UC Davis
UC Davis Previously Published Works

Title

Combining Traditional Mutagenesis with New High-Throughput Sequencing and Genome Editing to Reveal Hidden Variation in Polyploid Wheat

Permalink

Journal

ISSN

Authors

Uauy, Cristobal
Wulff, Brande BH
Dubcovsky, Jorge

Publication Date

DOI

Peer reviewed

ANNUAL REVIEWS

Annual Review of Genetics

Combining Traditional Mutagenesis with New High-Throughput Sequencing and Genome Editing to Reveal Hidden Variation in Polyploid Wheat

Cristobal Uauy,[1] Brande B.H. Wulff,[1] and Jorge Dubcovsky[2]

[1]John Innes Centre, Norwich Research Park, Norwich NR4 7UH, United Kingdom; email: cristobal.uauy@jic.ac.uk

[2]Howard Hughes Medical Institute and Department of Plant Sciences, University of California, Davis, California 95616, USA

Keywords

Abstract

Induced mutations have been used to generate novel variation for breeding purposes since the early 1900s. However, the combination of this old technology with the new capabilities of high-throughput sequencing has resulted in powerful reverse genetic approaches in polyploid crops. Sequencing genomes or exomes of large mutant populations can generate extensive databases of mutations for most genes. These mutant collections, together with genome editing, are being used in polyploid species to combine mutations in all copies of a gene (homoeologs), and to expose phenotypic variation that was previously hidden by functional redundancy among homoeologs. This redundancy is more extensive in recently formed polyploids such as wheat, which can now benefit from the deployment of useful recessive mutations previously identified in its diploid relatives. Sequenced mutant populations and genome editing have changed the paradigm of what is possible in functional genetic analysis of wheat.

POLYPLOIDY, DIPLOIDIZATION, AND THEIR EFFECT ON FORWARD AND REVERSE GENETIC APPROACHES

Our physical landscape is constantly modified by the relatively fast rise of mountains followed by a slower erosion process that over the years reduces the height and sharp edges of the original mountains. Similarly, the plant genome landscape has been punctuated by polyploidization events that result in the immediate duplication of most genes in the genome followed by long periods of diploidization that dilute and gradually erase the previous duplications (2, 40). The diploidization process includes the differential deletion or inactivation of duplicated genes in one of the genomes as well as the divergence of the multiple copies by sub- or neofunctionalization (8, 53).

Because diploidization is a continuous process, the older the polyploidization event is, the more advanced the diploidization process is. For example, the grass lineage shares an old polyploidization event that occurred roughly 70 Mya (63). The diploidization process was so extensive after this long period that many of the resulting species (e.g., rice, barley, sorghum) were considered true diploids until very recently. Only when genomic tools revealed extensive duplications among chromosome regions did it become evident that species such as rice were ancient polyploids (28, 63). Younger polyploidization events, such as the one in maize 5–15 Mya (7, 92), were discovered earlier—when colchicine treatments showed the formation of occasional cells with five groups of four paired chromosomes in meiosis (58)—and later confirmed by genomic studies (73). Finally, more recent polyploid species such as tetraploid wheat, which originated less than 700,000 years ago (55), have been recognized as a polyploid species for a long time based on extensive gene duplication (57). Bread wheat, a hexaploid species containing genomes AABBDD that originated less than 10,000 years ago from the hybridization of tetraploid wheat (genomes AABB) with the diploid species *Aegilops tauschii* (genome DD) (reviewed in 71), has had almost no time for diploidization, and most genes are expected to have overlapping functions. There is, of course, the divergence among the diploid genomes during the period from their last common ancestor up until their polyploidization, which in the case of the bread wheat genomes was less than seven million years (55).

Polyploids are frequently classified as allopolyploids when the diploid species involved in the polyploidization event are different and have diverged enough to result in clear diploid inheritance. By contrast, autopolyploids originate from whole genome duplications or duplication of hybrids within the same species or between closely related species, which results in a tetraploid inheritance. There is continuous gradation between these two extreme categories. We argue here that it is important to classify polyploids also by the age of the polyploidization event because young and old polyploids exhibit different characteristics that are relevant for strategies to improve polyploid crop species. For example, loss-of-function mutations in single copy genes in ancient polyploid species frequently result in observable changes, and forward mutation screens are very effective tools to identify mutations affecting particular traits (62). By contrast, the effects of gene mutations in a recent polyploid species such as bread wheat are more frequently masked by redundant gene copies of the other genomes (designated homoeologs) compared to diploid species. Therefore, the importance of reverse genetic tools relative to forward genetic tools increases in young polyploid species.

SELECTION OF DIFFERENT TYPES OF VARIATION IN DIPLOID AND POLYPLOID SPECIES

Since the domestication of barley and wheat in the Fertile Crescent approximately 10,000 years ago, both species expanded westward and eastward with the spread of agriculture (15, 71). This brought barley and wheat into new environments that expanded south and north of the original

range of the wild species, exposing these crops to different photoperiod and temperature oscillations. These changes imposed a strong selection pressure on genes and pathways critical for reproductive success in the new environments. Since ancient times, wheat growers have been selecting mutations in reproductive regulatory genes that are critical to obtain plants that flower at the optimal time and maximize grain production (**Figure 1**). Ancestral photoperiod-sensitive winter barley and wheat forms evolved to generate different types of photoperiod sensitivity and different degrees of vernalization requirements. Because wheat and barley expanded together (71) and are still frequently grown in similar locations, the adaptive changes in the reproductive systems of these two related species provide a unique opportunity to compare changes selected in diploid and polyploid species.

Allelic differences in *PPD-1* are responsible for most of the natural variation in the photoperiodic response in barley and wheat. In barley, a loss-of-function recessive mutation in *PPD-H1* (the H indicates the H genome of barley) results in limited acceleration of flowering during long days (97). By contrast, the deletions selected in the promoter regions of the wheat *PPD-A1* and *PPD-D1* genes result in the misexpression of these genes during the night (5, 107), which is sufficient to accelerate flowering during short days (64). Although natural polyploid wheats with simultaneous loss-of-function mutations in all *PPD-1* homoeologs have not been reported so far, the combination of loss-of-function mutations in the three *PPD-1* homoeologs in hexaploid wheat using molecular markers results in a reduction in the acceleration of flowering during long days, similar to the one reported in barley (64, 80).

The difference in the types of mutations selected for *PPD-1* between diploid and polyploid species is also evident in the *VRN2* vernalization gene, which acts as a flowering repressor (**Figure 1**). Natural variation in the vernalization requirement for both barley and wheat is frequently associated with deletions or mutations in the regulatory regions of the *VRN1* gene (22, 30, 43, 46, 60), which promotes the transition of the apical meristem between the vegetative and reproductive stages. However, a spring growth habit associated with natural loss-of-function mutations or deletions of the *VRN2* gene has been observed so far only in barley and diploid species of wheat (17, 100, 110) (**Figure 1**). The combination of nonfunctional or deleted *VRN2* genes in all three homoeologous *VRN2* loci in hexaploid wheat using molecular markers results in a spring growth habit (42), suggesting that the lack of similar combinations in nature was likely due to the inability to select phenotypically the individual mutations in the presence of other functional homoeologs.

These examples suggest that recessive mutations are more difficult to select in the young polyploid genomes than in diploid wheat and barley species. This is likely because of the extensive functional overlap among wheat homoeologous genes, which have had very limited time to diverge since the recent polyploidization events that gave rise to the polyploid wheat species. This has favored the selection of dominant or semidominant mutations in the polyploid wheat species during domestication and more recent breeding efforts. These mutations are usually epistatic over the other homoeologs, facilitating the rapid detection of favorable phenotypes. The selection of the *Q* gene (*APETALA2* on chromosome 5A) responsible for free-threshing grains in wheat (82) is another good example of these types of mutations. The free-threshing *Q* allele originated from a silent mutation in the microRNA miR172 binding site in the 5A homoeolog that reduces the cleavage of this specific transcript. This mutation results in high levels of AP2 thus conferring a dominant effect on the phenotype that is not dependent on additional mutations in the other *AP2* homoeologs (13).

A large group of well-studied genes for resistance to wheat and barley rusts provides an additional example of the different types of mutations selected in polyploid and diploid *Triticeae* species. In tetraploid and hexaploid wheat, almost all characterized resistance genes are dominant

or semidominant (19). Among the 135 rust resistance genes with inheritance information listed in the 2017 Catalogue of Gene Symbols for Wheat, only 9 genes are characterized as recessive (6.7%) (57). The proportion of reported recessive rust resistance alleles in barley is higher (26.3%, χ^2 $P = 0.0007$) (see **Supplemental Table 1**). In diploid wheat, *Triticum monococcum*, we recently mapped the recessive stem rust resistance gene *SrTm4* on a region of the long arm of homoeologous group 2 where no recessive rust resistance genes have been reported so far in polyploid wheat (9). The barley rust resistance gene *Mildew Locus O* (*MLO*) is also a good example of a mutation-induced recessive resistance gene selected only in a diploid species (10). Recently, Wang et al. (103) and Acevedo-Garcia et al. (1) demonstrated the feasibility of using either gene editing or induced mutants to generate triple mutants for the A, B, and D genome homoeologs of *MLO* in wheat by crossing, which conferred strong resistance to powdery mildew.

