, J.R. Ecker, J. Salinas, CBF2/DREB1C is a negative regulator of CBF1/DREB1B and CBF3/DREB1A expression and plays a central role in stress tolerance in Arabidopsis, Proc. Natl. Acad. Sci. U.S.A. 101 (2004) 3985–3990. doi:10.1073/pnas.0303029101.

[79]   Y. Shi, J. Huang, T. Sun, X. Wang, C. Zhu, Y. Ai, H. Gu, The precise regulation of different COR genes by individual CBF transcription factors in Arabidopsis thaliana, J Integr Plant Biol. 59 (2017) 118–133. doi:10.1111/jipb.12515.

[80]   Z. Jeknić, K.A. Pillman, T. Dhillon, J.S. Skinner, O. Veisz, A. Cuesta-Marcos, P.M. Hayes, A.K. Jacobs, T.H.H. Chen, E.J. Stockinger, Hv-CBF2A overexpression in barley accelerates COR gene transcript accumulation and acquisition of freezing tolerance during cold acclimation, Plant Mol. Biol. 84 (2014) 67–82. doi:10.1007/s11103-013-0119-z.

[81]   B. Trevaskis, D.J. Bagnall, M.H. Ellis, W.J. Peacock, E.S. Dennis, MADS box genes control vernalization-induced flowering in cereals, PNAS. 100 (2003) 13099–13104. doi:10.1073/pnas.1635053100.

[82]   A. Watson, S. Ghosh, M.J. Williams, W.S. Cuddy, J. Simmonds, M.D. Rey, M. Asyraf Md Hatta, A. Hinchliffe, A. Steed, D. Reynolds, N.M. Adamski, A. Breakspear, A. Korolev, T. Rayner, L.E. Dixon, A. Riaz, W. Martin, M. Ryan, D. Edwards, J. Batley, H. Raman, J. Carter, C. Rogers, C. Domoney, G. Moore, W. Harwood, P. Nicholson, M.J. Dieters, I.H. Delacy, J. Zhou, C. Uauy, S.A. Boden, R.F. Park, B.B.H. Wulff, L.T. Hickey, Speed breeding is a powerful tool to accelerate crop research and breeding, Nat. Plants. (2018). doi:10.1038/s41477-017-0083-8.

[83]   A. Tondelli, E. Francia, A. Visioni, J. Comadran, A.M. Mastrangelo, T. Akar, A. Al-Yassin, S. Ceccarelli, S. Grando, A. Benbelkacem, F.A. van Eeuwijk, W.T.B. Thomas, A.M. Stanca, I. Romagosa, N. Pecchioni, QTLs for barley yield adaptation to Mediterranean environments in the "Nure" × "Tremois" biparental population, Euphytica. 197 (2014). doi:10.1007/s10681-013-1053-5.

Tables

Table 1 Assembly metrics of PacBio RS II reads. Number of generated contigs, mean coverage, mean QV, and the maximum obtainable size are indicated for each assembled clone.

BAC clone	No. of contigs	Mean coverage	QV [a]	Assembly size (~~kbp~~kb)
Nure_13J22	1	262	48.87	181
Nure_186A1	1	502	48.49	142
Nure_387O11	1	377	48.69	109
NureBI_32K2	1	500	48.77	111
NureBI_21H6	1	139	48.75	107
Nure_42O19	1	269	48.53	139
Nure_6M4	1	312	48.5	127
Nure_63P18	1	280	48.72	117
Nure_224I6	1	310	48.65	128
Nure_65L15	1	247	48.88	120
Nure_105O11	1	400	48.25	111
Nure_80M6	1	219	48.78	171
Nure_333A13	1	430	48.43	126
Nure_25L23	1	201	48.49	115

[a]: Quality value (QV) basecall error probability: QV40=1 in 10,000, QV50=1 in 100,000.

Table 2 Sequence annotation for *Fr-H2* locus in 'Nure' vs 'Morex'.

