Creation and judicious application of a wheat resistance gene atlas

Amber N. Hafeez[1], Sanu Arora[1], Sreya Ghosh[1], David Gilbert[1], Robert L. Bowden[2], Brande B. H. Wulff[1]

[1]John Innes Centre, Norwich Research Park, Norwich, UK
[2]USDA-ARS, Hard Winter Wheat Genetics Research Unit, Manhattan, KS 66506, USA

Abstract

Disease resistance (*R*) gene cloning in wheat (*Triticum aestivum*) has been accelerated by the recent surge of genomic resources, facilitated by advances in sequencing technologies and bioinformatics. However, with the challenges of population growth and climate change ahead of us, it is vital not only to clone and functionally characterise a few handfuls *R* genes, but to do so at an agroecologically relevant scale. Pathogen populations are continually changing, and breeders must have tools and resources available to respond to those changes as quickly as possible if we are to safeguard our daily bread. To meet this challenge, we propose the creation of a wheat *R* gene atlas by an international community of researchers and breeders. The atlas would take the form of an online directory from which sources of resistance could be identified and deployed to achieve more durable resistance to the major wheat pathogens, such as wheat rusts, blotch diseases, powdery mildew and wheat blast. We present a costed proposal detailing how the interacting molecular components governing disease resistance could be captured from both the host and pathogen through biparental mapping, mutational genomics and whole-genome association genetics. We explore options for the effective configuration and genotyping of diversity panels of hexaploid and tetraploid wheat, as well as their wild relatives and major pathogens, and discuss how the atlas could inform a dynamic, durable approach to *R* gene deployment. Set against the current magnitude of wheat yield losses worldwide (recently estimated at 209 million tonnes each year; Savary *et al.*, 2019a), this endeavour presents one route for bringing *R* genes from the lab to the field at a considerable speed and quantity.

Introduction

The recent discovery of charred breadcrumbs in an ancient Natufian fireplace in north-eastern Jordan revealed that bread making was practised long before the dawn of agriculture (Arranz-Otaegui *et al.*, 2018). This early bread, made from wild grasses (Arranz-Otaegui *et al.*, 2018), would have required extensive labour to forage, peel and grind the grains. It took another 3,000 years before wild grasses were domesticated and cultivated, which gave farmers far greater control over food production and thereby laid the foundation for early civilisations in the Fertile Crescent. Today, wheat (*Triticum aestivum*) provides 18% of the calories and 19% of the protein consumed by humankind (FAOSTAT, 2017; http://www.fao.org/faostat/en/#data/FBS). The sudden disruption of this supply can lead to famine and civil unrest; for example, many observers have noted the correlation between poor wheat harvests around the world in 2010, the associated sharp rise in the price of bread in North Africa and the Middle East, and the Arab Spring, which led to the toppling of governments across the region (Zurayk and Khalidi, 2011).

Diseases and pests constitute a major limitation to wheat production, reducing the world's projected harvest by an average of 21% every year (Savary *et al.*, 2019a; **Fig. 1**). The top four pests and diseases account for over half of this loss (**Fig. 1**). These damaging agents are seldom evenly spread; their sudden and unexpected appearance can rapidly cause an epidemic leading to large-scale regional crop failure (Saari and Wilcoxson, 1974; Hovmøller, Walter and Justesen, 2010; Islam *et al.*, 2016). Control measures include pesticides, quarantine and breeding to combine genetic variation for enhanced resistance with other traits, such as grain yield. Genetic resistance constitutes an economical and environmentally friendly way to control disease and is especially important for those farmers who lack access to pesticides, whether due to prohibitive costs or regulations. Moreover, the frequent and excessive application of pesticides can lead to selection for pathogen

resistance to these chemical compounds (Omrane *et al.*, 2017), further demonstrating the need for genetic resistance in crops.

A series of polyploidisation events in the evolutionary history of wheat, followed by domestication 10,000 years ago, sampled only a fraction of the genetic diversity at each successive step (Lu, 2020; **Fig. 2**). Breeding over the last 200 years has further reduced this diversity (Pont *et al.*, 2019; Zhou *et al.*, 2020), with modern farming practices portioning the remaining diversity into monocultures prone to epidemics. It has long been recognised that the wild progenitors, landraces and domesticated relatives of bread wheat constitute a reservoir of genetic diversity that can be exploited to improve disease and pest resistance (Feuillet, Langridge and Waugh, 2008). Indeed, over the last 100 years, the bread wheat gene pool has been endowed with 204 exotic resistance (*R*) genes, constituting 44% of the 467 *R* genes designated in wheat (**Fig. 2**; **Table S1**). Notable successes include the stem rust resistance genes *Sr2*, *Sr25* and *Sr31* introduced from emmer wheat (*Triticum turgidum*), tall wheat grass (*Thinopyrum obtusiflorum*) and rye (*Secale cereale*), respectively (Pan, 1940; Friebe *et al.*, 1994; Lukaszewski, 2000). Unfortunately, the breeding timelines are typically long (15 to 20 years; Erickson, 1945) and suffer from linkage drag – the co-introduction of deleterious linked alleles (Olson *et al.*, 2010; Niu *et al.*, 2014). Moreover, when only single *R* genes are deployed, these are often rapidly overcome by the pest or pathogen (McDonald and Linde, 2002).

In this review, we argue that, combined with large-scale phenotyping, sequenced diversity panels of (i) wheat and wild progenitors and (ii) pathogen isolates causing major diseases of wheat could be employed to rapidly identify and clone large numbers of *R* genes and their corresponding pathogen effectors. Our vision is to create an atlas of characterised *R* genes from bread and durum wheat, as well as their direct tetraploid and diploid progenitors, to support wheat pre-breeding and breeding. Combined with the monitoring of effectors in pathogen populations, this would enable conventional breeders and genetic engineers to more efficiently and judiciously manage natural variation for disease resistance, for example by eliminating varieties reliant on a single *R* gene and creating and maintaining broad-spectrum *R* gene stacks effective against many strains of different pathogens.

Molecular components of the atlas

During infection, pathogens secrete proteins known as effectors or avirulence (Avr) proteins. Plant hosts have evolved surface and intracellular immune receptors to recognise these molecules. They are generally encoded by *R* genes, and often interact with effectors in a gene-for-gene relationship (Flor, 1971). Effector detection triggers the initiation of defence responses that can limit pathogen proliferation in the host. These interacting molecules are the essential components of the atlas. Their structure, function and evolution have been explored in more detail in other reviews (e.g., Jones and Dangl, 2006; Cook, Mesarich and Thomma, 2015).

Diversifying selection, acting on variation generated by recombination and mutation, has often resulted in great diversity between *R* gene loci (Kuang *et al.*, 2004; Van de Weyer *et al.*, 2019; Hatta *et al.*, 2020). The vast majority of major-effect dominant *R* genes cloned to date have been found to encode intracellular nucleotide-binding and leucine-rich repeat (NLR) immune receptors (Kourelis and Van Der Hoorn, 2018). NLR-mediated resistance is typically race-specific (i.e., gene-for-gene) and operates at all stages of plant growth (Periyannan *et al.*, 2017). Some dominant race-specific *R* genes encoding membrane-anchored surface receptors have also been identified, such as *Xa21* (Song *et al.*, 1995), conferring bacterial blight resistance in rice (*Oryza sativa*), and *Stb6* (Saintenac *et al.*, 2018), conferring resistance to *Septoria tritici* Blotch (STB). There are also some cases of cloned recessive *R* genes; for example, the *mlo* gene in barley (*Hordeum vulgare*) and wheat encodes a multi-transmembrane protein and confers resistance to powdery mildew (Wang *et al.*, 2014; Acevedo-Garcia *et al.*, 2017). Examples of non-NLR *R* genes in wheat include the cloned rust adult plant resistance (*APR*) genes, which encode proteins belonging to disparate structural

classes, including an ABC transporter (*Lr34*; Krattinger *et al.*, 2009), a protein kinase (*Yr36*; Fu *et al.*, 2009) and a hexose transporter (*Lr67*; Moore *et al.*, 2015). These non-NLRs are effective only at the adult plant stages, are not race- or pathogen-specific, and are more quantitative in nature, conferring only partial resistance. This may make them less vulnerable to the process of reciprocal pathogen co-evolution since they present a relatively weak selection pressure.

