

Next-generation sequencing of amplicons is a rapid and reliable method for the detection of polymorphisms relevant for barley breeding

Leonie Fritsch · Raphael Soeur ·
Claudia Hansen · Rainer Fischer ·
Stefan Schillberg · Florian Schröper

Received: 22 February 2016 / Accepted: 6 June 2016 / Published online: 15 June 2016
© Springer Science+Business Media Dordrecht 2016

Abstract Modern plant breeding is based on the detection of traits at the molecular level so that the selection process can be accelerated. Methods that allow the rapid and high-throughput analysis of such traits are therefore useful for plant breeders and farmers. We have shown that next-generation sequencing, applied to short PCR products spanning or flanking polymorphisms of interest in barley, can provide robust genotyping data that allow the rapid determination of genotype and zygosity. This method can be used to genotype large panels of plants because up to 80 million individual reads can be produced in one sequencing run, and samples from different lines and/or traits can be pooled. These findings are significant because plant

breeders may need to screen large populations for multiple traits in parallel. Our next-generation sequencing method therefore provides a simple and inexpensive approach for the rapid and accurate genotyping of natural polymorphisms in barley, which can also be applied in many other economically relevant crop species.

Keywords Candidate gene · Crops · High-throughput genotyping · Polymorphism · QTL · SNP

Electronic supplementary material The online version of this article (doi:[10.1007/s11032-016-0507-6](https://doi.org/10.1007/s11032-016-0507-6)) contains supplementary material, which is available to authorized users.

L. Fritsch · R. Soeur · C. Hansen · R. Fischer ·
S. Schillberg · F. Schröper (✉)
Fraunhofer Institute for Molecular Biology and Applied
Ecology IME, Forckenbeckstrasse 6, 52074 Aachen,
Germany
e-mail: florian.schroeper@ime.fraunhofer.de

R. Fischer
Institute for Molecular Biotechnology, RWTH Aachen
University, Worringerweg 1, 52074 Aachen, Germany

S. Schillberg
Phytopathology Department, Institute for Phytopathology
and Applied Zoology, Justus-Liebig University Giessen,
Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

Introduction

Barley is an important cereal crop which has been cultivated for thousands of years (Reets and Léon 2004). It ranks fourth in terms of production volume behind maize, rice and wheat and is primarily used for food, feed and the production of alcoholic beverages (FAO 2014). Barley is cultivated in different climates, soils and environments and is exposed to diverse forms of abiotic and biotic stress. Cultivars and varieties with yield, nutritional quality and agronomic performance optimized for different environments are therefore needed to supply the growing global population (UN 2008). Optimized characteristics can be achieved by conventional breeding, but this takes up to 8 years because phenotype-based testing requires fully grown plants (Borlaug 1983; ISAAA 2014). The speed and efficiency of plant breeding can be improved by

adopting technologies such as reverse breeding, marker-assisted selection (MAS) and genetic modification, all of which have advantages and disadvantages (He et al. 2014; Jonas and de Koning 2013; Nakaya and Isobe 2012; Varshney et al. 2009). These modern techniques are steadily replacing or augmenting classical breeding approaches (Xu et al. 2012).

Next-generation sequencing (NGS) generates large amounts of data in short time by producing thousands or even millions of reads in parallel. The increasing throughput and falling costs of NGS have encouraged multiple applications in different areas of the life sciences, including medicine (Metzker 2010) and agriculture (Elshire et al. 2011; Mascher et al. 2013; Teixeira et al. 2014; You et al. 2011). For example, the high-throughput sequencing of large numbers of amplicons has been used to genotype the human leukocyte antigen (HLA) locus (Bentley et al. 2009; Holcomb et al. 2011) and to determine zygosity in transgenic maize (Fritsch et al. 2015), but has yet to be applied in the plant breeding sector.

Using a multiplex PCR-based approach (Bentley et al. 2009; Fritsch et al. 2015), we generated amplicons representing naturally occurring polymorphisms in two barley genes, namely the flowering time habit locus *VrnH1* (Fu et al. 2005; Szucs et al. 2007; von Zitzewitz et al. 2005) and the grain protein content

locus *HvNAM1* (Cai et al. 2013; Uauy et al. 2006) (Fig. 1). Each locus represents a relevant trait for breeding programs: *VrnH1* in terms of adaptation to environmental conditions and *HvNAM1* in terms of grain quality. The coverage of more than 1000 reads per PCR product ensured that the sequencing data were statistically valid and reduced the impact of sequencing errors, especially in the hybrids where anticipated polymorphisms only represented half of the reads. We validated our assay using two different genes containing different kinds of polymorphisms, i.e., indels and single nucleotide polymorphisms (SNPs). We found that our genotyping method is easy to implement, requires no PCR optimization steps and is suitable for the high-throughput analysis of many different samples and/or polymorphisms simultaneously.