The fact that mutations in a single wheat homoeolog are frequently masked by redundancy from other homoeologs suggests that recessive variation in polyploid wheat has remained hidden, for the most part, from natural and human selection (8, 47). In many cases, single and double
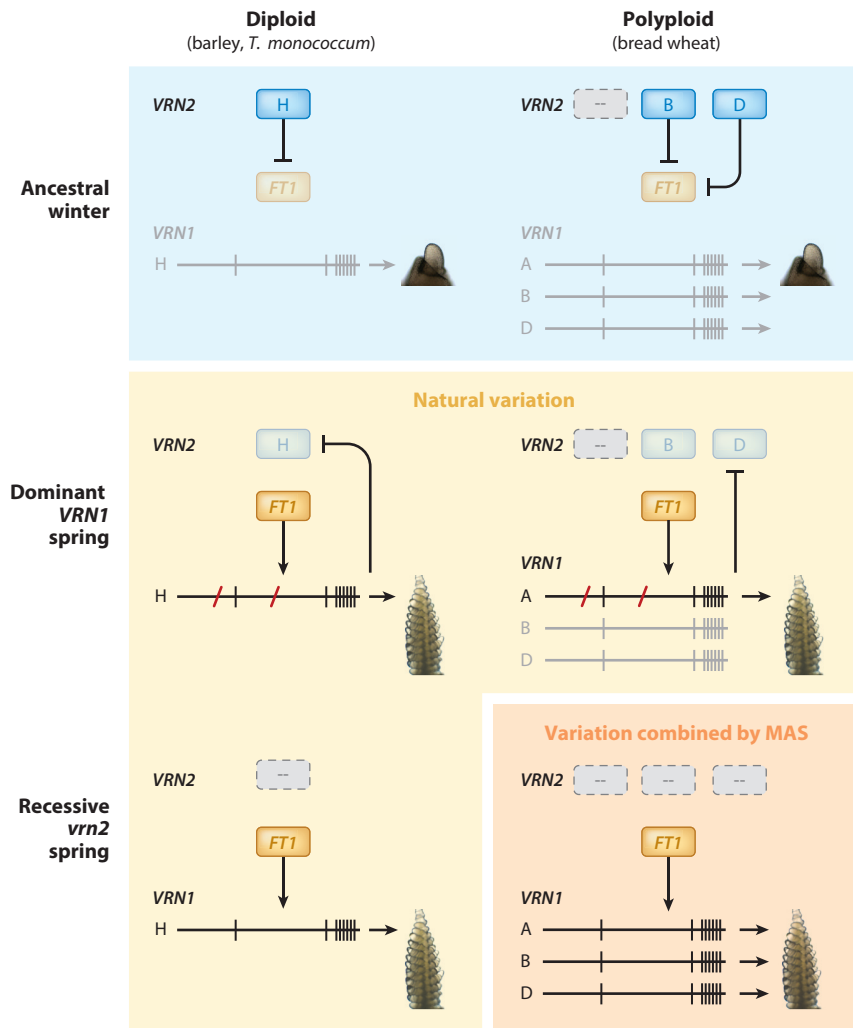
mutants in hexaploid wheat are indistinguishable from wild-type plants, making selection for allelic variation in any individual homoeolog largely irrelevant. This includes the *MLO* gene mentioned previously and variation in the *Starch Branching Enzyme II* (*SBEII*) gene. Here, single or double *SBEII* mutants have similar starch composition compared to wild-type plants, whereas triple mutants for all three homoeologs show a significant increase in amylose (75, 84). In other genes, dosage effects are sufficient to result in visible phenotypes that can be selected by breeders (81). This is the case for the *Grain Protein Content* (*GPC-1*) gene where combining an increasing number of mutant homoeologs leads to a progressive delay in leaf senescence, which is maximized in the triple mutant (4, 98).

In summary, recessive mutations are difficult to select in young polyploid crops given the subtle phenotypic differences that they confer compared to wild-type plants. However, the extensive catalog of sequenced mutations now available in wheat (47) provides the opportunity to combine loss-of-function mutations in multiple homoeologs to reveal this hidden variation. The tools are now available to transfer beneficial recessive mutations identified in diploid species to the commercial polyploid wheats.

GENERATING INDUCED MUTATIONS IN POLYPLOID WHEAT

Radiation has been used to generate variability in plant research since the late 1920s. Work by Stadler (85–87) between 1928 and 1930 showed the use of X-ray treatments to induce visible mutants in diploid and polyploid grasses such as barley, maize, oat, and wheat. Stadler's seminal work established that the rate of visible mutants recovered after mutagenesis in diploids was proportional to the radiation intensity. However, polyploid oat and wheat species yielded few

←——

Figure 1

Natural and induced variation for flowering in diploid barley and polyploid wheat. (*Top*) Winter cereals require long periods of cold temperatures (a process called vernalization) to become competent to flower. In these species, winter growth habit is the ancestral state and is determined by the presence of both functional *VERNALIZATION 2* (*VRN2*) alleles that repress flowering and *VERNALIZATION 1* (*VRN1*) alleles that promote flowering after vernalization. In the absence of vernalization, the shoot apical meristems remain in a vegetative stage (see pictures). *VRN2* acts as a repressor of *FLOWERING LOCUS 1* (*FT1*) (111), which encodes a mobile protein that carries the photoperiod signal from leaves to the shoot apical meristem (96). FT1 is required to form a floral activation complex that induces the meristem identity gene *VRN1* (49, 112), which is critical for the transition of the apical meristem to the reproductive stage. The natural upregulation of *VRN1* during the winter prevents the induction of *VRN2* in the spring, which results in the upregulation of *FT1* and the initiation of flowering. (*Middle*) In spring cereals, naturally occurring mutations or deletions in the *VRN1* promoter or first intron (*red diagonal lines*) result in the premature activation of *VRN1*, the repression of *VRN2*, and the upregulation of *FT1* in the absence of vernalization. This is sufficient to initiate the transition of the apical meristem to the reproductive stage (see pictures). Because these *VRN1* natural mutations are dominant, it is sufficient to mutate one of the three *VRN1* homoeologs to repress all the *VRN2* homoeologs and induce flowering. Dominant mutations are frequent in both diploid and polyploid species. (*Bottom*) A spring growth habit can also be determined by the elimination of all functional copies of *VRN2*. In the absence of *VRN2*, *FT1* is activated, inducing *VRN1* and the transition to the reproductive stage without vernalization. In barley, mutations or deletions in the single *VRN2* locus are sufficient to determine a spring growth habit, and this is frequently observed in this diploid species. By contrast, mutations or deletions in all three *VRN2* homoeologs need to be present to determine a spring growth habit in polyploid wheat. This is an unlikely event, and this combination has not been observed in natural polyploid wheats. However, a spring polyploid wheat was recently obtained using molecular marker assisted selection (MAS) to combine a nonfunctional *vrn-A2* allele from common wheat, a *vrn-B2* deletion from *Triticum turgidum* subsp. *dicoccon*, and a *vrn-D2* loss-of-function mutation from *Aegilops tauschii* (42). In general, recessive mutations are easier to detect in diploid species than in young polyploid species.

or no visible mutant phenotypes when subjected to otherwise lethal doses in diploid species. An increase in the ploidy level in wheat, from diploid to hexaploid, led to a progressive decrease in the number of visible mutants (86). Based on these observations and inspired by Nilsson-Ehle's (61) earlier work on seed color in wheat, Stadler assumed that the 21 chromosomes of hexaploid wheat represented three groups of 7 pairs each and that these groups shared a proportion of identical genes. Already at this early stage, Stadler (86, pp. 877–78) concluded that the presence of gene "reduplication" (i.e., homoeologs) would not allow recessive mutations to "appear" as visible effects in polyploids. This buffering effect of polyploidy was supported by later studies in diploid and polyploid *Avena* (oat) and additional wheat species (21, 26).

Chemical mutagenesis agents such as ethyl methanesulfonate (EMS) became commonplace in wheat research in the 1960s (78, 91). EMS was shown to produce point mutations in wheat (G to A and C to T) contrary to radiation, which produced large deletions and chromosome breaks (54), and was adopted both for developing new alleles for breeding (27, 45) and for functional studies (20, 23). Its use, however, was limited by the reduced number of visible mutations, similar to the challenge that Stadler confronted in the 1920s.

An important breakthrough came with the development of the TILLING (targeting induced local lesions in genome) approach (56), which provides a relatively simple strategy to identify mutants (or lesions) in a target sequence independently of its phenotypic effect. TILLING approaches require a population of, typically, EMS-induced mutants and a screening method to identify individuals with mutations in the target gene (reviewed in 101). Mutations are then prioritized based on their potential to disrupt protein function. Truncated proteins resulting from a premature termination codon or a change to a canonical splice site provide a high probability of functional disruption. However, nonsynonymous mutations resulting in amino acids with different chemical properties (31) and located in evolutionary conserved positions of the protein also have a good probability of altering protein function (48). The higher the number of mutant alleles recovered, the higher is the probability of identifying an allele leading to a nonfunctional protein.

