Patterns of nucleotide diversity for different domains of centromeric histone H3 (CENH3) gene in *Secale* L.

S.S. Gatzkaya, E.V. Evtushenko ✉

Institute of Molecular and Cellular Biology, SB RAS, Novosibirsk, Russia
✉ e-mail: evt@mcb.nsc.ru

Rye (*Secale*) is among staple cereals along with other members of the Triticeae tribe: wheat and barley. The genus *Secale* includes perennial and annual, cross-pollinating and self-pollinating species, and they can be donors of valuable genes in wheat and rye breeding programs. Studies of the structure of the gene for centromeric histone H3 (CENH3), essential for centromere functions, are relevant to the breeding of agronomically important crops. We have investigated the nucleotide diversity of sequences of two variants of the rye *CENH3* gene inside the N-terminal tail (NTT) and the conservative HFD (histone fold domain) domain in the genus *Secale*. The mean values of nucleotide diversity in the NTT and HFD of wild cross- and self-pollinating taxa are close in α*CENH3*: π_{tot} = 0.0176–0.0090 and 0.0136–0. 0052, respectively. In the case of β*CENH3*, the mean values for NTT (π_{tot} = 0.0168–0.0062) are lower than for HFD (π_{tot} = 0.0259–0.084). The estimates of nucleotide and haplotype diversity per site for the *CENH3* domains are considerably lower in taxa with narrow geographic ranges: *S. cereale* subsp. *dighoricum* and *S. strictum* subsp. *kuprijanovii*. Commercial breeding reduces the nucleotide sequence variability in α*CENH3* and β*CENH3*. Cultivated rye varieties have π values within 0.0122–0.0014. The nucleotide and haplotype diversity values in α*CENH3* and β*CENH3* are close in *S. sylvestre*, which is believed to be the oldest rye species. The results of this study prove that the frequency of single nucleotide polymorphisms and nucleotide diversity of sequences in genes for CENH3 in *Secale* species are influenced by numerous factors, including reproduction habits, the geographic isolation of taxa, breeding, and the evolutionary age of species.
Key words: *Secale* L.; centromeric histone CENH3; nucleotide diversity; single nucleotide polymorphism.

Паттерны нуклеотидного разнообразия различных доменов гена центромерного гистона H3 (CENH3) у *Secale* L.

С.С. Гацкая, Е.В. Евтушенко ✉

Институт молекулярной и клеточной биологии Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ e-mail: evt@mcb.nsc.ru

Рожь (*Secale*) входит в группу экономически важных злаков наряду с такими представителями трибы Triticeae, как пшеница и ячмень. Род *Secale* включает многолетние, однолетние, перекрестноопыляющиеся и самоопыляющиеся виды, которые используются как источник ценных генов для улучшения существующих сортов пшеницы и ржи. Исследования структуры гена центромерного гистона H3 (CENH3), определяющего функциональную центромеру, сейчас становится актуальным для агрономически важных растений. Мы изучили нуклеотидное разнообразие последовательностей двух вариантов гена *CENH3* ржи внутри N-терминального района (NTT) и консервативного домена (HFD) гена в роде *Secale*. Средние значения нуклеотидного разнообразия у диких перекрестно- и самоопыляющихся таксонов для доменов α*CENH3* были близки для NTT (π_{tot} = 0.0176–0.0090) и HFD (π_{tot} = 0.0136–0.0052), а для β*CENH3* средние значения были меньше в NTT (π_{tot} = 0.0168–0.0062), чем в HFD (π_{tot} = 0.0259–0.084). Значения нуклеотидного и гаплотипного разнообразия для доменов *CENH3* были существенно меньше у таксонов, занимающих узкую географическую нишу, *S. cereale* subsp. *dighoricum* и *S. strictum* subsp. *kuprijanovii*. К снижению изменчивости нуклеотидных последовательностей доменов α*CENH3* и β*CENH3* приводит действие селекции: у сортов культивируемой ржи значения π варьируют от 0.0122 до 0. 0014. Значения нуклеотидного и гаплотипного разнообразия поддерживаются на одном уровне в последовательностях α*CENH3* и β*CENH3* у *S. sylvestre*, считающегося наиболее древним видом ржи. Полученные результаты подтверждают, что на частоту однонуклеотидных полиморфизмов и нуклеотидное разнообразие последо-

S.S. Gatzkaya
E.V. Evtushenko

Patterns of nucleotide diversity for different domains
of centromeric histone H3 (CENH3) gene in *Secale* L.