Materials and methods

Barley seeds, plant cultivation and interbreeding

Barley (*Hordeum vulgare*) seeds of the winter barley line Strider and the spring barley line Morex were used to prepare genomic templates representing the flowering time locus *VrnH1*. The Strider *VrnH1* locus

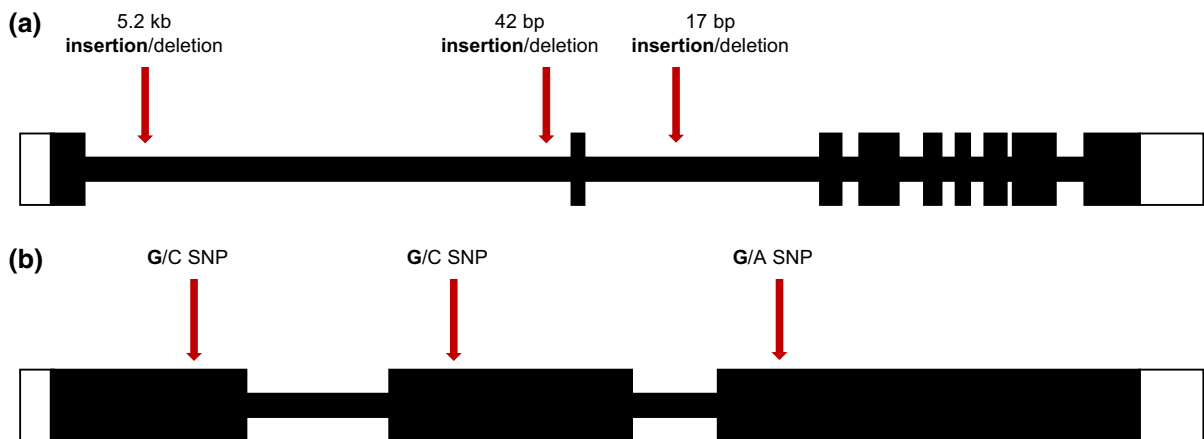


Fig. 1 **a** Schematic depiction of the *VrnH1* gene. The total size of the *VrnH1* gene is 14,776 bp (including the potential insertions). The genotypic features of the winter barley line Strider are shown in *bold*, whereas *normal font* shows the genotypic features of the spring barley line Morex. **b** Schematic depiction of the *HvNAM1* gene. The total size of the *HvNAM1* gene is 1585 bp. The genotypic features of the low grain protein

content barley line Karl are shown in *bold*, whereas *normal font* shows the genotypic features of the high grain protein content barley line Clipper. *Horizontal bars* show stretches of DNA with thick *black bars* representing exons, *thin black bars* representing introns and the *thick white bars* representing untranslated regions. The polymorphisms are indicated by *red arrows*. (Color figure online)

contains additional sequences that are not present in Morex; thus, the locus is characterized by an indel polymorphism (Fig. 1a). Similarly, the Karl and Clipper lines were used to prepare genomic templates representing the grain protein content locus *HvNAM1*. Karl is a low grain protein content line with guanidine residues at three single nucleotide polymorphisms (SNPs) in this locus, whereas Clipper is a high grain protein content line with cytidine residues at the SNPs in the first two exons and an adenosine residue at the SNPs in exon three (Fig. 1b). All seeds were provided by the National Small Grains Collection (Aberdeen, Idaho, USA) of the United States Department of Agriculture (USDA). To generate hybrids, the homozygous seeds were planted in Jiffy-7 pellets (Jiffy Products International BV, Moerdijk, Netherlands) and transferred to pots filled with Einheitserde classic substrate (Einheitserdewerke Werkverband e.V., Sinntal-Altengronau, Germany) after ~14 days. As soon as the awns started to grow out of the husks, the mother plant was emasculated by removing immature anthers from the flowers. To interbreed the plants, ripe anthers from the paternal line were transferred to the emasculated flower and placed on the stigma of the maternal plant 2–3 days after the emasculation. The resulting hybrid seeds were harvested after ~3 months, when the seeds and plants were completely dry (Cornelia Marthe and Dr. Jochen Kumlehn, IPK Gatersleben, Germany, personal communication). To extract DNA, the homozygous and heterozygous seeds were planted in Jiffy-7 pellets and leaves from 7- to 14-day-old plants were processed with the Nucleospin Plant II kit (Machery-Nagel, Düren, Germany) according to the manufacturer's protocol. All DNA samples were dissolved in elution buffer (50 mM Tris-HCl, pH 8.5) to a final concentration of 150–250 ng/μl.