Young polyploids are especially well suited for TILLING due to their tolerance of high densities of induced mutations compared to diploid species (101). We examined the mutation rates of 54 published studies that, in most cases, used the highest possible dose of EMS to develop TILLING populations (**Figure 2**) (see also **Supplemental Table 2**). We found a significant difference (Kruskal–Wallis $P < 0.001$) between the mutation rates in diploid species (4.3 ± 0.8 mutations per megabase) and those in young polyploid species (26 ± 3.8 mutations per megabase). Within polyploid species, tetraploids had significantly ($P < 0.02$) lower mutation rates (22.4 ± 6.1 mutations per megabase) than hexaploids (31.1 ± 2.8 mutations per megabase). Among diploids, two populations stand out for their high mutation rates: the flax *Linum usitatissimum* (24.4 mutations per megabase) (11) and the yellow sarson/mustard *Brassica rapa* (16.7 mutations per megabase) (88). However, these two species went through whole genome duplication (flax) (104) or triplication (yellow sarson) (102) events that occurred within the last nine million years. As in maize, these species should be better classified as polyploids in an intermediate stage of diploidization.

The comparison of mutation densities between diploid and polyploid species supports the hypothesis that loss-of-function mutations in polyploids are masked by genetic redundancy among homoeologs. The direct link between ploidy level and tolerance to increasing mutation densities was established in an elegant study by Tsai et al. (94). Here, mutation densities were compared between a diploid *Arabidopsis* and a newly synthetized autopolyploid from the same accession. The autopolyploid lines tolerated fivefold higher dosage of EMS treatments than the original diploid lines and had higher mutation densities.

The high density of mutations tolerated by polyploid species increases the number of mutant alleles recovered per plant and reduces the cost of implementing high-throughput sequencing
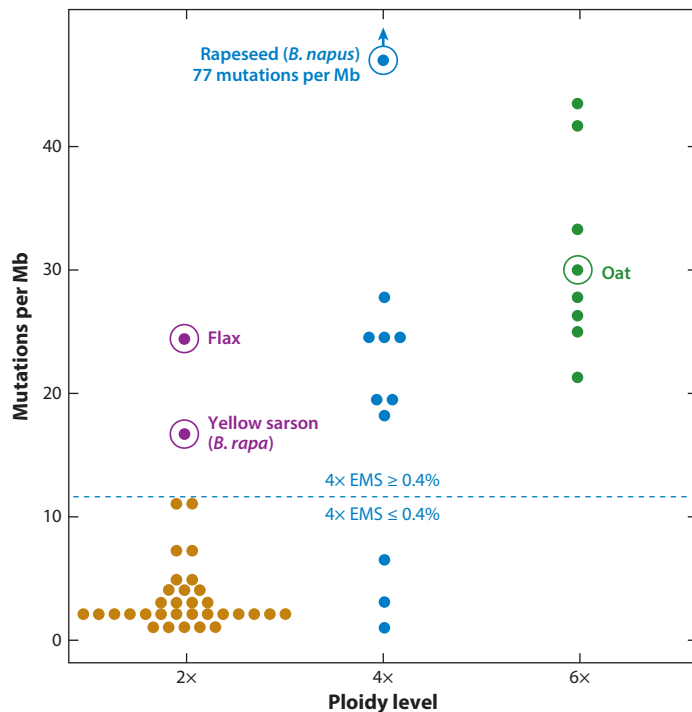
Supplemental Material

Figure 2

Mutation rates in mutant populations according to their ploidy level. Mutant populations from diploid species (2×, *gold*) have an average density of 4.3 mutations per megabase and differ from those of tetraploid species (4×, *blue*), which have an average density of 22.4 mutations per megabase. Flax and yellow sarson (*purple*), two diploid species that have undergone whole genome duplications within the last nine million years, have mutation rates comparable with those of tetraploid species and should be considered polyploids in an intermediate state of diploidization. The high spread of tetraploid values seems to originate in part by different dosages of the mutagen used in these species [see EMS threshold (*blue dashed line*)]. Hexaploid species (6×, *green*) have average densities of 31.1 mutations per megabase and include one oat and seven bread wheat mutant populations. References are provided in **Supplemental Table 2**. Abbreviations: EMS, ethyl methanesulfonate; Mb, megabase.

▶ Supplemental Material

approaches to identify them. However, the utilization of these strategies in wheat requires an intermediate step to reduce the complexity of the large wheat genomes.

COMPLEXITY REDUCTION TO IDENTIFY INDUCED, RANDOM MUTATIONS IN TARGETED GENOMIC REGIONS

Direct whole genome shotgun (WGS) sequencing to identify and catalog induced mutations is a feasible approach for species with small genomes (39, 50, 70, 95). However, accessing induced variation for a large number of individuals in wheat by WGS sequencing is not economically feasible because of the large size of the wheat genome (12 Gb in tetraploid and 16 Gb in hexaploid wheat) (6). Genes represent less than 2% of the wheat genome (38) and the rest mainly comprises highly repetitive sequences. Mutations in the repetitive sequences cannot be easily identified, uniquely mapped to a physical position, or assigned a biological function. The sheer size of the hexaploid wheat genome makes WGS sequencing of mutant lines a costly and computationally

intensive proposition. Even using a HiSeq X platform, today's cost of sequencing the genome of a single mutant hexaploid line is approximately $8,000.

With the exome capture technique, a smaller, specific portion of a plant genome can be captured for resequencing if the target sequences are known. This is achieved by the use of small biotinylated nucleotide probes, typically RNA, with homology to the target sequence. Hybridization of the probes to the WGS library followed by probe capture enriches for the target sequence prior to sequencing (36). The technology is well established, with several commercial service providers offering probe synthesis. Multiple laboratories around the world have designed and captured the protein coding regions of the wheat genome using exome capture to discover SNPs (3, 108) and catalog induced mutations (32, 41). Most recently, Krasileva and colleagues (47) described a wheat TILLING resource for pasta and bread wheat containing 2,735 lines with more than 10,000,000 EMS-induced mutations. In this study, DNA from four to eight mutant individuals was pooled, captured, and sequenced to obtain at least 20 million 100-base pair (bp) paired-end reads per sample. Although this would be equivalent to 0.4–0.6 times coverage across the whole genome, the complexity reduction using exome capture resulted in a median of 21 times coverage at mutation sites, a 40- to 60-fold enrichment.

Exome captures are biased by the gene space annotated in a reference sequence. Automated de novo gene-prediction programs can often miss real genes and include artificial ones. Prediction programs that incorporate expression data (e.g., RNA-Seq) perform better at predicting genes, but can still fail to detect those with low expression or with a very restricted expression profile. In addition to the mutations not detected in genes excluded from the capture assay, further mutations will be missed if a gene sequence present in the capture assay is absent from the reference genome used to map the reads. A single reference genome only represents a certain proportion of the pangenome space of a plant species (35, 51, 59, 90). Recent resequencing and annotation of 19 wheat genomes suggest that the reference genome of the Chinese Spring landrace is missing approximately 12,000 genes present among the pangenome of elite, commercial cultivars (59). Krasileva et al. (47) used RNA-Seq data from multiple varieties to enrich the exome capture assay, and further expanded the reference sequence by reassembling the reads that did not map to the original reference in each of the TILLING populations. The quality of the assembly of unmapped reads was improved by using unmapped reads from multiple mutant plants. This increases the proportion of targeted captured sequences relative to the nontargeted DNA sequences, which are variable across individual mutant lines. Resequencing of multiple wheat varieties (18) will ultimately improve future exome capture designs and generate a more complete reference with which to map the captured reads.

STRATEGIES TO LINK INDUCED VARIATION WITH PHENOTYPES

Sequenced Mutant Populations

Reverse genetic approaches, such as TILLING, are well suited to polyploid species because mutations in individual homoeologs can be identified and combined before analyzing the phenotype. However, the screening of TILLING polyploid populations by PCR amplification and sequencing requires the development of efficient genome-specific primers, a task that requires sequences from the different homoeologs and practical experience (93, 99). The advent of next-generation sequencing provides a new platform for detection of mutations in the different homoeologs and eliminates the need to generate genome-specific primers. The simultaneous sequencing of most of the mutations present in the different homoeologs also eliminates the need to access the DNAs of the TILLING populations to search for a mutation.
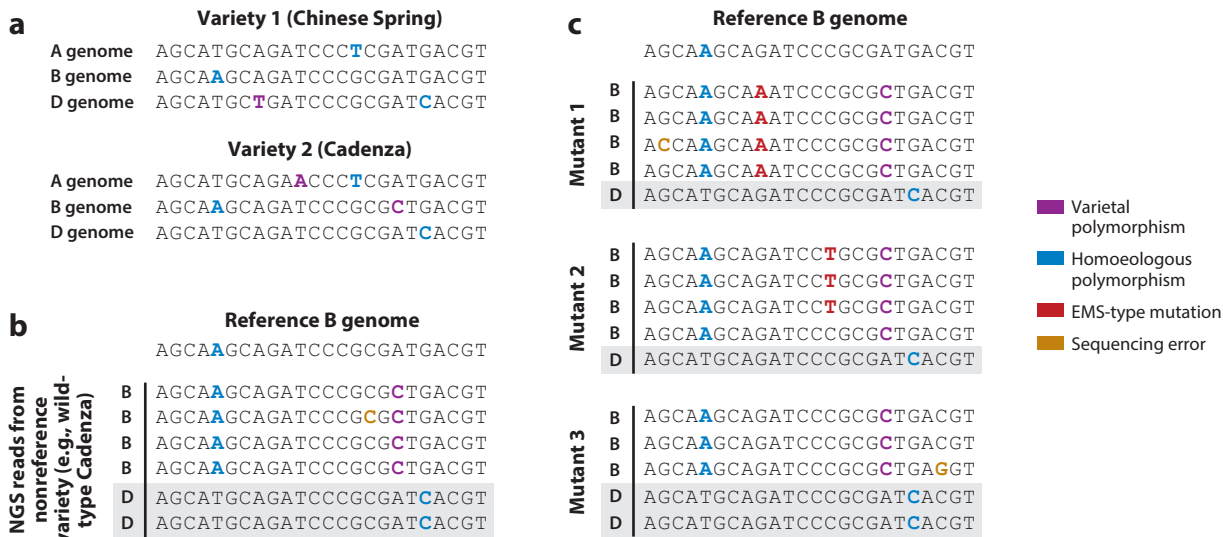
a