вательностей вариантов CENH3 у видов *Secale* влияет ряд факторов, включая способы размножения, степень географической изоляции таксона, действие селекции, эволюционный возраст видов.

Ключевые слова: *Secale* L.; центромерный гистон H3 (CENH3); нуклеотидное разнообразие; однонуклеотидный полиморфизм.

Introduction

Rye (*Secale*) belongs to the tribe Triticeae along with other grain crops, such as wheat (*Triticum* spp.) and barley (*Hordeum* spp.) According to the taxonomy proposed by S. Frederiksen and G. Petersen (1998) on the base of morphometrical analysis of rye species, the genus consists of three botanical species: the cross-pollinating perennial species *Secale strictum* Presl., cross-pollinating annual species *S. cereale* L., and self-pollinating annual species *S. sylvestre* Host. The cross-pollinating species *S. cereale* L. and *S. strictum* Presl. include self-pollinating subspecies *S. vavilovii* and *S. africanum*, respectively.

Cultivated and wild rye varieties can be donors of valuable traits, such as winter hardiness, high protein content, and disease resistance. They are used for improvement of existing rye and wheat cultivars and in interspecies crosses of rye and wheat (Tang et al., 2011). Interspecies hybridization is often accompanied by elimination of whole chromosomes or their parts. The incompatibility between centromeres and the centromere-specific histone H3 variant (CENH3) of parents may be one of the causes of chromosome elimination from interspecies hybrids (Sanei et al., 2011). Conversely, close similarity between CENH3 proteins of distant parents can secure the normal function of centromeres and formation of true hybrid plants bearing genomes of both parents (Ishii et al., 2015). The structure of CENH3 includes two domains: the variable N-terminal tail (NTT) and the more conservative C-terminal histone fold domain (HFD) (Roach et al., 2012). The latter interacts with centromeric DNA, whereas NTT is not required for localization on centromeric DNA but is essential for correct chromosome segregation in mitosis and meiosis (Maheshwari et al., 2015). Nevertheless, the segregating polymorphism of *CENH3* genes in grass species is poorly investigated. The multitude of rye species, including annual, perennial, self-pollinating, and cross-pollinating forms, allows assessment of the action of various factors on the genetic variation of *CENH3*.

The objectives of this study were: (1) estimation of the nucleotide and haplotype diversity of *CENH3* domains in three *Secale* species and (2) assessment of the influence of taxonomic and geographic factors and mating systems on nucleotide polymorphisms within the domains in two variants of the *CENH3* gene.

Materials and methods

Plant material and RNA isolation. Experiments were done with ten wild and cultivated rye accessions of three *Secale* species. Seeds of *Secale cereale* subsp. *cereale* (cvs. Otello, Imperial), *S. cereale* subsp. *vavilovii*, *S. cereale* subsp. *dighoricum*, *S. cereale* subsp. *afghanicum*, *S. strictum* subsp. *kuprijanovii*, *S. strictum* subsp. *strictum*, *S. strictum* subsp. *anatolicum*, *S. strictum* subsp. *africanum*, and *S. sylvestre* were supplied by the Leibniz Institute of Plant Genetics and Crop Plant Research (Germany), the US Department of Agriculture

(United States), and the N.I. Vavilov Research Institute of Plant Industry (Russia) from their germplasm collections. Total RNA isolation, synthesis of first-strand cDNA, and PCR were conducted as in (Evtushenko et al., 2017). Amplification primers specific to NTTs and HFDs of α*CENH3* and β*CENH3* genes from rye cDNA had been chosen in (Evtushenko et al., 2017). The amplification products were cloned and sequenced using BigDye Terminator Cycle Sequencing chemistry (v. 3.1) on an ABI3100 Genetic Analyzer (Applied Biosystems, CA, USA).