	'Nure'			'Morex'		
	No.	Length (bp)	% [a]	No.	Length (bp)	% [a]
Genes						
CBF	19	13,574	1.28	13	8,313	0.66
Pseudo-CBF	5	1,823	0.17	5	3,150	0.25
RAPT	2	649	0.06	3	741	0.06
Hypothetical	15	6,569	0.62	20	8,044	0.64
Other	12	10,596	0.10	6	7,706	0.62
Total	*53*	*33,211*	*2.23*	*48*	*27,954*	*2.24*
Class I retroelements						
Gypsy	135	289,127	27.20	136	491,373	39.40
Copia	104	315,332	29.70	82	287,178	23.00
LINE	10	5,411	0.50	9	4,670	0.37
Total	*249*	*609,870*	*57.40*	*227*	*783,221*	*62.77*
Class II DNA transposons						
Mariner	3	690	0.06	3	497	0.04
Mutator	9	2,472	0.23	6	2,927	0.23
CACTA	66	48,660	4.58	75	110,967	8.90
Harbinger	20	20,199	1.90	12	9,842	0.79
MITE	12	2,266	0.21	4	939	0.07
Total	*110*	*74,287*	*6.98*	*100*	*125,172*	*10.03*

[a]: Percentage values calculated with respect to the total *Fr-H2* sequence length.

Table 3 Putative regulatory elements identified in 1,000 bp upstream of the transcription starting site of the 'Morex' and 'Nure' *CBF* genes. Gene order follows the physical position within the 'Nure' *Fr-H2* locus. For *CBF4* and *CBF2A* elements, count is referred to a single copy and, in brackets, to the total number given by CNV.

Element(Factor)	Effect	**HvCBF gene**															
		9	2C	4[a]	2A[a]	2B[b]	12C	14	15A	12	16	6B	13[c]	3	10A	10B[b]	6
'Nure'																	
CM2(CAMTA3)	+	0	0	1(3)	0	0	0	0	1	1	2	0	1	0	0	0	0
MYC(ICE1)	+	6	8	6(18)	3(9)	2	2	9	2	4	5	2	8	6	5	5	5
G-box(1PIF/CESTA)	+/-	1	4	4(12)	0	1	0	2	0	0	1	0	2	0	0	0	0
MYB(MYB15)	-	1	2	0	1(3)	0	1	1	1	1	2	0	1	0	0	1	0
GATA(GATA)	+/-	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
CarG-box(MADS)	-	1	0	0	0	1	0	0	1	0	1	1	0	0	0	0	0
CRT-DRE(CBF)	+	0	2	0	0	1	1	1	1	1	1	0	0	0	0	3	0
'Morex'																	
CM2(CAMTA3)	+	0	0	1	0	NP	2	1	1	1	1	2	ψ	1	0	NP	2
MYC(ICE1)	+	6	10	4	3	NP	4	5	4	4	1	6	ψ	6	4	NP	6
G-box(1PIF/CESTA)	+/-	1	4	2	0	NP	1	1	0	0	1	0	ψ	0	0	NP	1
MYB(MYB15)	-	3	2	0	1	NP	1	1	1	2	0	1	ψ	1	4	NP	1
GATA(GATA)	+/-	0	0	0	0	NP	0	0	0	0	1	0	ψ	1	0	NP	0
CarG-box(MADS)	-	1	0	0	0	NP	0	0	1	0	0	0	ψ	0	1	NP	0
CRT-DRE(CBF)	+	0	2	0	0	NP	1	1	1	1	4	1	ψ	0	3	NP	0

[a]: In 'Nure' *CBF4* and *CBF2A* are present in three and two copies respectively;

[b]: *CBF2B* and *CBF10B* are not present (NP) in 'Morex';

[c]: *CBF13* in 'Morex' is a pseudogene.

Table 4 Matrix of correlations among relative expression levels (at 23°C/light and 6°C/dark conditions), copy number relative quantity (CNRQ) of *HvCBF2A*, *HvCBF4* and *HvCBF14* and frost resistance (FR) measured as photosystem II functionality (F_v/F_m; [24]). Pearson correlation *r* is followed by its two-tailed significance level (*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$).