Variety 1 (Chinese Spring)

A genome AGCATGCAGATCCC**T**CGATGACGT
B genome AGCA**A**GCAGATCCCGCGATGACGT
D genome AGCATGC**T**GATCCCGCGAT**C**ACGT

Variety 2 (Cadenza)

A genome AGCATGCAGA**A**CCC**T**CGATGACGT
B genome AGCA**A**GCAGATCCCGCG**C**TGACGT
D genome AGCATGCAGATCCCGCGAT**C**ACGT

c

Reference B genome

AGCA**A**GCAGATCCCGCGATGACGT

Mutant 1

B AGCA**A**GCA**A**ATCCCGCG**C**TGACGT
B AGCA**A**GCA**A**ATCCCGCG**C**TGACGT
B A**C**CA**A**GCA**A**ATCCCGCG**C**TGACGT
B AGCA**A**GCA**A**ATCCCGCG**C**TGACGT
D AGCATGCAGATCCCGCGAT**C**ACGT

Mutant 2

B AGCA**A**GCAGATCC**T**GCG**C**TGACGT
B AGCA**A**GCAGATCC**T**GCG**C**TGACGT
B AGCA**A**GCAGATCC**T**GCG**C**TGACGT
B AGCA**A**GCAGATCCCGCG**C**TGACGT
D AGCATGCAGATCCCGCGAT**C**ACGT

Mutant 3

B AGCA**A**GCAGATCCCGCG**C**TGACGT
B AGCA**A**GCAGATCCCGCG**C**TGACGT
B AGCA**A**GCAGATCCCGCG**C**TGA**G**GT
D AGCATGCAGATCCCGCGAT**C**ACGT
D AGCATGCAGATCCCGCGAT**C**ACGT

Legend:
- Varietal polymorphism (purple)
- Homoeologous polymorphism (blue)
- EMS-type mutation (red)
- Sequencing error (gold)

b

Reference B genome

AGCA**A**GCAGATCCCGCGATGACGT

NGS reads from nonreference variety (e.g., wild-type Cadenza)

B AGCA**A**GCAGATCCCGCG**C**TGACGT
B AGCA**A**GCAGATCCCG**C**G**C**TGACGT
B AGCA**A**GCAGATCCCGCG**C**TGACGT
B AGCA**A**GCAGATCCCGCG**C**TGACGT
D AGCATGCAGATCCCGCGAT**C**ACGT
D AGCATGCAGATCCCGCGAT**C**ACGT

Figure 3

Types of mutations found in a mutagenized polyploid plant. (*a*) The presence of homoeologous single-nucleotide polymorphisms (SNPs, *blue*) complicates the identification of ethyl methanesulfonate (EMS) mutations in polyploid wheat. Different wheat varieties (e.g., Chinese Spring and Cadenza) share the majority of these homoeologous SNPs but also have varietal SNPs (*purple*). (*b*) When next-generation sequencing (NGS) reads from one variety are mapped to the reference genome sequence, reads from a different homoeolog (e.g., D genome, *highlighted in gray*) may map to the same reference sequence in conserved regions or in regions with a missing homoeolog in the reference. This generates homoeologous SNPs that can confound the analysis (41). (*c*) The Mutation and Polymorphism Survey (MAPS) pipeline (32) solves this problem by eliminating variants that appear in more than one mutant included in each analyzed batch (a batch normally contains 24–30 mutants; 3 are shown for the sake of this example). This approach removes both varietal SNPs (*purple*) and homoeologous SNPs (*blue*). The residual variants present in a single mutant can be attributed to sequencing errors (*gold*) and real EMS-type mutations (*red*). The two can be distinguished by their frequency: Homozygous and heterozygous EMS mutations are expected to be present in all or half of the reads, whereas errors are randomly distributed in one read or exceptionally more. In addition, EMS generates only G to A and C to T mutations in wheat. The MAPS pipeline allows users to define the minimum coverage required to call a mutation homozygous (HomMC) or heterozygous (HetMC). Using HetMC = 5 and HomMC = 3 values resulted in >99.8% accuracy in the detection of wheat mutations (47).

In wheat, the coding regions of homoeologous genes are on average 97.2% identical (with a standard deviation of 1.8%) (76). Therefore, 100-bp paired-end reads are expected to include on average 5.6 SNPs, which is sufficient to map most of the reads from the mutant lines to unique places in the reference (**Figure 3a,b**). As the size of the paired-end reads continues to increase, the proportion of reads mapped to unique sequences will also increase, and it will be possible to expand this approach to polyploid species with even more similar genomes. Even for those individual reads that are mapped to more than a single genome location, it is still possible to identify mutations. A dedicated bioinformatics pipeline was developed by Krasileva et al. (47) to keep track of the multiple locations where those reads map (designated as multimap reads) and to provide users with those alternative locations.

Krasileva et al. (47) used an 84-Mb exome capture assay to sequence the coding regions of 1,535 tetraploid (cultivar Kronos) (99) and 1,200 hexaploid (cultivar Cadenza) (67) mutagenized lines. More than 10 million uniquely mapped high-confidence ($P > 0.998$) EMS-type mutations were identified using the MAPS bioinformatics pipeline (32) (**Figure 3c**). Across the complete populations, this represents a density of 35 mutations per kilobase in tetraploid wheat and

IN SILICO FUNCTIONAL GENOMIC RESOURCES

It seems likely that in the coming years multiple mutant populations will be sequenced to establish in silico functional genomic resources in crops. This approach requires an up-front investment but generates a lasting and easy-to-use community resource. The specific sequencing strategy will depend on the genome size, the cost-effectiveness of sequencing platforms and complexity reduction methods, the similarity of the genomes, and the objective of the resource (e.g., to generate a complete genome sequence versus the protein-coding exome). Recent examples include the complete genome sequence for 1,504 mutant lines of rice (50) and 256 mutant lines of sorghum (39), which have a relatively small diploid genome (389 Mb and 750 Mb, respectively) compared to the genome of polyploid wheat (16,000 Mb). Despite the small population size in the sorghum study, 57% of the genes harbored at least one disruptive mutation (truncation or missense SIFT < 0.05). In soybean (975 Mb), 12 mutant lines were sequenced to verify mutation density in a new mutant population that was then screened with amplicon resequencing (95). More recently, a population of over 1,000 *Brassica rapa* (a diploid genome of 485 Mb) mutant lines was exome captured and sequenced and made available online (70, 88). Together with the sequenced mutant resources in tetraploid and hexaploid wheat (47), these examples document a shift towards whole genome/exome-sequenced mutant populations in the major crop species.

40 mutations per kilobase in hexaploid wheat. Effects were predicted for 7.1 million mutations (67%) using the International Wheat Genome Sequencing Consortium gene models (38). Of these, 3% corresponded to truncations (premature termination codons or splice variants) and 38% were missense. The mutation density was sufficient to find an average of 23–24 predicted missense and truncation alleles per population per gene. Using the sorting intolerant from tolerant (SIFT) algorithm (48), we found that over 85% of genes captured had at least one putative reduced- or loss-of-function mutation (SIFT < 0.05). When truncation mutations were added, more than 90% of the captured wheat genes were expected to have mutations severely affecting their function (47). To facilitate the utilization of these mutations, we designed homoeolog-specific SNP assays (69, 77) for the majority of the mutations and seeds from all sequenced mutant lines were deposited in public repositories. These sequenced mutant populations are a valuable resource for functional studies, and similar strategies can be applied to other crop species (see the sidebar titled In Silico Functional Genomic Resources).

Although this review focuses on reverse genetic approaches, natural and induced variation at a single locus can lead to sufficient phenotypic differences between wild-type and mutant individuals in a polyploid species to allow traditional (forward) genetic mapping. The causal mutations of the phenotypes identified in forward screens can be identified by map-based cloning, which has been used successfully in wheat (98, 111, 114). However, these methods are laborious and time-consuming, limiting the rate at which genetic loci can be identified. Combining high-throughput sequencing and genomic technologies provides an opportunity to greatly accelerate the identification of causal mutations (68, 72, 74, 89). This process is also greatly accelerated when the mutant phenotypes are identified in the sequenced mapping populations. However, thus far these methods have been restricted to single locus loss-of-function mutations and have not been implemented for more complex quantitative traits in polyploid wheat.