Primers: general aspects

Primers were designed using CLC Main Workbench version 6.9.1 (CLC Bio, Qiagen, Venlo, Netherlands) and synthesized by MWG-Biotech (Ebersberg, Germany). Primers for single PCR and/or for multiplex PCR had a length between 19 and 21 bases and a melting temperature between 58 and 62 °C. The primer sequences are listed in Table S1 (online resource). The reference sequences for primer design were obtained from the NCBI nucleotide database

(<http://www.ncbi.nlm.nih.gov/nuccore/AY750993.1> for the *VrnH1* gene in the Strider line; <http://www.ncbi.nlm.nih.gov/nuccore/AY750995.1> for the *VrnH1* gene in the Morex line; and <http://www.ncbi.nlm.nih.gov/nuccore/EU368851.1> for the *HvNAM1* gene). The barcodes and adapters associated with each primer are listed in Tables S2 and S3 (online resources).

Multiplex PCR to generate NGS templates

NGS templates were prepared by PCR on a Veriti 96-well thermocycler (Life Technologies, Carlsbad, USA) using the Expand High Fidelity PCR System (Roche, Rotkreuz, Switzerland), Axygen eight-strip tubes (Thermo Fisher Scientific, Waltham, USA) and eight-lid flat-cap strips (Sarstedt, Nümbrecht, Germany). Each reaction comprised 3.5 U *Taq/Tgo* DNA polymerase enzyme mix, 500 μM dNTPs, 0.5 μM of each primer (Table S1, online resource) and 150–200 ng of template DNA, topped up to 50 μl with the buffer supplied in the kit. The template was denatured at 95 °C for 2 min and then amplified (30 cycles at 95 °C for 30 s, 55 °C for 30 s and 72 °C for 30 s) followed by a final elongation step for 4 min at 72 °C and indefinite storage at 8 °C. PCR products were sized and quantified by capillary electrophoresis using the Agilent DNA 1000 Kit on an Agilent 2100 Bioanalyzer according to the manufacturer's instructions (Life Technologies). The PCR products were pre-diluted to at least 100 pM according to the concentration determined by capillary electrophoresis (data not shown). The PCR products were diluted according to the concentration of the least concentrated amplicon because it was a multiplex reaction.

Next-generation sequencing

The pre-diluted PCR products were purified and diluted to the final concentration of 26 pM using the Ion Library Equalizer Kit (Life Technologies). Therefore, 2-μl aliquots from each multiplex PCR vessel were pooled and topped up to 50 μl with elution buffer (50 mM Tris-HCl, pH 8.5). The pool was processed with the Ion Library Equalizer Kit using 90 μl Ampure beads and otherwise following the manufacturer's protocol. The pooled and equalized PCR products were then sequenced on an Ion Torrent Sequencer (Life Technologies) using an Ion 316 chip (Life Technologies). The results were analyzed using

Lasergene Genomics Suite software (DNA Star, Madison, USA). The barcodes shown in Table S2 (online resource) were used to differentiate among the samples.

Results and discussion

We developed our NGS-based polymorphism detection procedure using two previously described quantitative trait loci (QTLs), namely the barley flowering time locus *VrnH1* (Fu et al. 2005; Szucs et al. 2007; von Zitzewitz et al. 2005) and the grain protein content locus *HvNAM1* (Cai et al. 2013; Uauy et al. 2006) (Fig. 1). Primers were designed to define amplicons of 146–268 bp spanning or flanking the polymorphisms of interest (Fig. 2). The 5.2-kb indel polymorphism at the *VrnH1* locus (Fig. 1a) was amplified by competitive PCR with two reverse primers: one specific for the insert and one specific for the downstream 3' sequence flanking the insert (Fig. 2a). Amplification of the entire insert with the forward and reverse flanking primers was prevented by limiting the PCR elongation time to 30 s, which is not enough to generate a full-size product of >5.4 kb because the polymerization rate of a standard PCR is approximately 1500 bp/min (Roche 2011). A similar approach was used by Fritsch et al. (2015) to verify the presence of transgenic events in maize. The 42- and 17-bp indels at the *VrnH1* locus were amplified using primers flanking the indels (Fig. 2b). The same strategy was used for the three SNPs at the *HvNAM1* locus. Primers with barcodes (Table S2 of online resource) and adapters (Table S3 of online resource) were used to reduce the number of sample preparation steps by generating short PCR products, covering the indels and SNPs, linked to two adapters and one sample-specific barcode that can be sequenced directly, without prior library preparation. The