The wheat *R* gene catalogue must be expanded to facilitate deployment in a variety of combinations, with the aim of slowing down pathogen co-evolution and improving durability. Cloning genes offers insight into their mechanism of function, which in turn allows *R* genes with complementary functions to be combined. Perfect markers can also be designed from cloned *R* genes, facilitating their tracking in breeding programs and integration with quantitative resistance. Combining this with information about the corresponding effectors, such as their frequencies in pathogen populations and their fitness costs, will help to predict combinations of *R* genes that provide effective and durable resistance.

Populating the atlas: *R* genes

Methods of *R* gene cloning

To date, 38 wheat *R* genes have been cloned, more than half of which have been contributed by the species other than bread and durum wheat (**Table S2**). The process of isolating *R* genes from wider gene pools is challenging, however, and it is progressively more difficult to transfer genetic material to wheat from the secondary and tertiary gene pools (Harlan and de Wet, 1971). Conventionally, a biparental mapping population would be generated, using parents with contrasting phenotypes followed by the selfing of the offspring to generate a segregating population (**Fig. 2**). The mapping population is genotyped and phenotyped to facilitate the coarse mapping of the *R* gene locus, followed by high-resolution mapping and positional cloning. This resource-intensive, time-consuming process is often hampered by suppressed recombination between the mapping parents – particularly when the target *R* gene resides in chromatin from the secondary or tertiary gene pool.

Mutational genomics can overcome the limitation of recombination. This strategy is based on mutagenesis using a mutagen, such as ethyl methanesulfonate, followed by the phenotypic screening and sequencing of multiple independently-derived susceptible mutants to identify the candidate gene. Mutagenesis can be combined with capture-based methods to reduce the complexity of the sequence data gathered from genomes. The NLR gene family shares distinct protein motifs that can be used to guide the design of specific probes for *R* gene enrichment sequencing (RenSeq; Jupe *et al.*, 2013). Mutational genomics combined with RenSeq (MutRenSeq) was used to isolate the wheat rust resistance genes *Sr22*, *Sr26*, *Sr45*, *Sr61, Yr5a*, *Yr5b* and *Yr7*, and the powdery mildew resistance gene *Pm21* (**Table S2**; Steuernagel *et al.*, 2016; Marchal *et al.*, 2018; Xing *et al.*, 2018; Zhang *et al.*, 2020); however, capture-based methods are biased towards the annotated genes for which probes have been designed. To reduce genome complexity in an unbiased and lossless manner, specific chromosomes can be flow sorted and sequenced from the wild type and several independent mutants, followed by a sequence comparison for the identification of a candidate gene (MutChromSeq; Sánchez-Martín *et al.*, 2016). With declining sequencing costs, such reduced representation methods may not be necessary even for large Triticeae genomes.

These methods require the genetic isolation of the *R* gene followed by further crossing or mutagenesis, which is time consuming and generally restricted to the cloning of one gene per population. Performing mutagenesis on wild species can be particularly challenging due to their poor agronomy and, in the case of diploids, limited genetic redundancy, which necessitates the generation of mutant populations more than ten times larger than those required for hexaploid wheat (Uauy, Wulff and Dubcovsky, 2017). An alternative is to take advantage of Mother Nature's own genetically structured populations that have accumulated recombinations and mutations over thousands of generations. This can be achieved through association mapping with diversity panels,

which can determine genotype-trait correlations to permit the identification of markers tightly linked to the trait of interest (Togninalli *et al.*, 2018), providing the opportunity to clone many genes of interest from one sequence-configured population.

Building diversity panels

For the successful implementation of association genetics, effort must be made to develop suitable diversity panels – collections of individual accessions representing the genetic and phenotypic diversity of a population or species. There are several diversity panels of wheat and wild wheat progenitors available from different gene banks around the globe; however, many gene banks lack sufficiently informative catalogues, leading to redundancy between collections. Singh *et al.* (2019) found that over 50% of *Aegilops tauschii* accessions had been duplicated both within and between gene banks. Similarly, Bonman *et al.* (2015) explored genetic diversity among wheat accessions from the USDA National Small Grains Collection and found that most of the accessions were nearly identical to at least one other accession in the core subset. Given the cost of managing germplasm resources, it would be useful to compare the stocks held in different gene banks and weed out duplicate accessions. This would permit the more efficient utilisation of the germplasm and lower the barrier to mobilising novel genetic variation into breeding programs. Such a curation effort has been undertaken by the Seeds of Discovery project (CIMMYT, Mexico), wherein ~80,000 wheat accessions have been genotyped and core lines representing the genetic diversity of wheat have been identified (Sansaloni *et al.*, 2020).

A good diversity panel evenly captures the genetic diversity of the species across its distributional range. Uneven sampling may increase the effect of population demographic history on genome-wide association studies (GWAS); for example, the presence of accessions that are highly genetically and phenotypically similar can potentially inflate the allelic frequencies and linkage disequilibrium (LD) among unlinked loci, which can cause false positive associations (Arora *et al.*, 2017). This can be mitigated through the removal of such accessions and the application of statistical methods to control the impact of population structure (Zhang *et al.*, 2010; Bartoli and Roux, 2017). Additionally, if a phenotype is controlled by several genes or alleles which are each specific to different genetic groups within the population, the power to detect a candidate gene is reduced. The effects of genetic and allelic heterogeneity can be effectively controlled by performing analyses separately for each genetic group. This can result in greater power and resolution, despite the reduction of panel size. Of course, association genetics is not a viable approach in all cases; for example, some genes may be extremely rare, such as the lettuce (*Lactuca sativa*) downy mildew resistance gene *Dm3*, which was found to be present in only 1 of 1033 wild accessions (Kuang *et al.*, 2006). It would also be difficult to identify candidate genes through GWAS when the LD blocks of the associated region are very large, such as in non-recombining alien chromosome segments. Through characterising 'low-hanging fruit' via association mapping, those accessions containing unexplained resistance sources can be interrogated by generating biparental mapping or mutant populations. Thus, this approach can help to ensure that the best cloning method is applied to each gene.

In short, while a large and diverse panel increases the power of association, care must be taken to avoid a skewed population structure. Phenotyping a panel can be both difficult and expensive, and germplasm may be limited; therefore, it is advisable to first evaluate the phenotypic variation in a reduced core set of accessions that captures the majority of allelic diversity. Phenotypes exhibiting sufficient variation in the reduced core set are more likely to be worth carrying forward on the whole panel.

Genotyping diversity panels

Most association studies to date have used single-nucleotide polymorphisms (SNPs) as genetic markers, since they are widely distributed across the genome and can be cost-effectively scored using arrays containing probes for thousands of SNPs. These arrays typically only capture a fraction of the genetic variation within a population however, since they fail to capture structural variants

such as copy number variations (CNVs) and insertion–deletions (indels). Furthermore, the pre-selection of SNPs with high minor allele frequencies may reduce the likelihood of finding markers in LD with rare resistances.