	FR (F_v/F_m)	Relative expression 23°C/light			Relative expression 6°C/dark			CNRQ	
		HvCBF2A	*HvCBF4*	*HvCBF14*	*HvCBF2A*	*HvCBF4*	*HvCBF14*	*HvCBF2A*	*HvCBF4*
Relative expression 23°C/light									
HvCBF2A	0.661								
HvCBF4	0.711	**0.837 ***							
HvCBF14	0.468	**0.908 ***	0.717						
Relative expression 6°C/dark									
HvCBF2A	0.394	0.259	0.439	0.092					
HvCBF4	-0.017	-0.274	0.119	-0.318	-0.041				
HvCBF14	**0.871 ***	0.600	0.452	0.475	0.277	-0.046			
Copy Number Relative Quantity									
HvCBF2A	**0.814 ***	**0.838 ***	**0.958 ****	0.630	0.596	0.069	0.603		
HvCBF4	**0.831 ***	**0.862 ***	**0.920 ***	0.612	0.516	-0.015	0.641	**0.980 ***	
HvCBF14	0.421	0.372	0.704	0.165	0.405	0.721	0.260	0.702	0.667

Supplemental Tables

Table S1 Primer pairs used to identify BAC clones of 'Nure' *Fr-H2* locus.

Name	Primer sequences (5'-3')	Ta (°C)
Primer pairs derived from [30]		
HvCBF2A	F: CCACAACGCACTCTCGACGC	61
	R: GCATATTCATGGTTTGAGATTG	
HvCBF3	F: CTTGATTTGATTCCGCCTTG	59
	R: TGCTGAATGAGGTCGTCTTG	
HvCBF4	F: ACGAGGAGCAGTGGTTTAGA	55
	R: TTAGCAGTCGAACAAATAGCT	
HvCBF6	F: CGACTCGAGCAACCATAACAAG	60
	R: CCCCAATTTACACCATCACATAC	
HvCBF9	F: CTACTCCACACCTCTCACGAG	61
	R: TCCTTGATTTCGATTCATGGAGTGC	
HvCBF10B	F:TGTACTACTCTACTACTCCCTCCGTTC	58
	R:TGTGCTCCTTTTTACGGATTG	
HvCBF12	NU-F: CATAAAGGGCTAATTGCGGTT	59
	R: CCGGCCTTCTCATCTATCTG	
HvCBF14	F: CTCTCCAGCATCCATCTCTCC	61
	R: AAGCTGTGACGCCGAAAGTTC	
HvCBF15A	F: CTCCCCAAGACACTCAATCG	60
	R: GCGTCTCCTTGAACTTGGTG	
HvCBF15B	F: GAACAGCTTCGGTTTGTTCC	58
	R: AGCTCAATCCTCACCTGCTC	
CBF12c	F:GTGTTTCAGGTTCCGGCTTT	60
	R:ATTCTCCGTTCTCCCAGTGC	
MatE	F: CCATGATCACACTCGCTGTATT	59
	R: ATAACCTCCTCCTCATTGCTGA	
XPG-I	F: AGAGGCTGAAAGATGCAAAATC	58
	R: GGCCCAAGAGAAACACTAACAC	
Put-BRTC	F:CTTTCCATGGCGTAGGGG	59
	R:CGAGGGGAAGAAGTACGACA	
Primer pairs developed in this work		
13J22-186A1	FW: TGCGAGCTATTGTAGTATCCTCA	59
	RV: GAGTTTGTCACGCACCTACC	
6M4	FW: ACCAGTATCTCCTTTGGGCC	59
	RV: CCGCCTGTGTAGTCCATGAT	
63P18	FW: CGCCAATCAAACAACTCCGA	59

	RV: GCCGATGAAGGAAACCGTTT	
25L23	FW: AGCCACTAATCTGCTTACGC	58
	RV: CACTCACTTTCTTGCGGCTT	

Table S2 Primer sequences for *HvCBF* genes selected for expression analysis.

Name	Sequence (5'-3')	Amplicon size (bp)	Reference
EF1 α	ATGATTCCCACCAAGCCCAT ACACCAACAGCCACAGTTTGC	101	[46]
HvCBF4	AGCGCCGCTCTGTTTTACA AGCAGTCGAACAAATAGCTCCA	208	[24]
HvCBF2 [a]	GCAAGGTCGGGCAGTGG<u>A</u>T**G** GCGCCGCCATCTCGGGGTT	106	**This paper**
HvCBF14	AGCCGTTGACGAGAAGGAAGTC GTAGCATGATCCGGCATCCAT	112	[24]

[a] CBF2 forms specific primer: bold letter indicates the SNP site and underlined letter indicates the mismatch site.