preparation of a library involves fragmentation of the target DNA followed by end repair, adapter ligation, purification and size selection (Fig. 3) (Life Technologies 2015; Thermo Fisher Scientific 2014) and in our experience is time-consuming and expensive, especially when processing a large number of samples. We furthermore carried out three PCRs to amplify the three polymorphisms at each locus (Fig. 1) simultaneously in a multiplex reaction, rather than individually in three singleplex reactions, to limit the number of pipetting steps and therefore reduce consumables expenditure. Potentially, if more genetic loci would be investigated in a specific plant line, the PCR could even be multiplexed to a higher extent by adding more primer pairs to a reaction vessel. Certainly, the multiplexing capacity is limited to the extent where all target PCR products are still successfully amplified, which would have to be individually tested for each multiplex reaction. Capillary electrophoresis was carried out using an Agilent 2100 Bioanalyzer to confirm the success of the reactions and to quantify the products. This is particularly important because the PCR products need to be diluted to exactly 26 pM for a successful sequencing (Life Technologies 2013). Capillary electrophoresis showed that the PCRs were successful (data not shown) and bands with the anticipated sizes were observed (data not shown). The samples were pre-diluted to 100 pM according to the capillary electrophoresis results. To reduce the number and cost of pipetting steps even further, 2 μ l of each pre-diluted sample was pooled and processed collectively. The sample pool was diluted to the required concentration and thereby also purified, using a magnetic bead-based procedure with the Ion Library Equalizer Kit, therefore requiring a single reaction tube and set of reagents per sample.

The reads generated by NGS were aligned to the *VrnH1* and *HvNAM1* reference sequences, allowing the genotypes to be clearly distinguished (Table S4 of

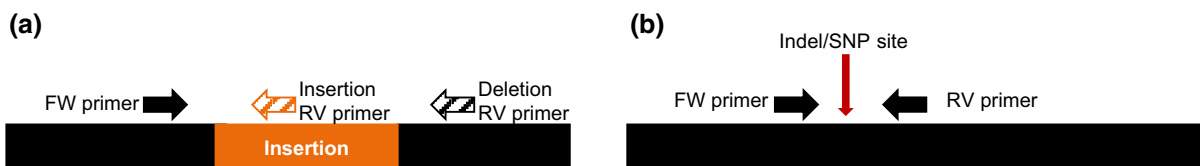


Fig. 2 PCR strategies to amplify polymorphisms of interest. *Horizontal bars* show the DNA; *horizontal arrows* represent primers and the direction of elongation. **a** PCR strategy to amplify the flanking sequence of the 5.2-kb insertion at the

VrnH1 locus. The *orange and black striped arrows* illustrate the use of two distinct reverse primers in a competitive PCR. **b** PCR strategy to amplify the small indels in the *VrnH1* locus and the SNPs in the *HvNAM1* locus. (Color figure online)

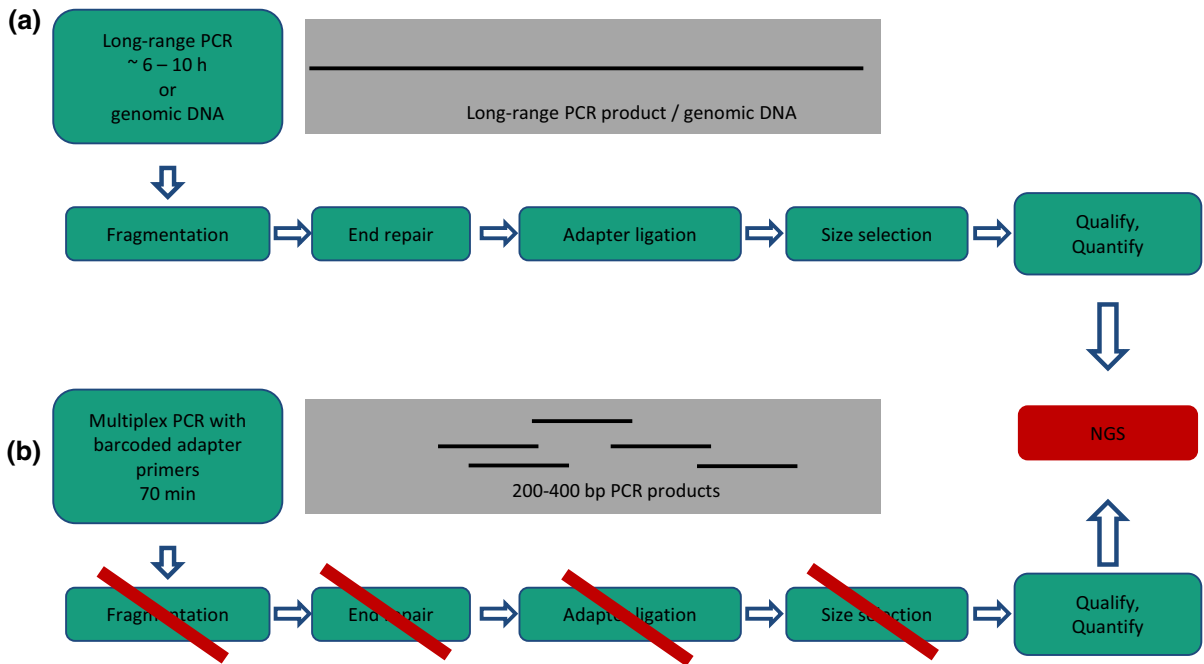


Fig. 3 Schematic depiction and comparison of the workflow with either (a) genomic DNA/long PCR products or (b) short PCR products of 200–400 bp with attached barcodes and adapters. Working steps are shown as required to apply next-