Another option is to perform reduced representation sequencing, such as exome or gene family capture sequencing, or whole-genome shotgun sequencing. Rather than using SNPs, reference bias can be mitigated by performing associations directly on *k*-mers, sub-sequences of length *k*, which can be generated directly from raw sequence reads (Audano, Ravishankar and Vannberg, 2018; Rahman *et al.*, 2018; Arora *et al.*, 2019; Voichek and Weigel, 2020). If *k*-mer variation in the panel is correlated with phenotypic variation, then *k*-mers significantly associated with the trait of interest can be used to assemble the reads from which they were derived or mapped directly to a *de novo* assembly generated from an accession expressing the trait. For example, association genetics in combination with RenSeq (AgRenSeq) was successfully applied to an *Aegilops tauschii* diversity panel, which allowed the rapid identification of four stem rust *R* genes (Arora *et al.*, 2019).

When the *de novo* assembly used as a reference for *k*-mer mapping is fragmented, it can be difficult to infer the complete sequence of the target gene. The situation can become more complex when a candidate gene is in LD with other genes in the vicinity of the true candidate; this may result in distinct association signals, implying several candidate genes are responsible for the phenotype. These problems can be mitigated through the use of a better-quality assembly to fetch the haplotype block surrounding the candidate gene. The continued drop in the cost of DNA sequencing now makes it economically feasible to generate a library of high-quality *de novo* assemblies for the core accessions that represent most of the genetic diversity in the panel; for example, 38 accessions are enough to represent 95% of the pan-NLR repertoire in *Arabidopsis thaliana* (Van de Weyer *et al.*, 2019). In this review, we have classified Triticeae assemblies into four categories – gold+, gold, silver and bronze—by the quality of the information that they provide (**Fig. 4**). By implementing a tiered sequencing approach, the maximum amount of genetic diversity can be captured while limiting computational and sequencing costs.

The first step in sequence-configuring a diversity panel is to perform cost-effective genotyping (e.g., genotyping-by-sequencing (Elshire *et al.*, 2011) or low-coverage whole-genome skim sequencing (Huang *et al.*, 2009)) to permit quality checking for the removal of duplicates and accessions with high levels of residual heterogeneity. This facilitates a population structure analysis for the identification of major genetic groups and genetic relationships among accessions in the population. Following this survey, accessions representing each genetic group can be selected to generate gold+ assemblies consisting of chromosome-level pseudomolecules (IWGSC *et al.*, 2018); for example, eight to 11 accessions to represent the sub-groups in wheat (Wingen *et al.*, 2014; Balfourier *et al.*, 2019) or two to represent the main *Aegilops tauschii* sub-groups (Mizuno *et al.*, 2010). Gold+ assemblies provide the position of all genes in their physical context, allowing the position of a new candidate gene to be compared with known genes, the anchoring of contigs from fragmented assemblies to infer their position, and the calling of SNP variation from shallow sequencing data.

Gold-level assemblies can be used to define the haplotype blocks in the genome. Depending on the size of the LD blocks, the scaffold size of these assemblies should be in the range of ~100 kb to several Mb. This allows the interrogation of all genes in the LD block of an associated region. Moreover, gold-level assemblies allow the identification of structural variants, which are important determinants of plant phenotype (Alonge *et al.*, 2020; Liu *et al.*, 2020) including disease resistance (Cook *et al.*, 2012). The gene(s) most strongly correlated with the phenotype can then be identified as the most likely candidate(s) for functional validation. To maximise the representation of the haplotype diversity within a panel, 20–30 geographically and genetically distinct accessions representing genetic sub-groups within each major group could be assembled to gold-level standard.

More cost-effective silver-level assemblies can be generated for additional tens of accessions, with scaffold sizes ranging from 5 to 15 kb. They provide gene-level resolution, including regulatory regions, which can be useful for designing and engineering gene constructs for the functional

validation of candidate genes. They can also be anchored to the reference assemblies for the identification of small-scale structural variation.

The gold+, gold and silver assemblies could be used to construct an inter-connected graph-based pangenome assembly containing the majority of the structural and sequence variation, as recently demonstrated in soybean (*Glycine max*) (Liu *et al.*, 2020). Pangenomes generated for other *Triticum* species sequenced as part of the atlas project could also be combined to form a genus-wide "super-pangenome" (Khan *et al.*, 2020). Following this, bronze-level sequencing generated from low-coverage Illumina reads can be carried out cost-effectively to genotype hundreds of individuals; one-fold coverage or less is likely sufficient to impute their haplotype structure for use in GWAS (Gilly *et al.*, 2019).

Additional features could be included, such as the epigenetic status of gold-level accessions (Gardiner *et al.*, 2015), which, when combined with RNA-seq data for all accessions in the panel (Harper *et al.*, 2012), could provide insight into the regulation of gene expression (Kawakatsu *et al.*, 2016). This may help to explain disease resistance or virulence that cannot be elucidated through the interrogation of sequence variation alone; for example, epigenetics and small RNAs may play a role in the pathogenicity of the wheat leaf rust pathogen *Puccinia triticina* (Wu *et al.*, 2017).

Populating the atlas: Effectors

Not all combinations of *R* genes in a stack will be successful. Functional suppression is a common problem in wheat breeding, especially following the introgression of chromosomes from wild relatives with a lower ploidy level; resistance from these sources may become less potent, or even ineffective, once in a hexaploid wheat background (Bai and Knott, 1992; Innes and Kerber, 1994; McIntosh *et al.*, 2011; Chen, Liu and Gao, 2013; Hiebert *et al.*, 2020). One example is the post-translational suppression of the rye-derived powdery mildew NLR *Pm8* by the orthologous wheat gene *Pm3*, likely through the formation of non-functional protein complexes (Hurni *et al.*, 2014). It is therefore important to rigorously test the function of individual *R* genes in stacks to ensure that single *R* genes do not become exposed to pathogens in isolation due to non-functioning stack components. This is difficult or impossible to determine through pathogen assays due to epistasis from the recognition of multiple effectors; however, with cloned effectors, it is possible to probe the function of individual stack components using heterologous delivery systems. For example, the functional verification of the *Stagonospora nodorum* host-selective toxin *SnTox1* was carried out by producing the effector protein in yeast then infiltrating the culture filtrate into wheat leaves to induce necrosis (Liu *et al.*, 2012). The improvement of transient expression systems in wheat, such as virus-mediated overexpression (VOX) (Bouton *et al.*, 2018) and transient expression in wheat protoplasts (Saur *et al.*, 2019), has facilitated the rapid functional analysis of intracellular effectors.

Cloning effectors

There are several approaches to cloning effector genes. Variation can be generated through mutagenesis (Salcedo *et al.*, 2017; Kangara *et al.*, 2020), but there are challenges associated with using mutant populations to clone effector genes: the presence of more than one nucleus in some pathogen species (e.g. dikaryotic *Puccinia* spp. and stinking smut, *Tilletia tritici*), labour-intensive bulking, applying an appropriate mutagen dose, and the difficulty of accurately identifying mutant variants. Positional mapping has been used to clone effectors, such as from powdery mildew (Praz *et al.*, 2017) and Hessian fly (Aggarwal *et al.*, 2014); however, the sexual recombination required for creating a mapping population may not be straightforward for all pathogens and pests. The sexual cycle is unknown for some, while others have complex heteroecious lifecycles involving multiple hosts; for example, the wheat stem rust pathogen (*Puccinia graminis*) requires an alternate host (usually barberry, *Berberis vulgaris*) to complete its sexual lifecycle.