Targeted Mutagenesis via Genome Editing

Genome editing technologies provide new routes to induce variation in a more targeted manner than random chemical mutagenesis. The first examples of gene editing in wheat aimed to generate

double-strand breaks in precise genomic positions to induce small deletions. Early work in wheat used the transcription activator-like effector nucleases to simultaneously target the three homoeol-ogous copies of *MLO* (103). This method was rapidly superseded by Cas9-mediated gene editing, which was shown to be effective for inducing small deletions in wheat protoplast (79) and in stably transformed plants when a single homoeolog was targeted (103). Since then, several studies have shown the feasibility of generating transgene-free edited wheat using the Cas9 systems. These include transiently expressed plasmid DNA, in vitro synthesized mRNA of the Cas9 endonuclease and guide RNAs, and preassembled ribonucleoprotein complexes of purified Cas9 and in vitro transcribed guide RNAs (52, 114). The latter two methods avoid transgene integration altogether because no DNA is introduced into the plant. Between 1% and 3% of bombarded embryos show an on-target deletion event using these methods, resulting almost exclusively in frameshift alleles and truncated proteins. Across these studies (52, 103, 114), seven genes have been edited in both tetraploid and hexaploid wheat.

Most recently, base editing was documented in wheat using a mutant nickase version of Cas9 fused to a cytidine deaminase (116). In this system, cytosine residues in a fixed position within the protospacer sequence [13–18 bp from the protospacer adjacent motif (PAM)] were substituted to thymine without double-strand breaks. Although only a single homoeolog was targeted in this study, it expands the applications of genome editing in wheat by shifting from targeted deletions into targeted base editing. Imminently, allele replacement (115), targeted gene insertion (25), and additional applications (66) are likely to be implemented in wheat. This raises the possibility of developing dominant alleles in single homoeologs in a targeted manner.

The sequenced mutant populations and the Cas9-mediated gene editing technologies provide complementary approaches to generate and utilize induced variation in polyploid wheat. Each approach has its own merits and demerits relating to the initial investment by researchers, access to the technology, the range of mutations that are identified, and the use of these mutations in breeding (47). These are summarized in **Table 1**.

STRATEGIES FOR THE USE OF MUTANTS IN GENE ANALYSIS AND BREEDING

To overcome the functional overlap among homoeologs, researchers and breeders need to identify loss-of-function mutations in the A and B genome copies in tetraploid wheat and in the A, B, and D genome copies in hexaploid wheat, and combine them through crossing to select homozygous double or triple mutants in the F_2 generation. Because these crosses take time, it is critical to select mutations with a strong effect on gene function. As discussed previously, use of truncation muta-tions gives the best chance of achieving this objective, followed by use of mutations in conserved regions of the protein. However, using missense mutations is riskier as it is difficult to predict a priori if the change in amino acid will completely eliminate protein function (see the sidebar titled Missense Mutations, Deletions, and Natural Mutants). If no phenotype is observed after multiple crosses, it is difficult to determine if this is due to the ineffectiveness of the missense allele, a different biological function of the gene in question, or redundant function between paralogous genes.

The high frequency of truncations and deleterious mutations detected in sequenced poly-ploid wheat mutant populations is advantageous for selecting mutations with a high likelihood of disrupting protein function. It also provides an extended allelic series for detailed functional char-acterization (e.g., *WKS1*) (23) or for more subtle modifications of the phenotype. However, once the desired mutant is identified, this advantage becomes a drawback because the high number of background mutations complicates the phenotypic analysis. For example, in the sequenced wheat

Table 1 Comparison between sequenced mutant populations and gene editing

Features	Sequenced mutant population	Gene editing
Ease of getting started	Mutations are searchable online (16, 106); immediate access to mutant seed	Requires additional time for construct design and optimization; delivery into wheat is dependent on access to technology; relatively high cost
Achieving specificity	Specific for C to T and G to A transitions; local sequence-dependent bias affects the probability that C/G positions will be mutated (32, 47)	Dependent on presence of PAM (5′-NGG-3′); new Cas9 specificities have been published (44); new nucleases [Cpf1 (113)] have a different range of PAMs (24, 109)
Off-target effects	Thousands of mutations outside gene of interest with many potential deleterious mutations	Very specific with more limited off-target effects
Developing triple mutants	Mutants in individual homoeologs can be combined through traditional crossing and marker-assisted selection	Triple mutants in first generation not likely (approximately 0.5%) (114); requires crossing of single homoeologs
Range of varieties	Original mutants restricted to sequenced populations; can be transferred to locally relevant germplasm by crossing	Dependent on transformation efficiency of variety; requires crossing to locally relevant germplasm to deploy in agriculture
Use in breeding	Currently deployed and not subject to regulation	Nontransgenic classification is still uncertain in many countries; may be problematic for globally traded crops

Abbreviation: PAM, protospacer adjacent motif.

TILLING populations an average tetraploid mutant line has 2,705 high-confidence EMS-type mutations that, on average, result in approximately 50 truncation alleles and over 670 missense mutations per individual mutant line in currently annotated genes. The expectation for Cadenza, which has a higher mutation load, is for there to be approximately 110 truncation alleles and close to 1,400 missense mutations per individual in annotated genes. Put another way, between 1.5% and 2% of the genes in any given wheat mutant line will have a truncation or missense allele (47).

MISSENSE MUTATIONS, DELETIONS, AND NATURAL MUTANTS

Given the time required in wheat to combine mutants in all homoeologs, initial efforts to determine the effect of individual missense mutations on the activity of the protein are justified. For well-characterized protein domains, one can select mutations that affect known active sites (12, 23). However, for most proteins, computational methods are used to predict the effect of mutations on protein function (e.g., 48). The majority of these methods rely on the principle that proteins of related species performing common functions have evolutionary conservation in key amino acid residues. Mutations affecting these positions are predicted to be deleterious. However, recent methods, which account for residue dependencies between all positions in a protein, have been shown to outperform traditional single-position models (37). In some cases, natural variation may be available to complement induced mutants (4, 23, 83, 98), whereas, in other cases, one of the homoeologs may not be expressed, simplifying the development of null mutants (14, 65). Gamma-irradiated deletion populations can also be used to complement natural and induced variation. These large deletions are particularly useful when closely related genes are duplicated in tandem. Ultimately, the best strategy will depend on the availability of induced mutants and deletions, natural variation, and the specific gene being targeted.
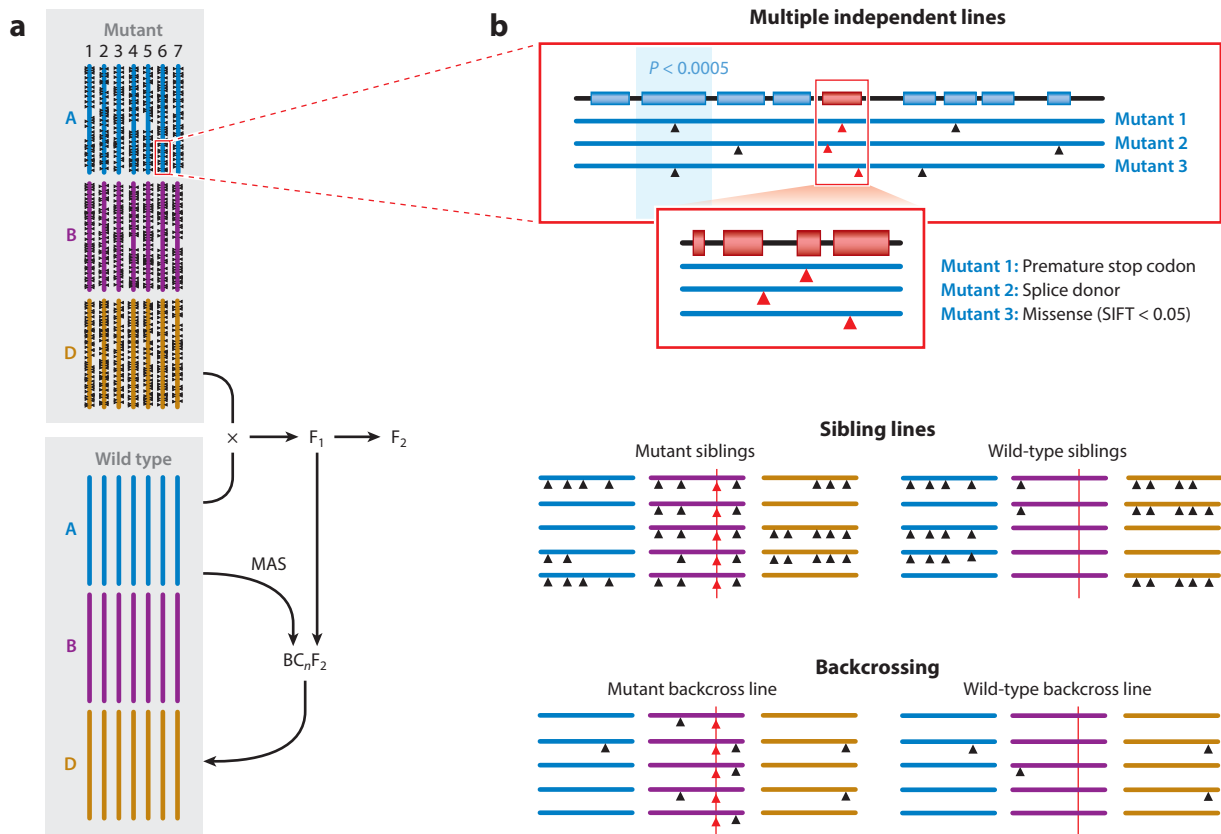
Figure 4

Strategies for the use of mutants in gene analysis and breeding. (*a*) Mutant lines carrying large numbers of mutations can confound phenotypic analyses. Depicted are chromosome groups 1–7 for genomes *A* (*blue*), *B* (*purple*), and *D* (*gold*) with multiple mutations (*top*) versus the wild type (*bottom*). (*b*) Three strategies can be used to minimize the risk of confounding effects: the use of multiple independent lines, the generation of sibling lines, and backcrossing. (*Top*) Multiple independent lines that carry mutations in a target gene (*filled red rectangle*) are unlikely to carry independent mutations affecting a second gene (*light blue background*, $P < 0.0005$ for two independent lines). (*Middle*) Mutant lines can be crossed to the wild-type parent to generate a segregating F_2 population. Using markers, F_2 siblings that differ in the target gene (*red triangle*) but share background mutations (*black triangles*) can be compared. (*Bottom*) Mutant lines also can be backcrossed (BC) for *n* generations to the wild-type parent, and progeny carrying the mutation can be selected using marker assisted selection (MAS). On average, each BC generation reduces unlinked background mutations by half. When the desired reduction of background mutations is achieved, BC_nF_2 siblings differing for the target mutation (*red triangle*) are selected. Abbreviation: SIFT, sorting intolerant from tolerant algorithm.