generation sequencing to the samples of interest, based on the designated PCR strategy and starting material. The crossed out steps in (b) are omitted when PCR is carried out with barcoded primers amplifying 200- to 400-bp fragments

online resource). For the *VrnH1* gene, the three anticipated insertions were detected in all reads representing the Strider line, revealing the winter growth genotype of this QTL. The anticipated deletions were detected in all reads representing the Morex line, confirming the spring growth genotype of the QTL in this line. Insertions at these three locations correspond to winter barley alleles, whereas deletions correspond to spring barley alleles (Fu et al. 2005; Szucs et al. 2007; von Zitzewitz et al. 2005). In the Morex × Strider hybrid, the insertions and deletions were distributed in approximately equal shares among the reads (Table S4 of online resource) and the loci could therefore be identified as heterozygous. The genotype of the 5.2-kb indel could be detected by counting the reads aligned either to the 5' part of the insertion or to the downstream flanking sequence of the anticipated insertion (Fig. 4) because competitive PCR was carried out with three primers (Fig. 2a). The Strider line exclusively generated reads matching the 5' sequence of the insertion, whereas the Morex line exclusively generated reads matching the downstream flanking sequence. These results show that the Strider

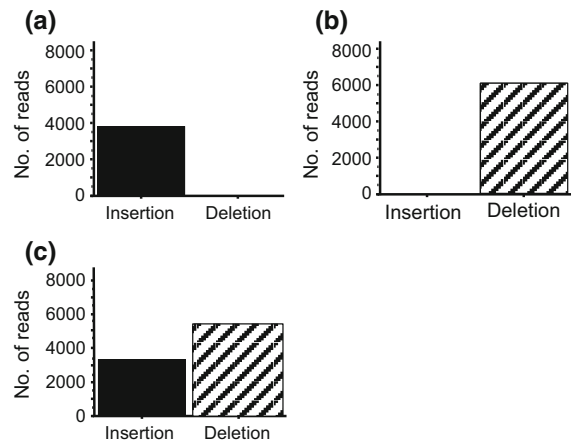


Fig. 4 Allelic identification at the 5.2-kb indel site of the *VrnH1* gene, showing the number of reads aligned to the 5' sequence of the insertion, or to the flanking sequence indicating a deletion. **a** Reads on the Strider template. **b** Reads on the Morex template. **c** Reads on the heterozygous Morex × Strider template. *Insertion*: reads aligned to the 5'-insert sequence. *Deletion*: reads aligned to the sequence flanking the insertion site

line contains the 5.2-kb insertion which is not present in the Morex line. The Morex × Strider hybrid generated reads aligning to both reference sequences,

confirming that both indel alleles are present. The 42- and 17-bp indels were detected as gaps in the sequence. The sequencing results for the 17-bp indel are shown as an example in Table 1.

The three SNPs of interest in the *HvNAM1* gene (Fig. 1b) were also detected by sequencing. The Clipper line contains guanidine residues at nucleotide positions 234 in the first exon and 544 in the second exon, whereas cytidine residues occupy both positions in the Karl line. In the third exon, the Clipper line contains a guanidine residue at nucleotide position 1433, whereas the Karl line contains an adenosine residue at this site. The Clipper alleles correspond to a high grain protein content, whereas the Karl alleles correspond to a low grain protein content phenotype (Cai et al. 2013; Uauy et al. 2006). The Karl \times Clipper hybrid showed a near-equal distribution of the two alternative nucleotides in each position (Table S4 of online resource), confirming the heterozygosity of the hybrid at locus. The sequencing results for the SNP at nucleotide position 234 are shown as an example in Table 1. An overview of the sequencing results and read numbers is given in the online resource (Table S4).