Another approach would be to screen a genetically diverse set of pathogen isolates against a differential set of wheat lines that contain known *R* genes – the complement to association genetic studies focussed on the cloning of host *R* genes. While improvement on the host side could focus

on capturing lost genetic solutions to the problem of disease resistance (perhaps in wild grasses with smaller genomes), on the pathogen side, the opportunity exists to sample swathes of the standing genetic variation that contributes to virulence in the field, which is most directly relevant to agriculture (Talas *et al.*, 2016). Association studies have so far been used to clone effectors from the pathogens *Magnaporthe oryzae* (Yoshida *et al.*, 2009), *Zymoseptoria tritici* (Zhong *et al.*, 2017) and *Stagonospora nodorum* (Gao *et al.*, 2016). Given the small size of pathogen genomes (in the range of 30 to 60 Mbp for fungal plant pathogens (Aylward *et al.*, 2017)), whole-genome association genetics is feasible. Fewer than 100 sequenced pathogen isolates would likely be sufficient to clone the major genes that have a gene-for-gene interaction with the host (Sánchez-Vallet *et al.*, 2018), although for some pathogens a lack of sexual recombination and gene flow could limit the power of association studies. Furthermore, by compiling pathogen diversity panels, we can also elucidate the genetic structure of pathogen populations. This is a gateway to exploring the evolutionary history and potential of pathogens, and therefore to better understanding the durability of host resistance, which would aid in planning for future epidemics (McDonald and Linde, 2002).

A two-way association genetics approach, where the analysis is carried out on the host and pathogen in tandem, could increase the analytical power enough to identify components in host–pathogen interactions with minor effects. A phenotyping-free strategy was adopted by Bartha *et al.* (2013) for analysis of the human–HIV pathosystem, revealing that human SNPs were associated with amino acid variations in the virus HIV-1 protein, providing stronger signals than when human SNPs were associated with viral load. This genome-to-genome analysis could be adapted to detect signals of coevolution between plants and their pathogens through an 'ecological genomics approach', where the genome sequences of a host diversity panel and its adapted pathogen strains are linked (Bartoli and Roux, 2017). There is potential for this method to identify variation in *R* or *Avr* loci in both the host and pathogen, respectively, that may be difficult to derive from phenotypic data alone.

Reverse genetics can also be applied to identify effectors. Putative effectors can be elucidated from whole-genome sequences of pathogens *in silico* through the identification of effector characteristics such as small size (<30 kDa), high cysteine content, proximity to repetitive sequences, expression patterns, and the presence/absence in known virulent and avirulent isolates (Liu *et al.*, 2012).

To date, only eight pathogen effectors corresponding to cloned wheat *R* genes have been identified (**Table S3**); however, as more pathogen pangenomes are sequenced, the power to identify effector genes will increase, whether through association genetics or *in silico* identification and confirmation. One example is *SnTox1,* as the properties of the previously cloned *SnToxA* and *SnTox3* genes contributed to its cloning (Liu *et al.*, 2012).

Capturing disease response phenotypes

An important consideration when screening pathogen diversity panels is the laborious handling and calibration of inoculum for so many isolates. High-throughput procedures and equipment would be required, while more standardised and precise methods could help to provide the clear phenotypes that facilitate association studies. The sensitivity of some pathogens to environmental conditions can render interactions incomparable between institutions or even particular growth chambers. Techniques such as detached leaf assays (Arraiano, Brading and Brown, 2001) can help to minimise environmental factors, facilitate the testing of multiple pathogen isolates at once without cross-contamination and also reduce the risk of pathogen escape (therefore removing the need for high-level containment facilities). The precision and efficiency of phenotyping could be improved through techniques such as automated image analysis, which potentially allow a greater array of virulence phenotypes to be captured in each experiment, such as pycnidia or pustule number, size, density and other quantitative measures (Stewart and McDonald, 2014; Stewart *et al.*, 2016; Xu *et al.*, 2017; Bueno-Sancho *et al.*, 2019). Quantitative trait loci that account for host resistance to these components of disease aggressiveness can then be identified from diversity panels using association genetics (Yates *et al.*, 2019).

The use of purified or transiently expressed effectors for phenotyping could overcome several limitations. For apoplastic pathogens such as *Zymoseptoria tritici* and *Stagonospora nodorum*, culture filtrates of pathogen or yeast strains expressing effectors of interest can be injected into leaves to cause an interaction (Liu *et al.*, 2012; Ben M'Barek *et al.*, 2015); however, such methods are not effective for intracellular haustoria-forming pathogens. Deploying transient expression systems such as VOX across entire diversity panels would be a great challenge, but the use of purified or transiently expressed effectors would alleviate containment demands as well as potentially provide clearer phenotypes. These methods could also increase the number of host diversity panels available for cloning *R* genes, since some pathogen–wild grass combinations may result in incompatible or non-host interactions (e.g., as was observed between *Aegilops tauschii* and *Zymoseptoria tritici;* Seifbarghi *et al.*, 2009) that could potentially be overcome through the introduction of effector proteins alone.

Due to the need for well-developed pathogen diversity collections, expertise and equipment, phenotyping would likely be restricted to a few specialised labs. As such, there would be a need for global co-operation to achieve the atlas as efficiently as possible. Where multiple labs are involved, standardised processes should be implemented to reduce the impact of environmental factors. Ultimately, as phenotyping improves, so does our ability to clone both host *R* and pathogen *Avr* genes.

Applying the atlas to wheat breeding

With the cost of genetic research continually falling, especially for wild relatives with comparatively small genomes, the ideal of a wheat *R* gene atlas seems ever more realisable. Previous GWAS in *Aegilops tauschii* ssp. *strangulata* identified all of the stem rust resistance genes that had been designated in this subspecies (Arora *et al.*, 2019). Extrapolating from this, and assuming the atlas would explore 10 hosts, this could potentially lead to the cloning of ~100 genes for the three rusts alone. The 10 host panels would include: winter, spring and landrace panels of bread wheat; elite and landrace panels of durum wheat; the direct progenitor species; and key wild or domesticated relatives that have made the largest contributions to *R* gene diversity in wheat, such as *Thinopyrum obtusiflorum* and rye (**Table S1**). Major pathogens to target first could include the three wheat rusts, powdery mildew and the blotch diseases (STB, SNB, spot blotch and tan spot). Sequencing all this material would cost $7.6 million and 75 full-time equivalents over five years to carry out the bulking, pathology, molecular biology and bioinformatics work (**Table S4**). This is on par with the staff numbers at The Sainsbury Laboratory, Norwich, UK, for which the running costs are ~$10 million per year (https://find-and-update.company-information.service.gov.uk/company/03346853). The atlas could therefore be produced for $58.6 million, or $2.9 million per G20 country, over five years. This would include the development of a web-based portal where phenotypes can be plugged in and analysis automated through the use of cloud computing resources, as well as a hub from which information on genes cloned through the project could be accessed. Given the gross value of wheat production worldwide ($114.0 billion; FAOSTAT, 2018; http://www.fao.org/faostat/en/#data/QV), the loss of 21% of yield to pests and diseases (Savary *et al.*, 2019) could be valued at around $31.2 billion dollars. It is therefore clear that the cost of generating the atlas is negligible compared with the magnitude of the problem. Even if the atlas were only to reduce loss to pathogens by one percentage point, this would still be equivalent to a 31-fold return on the investment (not taking factors such as inflation into account).