Table S3 Estimated composition of the two non-gridded BAC libraries constructed for cv. 'Nure'.

Library code	No. of pools	No. of clones	No. of clones per pool	Avg. insert size (~~kbp~~kb) [a]	Genome coverage [b]
Library Hvu-B-Nure					
Fraction A	96	224,270	2,336	110.0	4.8
Fraction A	96	310,500	3,234	100.0	6.2
Fraction B	64	108,855	1,701	121.3	2.7
Fraction B	92	64,700	703	124.0	1.5
Fraction B	96	77,130	803	124.0	1.5
Fraction C	32	33,005	1,031	137.0	0.9
	476	*814,460*	*1,711*	*119.4*	*17.6*
Library Hvu-B-NureBI					
Fraction A	64	109,675	1,714	109.0	2.4
Fraction A	64	112,750	1,762	109.0	2.5
Fraction A	48	133,813	2,788	108.0	2.9
Fraction A	48	145,242	3,026	108.0	3.1
Fraction B	32	40,230	1,257	138.0	1.1
Fraction C	32	7,688	240	177.0	0.3
	288	*550,398*	*1,911*	*124.8*	*12.3*

[a]: Average insert size estimated with randomly selected BAC clones per fraction;

[b]: Calculated from the estimated number of white colonies and their average insert size

Table S4 Nucleotide and protein similarity of CBF coding sequence between 'Morex' and 'Nure'.