Several strategies can be used to minimize the confounding effects of high numbers of background mutations in functional genetic studies. These strategies include the analysis of independent mutations, the use of sibling mutant lines, and backcrossing to the nonmutagenized line (**Figure 4**).

Multiple Independent Lines

The use of multiple independent mutants is a good practice, especially in basic research projects, to avoid the potentially confounding effect of other mutations present in the same plant. In polyploid

wheat, the mutants in different homoeologs represent a first level of replication. If the phenotype is observed only in plants combining mutation in all homoeologs, it is very unlikely that the effect will be caused by independent mutations in the same separate gene. For example, the probability of two mutant lines sharing by chance a truncation or missense mutation in the same second gene is $P < 0.0003$ in Kronos and $P < 0.0005$ in Cadenza. This drops dramatically to a probability of less than 1 in 100,000 or 1 in 5 million when three or four independent lines are examined (**Supplemental File 3**).

Sibling Lines

In tetraploid wheat, homozygous null mutants and wild-type sibling plants can be selected using molecular markers from segregating F_2 individuals from the cross between A and B genome mutants. These F_2 sibling plants differ at the gene of interest but share many of the same background mutations, resulting in a more valid comparison. The analysis of the segregating populations can be also used to rule out effects caused by unlinked mutations. In hexaploid wheat, the same strategy is effective although an additional cross is required to generate the F_2 triple mutants. This strategy, alongside the use of multiple independent lines, is best suited for basic research projects.

Backcrossing

Background mutations can not only confound gene–trait associations, but can also have detrimental effects on overall plant performance. Therefore, for quantitative traits, breeding applications, or when traits need to be assessed in plants grown under field conditions, it is advisable to backcross the mutants to the original parent one or two times to reduce the mutation load. The trade-off between reducing mutation load and time required for the backcrossing will ultimately depend on the objective of the experiment and the phenotype being studied.

The time required to reduce the mutation load to acceptable levels can be shortened by selecting plants with the minimum number of background mutations in each generation. Because the positions of the mutations in the genome are known (16, 106), background selection can be used to minimize the length of the chromosome segments from the mutant line. Additionally, the use of accelerated growth conditions (33, 34) can further reduce generation time to approximately 10 weeks in Kronos and Cadenza (105). This accelerated growth cycle is particularly useful for breeding applications, where more extensive backcrossing is justified to further reduce the load of background mutations. When mutations across multiple homoeologs need to be combined, it is prudent to backcross the single mutants separately and then intercross them to generate the desired double or triple mutants (29, 81).

SUMMARY POINTS

1. Functional redundancy among homoeologs is more extensive in recently formed (young) polyploid species compared to older polyploids with more extensive diploidization.

2. A large number of recessive mutations are not detected by phenotypic selection in young polyploid species due to functional redundancy among homoeologs.

3. Domestication and previous breeding efforts have favored selection of dominant or semidominant mutations in polyploid wheat because these are usually epistatic over the other homoeologs, allowing the rapid detection of favorable phenotypes.

4. Young polyploid species tolerate higher mutation densities than diploids, making them especially well suited to the development of sequenced mutant populations. This tolerance increases the number of mutant alleles recovered per plant, reducing the cost to identify them by high-throughput sequencing.

5. Useful recessive mutations identified in diploid relatives are valuable targets to combine mutant alleles in all homoeologs in the polyploid crop. Sequenced mutant populations and genome editing facilitate the identification of these loss-of-function mutant alleles in the polyploid crop species.

6. The combination of traditional mutagenesis with new high-throughput sequencing and genomics approaches has resulted in new powerful tools to analyze gene function in wheat. Alongside genome editing, these tools are changing the paradigm of what is possible in functional genetic analysis in wheat.

7. The strategies developed for the generation of sequenced mutant populations in wheat can be transferred to other polyploid crops.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

LITERATURE CITED

1. Acevedo-Garcia J, Spencer D, Thieron H, Reinstädler A, Hammond-Kosack K, et al. 2017. *mlo*-based powdery mildew resistance in hexaploid bread wheat generated by a non-transgenic TILLING approach. *Plant Biotechnol. J.* 15:367–78

2. Adams KL, Wendel JF. 2005. Polyploidy and genome evolution in plants. *Curr. Opin. Plant Biol.* 8:135–41

3. Allen AM, Barker GLA, Wilkinson P, Burridge A, Winfield M, et al. 2013. Discovery and development of exome-based, co-dominant single nucleotide polymorphism markers in hexaploid wheat (*Triticum aestivum* L.). *Plant Biotechnol. J.* 11:279–95

4. Avni R, Zhao R, Pearce S, Jun Y, Uauy C, et al. 2014. Functional characterization of *GPC-1* genes in hexaploid wheat. *Planta* 239:313–24

5. Beales J, Turner A, Griffiths S, Snape JW, Laurie DA. 2007. A *Pseudo-Response Regulator* is misexpressed in the photoperiod insensitive *Ppd-D1a* mutant of wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 115:721–33

6. Bennett MD, Smith JB. 1976. Nuclear DNA amounts in angiosperms. *Philos. Trans. R. Soc. B* 274:227–74

7. Blanc G, Wolfe KH. 2004. Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *Plant Cell* 16:1667–78

8. Borrill P, Adamski N, Uauy C. 2015. Genomics as the key to unlocking the polyploid potential of wheat. *New Phytol.* 208:1008–22

9. Briggs J, Chen S, Zhang W, Nelson S, Dubcovsky J, Rouse MN. 2015. Mapping of *SrTm4*, a recessive stem rust resistance gene from diploid wheat effective to Ug99. *Phytopathology* 105:1347–54

10. Büschges R, Hollricher K, Panstruga R, Simons G, Wolter M, et al. 1997. The barley *Mlo* gene: a novel control element of plant pathogen resistance. *Cell* 88:695–705

11. Chantreau M, Grec S, Gutierrez L, Dalmais M, Pineau C, et al. 2013. PT-Flax (phenotyping and TILLinG of flax): development of a flax (*Linum usitatissimum* L.) mutant population and TILLinG platform for forward and reverse genetics. *BMC Plant Biol.* 13:159

12. Chen A, Li C, Hu W, Lau MY, Lin H, et al. 2014. *PHYTOCHROME C* plays a major role in the acceleration of wheat flowering under long-day photoperiod. *PNAS* 111:10037–44

13. Dibernardi JM, Lin H, Chuck G, Faris JD, Dubcovsky J. 2017. microRNA172 plays a crucial role in wheat spike morphogenesis and grain threshability. *Development* 144:1966–75

14. Dobrovolskaya O, Pont C, Sibout R, Martinek P, Badaeva E, et al. 2015. *FRIZZY PANICLE* drives supernumerary spikelets in bread wheat. *Plant Physiol.* 167:189–99

15. Dubcovsky J, Dvorak J. 2007. Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* 316:1862

16. Dubcovsky Lab. 2017. *Wheat TILLING*. Database, Univ. Calif., Davis, CA, updated Jan. 17, 2017. **http://dubcovskylab.ucdavis.edu/wheat-tilling**

17. Dudnikov AJ. 2003. Allozymes and growth habit of *Aegilops tauschii*: genetic control and linkage patterns. *Euphytica* 129:89–97

18. Earlham Inst. 2017. *Grassroots Genomics*. Proj., Norwich, UK. **http://www.earlham.ac.uk/grassroots-genomics**

19. Ellis JG, Lagudah ES, Spielmeyer W, Dodds PN. 2014. The past, present and future of breeding rust resistant wheat. *Front. Plant Sci.* 5:641

20. Feuillet C, Travella S, Stein N, Albar L, Nublat A, Keller B. 2003. Map-based isolation of the leaf rust disease resistance gene *Lr10* from the hexaploid wheat (*Triticum aestivum* L.) genome. *PNAS* 100:15253–58

21. Fröier K, Gustafsson Å, Tedin O. 1942. The relation of mitotic disturbances to X-ray dosage and polyploidy. *Hereditas* 28:165–70