The wheat *R* gene atlas could take the form of an online directory providing free access to *R* genes, their allelic variation, molecular markers and sequences, as well as the prevalence of *Avr* effectors in locally sampled pathogen populations. This information could then be fed into a model which would select the optimal number of *R* genes required to provide broad-spectrum resistance to pathogen populations in target regions, while maximising the potential for durability by ensuring

each pathogen isolate is recognised by multiple *R* genes (**Fig. 5a**). This would reduce the selection pressure on individual *R* genes, prolonging their efficacy in the field.

Tracking effectors in the environment

Plant pathogens evolve rapidly (Oliver, 2012), so their populations would ideally be surveyed constantly to identify changes to their effector complement that could impact the efficacy of existing *R* genes in the field. Optimally, trap plots could be used to pick up rare variants; if placed next to sexual hosts, single-*R*-gene host differentials can be used to sample the sexual recombinants (Wellings *et al.*, 2009). This could enhance early-warning systems, such as those in Ethiopia, which facilitate the optimisation of variety choice and fungicide use in response to outbreaks – particularly pertinent concerns for smallholder farmers (Allen-Sader *et al.*, 2019). To do this effectively however, in-field pathogen surveillance infrastructure must be co-ordinated at the local, national and global scale (Carvajal-Yepes *et al.*, 2019).

There are several methods under development for the monitoring of pathogen populations and their virulence. One example is "field pathogenomics", where RNA-seq data is generated from pathogen-infected leaf samples collected in the field. Phylogenetic and structural analyses based on the most variable SNPs within the pathogen population are employed to gain insight into the genetic groups present (Hubbard *et al.*, 2015; Bueno-Sancho *et al.*, 2017). This has been used to demonstrate population shifts in wheat yellow rust (*Puccinia striiformis* f. sp. *tritici*) in the UK (Hubbard *et al.*, 2015) and to rapidly determine that a wheat blast outbreak in Bangladesh in 2016 was caused by a South American lineage of the blast fungus (*Magnaporthe orzae*; Islam *et al.*, 2016). The method can work with portable real-time sequencing devices such as the MinION (Oxford Nanopore Technologies) for full analysis at the site of collection (Radhakrishnan *et al.*, 2019). Since there can be considerable variation in virulence even between closely related pathogen strains, methods which deliver phylogenetic readouts may not provide sufficient resolution to detect shifts in pathogen virulence based on subtle changes in effector complement. With an atlas of cloned effectors, however, existing surveillance technologies could be tailored to detect pathogen effector complements. A case in hand is PathoTracer, which monitors sequences in pathogen TAL effectors to inform variety choice and the editing of recessive susceptibility genes in rice, ensuring resistance to bacterial blight (*Xanthomonas oryzae* pv. *oryzae*) (Eom *et al.*, 2019).

Developing effective *R* gene stacks

Although wild relatives and landraces may provide novel sources of genetic resistance that can be unlocked through genomics, the challenge to mobilise these resistances into elite wheat varieties still remains. Introducing genes through traditional breeding often results in linkage drag, while some species may not be sexually compatible with wheat or may generate progeny that suffer from abnormalities that hinder development, such as hybrid necrosis (Bomblies and Weigel, 2007). It may also be difficult to mobilise genes that exist in repulsion, as exemplified by *Pm64* and *Yr5* (Zhang *et al.*, 2019). Moreover, it would take many generations to combine multiple genes in the same background; for example, combining four broad-spectrum *R* genes for each of the three rusts (i.e., 12 genes) in an elite background would require 19 generations using a crossing approach which (i) selects for recombinants around the genes of interest in the first two backcrosses, followed by (ii) four generations of backcrossing and marker-assisted selection for the *R* gene and the elite background to generate near-isogenic lines with a >98% elite background. Next, (iii) a stepwise crossing approach would be used to combine the genes in twos, fours, sixes, eights, tens, and finally 12 (**Figure 6**). The early incorporation of a male sterile mutation such as *Ms3* would reduce the number of emasculations required (Guttieri, 2020). The strategy of starting with near-isogenic lines would minimise the accumulation of linkage drag from agronomically diverse and poor backgrounds in the final stack. Under optimal speed-breeding conditions in a spring wheat background (Ghosh *et al.*, 2018; Watson *et al.*, 2018), this could be achieved in under four years; however, if this were reduced to two diseases with selection for three genes per disease, the timeline could be reduced to 14 generations over around 2.5 years. Introgressing this many *R* genes into winter wheat varieties would take much longer unless vernalisation requirements can be

circumvented (Hickey *et al.*, 2019). Targeting three diseases (stem rust, stripe rust and STB) would be required in large parts of sub-Saharan Africa, while in Europe targeting stripe rust and STB would be sufficient to substantially reduce reliance on fungicides (Savary *et al.*, 2019). An advantage of this scheme is that it allows background quantitative resistance to be preserved, but it would be a Sisyphean struggle to retain these genetically separated stack components when cultivars are often swiftly turned over to maintain genetic gains in yield.

As an alternative to traditional crossing, multi-*R* gene stacks can be generated through DNA engineering and transformation. This has been demonstrated in cultivated potato (*Solanum tuberosum*) using three *R* genes against late blight from wild *Solanum* relatives (Ghislain *et al.*, 2019). Transgene stacking allows all of the *R* genes to be combined at a single locus, which ensures that the stack remains intact in downstream breeding programs, avoiding separation and therefore the exposure of single *R* genes to the pathogen (Wulff and Moscou, 2014; **Fig. 5b**). DNA synthesis and assembly technology now allows for the rapid generation of constructs from tens to several hundreds of kbp (Gibson *et al.*, 2008; Petersen and Stowers, 2011; Weber *et al.*, 2011; Annaluru *et al.*, 2014), while improvements to transformation will make it easier to deliver large constructs into the wheat genome (Richardson *et al.*, 2014; Lowe *et al.*, 2016; Hayta *et al.*, 2019; Debernardi *et al.*, 2020).

Quantitative and/or broad-spectrum *R* genes can play an important role in stacks due to their durability and ability to boost the function of other *R* genes; for example, resistance to leaf rust can be enhanced through a combination of *Lr34/Yr18/Sr57* and *Lr68*, and resistance to all three rusts can be increased significantly through a combination of *Sr2/Yr30* and *Lr34/Yr18/Sr57* (Randhawa *et al.*, 2018). The as-yet uncloned rust *R* gene *Sr2* provides partial *APR*, resulting in a slow-rusting phenotype that has remained effective in commercial wheat varieties for ~100 years (Ellis *et al.*, 2014). This demonstrates the importance of maintaining strong background resistance through breeding. Indeed, the combination of multiple minor *APR* genes forms the basis of the CIMMYT breeding programme for rust resistance (Bhavani *et al.*, 2019). Although originally linked in repulsion on chromosome 3B, the *Fhb1* and *Sr2* genes have been combined to provide a baseline of resistance to both Fusarium head blight and stem rust (Zhang *et al.*, 2016) and the coupled genes have been mobilised in CIMMYT germplasm (He *et al.*, 2020). The use of genetic modification (GM) cassettes would avoid the laborious process of developing lines in which genes in repulsion become coupled. Quantitative genes can also be combined with race-specific *R* genes to provide more durable resistance; in near-isogenic lines of oilseed rape (*Brassica napus*) all containing the *R* gene *Rlm6* but differing in quantitative resistance to black leg disease, *Rlm6* resistance was overcome within three growing seasons when deployed on its own, but remained effective after five years when combined with quantitative resistance (Brun *et al.*, 2010).

Gene stewardship

R gene stewardship can be defined as the careful and responsible management of *R* genes to remain effective during prolonged use. Gene stewardship may be directed towards decreasing the probability that virulent mutants arise, or towards reducing selection pressure for new virulent mutants after they arise. Although gene stewardship is often associated with gene deployment strategies, it also encompasses a broad range of activities including understanding *R* gene function and specificity in the host, *Avr* effector gene function and specificity in the pathogen, the identification of *R* genes in cultivars, and pathogen virulence frequency surveys.