The different numbers of reads aligned to the reference sequences of each locus (Fig. 4; Table S4 of online resource) may reflect the uneven amplification efficiency of the multiplex PCR, which can be caused by differences in primer binding efficiency, the favored amplification of a specific target or the formation of primer dimers that inhibit amplification (Le et al. 2009). Furthermore, read errors/low-quality reads are excluded from the final dataset. This often occurs when reads are automatically trimmed or filtered out, e.g., when they are polyclonal or produce an off-scale signal on the Ion Torrent server (Life Technologies 2014). However, there was no need to normalize or equalize the PCRs in our method because a few reads are theoretically sufficient to confirm the presence of a given allele by mapping to a unique reference sequence. In heterozygous samples, those reads should be distributed in a near-equal manner. Although specific limits have not been proposed, higher read numbers are known to reduce the error frequency significantly (Sims et al. 2014), and we therefore propose that a coverage of at least 30 reads per PCR product is desirable.

We have demonstrated that single nucleotide polymorphisms and indel polymorphisms of different sizes can be characterized by NGS in terms of genotype and

zygosity. Our assay is therefore useful in the context of barley breeding because diverse polymorphisms in several genes can be investigated simultaneously. The duration and cost of the assay are reduced by the multiplex PCR with barcoded adapters, so that all samples, from different plants and with distinct adapters, can be pooled for parallel dilution, purification and sequencing. So far, 96 different barcodes are available and have been described (Elshire et al. 2011; Life Technologies 2015), enabling the analysis of 96 individual samples per gene of interest in one sequencing run. Therefore, when a large sequencing chip such as the Ion 318 (Life Technologies) with a capacity of 80 million reads (Life Technologies 2015) is used, and the read number is optimized to 30 parallel reads per PCR product, it would be possible to screen more than 27,000 individual PCR products representing polymorphisms in specific loci using 96 different samples or plant lines. These numbers have been theoretically calculated by considering that the capacity of reads divided by the number of barcodes and desired read depth gives the number potentially screenable PCR products ($80 \text{ million} \div 96 \div 30 = 27,777.78$). This large capacity makes the assay suitable for the high-throughput analysis of genotypes. It should also be possible to increase the number of barcodes as soon as they are defined in the sequencing software, so that even more samples could be processed in a single run. In general, our method can be carried out on any sequencing device that allows amplicon sequencing. Therefore, the number of screenable PCR products depends on the capacity of the used sequencer and its properties. As demonstrated by Campbell et al. (2015), who established an NGS assay with 100-bp reads to genotype rainbow trout *Oncorhynchus mykiss* for SNP markers, a different PCR strategy, with two thermal cycling steps, could also be used to increase the number of samples.

Many parallel reads of the polymorphisms in each of the finely mapped genes of interest provide robust data about the genotype and therefore allow the phenotype to be predicted. Plants can be selected as soon as they have grown enough to extract DNA, usually 1–2 weeks after planting. It has yet to be determined whether this assay is suitable for polyploid crops such as potato and wheat, and whether it can be implemented in plant species with large genomes and/or large numbers of transposable elements (Choulet et al. 2010; Schnable et al. 2009). To

Table 1 Sequencing results for (a) the 17-bp indel in the *VrmH1* gene and (b) the SNP at nucleotide position 243 in exon 1 for the *HvNAM1* gene. (Color figure online)

(a)			
Target	Barcode	Plant line	Sequence (5' → 3')
<i>VrmH1</i> 17 bp indel	Reference	Strider	...ggaacgacgcccctcaggcgctcaacgcccggccatg gcccgccagg gatttcgcccgaaac...
<i>VrmH1</i> 17 bp indel	A04	Morex	...ggaacgacgcccctcaggcgctcaacgc ————— cagggatttcgcccgaaac...
<i>VrmH1</i> 17 bp indel	A05	Strider	...ggaacgacgcccctcaggcgctcaacgcccggccatg gcccgccagg gatttcgcccgaaac...
<i>VrmH1</i> 17 bp indel	A06	Morex x Strider	...ggaacgacgcccctcaggcgctcaacgcccggccatg gcccgccagg gatttcgcccgaaac... ...ggaacgacgcccctcaggcgctcaacgc ————— cagggatttcgcccgaaac...
(b)			
Target	Barcode	Plant line	Sequence (5' → 3')
<i>HvNAM1</i> SNP exon 1	Reference	Clipper	...cgaggTggacctctacaagttcgaccatg cgagctccccgg tatgactactagtagtagtact...
<i>HvNAM1</i> SNP exon 1	A01	Karl	...cgaggTggacctctacaagttcgaccatg ggagctccccgg taTgactactagtagtagtact...
<i>HvNAM1</i> SNP exon 1	A02	Clipper	...cgaggTggacctctacaagttcgaccatg cgagctccccgg tatgactactagtagtagtact...
<i>HvNAM1</i> SNP exon 1	A03	Karl x Clipper	...cgaggTggacctctacaagttcgaccatg ggagctccccgg taTgactactagtagtagtact... ...cgaggTggacctctacaagttcgaccatg cgagctccccgg tatgactactagtagtagtact...