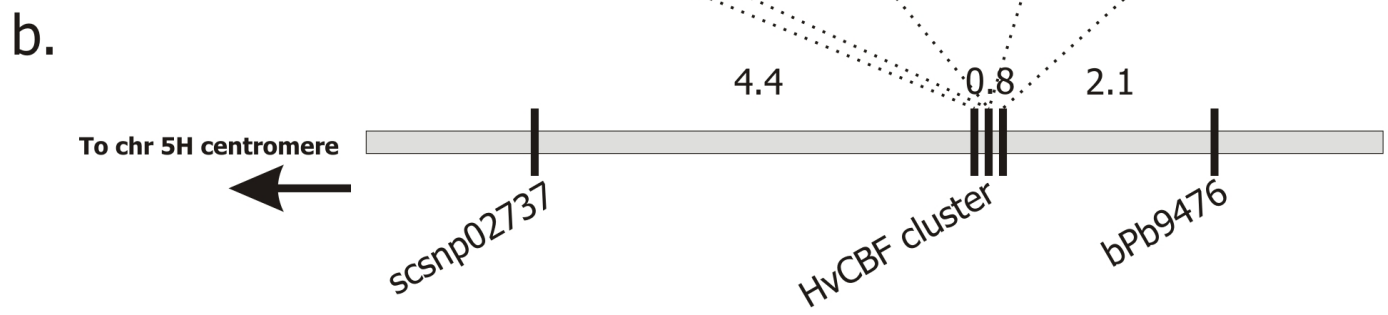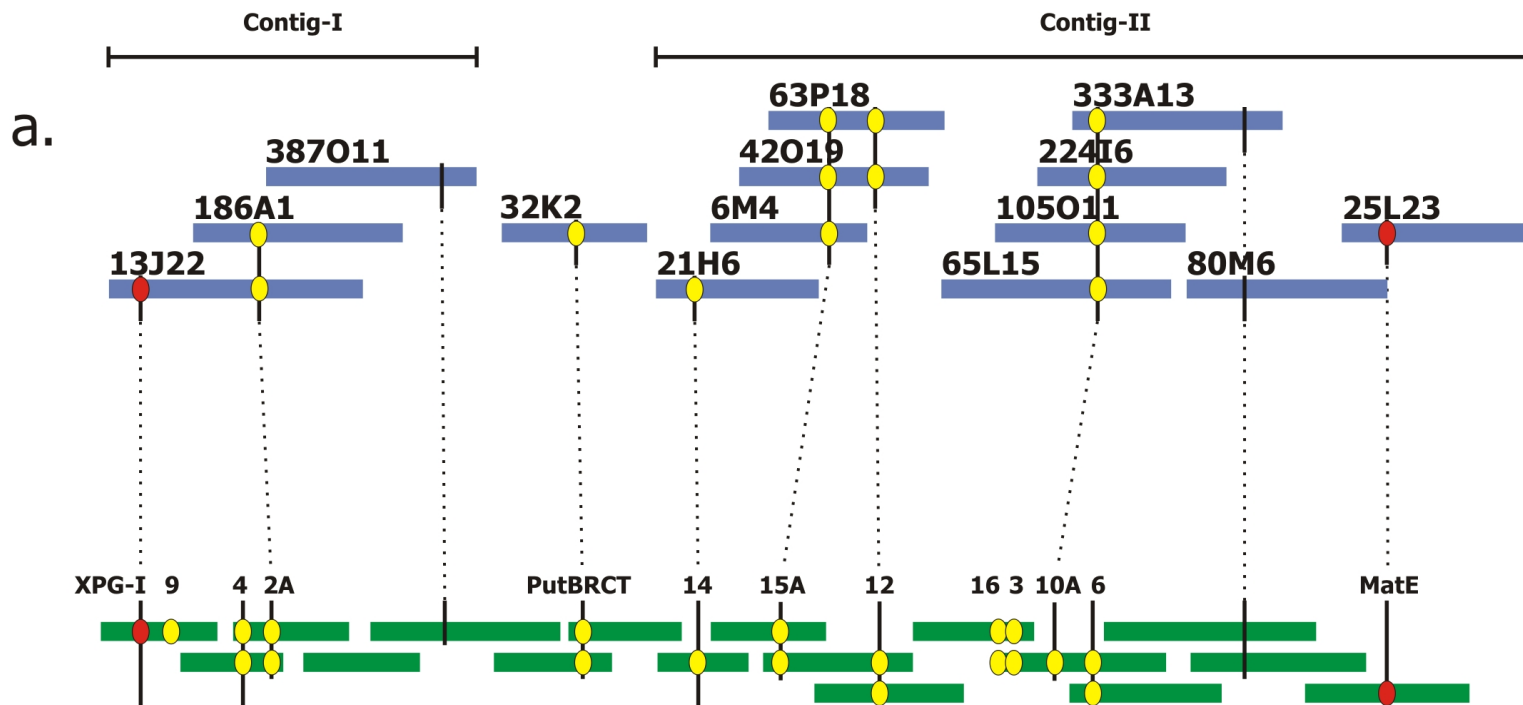| Gene	DNA similarity		Protein similarity		Amino acid substitution	
	Identities	Gaps	Identities	Gaps	Mx→Nu	Replacement features [a]
HvCBF2A	665/666 (99.8%)	0/666	221/221 (100%)	0/221	–	–
HvCBF2C	672/675 (99.6%)	0/675	223/224 (99.6%)	0/224	W19C	Rad, Nt
HvCBF3	743/750 (99.1%)	0/750	246/249 (98.8%)	0/249	S16P	Rad, Nt, Dis, Pol
					T102A	Con, AP2
					V184E	Rad, Ct
HvCBF4	675/678 (99.6%)	0/678	225/225 (100.0%)	0/225	–	–
HvCBF6	726/735 (98.8%)	0/735	239/244 (98.0%)	0/244	C168G	Con, Ct, Dis, Pol
					L177I	Con, Ct, Dis, Pol
					E181D	Con, Ct, Dis, Pol
					A188D	Con, Ct
					A227G	Con, Ct
HvCBF6B	760/762 (99.7%)	0/762	253/253 (100.0%)	0/253	–	–
HvCBF9	874/876 (99.8%)	0/876	289/291 (99.3%)	0/291	S185A	Con, Ct
					N291D	Con, Ct
CBF10A	720/726 (99.2%)	0/729	239/241 (99.2%)	0/241	P102S	Rad, AP2
					A129S	Con, Ct
HvCBF12	733/735 (99.7%)	0/735	243/244 (99.6%)	0/244	A175S	Con, Ct
HvCBF12C	694/702 (98.9%)	0/702	228/233 (97.9%)	0/233	A127V	Con, Ct
					M143L	Con, Ct
					A176T	Con, Ct
					M192V	Con, Ct
					A225G	Con, Ct
HvCBF14	644/645 (99.8%)	0/645	214/214 (100.0%)	0/214	–	–
HvCBF15A	714/714 (100.0%)	0/714	237/237 (100.0%)	0/237	–	–
HvCBF16	679/684 (99.3%)	0/684	225/227 (99.1%)	0/227	G174C	Con, Ct
					P224L	Rad, Ct

[a] Abbreviations used for features' description: conservative or radical replacement, Con or Rad; N- or C-terminal region, Nt or Ct; AP2/ERF-binding domain, AP2; disorder signature, Dis; polar signature, Pol.