The history of a gene's exposure to pathogens may account for much of its effect in the field; genes which come from a background that has never been exposed to particular pathogens are likely to be more effective. In fact, 'broad-spectrum' resistance refers to the ability of an *R* gene to recognise all corresponding *Avr* alleles present in many isolates at a particular point in time, and does not guarantee long-term durability (Ellis *et al.*, 2014). Even non-host resistance, in which a host species

is immune to all races of a particular pathogen species, opens up opportunities for pathogens to make a host jump once these genes are overcome. The wheat blast pathogen, for example, emerged after rye-infecting blast strains infected wheat lines lacking functional *Rwt3* resistance. Selection for the loss of the *PWT3* effector then allowed the widespread infection of *Rwt3* wheat lines (Inoue *et al.* 2017). It is therefore imperative to closely monitor the distribution of *R* genes in agricultural systems worldwide to prevent the development of cultivars with compromised resistance profiles, such as those with *R* gene stacks necessitating only a few stepwise mutations in the pathogen to be overcome. Releasing novel sources of resistance as part of stacks and employing good gene stewardship practices could reduce the occurrence of such gene collapses.

R genes can be divided into three groups based on their need for stewardship. Group A genes do not require stewardship. These include certain genes proven to be durable in the field, such as the *APR* genes *Lr34/Yr18/Sr57*, *Lr46/Yr29/Sr58*, and *Sr2/Yr30/Lr27*. It also includes newly cloned genes such as *Lr67/Yr46/Sr55* or *Yr36* that lack classic recognition-triggered immunity motifs and are therefore predicted to be inherently more durable than other *R* genes. Non-race-specific recessive *R* genes, such as *mlo*, are also expected to be inherently durable (Wang *et al.*, 2014; Acevedo-Garcia *et al.*, 2017). Group A would also include genes that are already widely deployed, genes that are already widely defeated by virulent pathogen races, and genes that have relatively minor effects on disease severity.

Group B genes are strongly recommended for stewardship. They include resistance genes that are undefeated or defeated by only one or a few races. Group B genes are still very useful, but they are expected to be vulnerable to defeat by new virulent races. Some examples would be *Yr5, Yr15, Sr22* and *Sr26*. Group B genes should be deployed only in combinations with other major-effect and quantitative genes. Combinations should be designed to withstand one or more new mutations to virulence in the pathogen while still preventing pathogen reproduction. For example, the stem rust isolate TTRTF is virulent to three of five genes included in a recently-developed wheat *R* gene stack, which could limit the efficacy and durability of the stack in regions where this isolate is present (Luo *et al.*, 2021). The atlas would be an invaluable resource for analysing the vulnerability of *R* gene combinations to extant pathogen populations and maximising the number of independent mutations required in each pathogen race to defeat the combination. The combinations could be dynamically updated, with old inadequate combinations being removed from production. Group B gene stewardship would require the co-operation of breeding programs to achieve the common good of preserving these genes over the long term. Existing cultivar release programmes (such as the UK Recommended List) could be leveraged by awarding “extra points” to cultivars carrying multiple *R* genes to regionally prevalent pathogen populations, boosting the incentive to breed for durable resistance.

Group C is for formal gene stewardship programs. Genes from uncrossable alien species or newly designed synthetic genes are good candidates for Group C. The gene sequences will be known and can be patented for use in transgenic cassettes, preventing the genes from being deployed irresponsibly elsewhere. Currently, there are active patents associated with ten wheat *R* genes (**Table S5**). Combinations should be designed to withstand at least three independent mutations to virulence and still prevent all reproduction in the pathogen. Again, the atlas will inform the selection of *R* genes to maximise durability and gene stewardship. If the pyramid is going to face a sexual population of the pathogen, a larger stack would be needed. Knowledge of pre-existing virulence frequencies would aid the design of appropriate combinations. Further, stacking *R* genes operating through different defence pathways would ensure diverse sources of resistance are acting against pathogens, which could enhance durability.

Reducing the exposure of *R* genes in the field

Novel sources of resistance could increase the durability of stacks, but their deployment in monoculture may not be enough to protect crops in the long term. The rate of evolution of pathogens combined with their huge volume in agricultural systems increases the likelihood that the multiple mutations required to overcome a stack will occur. This is especially the case if the pathogen is

reproducing sexually and is not eliminated completely in resistant lines (which is often the case with apoplastic pathogens such as *Zymoseptoria tritici*; Stotz *et al.*, 2014), or if virulence to stack components is already present in the population (Stam and McDonald, 2018). In addition, the development of new stacks is costly. It is therefore important that steps are taken to keep the pathogen effective population size as small as possible, reducing the likelihood that virulence would evolve.

One way to reduce the rate of evolution of virulence to newly deployed *R* genes is to employ 'dynamic diversity' (McDonald, 2014). Strong selection pressure from static monoculture crop fields enhances the efficiency of directional selection to overcome host resistance. A dynamic diversity approach would involve strategic and frequent changes to the resistance repertoire present in the field at any one time, disrupting selection for *Avr* evolution in the pathogen. This approach can be applied to any cropping system or genotype (**Fig. 5c**). Furthermore, the progress made by pathogens towards virulence within a growing season may hamper fitness in the succeeding season when different host genotypes are introduced. In this vein, it is also important to minimise the occurrence of "green bridges" between growing seasons, which can facilitate epidemics by increasing the pathogen effective population size. This could be a factor in stem rust epidemics in the Ethiopian Highlands, where conditions allow for two wheat seasons per year (Allen-Sader *et al.*, 2019).

The development of multilines, which differ in their *R* genes but are otherwise isogenic, could ensure that uniformity in heading date and grain quality is maintained while the *R* genes being presented to pathogens are diversified. This strategy has proven successful in the past; for example, in the deployment of barley multilines for control of mildew in former East Germany (Wolfe *et al.*, 1992). However, contemporary breeding programmes are unlikely to consider such a strategy due to regulatory bottlenecks, such as the capacity of seed producers to handle many varieties and the need to keep track of varieties in terms of Plant Breeder's Rights and to maintain uniformity within varieties (Louwaars, 2018). Another option for diversifying resistance profiles in the field is the use of cultivar mixtures. Susceptible varieties with more desirable backgrounds could be released in mixtures where 25% of plants are resistant, which has been successfully employed in durum wheat to reduce STB disease severity by 48% compared with susceptible pure stands (Ben M'Barek *et al.*, 2020). This strategy lacks the advantage of preserving uniformity in the field, but would allow the benefits of durable *R* gene stacks to be reaped even if they have not yet been mobilised in locally adapted germplasm.

When new sources of major genetic resistance are identified, there can be a temptation to immediately release them to the market to maximise yields and profit in the short term. The commercial cultivation of resistant varieties can lead to rapid adaptation of the pathogen and the breakdown of resistance in as little as five years, as was demonstrated by the case of the wheat variety 'Gene' in response to STB; this virulence can persist in the pathogen population even after use of the cultivar diminishes (Cowger, Hoffer and Mundt, 2000). It is therefore important to ensure that the immense investment of time and resources required for the deployment and/or cloning of *R* genes is not wasted through an immediate exposure to pathogen populations. Gene squandering could be avoided by empowering breeders with tools such as the atlas and supporting their training in molecular biology and bioinformatics – a growing need that was made clear through recent personal communication with wheat breeders.