Sequence analysis. Alignments of *CENH3* coding sequences were performed using online Clustal Omega (Sievers et al., 2011) at http://www.ebi.ac.uk/Tools/msa/clustalo. The DnaSP version 5.10.01 (Librado, Rozas, 2009) was used to estimate the levels of nucleotide diversity for each domain individually in all subspecies with regard to different functions of CENH3 domains. The levels of genetic variation within *CENH3* were estimated as nucleotide diversity π, haplotype diversity H_d, and θ_W, the last index being the relationship between segregating sites and alleles. The Watterson estimator θ_W is based on the number of polymorphic sites in a sample of sequences drawn at random from a population (Watterson, 1975), whereas nucleotide diversity π represents the average sequence divergence of all homologous sequences among all individuals in a given set for comparison (Nei, Li, 1979).

Results

We sequenced the NTT and HFD domains of *CENH3* from 10 to 25 samples per domain for each accession. Formerly, we had shown that the main forms of rye α*CENH3* are β*CENH3* were 501- and 456-bp long, respectively, in all *Secale* species and subspecies (Evtushenko et al., 2017). The lengths of NTTs and HFDs in α*CENH3* and β*CENH3* analyzed in this study are shown in Table 1.

In α*CENH3*, 1 to 21 single-nucleotide polymorphisms (SNPs), or segregating sites, were found inside the domains (Table 2). In the perennial *S. strictum* Presl. subspecies, the number of SNPs in HFD was greater than in NTT, whereas in annual *S. cereale* L. subspecies NTTs had more segregating sites. Among perennial subspecies, the highest genetic variation in the *CENH3* domains was detected in wild cross-pollinating forms: *S. strictum* ssp. *strictum*, hereafter referred to as *S. strictum* and *S. strictum* ssp. *anatolicum* (hereafter *S. anatolicum*). There, π ranged from 0.0131 to 0.0140 in NTT

Table 1. Lengths of coding sequences (CDS) of *CENH3* genes in *Secale*

Full gene/domain	α*CENH3*	β*CENH3*
Full-length CDS (bp)	501	456
N-terminal tail	213	165
HFD	288	291

С.С. Гацкая
Е.В. Евтушенко

Паттерны нуклеотидного разнообразия различных
доменов гена центромерного гистона H3 (CENH3) у *Secale* L.