22. Fu D, Szűcs P, Yan L, Helguera M, Skinner JS, 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

23. Fu D, Uauy C, Distelfeld A, Blechl A, Epstein L, et al. 2009. A *Kinase-START* gene confers temperature-dependent resistance to wheat stripe rust. *Science* 323:1357–60

24. Gao L, Cox DBT, Yan WX, Manteiga J, Schneider M, et al. 2017. Engineered Cpf1 enzymes with altered PAM specificities. *Nat. Biotechnol.* 35:789–92

25. Gil-Humanes J, Wang Y, Liang Z, Shan Q, Ozuna CV, et al. 2017. High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. *Plant J.* 89:1251–62

26. Gustafsson Å. 1947. Mutations in agricultural plants. *Hereditas* 33:1–100

27. Guttieri M, Bowen D, Dorsch JA, Raboy V, Souza E. 2004. Identification and characterization of a low phytic acid wheat. *Crop Sci.* 44:418–24

28. Guyot R, Keller B. 2004. Ancestral genome duplication in rice. *Genome* 47:610–14

29. Hazard B, Zhang X, Colasuonno P, Uauy C, Beckles DM, Dubcovsky J. 2012. Induced mutations in the *Starch Branching Enzyme II* (*SBEII*) genes increase amylose and resistant starch content in durum wheat. *Crop Sci.* 52:1754–66

30. Hemming MN, Fieg S, Peacock WJ, Dennis ES, Trevaskis B. 2009. Regions associated with repression of the barley (*Hordeum vulgare*) *VERNALIZATION1* gene are not required for cold induction. *Mol. Genet. Genom.* 282:107–17

31. Henikoff S, Henikoff JG. 1992. Amino acid substitution matrices from protein blocks. *PNAS* 89:10915–19

32. Henry IM, Nagalakshmi U, Lieberman MC, Ngo KJ, Krasileva KV, et al. 2014. Efficient genome-wide detection and cataloging of EMS-induced mutations using exome capture and next-generation sequencing. *Plant Cell* 26:1382–97

33. Hickey LT, Germán SE, Pereyra SA, Diaz JE, Ziems LA, et al. 2017. Speed breeding for multiple disease resistance in barley. *Euphytica* 213:64

34. Hickey LT, Wilkinson PM, Knight CR, Godwin ID, Kravchuk OY, et al. 2012. Rapid phenotyping for adult-plant resistance to stripe rust in wheat. *Plant Breed.* 131:54–61

35. Hirsch CN, Foerster JM, Johnson JM, Sekhon RS, Muttoni G, et al. 2014. Insights into the maize pan-genome and pan-transcriptome. *Plant Cell* 26:121–35

36. Hodges E, Xuan Z, Balija V, Kramer M, Molla MN, et al. 2007. Genome-wide in situ exon capture for selective resequencing. *Nat. Genet.* 39:1522–27

37. Hopf TA, Ingraham JB, Poelwijk FJ, Scharfe CPI, Springer M, et al. 2017. Mutation effects predicted from sequence co-variation. *Nat. Biotechnol.* 35:128–35

38. IWGSC (Int. Wheat Genome Seq. Consort.). 2014. A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science* 345:1251788

39. Jiao Y, Burke JJ, Chopra R, Burow G, Chen J, et al. 2016. A sorghum mutant resource as an efficient platform for gene discovery in grasses. *Plant Cell* 28:1551–62

40. Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, et al. 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* 473:97–100

41. King R, Bird N, Ramirez-Gonzalez R, Coghill JA, Patil A, et al. 2015. Mutation scanning in wheat by exon capture and next-generation sequencing. *PLOS ONE* 10:e0137549

42. Kippes N, Chen A, Zhang X, Lukaszewski AJ, Dubcovsky J. 2016. Development and characterization of a spring hexaploid wheat line with no functional *VRN2* genes. *Theor. Appl. Genet.* 129:1417–28

43. Kippes N, Debernardi JM, Vasquez-Gross HA, Akpinar BA, Budak H, et al. 2015. Identification of the *VERNALIZATION 4* gene reveals the origin of spring growth habit in ancient wheats from South Asia. *PNAS* 112:E5401–10

44. Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, et al. 2015. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* 523:481–85

45. Knott DR. 1980. Mutation of a gene for yellow pigment linked to *Lr19* in wheat. *Can. J. Genet. Cytol.* 22:651–54

46. Konopatskaia I, Vavilova V, Kondratenko EY, Blinov A, Goncharov NP. 2016. *VRN1* genes variability in tetraploid wheat species with a spring growth habit. *BMC Plant Biol.* 16:244

47. Krasileva KV, Vasquez-Gross HA, Howell T, Bailey P, Paraiso F, et al. 2017. Uncovering hidden variation in polyploid wheat. *PNAS* 114:E913–21

48. Kumar P, Henikoff S, Ng PC. 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* 4:1073–81

49. Li C, Lin H, Dubcovsky J. 2015. Factorial combinations of protein interactions generate a multiplicity of florigen activation complexes in wheat and barley. *Plant J.* 84:70–82

50. Li G, Jain R, Chern M, Pham NT, Martin JA, et al. 2017. The sequences of 1,504 mutants in the model rice variety Kitaake facilitate rapid functional genomic studies. *Plant Cell* 29:1218–31

51. Li Y-H, Zhou G, Ma J, Jiang W, Jin L-G, et al. 2014. *De novo* assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. *Nat. Biotechnol.* 32:1045–52

52. Liang Z, Chen K, Li T, Zhang Y, Wang Y, et al. 2017. Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. *Nat. Commun.* 8:14261

53. Lynch M, Conery JS. 2000. The evolutionary fate and consequences of duplicate genes. *Science* 290:1151–55

54. MacKey J. 1968. Mutagenesis in vulgare wheat. *Hereditas* 59:505–17

55. Marcussen T, Sandve SR, Heier L, Spannagl M, Pfeifer M, et al. 2014. Ancient hybridizations among the ancestral genomes of bread wheat. *Science* 345:1250092

56. McCallum CM, Comai L, Greene EA, Henikoff S. 2000. Targeted screening for induced mutations. *Nat. Biotechnol.* 18:455–57

57. McIntosh R, Dubcovsky J, Rogers W, Morris C, Xia X. 2017. *Catalogue of gene symbols for wheat.* Komugi Wheat Genet. Resour. Database, Comm. Natl. BioResour. Proj., Kyoto, Japan, updated Jan. 11. **http://shigen.nig.ac.jp/wheat/komugi/genes/symbolClassList.jsp**

58. Molina M del C, Naranjo CA. 1987. Cytogenetic studies in the genus *Zea*. *Theor. Appl. Genet.* 73:542–50

59. Montenegro JD, Golicz AA, Bayer PE, Hurgobin B, Lee H, et al. 2017. The pangenome of hexaploid bread wheat. *Plant J.* 90:1007–13

60. Muterko A, Kalendar R, Salina E. 2016. Allelic variation at the *VERNALIZATION-A1*, *VRN-B1*, *VRN-B3*, and *PHOTOPERIOD-A1* genes in cultivars of *Triticum durum* Desf. *Planta* 244:1253–63

61. Nilsson-Ehle H. 1909. Kreuzungsuntersuchungen an Hafer und Weizen. *Lunds Univ. Årsskr. N.F.* 5:1–122

62. Page DR, Grossniklaus U. 2002. The art and design of genetic screens: *Arabidopsis thaliana*. *Nat. Rev. Genet.* 3:124–36

63. Paterson AH, Bowers JE, Chapman BA. 2004. Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *PNAS* 101:9903–8

64. Pearce S, Shaw LM, Lin H, Cotter JD, Li C, Dubcovsky J. 2017. Night-break experiments shed light on the Photoperiod1-mediated flowering. *Plant Physiol.* 174:1139–50

65. Poursarebani N, Seidensticker T, Koppolu R, Trautewig C, Gawroński P, et al. 2015. The genetic basis of composite spike form in barley and 'Miracle-Wheat.' *Genetics* 201:155–65

66. Puchta H. 2017. Applying CRISPR/Cas for genome engineering in plants: The best is yet to come. *Curr. Opin. Plant Biol.* 36:1–8

67. Rakszegi M, Kisgyörgy BN, Tearall K, Shewry PR, Láng L, et al. 2010. Diversity of agronomic and morphological traits in a mutant population of bread wheat studied in the Healthgrain program. *Euphytica* 174:409–21

68. Ramirez-Gonzalez RH, Segovia V, Bird N, Fenwick P, Holdgate S, et al. 2015. RNA-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat. *Plant Biotechnol. J.* 13:613–24

69. Ramirez-Gonzalez RH, Uauy C, Caccamo M. 2015. PolyMarker: a fast polyploid primer design pipeline. *Bioinformatics* 31:2038–39

70. RevGenUK. 2017. *RevGenUK*. TILLING Serv., John Innes Cent., Norwich, UK, updated Jan. 20, 2017. **http://revgenuk.jic.ac.uk**

71. Salamini F, Ozkan H, Brandolini A, Schafer-Pregl R, Martin W. 2002. Genetics and geography of wild cereal domestication in the near east. *Nat. Rev. Genet.* 3:429–41