Outlook

Here, we have discussed using genomics-assisted methods, association mapping in particular, for the rapid and widespread discovery and cloning of both *R* and *Avr* genes. Once diversity panels have been established, they can be evaluated for practically any trait of interest. New sources of variation can also be incorporated by increasing the panel size or the genetic information captured,

potentially increasing the power of GWAS to identify the genomic features underlying traits of interest. In this model, value can be continually added to diversity panels as and when motivation, time and funds allow.

Through an international effort, we envisage compiling a catalogue of *R* genes and corresponding *Avr* genes—the atlas. This would be a revolutionary tool with which breeders could select bespoke *R* gene combinations that address disease problems in response to regional pressures. Once stacks are introduced into wheat, the functionality of individual components could be validated with cloned *Avr* effectors. The atlas will also provide a unique opportunity to study the population genetic factors involved in the evolution and spread of disease resistance, such as the effects of polyploidisation, domestication and intensive breeding.

This mass-cloning and characterisation project will also create knowledge about the structure–function relationships of *R* gene recognition specificity. This can be exploited to expand *R* gene recognition through single amino acid changes, as in the case of R2a, which was engineered to recognise both variants of Avr3a in the potato late blight oomycete *Phytophthora infestans* (Segretin *et al.*, 2014), and PikP, which was engineered to recognise multiple variants of AvrPik from the rice blast fungus *Magnaporthe oryzae* (De La Concepcion *et al.*, 2019). *R* gene activation can also be manipulated to enhance the speed of response for more effective resistance (Harris *et al.*, 2013). These strategies could be deployed through gene editing; thus, gene stacks of the future could consist of major, minor and synthetic *R* genes.

Currently, the designation of new *R* genes requires genetic proof that the gene is in a novel location and/or has a novel specificity. This is complicated by the need to send seed for verification, often across international borders, which can present difficulties for some countries due to strict quarantine requirements. Lengthy genetic testing involving several generations of crossing can add further years to the process. Because of these barriers, there is a backlog of genes that have been given temporary designations. With cloned genes, high-quality assemblies, and knowledge of the particular *Avr*s recognised by the *R* gene, the science of gene designation will become easier, more precise and potentially faster. The need to send seed across international borders could be circumvented by simply sharing DNA sequences. It would also allow the community to integrate the historic cataloguing of genes based on genetics with sequence data and physical genome location.

Climate change is likely to exacerbate the current losses caused by crop pathogens (Caubel *et al.*, 2012). By the middle of the 21st century, many countries with high crop production could become fully saturated with pests (Bebber, Holmes and Gurr, 2014) as the ranges of pests and pathogens expand (Roos *et al.*, 2011) and climate and weather changes have unexpected effects on pathogen lifecycles (Shaw *et al.* 2018). At the same time, population expansion is increasing the demand on food production and its environmental impact (Foley *et al.*, 2011), and although changes in diet could allow more calories to be produced with less land (Poore and Nemecek, 2018), such large-scale cultural shifts are slow and perhaps unrealistic. This presses a continually growing demand on scientists and breeders to find faster and more efficient means of introducing novel genetic disease resistance into the field.

In Europe, many pesticides and fungicides have lost their efficacy, while many of the remaining effective compounds are heavily regulated or have been banned (Hillocks, 2012). Robust genetic resistance is therefore also important for reducing disease pressures to prolong the life of chemical controls, making it the most important tool for combating disease in the long term. If the perception of GM is tackled in Europe, there could be increased implementation in other countries that depend on trade with the European Union, such as nations in Africa and Asia, which more urgently need to use new technologies to ensure their food security (Zaidi *et al.*, 2019). There have been exciting developments already; in Bangladesh, *Bt* brinjal (aubergine, *Solanum melongena*) and late blight–resistant potato have been developed, while researchers at the Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia, and the 2Blades Foundation, USA, have

field-trialled a five-gene stem rust resistance gene stack (Luo *et al.*, 2021). It is unsustainable to continue feeding 20% of our wheat production to pathogens (**Fig. 1**). Epidemics are unpredictable, so we must strive for immunity across our cultivated lands. Also of importance is the need to ensure new resistances are deployed ethically and can benefit as many countries as possible. This could be achieved through public–private partnership and infrastructure building to create hubs in more advanced developing countries for the innovation and dissemination of improved crops to nearby less-developed countries (Zaidi *et al*., 2019).

Treating diversity panels as libraries of *R* genes means scientists and breeders have the necessary genetic resources at hand to work together to achieve the *R* gene machine: a scenario where bioinformatics pipelines facilitate a swift movement from phenotype to full-length candidate gene, followed by the gene's synthesis, testing and deployment in an appropriate stack (determined by the atlas algorithm). The ultimate goal would be to one day turn wheat into a non-host for its major pathogens.

Acknowledgements

The authors wish to thank the support of the UK Biotechnology and Biological Sciences Research Council (BBSRC) through the cross-institute strategic programme Designing Future Wheat (BB/P016855/1); the 2Blades Foundation, USA; a UKRI-BBSRC Norwich Research Park Biosciences Doctoral Training Partnership fellowship to ANH; a BBSRC/RAGT Industrial-Collaborative Award in Science and Engineering fellowship to DG; and a Monsanto's Beachell-Borlaug International Scholars Program fellowship (06-400258-12580) and John Innes Centre International Scholarship Scheme award to SG. We are grateful to Michiel van Slageren (Royal Botanic Garden, Kew, UK) for help with taxonomic classifications, Tobin Florio (http://www.flozbox.com/illu_sci.htm) for the artwork, Mark Luterbacher (John Innes Centre, UK) for help with budgeting the wheat *R* gene atlas, Simon Aspland (John Innes Centre) for help with *R* gene patent searches, Simon Krattinger for feedback on an earlier draft of the manuscript, and the following colleagues for discussions on *R* gene stewardship: Brian Steffenson (University of Minnesota, USA), Caixia Lan (Huazhong Agricultural University, China), Filippo Bassi (ICARDA, Morocco), Hannah Robinson (Intergrain, Australia), Moisés Burachik and Francisco Ayala (Bioceres, Argentina), Richard Summers and Ruth Bryant (RAGT, UK), Pierre Hucl (University of Saskatchewan, Canada), Roger Freedman (2Blades Foundation, USA), Ruth Wanyera (KALRO, Kenya), Simon Berry (Limagrain, UK), Sridhar Bhavani (CIMMYT, Mexico), Viktor Korzun (KWS, Germany), Nick Bird (KWS, UK), Willem Boshoff and Zakkie Pretorius (University of the Free State, South Africa), and Xiue Wang (Nanjing Agricultural University, China).

References

Acevedo-Garcia, J. *et al.* (2017) 'mlo-based powdery mildew resistance in hexaploid bread wheat generated by a non-transgenic TILLING approach', *Plant Biotechnology Journal*, 15(3), pp. 367–378. doi: 10.1111/pbi.12631.

Aggarwal, R. *et al.* (2014) 'Avirulence effector discovery in a plant galling and plant parasitic arthropod, the Hessian fly (Mayetiola destructor)', *PLoS ONE*. Edited by P. Castagnone-Sereno. Public Library of Science, 9(6), p. e100958. doi: 10.1371/journal.pone.0100958.

Allen-Sader, C. *et al.* (2019) 'An early warning system to predict and mitigate wheat rust diseases in Ethiopia', *Environmental Research Letters*. IOP Publishing, 14(11). doi: 10.1088/1748-9326/ab4034.

Alonge, M. *et al.* (2020) 'Major Impacts of Widespread Structural Variation on Gene Expression and Crop Improvement in Tomato', *Cell*. Cell Press, 182(1), pp. 145-161.e23. doi: 10.1016/j.cell.2020.05.021.