2019
23•2

Table 2. Estimates of nucleotide diversity in the NTT and HFD of α*CENH3*

Species and subspecies	Growth habit	*CENH3* domain	S_n	$H_d \pm$ SD	Polymorphism ($\times 10^{-2}$)			
					$\theta_W \pm$ SD	$\pi_{tot} \pm$ SD	π_{syn}	π_{nonsyn}
S. strictum ssp. *strictum*	O, P, W	NTT	10	0.972 ± 0.064	1.70 ± 0.85	1.31 ± 0.22	0.985	1.43
		HFD	14	0.805 ± 0.069	1.87 ± 0.78	1.16 ± 0.26	1.80	2.35
S. strictum ssp. *anatolicum*	O, P, W	NTT	15	0.879 ± 0.052	1.95 ± 0.81	1.40 ± 0.26	2.41	1.02
		HFD	21	0.942 ± 0.048	2.84 ± 1.13	1.36 ± 0.21	2.59	1.01
S. strictum ssp. *africanum*	S, P, W	NTT	8	0.867 ± 0.107	1.31 ± 0.67	1.28 ± 0.25	3.21	0.53
		HFD	13	0.881 ± 0.047	1.71 ± 0.27	1.01 ± 0.16	1.98	0.73
S. strictum ssp. *kuprijanovii*	O, P, W	NTT	3	0.700 ± 0.218	0.66 ± 0.47	0.65 ± 0.26	1.69	0.25
		HFD	5	0.380 ± 0.125	0.63 ± 0.33	0.19 ± 0.07	0.35	0.15
S. cereale ssp. *cereale* (Otello)	O, A, Cv	NTT	7	0.736 ± 0.109	1.01 ± 0.51	0.81 ± 0.21	1.66	0.49
		HFD	4	0.476 ± 0.155	0.58 ± 0.34	0.25 ± 0.09	0.84	0.08
S. cereale ssp. *cereale* (Imperial)	O, A, Cv	NTT	14	0.801 ± 0.088	1.82 ± 0.75	0.79 ± 0.16	0.95	0.76
		HFD	4	0.806 ± 0.120	0.64 ± 0.39	0.51 ± 0.12	0.97	0.37
S. cereale ssp. *afghanicum*	O, A, W	NTT	5	0.833 ± 0.222	1.26 ± 0.82	1.23 ± 0.40	2.50	0.74
		HFD	6	0.727 ± 0.144	0.97 ± 0.52	0.52 ± 0.15	1.14	0.34
S. cereale ssp. *vavilovii*	S, A, W	NTT	14	0.924 ± 0.057	2.14 ± 0.97	1.76 ± 0.47	3.56	1.08
		HFD	13	0.758 ± 0.106	1.79 ± 0.77	0.96 ± 0.26	3.55	0.20
S. cereale ssp. *dighoricum*	O, A, W	NTT	3	0.524 ± 0.209	0.56 ± 0.38	0.39 ± 0.19	0.48	0.36
		HFD	1	0.111 ± 0.096	0.14 ± 0.13	0.053 ± 0.046	0.0000	0.068
S. sylvestre	S, A, W	NTT	9	0.879 ± 0.079	1.31 ± 0.62	0.90 ± 0.16	1.87	0.53
		HFD	16	0.859 ± 0.066	2.03 ± 0.81	0.86 ± 0.14	2.24	0.46

N o t e : O – open-pollinated; S – self-pollinated; A – annual; P – perennial; W – wild/weedy; Cv – cultivar; S_n – number of segregating sites; H_d – haplotype diversity; θ_W – Watterson estimator; SD – standard deviation.
Average nucleotide pairwise diversity: π_{tot} – total (synonymous, nonsynonymous); π_{syn} – synonymous; π_{nonsyn} – nonsynonymous.

and from 0.0116 to 0.0136 in HFD. Similarly, the estimates of nucleotide diversity for wild annual cross-pollinating *S. cereale* ssp. *afghanicum* (hereafter *S. afghanicum*) in both α*CENH3* domains were higher than in annual rye cultivars. High nucleotide diversity levels were found in all the three self-pollinating rye accessions studied: *S. cereale* ssp. *africanum* (hereafter *S. africanum*), *S. cereale* ssp. *vavilovii* (hereafter *S. vavilovii*), and *S. sylvestre* Host., regardless of their belonging to annual or perennial *Secale* species. The overall number of segregating sites in both α*CENH3* domains of self-pollinating subspecies varied from 21 in in *S. africanum* to 27 in *S. vavilovii*, and the π values for these subspecies ranged from 0.0128 to 0.0176 in NTT (see Table 2). The estimates of nucleotide diversity for the self-pollinating rye species *S. sylvestre* were somewhat lower, but the overall number of segregating sites was 25 for both domains. The θ_W values in *S. sylvestre* were also higher than in some cross-pollinating subspecies: 1.31 (NTT) and 2.03 (HFD), and the π values for NTT and HFD were close: 0.0090 and 0.0086. Thus, the nucleotide diversity of *CENH3* is equally high in rye accessions from wild cross- and self-pollinating populations.