Figures Captions

Fig. 1. Minimum Tiling Path of 'Nure' *Fr-H2*. a. Alignment of the physical maps of 'Nure' (above) and 'Morex' (below). *HvCBF* paralogs in the cluster are indicated by their corresponding number [30], while dashed lines connect consistent physical positions in the two MTPs. **b.** Barley chromosome 5H location and genetic anchoring of the physical map. Distances are in Kosambi cM and derive from [9] and [83].

Fig. 2. Expression patterns of *HvCBF2A*, *HvCBF4* and *HvCBF14* genes in seven barley varieties with different levels of frost resistance at 23°C with light stimulus and in the dark at 6°C. a. Experimental design. Samples were collected just before dark and after 10 h at 6°C (gray triangles). **b.** Expression analysis. Transcript levels were calculated with the ΔCt method. Error bars indicate the range defined by the standard error of the ΔCt. **c.** Genotypic and phenotypic characteristics (growth habit and allelic composition at *VRN* genes, copy number relative quantity of *HvCBF4* and *HvCBF2A*, and frost resistance; data from [24,40]) of the seven cultivars. *: 'Tremois' harbors a fused *HvCBF2A/B* form.

Supplemental Fig. 1. Structural comparison of annotated 'Nure' and 'Morex' *Fr-H2* loci. Major features like genes and repetitive elements are drawn as oriented arrows of different shape and color. Genes at *Fr-H2* are in blue and, when predicted, 5'UTR and 3'UTR flanking the open reading frame are superimposed in orange and brown, respectively. Class I retroelements and Class II DNA transposons are in coral and pink, respectively. Names for *XPG-I* and *MatE* – identified as flanking gene elements of the *CBF* cluster in several *Poaceae*-sequenced genomes – are in red. The two exact copies of the 22 ~~kbp~~kb *HvCBF4-HvCBF2A* region and the single *HvCBF4-HvCBF2B* segment are highlighted by green and yellow rectangles, respectively. Dotted line segments designate major sequence differences between 'Nure' (Nu) and 'Morex' (Mx), while dashed lines indicate existing gaps of unknown size in 'Nure'.

a.

0		8		18	24 h
	23°C		6°C		

b.

c.

	Pamina	**Lunet**	**Nure**	**Mellori**	**Ponente**	**Tremois**	**Morex**
Growth habit	F	F	W	W	W	S	S
***VRN-H1/VRN-H2* haplotype**	*vrn1-5200 vrn-H2*	*vrn1-5300 vrn-H2*	*vrn1-5200 Vrn-H2*	*VRN1-6 Vrn-H2*	*VRN1-6 Vrn-H2*	*VRN1-7 vrn-H2*	*VRN1-1 vrn-H2*
CNV at *CBF4* / *CBF2A*	10 / 11	9 / 10	4 / 4	6 / 6	6 / 4	2 / 2*	2 / 2
Frost Res (F$_v$/F$_m$)	0.759*a*	0.748*a*	0.736*ab*	0.721*ab*	0.703*ab*	0.663*bc*	0.620*c*

Reggio Emilia, 02/08/2019

Subject: Declaration of interest – manuscript PSL_2019_951

Ref: PSL_2019_951

Title: Influence of CNV on transcript levels of HvCBF genes at Fr-H2 locus revealed by resequencing in resistant barley cv. 'Nure' and expression analysis

Journal: Plant Science

Authors: Lavinia Mareri, Justyna Milc, Luca Laviano, Matteo Buti, Sonia Vautrin, Stéphane Cauet, Flavia Mascagni, Lucia Natali, Andrea Cavallini, Hélène Bergès, Nicola Pecchioni, Enrico Francia

Corresponding author: Enrico Francia

Conflict of Interest Declaration: The authors declare that they have no conflict of interest