72. Sánchez-Martín J, Steuernagel B, Ghosh S, Herren G, Hurni S, et al. 2016. Rapid gene isolation in barley and wheat by mutant chromosome sequencing. *Genome Biol.* 17:221

73. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, et al. 2009. The B73 maize genome: complexity, diversity, and dynamics. *Science* 326:1112–15

74. Schneeberger K. 2014. Using next-generation sequencing to isolate mutant genes from forward genetic screens. *Nat. Rev. Genet.* 15:662–76

75. Schönhofen A, Hazard B, Zhang X, Dubcovsky J. 2016. Registration of common wheat germplasm with mutations in *SBEII* genes conferring increased grain amylose and resistant starch content. *J. Plant Regist.* 10:200–5

76. Schreiber AW, Hayden MJ, Forrest KL, Kong SL, Langridge P, Baumann U. 2012. Transcriptome-scale homoeolog-specific transcript assemblies of bread wheat. *BMC Genom.* 13:492

77. Semagn K, Babu R, Hearne S, Olsen M. 2014. Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. *Mol. Breed.* 33:1–14

78. Shama Rao HK, Sears ER. 1964. Chemical mutagenesis in *Triticum aestivum*. *Mutat. Res./Fundam. Mol. Mech. Mutagen.* 1:387–99

79. Shan Q, Wang Y, Li J, Gao C. 2014. Genome editing in rice and wheat using the CRISPR/Cas system. *Nat. Protoc.* 9:2395–410

80. Shaw LM, Turner AS, Herry L, Griffiths S, Laurie DA. 2013. Mutant alleles of *Photoperiod-1* in wheat (*Triticum aestivum* L.) that confer a late flowering phenotype in long days. *PLOS ONE* 8:e79459

81. Simmonds J, Scott P, Brinton J, Mestre TC, Bush M, et al. 2016. A splice acceptor site mutation in *TaGW2-A1* increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. *Theor. Appl. Genet.* 129:1099–112

82. Simons KJ, Fellers JP, Trick HN, Zhang Z, Tai Y-S, et al. 2006. Molecular characterization of the major wheat domestication gene *Q*. *Genet.* 172:547–55

83. Slade AJ, Fuerstenberg SI, Loeffler D, Steine MN, Facciotti D. 2005. A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nat. Biotechnol.* 23:75–81

84. Slade AJ, McGuire C, Loeffler D, Mullenberg J, Skinner W, et al. 2012. Development of high amylose wheat through TILLING. *BMC Plant Biol.* 12:69

85. Stadler LJ. 1928. Mutations in barley induced by X-rays and radium. *Science* 68:186–87

86. Stadler LJ. 1929. Chromosome number and the mutation rate in *Avena* and *Triticum*. *PNAS* 15:876–81

87. Stadler LJ. 1930. Some genetic effects of X-rays in plants. *J. Hered.* 21:3–19

88. Stephenson P, Baker D, Girin T, Perez A, Amoah S, et al. 2010. A rich TILLING resource for studying gene function in *Brassica rapa*. *BMC Plant Biol.* 10:62

89. Steuernagel B, Periyannan SK, Hernandez-Pinzon I, Witek K, Rouse MN, et al. 2016. Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nat. Biotechnol.* 34:652–55

90. Sun C, Hu Z, Zheng T, Lu K, Zhao Y, et al. 2017. RPAN: rice pan-genome browser for ∼3000 rice genomes. *Nucleic Acids Res.* 45:597–605

91. Swaminathan M, Chopra V, Bhaskaran S. 1962. Chromosome aberrations and the frequency and spectrum of mutations induced by ethylmethane sulphonate in barley and wheat. *Indian J. Genet. Plant Breed.* 22:192–207

92. Swigoňová Z, Lai J, Ma J, Ramakrishna W, Llaca V, et al. 2004. Close split of sorghum and maize genome progenitors. *Genome Res.* 14:1916–23

93. Tsai H, Howell T, Nitcher R, Missirian V, Watson B, et al. 2011. Discovery of rare mutations in populations: TILLING by sequencing. *Plant Physiol.* 156:1257–68

94. Tsai H, Missirian V, Ngo KJ, Tran RK, Chan SR, et al. 2013. Production of a high-efficiency TILLING population through polyploidization. *Plant Physiol.* 161:1604–14

95. Tsuda M, Kaga A, Anai T, Shimizu T, Sayama T, et al. 2015. Construction of a high-density mutant library in soybean and development of a mutant retrieval method using amplicon sequencing. *BMC Genom.* 16:1014

96. Turck F, Fornara F, Coupland G. 2008. Regulation and identity of florigen: FLOWERING LOCUS T moves center stage. *Annu. Rev. Plant Biol.* 59:573–94

97. Turner A, Beales J, Faure S, Dunford RP, Laurie DA. 2005. The *Pseudo-Response Regulator Ppd-H1* provides adaptation to photoperiod in barley. *Science* 310:1031–34

98. 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* 314:1298–301

99. Uauy C, Paraiso F, Colasuonno P, Tran RK, Tsai H, et al. 2009. A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat. *BMC Plant Biol.* 9:115

100. von Zitzewitz J, Szűcs P, Dubcovsky J, Yan L, Francia E, et al. 2005. Molecular and structural characterization of barley vernalization genes. *Plant Mol. Biol.* 59:449–67

101. Wang TL, Uauy C, Robson F, Till B. 2012. TILLING *in extremis*. *Plant Biotechnol. J.* 10:761–72

102. Wang X, Wang H, Wang J, Sun R, Wu J, et al. 2011. The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* 43:1035–39

103. Wang Y, Cheng X, Shan Q, Zhang Y, Liu J, et al. 2014. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat. Biotechnol.* 32:947–51

104. Wang Z, Hobson N, Galindo L, Zhu S, Shi D, et al. 2012. The genome of flax (*Linum usitatissimum*) assembled *de novo* from short shotgun sequence reads. *Plant J.* 72:461–73

105. Watson A, Ghosh S, Williams M, Cuddy WS, Simmonds J, et al 2017. Speed breeding: a powerful tool to accelerate crop research and breeding. *bioRxiv*. **https://doi.org/10.1101/161182**

106. Wheat TILLING. 2017. *In silico wheat Target Induced Local Lesions In Genome (TILLING) website*. Univ. Calif. Davis, Rothamsted Res., Earlham Inst., and John Innes Cent. **http://www.wheat-tilling.com**

107. Wilhelm EP, Turner AS, Laurie DA. 2009. Photoperiod insensitive *Ppd-A1a* mutations in tetraploid wheat (*Triticum durum* Desf.). *Theor. Appl. Genet.* 118:285–94

108. Winfield MO, Wilkinson PA, Allen AM, Barker GLA, Coghill JA, et al. 2012. Targeted re-sequencing of the allohexaploid wheat exome. *Plant Biotechnol. J.* 10:733–42

109. Xu R, Qin R, Li H, Li D, Li L, et al. 2017. Generation of targeted mutant rice using a CRISPR-Cpf1 system. *Plant Biotechnol. J.* 15:713–17

110. Yan L, Helguera M, Kato K, Fukuyama S, Sherman J, Dubcovsky J. 2004. Allelic variation at the *VRN-1* promoter region in polyploid wheat. *Theor. Appl. Genet.* 109:1677–86

111. Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, et al. 2004. The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. *Science* 303:1640–44

112. Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J. 2003. Positional cloning of the wheat vernalization gene *VRN1*. *PNAS* 100:6263–68

113. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, et al. 2015. Cpf1 is a single RNA-guided endonuclease of a Class 2 CRISPR-Cas system. *Cell* 163:759–71

114. Zhang Y, Liang Z, Zong Y, Wang Y, Liu J, et al. 2016. Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9 DNA or RNA. *Nat. Commun.* 7:12617

115. Zhao Y, Zhang C, Liu W, Gao W, Liu C, et al. 2016. An alternative strategy for targeted gene replacement in plants using a dual-sgRNA/Cas9 design. *Sci. Rep.* 6:23890

116. Zong Y, Wang Y, Li C, Zhang R, Chen K, et al. 2017. Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. *Nat. Biotechnol.* 35:438–40

ANNUAL REVIEWS
Connect With Our Experts
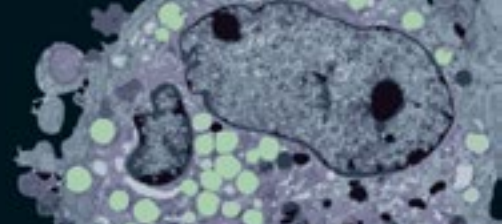
New From Annual Reviews:

Annual Review of Cancer Biology
cancerbio.annualreviews.org • Volume 1 • March 2017

ONLINE NOW!

Co-Editors: **Tyler Jacks**, *Massachusetts Institute of Technology*

Charles L. Sawyers, *Memorial Sloan Kettering Cancer Center*

The *Annual Review of Cancer Biology* reviews a range of subjects representing important and emerging areas in the field of cancer research. The *Annual Review of Cancer Biology* includes three broad themes: Cancer Cell Biology, Tumorigenesis and Cancer Progression, and Translational Cancer Science.

ANNUAL REVIEWS | CONNECT WITH OUR EXPERTS

650.493.4400/800.523.8635 (US/CAN)
www.annualreviews.org | service@annualreviews.org

Contents

Errata

An online log of corrections to *Annual Review of Genetics* articles may be found at
http://www.annualreviews.org/errata/genet