Perennial *S. strictum* ssp. *kuprijanovii* (*S. kuprijanovii*) and annual *S. cereale* ssp. *dighoricum* (*S. dighoricum*) showed

the lowest nucleotide and haplotype diversities of α*CENH3* among the subspecies: 0.0065 to 0.0039 in NTT and 0.0019 to 0.005 in HFD. Accessions of these subspecies may have originated from a small geographic range, where they had higher inbreeding coefficients to limit geneflow within the taxa (Hagenblad et al., 2016). The comparison of nucleotide diversity values (π_{tot}) for α*CENH3* domains of two rye cultivars and wild rye subspecies showed that the estimates of diversity were low in cultivated rye, which might have resulted from specific breeding features. Haplotype diversity values (H_d) were uniform across all domains of rye α*CENH3*, ranging from 0.972 (*S. strictum*) to 0.736 (*S. cereale* cv. Otello) in NTT and from 0.942 (*S. anatolicum*) to 0.727 (*S. afghanicum*) in HFD with the exception of low H_d values in *S. kuprijanovii* and *S. dighoricum*.

The average estimates of nucleotide diversity π for β*CENH3* were no lower than for α*CENH3* (Table 3). High π values were observed for the wild cross-pollinating subspecies *S. strictum* (NTT, 0.0137; HFD, 0.0120) and *S. afghanicum* (NTT, 0.0134; HFD, 0.0112). In the self-pollinating subspecies *S. africanum* and *S. vavilovii*, we also noted large numbers of segregating sites (21 and 15 in both *CENH3* domains) and high nucleotide diversity (0.0168 and 0.0095 in NTT, 0.0295 and 0.0084 in

S.S. Gatzkaya
E.V. Evtushenko

Patterns of nucleotide diversity for different domains
of centromeric histone H3 (CENH3) gene in *Secale* L.

Table 3. Estimates of nucleotide diversity in the NTT and HFD of β*CENH3*

Species and subspecies	Growth habit	*CENH3* domain	S_n	$H_d \pm SD$	Polymorphism ($\times 10^{-2}$)			
					$\theta_W \pm SD$	$\pi_{tot} \pm SD$	π_{syn}	π_{nonsyn}
S. strictum ssp. *strictum*	O, P, W	NTT	10	0.790 ± 0.105	1.83 ± 0.85	1.37 ± 0.36	2.22	1.06
		HFD	10	0.786 ± 0.151	1.35 ± 0.69	1.20 ± 0.35	3.86	0.34
S. strictum ssp. *africanum*	S, P, W	NTT	11	0.933 ± 0.077	2.35 ± 1.14	1.68 ± 0.35	2.09	1.53
		HFD	10	1.000 ± 0.045	3.69 ± 1.60	2.59 ± 0.59	6.11	1.39
S. strictum ssp. *kuprijanovii*	O, P, W	NTT	6	0.731 ± 0.133	1.15 ± 0.61	0.69 ± 0.17	0.69	0.68
		HFD	8	0.800 ± 0.010	1.01 ± 0.52	0.95 ± 0.17	1.67	0.74
S. cereale ssp. *cereale* (Otello)	O, A, Cw	NTT	11	0.956 ± 0.045	2.05 ± 0.95	1.22 ± 0.22	1.25	1.20
		HFD	9	0.933 ± 0.077	1.11 ± 0.56	0.87 ± 0.08	1.09	0.80
S. cereale ssp. *cereale* (Imperial)	O, A, Cw	NTT	7	0.657 ± 0.138	1.59 ± 0.63	0.54 ± 0.16	0.89	0.42
		HFD	3	0.257 ± 0.142	0.31 ± 0.20	0.14 ± 0.08	0.00	0.18
S. cereale ssp. *afghanicum*	O, A, W	NTT	15	0.875 ± 0.081	2.05 ± 0.88	1.34 ± 0.27	1.67	1.22
		HFD	8	0.900 ± 0.161	1.34 ± 0.78	1.12 ± 0.33	2.89	0.55
S. cereale ssp. *vavilovii*	S, A, W	NTT	9	0.699 ± 0.117	1.55 ± 0.72	0.95 ± 0.28	1.16	0.87
		HFD	11	0.694 ± 0.021	1.41 ± 0.60	0.84 ± 0.17	1.96	0.46
S. sylvestre	S, A, W	NTT	6	0.571 ± 0.014	1.11 ± 0.58	0.62 ± 0.21	0.56	0.64
		HFD	15	0.857 ± 0.090	1.60 ± 0.69	0.90 ± 0.16	2.61	0.37

N o t e: Designations follow Table 2.