The polymorphisms and flanking sequences are aligned to show the different alleles as well as their zygosity. Polymorphic sites are highlighted using bold letters. Red lines and red font indicate polymorphisms that differ from the reference sequence

design ideal primers enabling the PCR amplification of desired polymorphic regions, preliminary investigations of relevant traits and associated gene loci as well as polymorphisms affecting the phenotype are necessary. PCRs should be designed based on these reference sequences of choice, so that PCR products can be generated in most of the cases, but at least in individuals carrying the desired loci variations. Further polymorphisms especially in the primer binding regions occurring in more diverse barley lines will not be addressed for the selection process. Thus, a failed PCR product will just lead to no sequencing result and such plant lines will be sorted out from the selection process. To reduce the number of failed PCRs in such cases, it would be possible to design several primers in close proximity to the original primer binding site to increase the probability of a successful primer annealing. Also, the use of wobble primers would be possible if certain mutations within the primer binding sites are known. However, the straightforward implementation of our assay and its consistent results address the limitations of other modern plant breeding techniques such as MAS (Jonas and de Koning 2013; Nakaya and Isobe 2012). Sample preparation is simple, and library preparation steps such as fragmentation, sizing and adapter ligation can be omitted, thus providing an advantage over other genotyping-by-sequencing approaches (Deschamps et al. 2012; Elshire et al. 2011; He et al. 2014). Although the assay relies on prior knowledge concerning the sequence of polymorphic sites and the distribution of alleles linked to certain phenotypes, as long as association studies and the fine mapping of traits continue to be used in plant breeding, our assay could nevertheless be adopted widely as a new tool for high-throughput selection in many species of crops and other plants.

Acknowledgments We acknowledge Harold E. Bockelman from the United States Department of Agriculture (USDA) for providing the barley seeds. We acknowledge Cornelia Marthe and Dr. Jochen Kumlehn (IPK Gatersleben, Group Plant Reproductive Biology) for sharing their expertise on the crossbreeding of barley. We acknowledge Dr. Richard M Twyman for editing the manuscript. This work is supported by the BMBF (0316075B).

References

- Bentley G, Higuchi R, Hoglund B, Goodridge D, Sayer D, Trachtenberg EA, Erlich HA (2009) High-resolution, high-throughput HLA genotyping by next-generation sequencing. *Tissue Antigens* 74:393–403. doi:10.1111/j.1399-0039.2009.01345.x
- Borlaug NE (1983) Contributions of conventional plant breeding to food production. *Science* 219:689–693
- Cai S, Yu G, Chen X, Huang Y, Jiang X, Zhang G, Jin X (2013) Grain protein content variation and its association analysis in barley. *BMC Plant Biol* 13:35. doi:10.1186/1471-2229-13-35
- Campbell N, Harmon S, Narum S (2015) Genotyping-in-thousands by sequencing (GT-seq): a cost effective SNP genotyping method based on custom amplicon sequencing. *Mol Ecol* 15:855–867. doi:10.1111/1755-0998.12357
- Choulet F et al (2010) Megabase level sequencing reveals contrasted organization and evolution patterns of the wheat gene and transposable element spaces. *Plant Cell* 22:1686–1701. doi:10.1105/tpc.110.074187
- Deschamps S, Llaca V, May GD (2012) Genotyping-by-sequencing in plants. *Biology* 1:460–483. doi:10.3390/biology1030460
- Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6:e19379. doi:10.1371/journal.pone.0019379
- Food and Agriculture Organization of the United Nations (2014) Food and nutrition in numbers. <http://www.fao.org/statistics/en/>. Accessed 04 March 2015
- Fritsch L, Fischer R, Wambach C, Dudek M, Schillberg S, Schröper F (2015) Next-generation sequencing is a robust strategy for the high-throughput detection of zygosity in transgenic maize. *Transgenic Res* 4:615–623. doi:10.1007/s11248-015-9864-x
- Fu D et al (2005) Large deletions within the first intron in VRN-1 are associated with spring growth habit in barley and wheat. *Mol Genet Genom* 273:54–65
- He J, Zhao X, Laroche A, Lu ZX, Liu H, Li Z (2014) Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. *Front Plant Sci* 5:484. doi:10.3389/fpls.2014.00484
- Holcomb CL et al (2011) A multi-site study using high-resolution HLA genotyping by next generation sequencing. *Tissue Antigens* 77:206–217. doi:10.1111/j.1399-0039.2010.01606.x
- ISAAA (2014) International service for the acquisition of agribiotech applications. <http://isaaa.org/contactus/default.asp>. Accessed 27 June 2014
- Jonas E, de Koning DJ (2013) Does genomic selection have a future in plant breeding? *Trends Biotechnol* 31:497–504. doi:10.1016/j.tibtech.2013.06.003
- Le T, Hidalgo Ashrafi E, Paul N (2009) Enhancing multiplex PCR efficiency using Hot Start dNTPs. *Biotechniques* 47:972–973
- Life Technologies (2013) Ion PGM™ Template OT2 400 Kit, 3rd edn
- Life Technologies (2014) Trimming and filtering. Life Technologies. <http://137.226.86.55/ion-docs/Technical-Note—Filtering-and-Trimming>. Accessed 27 Nov 2014
- Life Technologies (2015) Ion Torrent™ next-generation sequencing <http://www.lifetechnologies.com/de/de/home/brands/ion-torrent.html?icid=fr-iontorrent-1>. Accessed 16 May 2016