HFD). In most accessions, the levels of nucleotide polymorphism were high in both β*CENH3* domains. In the HFD domains of *S. africanum* and *S. sylvestre*, the indices S_n (the number of segregating sites), θ_W, π, and H_d were higher than in NTTs. Of the two paralogous *CENH3* genes of rye, β*CENH3* appears to be the younger (Evtushenko et al., 2017), and this younger age may be responsible for the high genetic variation in HFD, the more conservative *CENH3* region. Both α*CENH3* and β*CENH3* show higher nucleotide diversities at synonymous sites than at nonsynonymous except for the nucleotide diversity values in NTT and HFD of α*CENH3* in *S. strictum* and two cases with $\pi_{syn} = 0.0000$: α*CENH3* of *S. dighoricum* and β*CENH3* of *S. cereale* cv. Imperial. These estimates confirm the effect of purifying selection on rye CENH3 and the possibility of adaptive selection for individual codons, formerly demonstrated by E.V. Evtushenko et al. (2017).

Discussion

We compare nucleotide diversity patterns in domains of the coding sequence of the gene encoding centromeric histone H3 (CENH3), which is one of the epigenetic tags of an active centromere. Centromeres, to which microtubules are attached, define the proper cell segregation in mitosis and meiosis (Comai et al., 2017). Nucleotide diversity comparisons are performed within the genus *Secale*, whose accessions represent annual, perennial, cross-pollinating, and self-pollinating forms; subspecies that experienced geographic isolation; and cultivated varieties. Levels of genetic diversity assessed as the numbers of segregating sites S_n, the Watterson estimator θ_W, nucleotide diversity π, and haplotype diversity H_d in the

sequences of the α*CENH3* and β*CENH3* domains are higher in wild cross-pollinating subspecies, be they perennial or annual. However, the cross-pollinating reproductive habit is by no means the only factor increasing nucleotide diversity in rye *CENH3*. The same trend is observed in self-pollinating perennial *S. africanum* and in the annual subspecies *S. vavilovii* and *S. sylvestre*. In contrast, significantly lower nucleotide diversities π and numbers of segregating sites S_n are found in cross-pollinating subspecies populating small geographic ranges: *S. dighoricum* and *S. kuprijanovii*. The lower *CENH3* nucleotide diversity in rye cultivars in comparison to wild accessions may be related to the effects of different genotypes involved in breeding and the breeding process itself. The estimates of nucleotide diversity for rye *CENH3* are higher than for the *ScVrn1* gene, which controls vernalization sensitivity in rye (Li et al., 2011), but are close to estimates for *HTR12* (*CENH3* analog) in *Arabidopsis lyrata* and *Arabidopsis lyrata* ssp. *petraea* (Kawabe et al., 2006).

Conclusion

The nucleotide variation, or sequence diversity, in rye *CENH3* is elevated by both cross-pollination and self-pollination in large natural populations. Geographic isolation of cross-pollinating rye forms, as well as breeding processes in case of cultivated rye reduce the nucleotide diversity of *CENH3*.

Nowadays, operations with the CENH3 structure are in broad use in breeding programs for production of haploid lines in important crops, capture of heterosis (Karimi-Ashtiyani et al., 2015). Further advance in this field demands identification and knowledge of CENH3 features in commercially important

С.С. Гацкая
Е.В. Евтушенко

Паттерны нуклеотидного разнообразия различных
доменов гена центромерного гистона Н3 (CENH3) у *Secale* L.

2019
23•2

species. Data on the nucleotide diversity of rye *CENH3* may be useful in choosing parents for crosses between wild and cultivated species.

References

Comai L., Maheshwari S., Marimuthu M.P.A. Plant centromeres. Curr. Opin. Plant Biol. 2017;36:158-167. DOI 10.1016/j.pbi.2017.03.003.

Evtushenko E.V., Elisafenko E.A., Gatzkaya S.S., Lipikhina Y.A., Houben A., Vershinin A.V. Conserved molecular structure of the centromeric histone CENH3 in *Secale* and its phylogenetic relationships. Sci. Rep. 2017;7:17628. DOI 10.1038/s41598-017-17932-8.

Frederiksen S., Petersen G. A taxonomic revision of Secale (Triticeae, Poaceae). Nord. J. Bot. 1998;18:399-420. DOI 10.1111/j.1756-1051.1998.tb01517.x.

Hagenblad J., Oliveira H.R., Forsberg N.E.G., Leino M.W. Geographical distribution of genetic diversity in *Secale landrace* and wild accessions. BMC Plant Biol. 2016;16:23. DOI 10.1186/s12870-016-0710-y.

Ishii T., Sunamura N., Matsumoto A., Eltayeb A.E., Tsujimoto H. Preferential recruitment of the maternal centromere-specific histone H3 (CENH3) in oat (*Avena sativa* L.)×pearl millet (*Pennisetum glaucum* L.) hybrid embryos. Chromosome Res. 2015;23:709-718. DOI 10.1007/s10577-015-9477-5.

Karimi-Ashtiyani R., Ishii T., Niessen M., Stein N., Heckmann S., Gurushidze M., Banaei-Moghaddam A.M., Fuchs J., Schubert V., Koch K., Weiss O., Demidov D., Schmidt K., Kumlehn J., Houben A. Point mutation impairs centromeric CENH3 loading and induces haploid plants. Proc. Nat. Acad. Sci. USA. 2015;112:11211-11216. DOI 10.1073/PNAS.1504333112.

Kawabe A., Shuhei Nasuda S., Charlesworth D. Duplication of centromeric histone H3 (*HTR12*) gene in *Arabidopsis halleri* and *A. lyrata*, plant species with multiple centromeric satellite sequences. Genetics. 2006;174(4):2021-2032. DOI 10.1534/genetics.106.063628.

Li Y., Haseneyer G., Schön C.-C., Ankerst D., Korzun V., Wilde P., Bauer E. High levels of nucleotide diversity and fast decline of linkage disequilibrium in rye (*Secale cereale* L.) genes involved in frost response. BMC Plant Biol. 2011;11:6. DOI 10.1186/1471-2229-11-6.

Librado P., Rozas J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009;25:1451-1452. DOI 10.1093/bioinformatics/btp187.

Maheshwari S., Tan E.H., West A., Franklin F.C.H., Comai L., Chan S.W.L. Naturally occurring differences in CENH3 affect chromosome segregation in zygotic mitosis of hybrids. PLoS Genet. 2015;11(2):e1004970. DOI 10.1371/journal.pgen.1004970.

Nei M., Li W.H. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA. 1979;76(4):5269-5273. DOI 10.1073/pnas.76.10.5269.

Roach K.C., Ross B.D., Malik H.S. Rapid evolution of centromeres and centromeric/kinetochore proteins. Eds. R. Singh, J. Xu, R. Kulathinal. Evolution in the Fast Lane: Rapidly Evolving Genes and Genetic Systems. Oxford University Press, 2012;83-93.

Sanei M., Pickering R., Kumke K., Nasuda S., Houben A. Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. Proc. Natl. Acad. Sci. USA. 2011;108(33):498-505. DOI 10.1073/pnas.1103190108.

Sievers F., Wilm A., Dineen D., Gibson T.J., Karplus K., Li W., Lopez R., McWilliam H., Remmert M., Söding J., Thompson J., Higgins D.G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 2011;7:539. DOI 10.1038/msb.2011.75.

Tang Z.X., Ross K., Ren Z.L., Yang Z.J., Zhang H.Y., Chikmawati T., Miftahudin M., Gustafson J.P. *Secale*. Ed. C. Kole. Wild Crop Relatives: Genomic and Breeding Resources. New York: Springer, 2011. DOI 10.1007/978-3-642-14228-4_8.

Watterson G.A. On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. 1975;7:256-276. DOI 10.1016/0040-5809(75)90020-9.

ORCID ID

E.V. Evtushenko orcid.org/0000-0003-2468-5924