- Mascher M, Wu S, Amand PS, Stein N, Poland J (2013) Application of genotyping-by-sequencing on semiconductor sequencing platforms: a comparison of genetic and reference-based marker ordering in barley. *PLoS ONE* 8:e76925. doi:[10.1371/journal.pone.0076925](https://doi.org/10.1371/journal.pone.0076925)
- Metzker ML (2010) Sequencing technologies—the next generation. *Nat Rev Genet* 11:31–46. doi:[10.1038/nrg2626](https://doi.org/10.1038/nrg2626)
- Nakaya A, Isobe SN (2012) Will genomic selection be a practical method for plant breeding? *Ann Bot* 110:1303–1316. doi:[10.1093/aob/mcs109](https://doi.org/10.1093/aob/mcs109)
- Reets T, Léon J (2004) Die Erhaltung der genetischen Diversität bei Getreide. Auswahl einer Gersten Core-Collection aufgrund geographischer Herkunft, Abstammung, Morphologie, Qualität, Anbaubedeutung und DNA Markeranalysen. In: Lehr- und Forschungsschwerpunkt „Umweltverträgliche und Standortgerechte Landwirtschaft“, vol 119. Rheinische Friedrich-Wilhelm-Universität Bonn
- Roche (2011) Expand High Fidelity PCR System, 20th edn
- Schnable PS et al (2009) The B73 maize genome: complexity, diversity, and dynamics. *Science* 326:1112–1115. doi:[10.1126/science.1178534](https://doi.org/10.1126/science.1178534)
- Sims D, Sudbery I, Ilott NE, Heger A, Ponting CP (2014) Sequencing depth and coverage: key considerations in genomic analyses. *Nat Rev Genet* 15:121–132. doi:[10.1038/nrg3642](https://doi.org/10.1038/nrg3642)
- Szucs P et al (2007) Validation of the VRN-H2/VRN-H1 epistatic model in barley reveals that intron length variation in VRN-H1 may account for a continuum of vernalization sensitivity. *Mol Genet Genomics* 277:249–261
- Teixeira RT, Fortes AM, Pinheiro C, Pereira H (2014) Comparison of good- and bad-quality cork: application of high-throughput sequencing of phellogenetic tissue. *J Exp Bot* 65:4887–4905. doi:[10.1093/jxb/eru252](https://doi.org/10.1093/jxb/eru252)
- Thermo Fisher Scientific (2014) Thermo Scientific™ ClonSeq™ Library Preparation Kit, Ion Torrent™ compatible
- Uauy C, Distelfeld A, Fahima T, Blechl A, Dubcovsky J (2006) A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. *Science* 314:1298–1301. doi:[10.1126/science.1133649](https://doi.org/10.1126/science.1133649)
- UN (2008) Population of the world, major development groups and major areas, 1950, 1975, 2009 and 2050, according to different variants
- Varshney RK, Nayak SN, May GD, Jackson SA (2009) Next-generation sequencing technologies and their implications for crop genetics and breeding. *Trends Biotechnol* 27:522–530. doi:[10.1016/j.tibtech.2009.05.006](https://doi.org/10.1016/j.tibtech.2009.05.006)
- von Zitzewitz J et al (2005) Molecular and structural characterization of barley vernalization genes. *Plant Mol Biol* 59:449–467. doi:[10.1007/s11103-005-0351-2](https://doi.org/10.1007/s11103-005-0351-2)
- Xu Y, Lu Y, Xie C, Gao S, Wan J, Prasanna BM (2012) Whole-genome strategies for marker-assisted plant breeding. *Mol Breeding* 29:833–854. doi:[10.1007/s11032-012-9699-6](https://doi.org/10.1007/s11032-012-9699-6)
- You FM et al (2011) Annotation-based genome-wide SNP discovery in the large and complex *Aegilops tauschii* genome using next-generation sequencing without a reference genome sequence. *BMC Genomics* 12:59. doi:[10.1186/1471-2164-12-59](https://doi.org/10.1186/1471-2164-12-59)