Annaluru, N. *et al.* (2014) 'Total synthesis of a functional designer eukaryotic chromosome', *Science*. American Association for the Advancement of Science, 344(6179), pp. 55–58. doi: 10.1126/science.1249252.

Arora, S. *et al.* (2017) 'Genome-Wide Association Study of Grain Architecture in Wild Wheat *Aegilops tauschii*', *Frontiers in Plant Science*, 8(May), pp. 1–13. doi: 10.3389/fpls.2017.00886.

Arora, S. *et al.* (2019) 'Resistance gene cloning from a wild crop relative by sequence capture and association genetics', *Nature Biotechnology*, pp. 139–143. doi: 10.1038/s41587-018-0007-9.

Arraiano, L. S., Brading, P. A. and Brown, J. K. M. (2001) 'A detached seedling leaf technique to study resistance to *Mycosphaerella graminicola* (anamorph *Septoria tritici*) in wheat', *Plant Pathology*, 50(3), pp. 339–346. doi: 10.1046/j.1365-3059.2001.00570.x.

Arranz-Otaegui, A. *et al.* (2018) 'Archaeobotanical evidence reveals the origins of bread 14,400 years ago in northeastern Jordan', *Proceedings of the National Academy of Sciences*. National Academy of Sciences, 115(31), pp. 7925–7930. doi: 10.1073/pnas.1801071115.

Audano, P. A., Ravishankar, S. and Vannberg, F. O. (2018) 'Mapping-free variant calling using haplotype reconstruction from k-mer frequencies', *Bioinformatics*. Oxford University Press, 34(10), pp. 1659–1665. doi: 10.1093/bioinformatics/btx753.

Aylward, J. *et al.* (2017) 'A plant pathology perspective of fungal genome sequencing', *IMA Fungus*. International Mycological Association, 8(1), pp. 1–15. doi: 10.5598/imafungus.2017.08.01.01.

Bai, D. and Knott, D. R. (1992) 'Suppression of rust resistance in bread wheat (*Triticum aestivum* L.) by D-genome chromosomes', *Genome*. NRC Research Press Ottawa, Canada , 35(2), pp. 276–282. doi: 10.1139/g92-043.

Balfourier, F. *et al.* (2019) 'Worldwide phylogeography and history of wheat genetic diversity', *Science Advances*, 5(5). doi: 10.1126/sciadv.aav0536.

Bartoli, C. and Roux, F. (2017) 'Genome-Wide Association Studies In Plant Pathosystems: Toward an Ecological Genomics Approach', *Frontiers in Plant Science*. Frontiers, 8, p. 763. doi: 10.3389/fpls.2017.00763.

Bhavani, S. *et al.* (2019) 'Progress in breeding for resistance to Ug99 and other races of the stem rust fungus in CIMMYT wheat germplasm', *Frontiers of Agricultural Science and Engineering*, 6(3), pp. 210–224. doi: 10.15302/J-FASE-2019268.

Bomblies, K. and Weigel, D. (2007) 'Hybrid necrosis: Autoimmunity as a potential gene-flow barrier in plant species', *Nature Reviews Genetics*, pp. 382–393. doi: 10.1038/nrg2082.

Bonman, J. M. *et al.* (2015) 'Genetic Diversity among Wheat Accessions from the USDA National Small Grains Collection', *Crop Science*. Crop Science Society of America, 55(3), pp. 1243–1253. doi: 10.2135/cropsci2014.09.0621.

Bouton, C. *et al.* (2018) 'Foxtail mosaic virus: A Viral Vector for Protein Expression in Cereals 1[CC-BY] the original idea', *Plant Physiology ®*, 177, pp. 1352–1367. doi: 10.1104/pp.17.01679.

Brun, H. *et al.* (2010) 'Quantitative resistance increases the durability of qualitative resistance to Leptosphaeria maculans in Brassica napus', *New Phytologist*, 185(1), pp. 285–299. doi: 10.1111/j.1469-8137.2009.03049.x.

Bueno-Sancho, V. *et al.* (2017) 'Field Pathogenomics: An Advanced Tool for Wheat Rust Surveillance', in *Methods in molecular biology (Clifton, N.J.)*, pp. 13–28. doi: 10.1007/978-1-4939-7249-4_2.

Bueno-Sancho, V. *et al.* (2019) *K-PIE: using K-means algorithm for Percentage Infection symptoms Estimation*.

Carvajal-Yepes, M. *et al.* (2019) 'A global surveillance system for crop diseases: Global preparedness minimizes the risk to food supplies', *Science*. American Association for the Advancement of Science, 364(6447), pp. 1237–1239. doi: 10.1126/science.aaw1572.

Chen, W., Liu, T. and Gao, L. (2013) 'Suppression of stripe rust and leaf rust resistances in interspecific crosses of wheat', *Euphytica*, 192(3), pp. 339–346. doi: 10.1007/s10681-012-0854-2.

Cook, D. E. *et al.* (2012) 'Copy number variation of multiple genes at Rhg1 mediates nematode resistance in soybean', *Science*. doi: 10.1126/science.1228746.

Cook, D. E., Mesarich, C. H. and Thomma, B. P. H. J. (2015) 'Understanding Plant Immunity as a Surveillance System to Detect Invasion', *Annual Review of Phytopathology*. Annual

Reviews, 53(1), pp. 541–563. doi: 10.1146/annurev-phyto-080614-120114.
Debernardi, J. M. *et al.* (2020) 'A chimera including a GROWTH-REGULATING FACTOR (GRF) and its cofactor GRF', *bioRxiv*. Cold Spring Harbor Laboratory, p. 2020.08.23.263905. doi: 10.1101/2020.08.23.263905.
Ellis, J. G. *et al.* (2014) 'The past, present and future of breeding rust resistant wheat', *Frontiers in Plant Science*. Frontiers, 5, p. 641. doi: 10.3389/fpls.2014.00641.
Elshire, R. J. *et al.* (2011) 'A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species', *PLoS ONE*. doi: 10.1371/journal.pone.0019379.
Eom, J. S. *et al.* (2019) 'Diagnostic kit for rice blight resistance', *Nature Biotechnology*. Nature Publishing Group, 37(11), pp. 1372–1379. doi: 10.1038/s41587-019-0268-y.
Erickson, A. W. (1945) *McFadden's Hope: Fighting plant breeders win battle for bread.* Minneapolis, Minnesota, USA: Field Notes Crop Reporting Service.
Feuillet, C., Langridge, P. and Waugh, R. (2008) 'Cereal breeding takes a walk on the wild side', *Trends in Genetics*. doi: 10.1016/j.tig.2007.11.001.
Flor, H. H. (1971) 'Current Status of the Gene-For-Gene Concept', *Annual Review of Phytopathology*. Annual Reviews, 9(1), pp. 275–296. doi: 10.1146/annurev.py.09.090171.001423.
Foley, J. A. *et al.* (2011) 'Solutions for a cultivated planet', *Nature*, 478(7369), pp. 337–342. doi: 10.1038/nature10452.
Friebe, B. *et al.* (1994) 'Compensation indices of radiation-induced wheat-*Agropyron elongatum* translocations conferring resistance to leaf rust and stem rust', *Crop Science*. doi: 10.2135/cropsci1994.0011183X003400020018x.
Fu, D. *et al.* (2009) 'A kinase-START gene confers temperature-dependent resistance to wheat stripe rust', *Science*. American Association for the Advancement of Science, 323(5919), pp. 1357–1360. doi: 10.1126/science.1166289.
Gao, Y. *et al.* (2016) 'Validation of Genome-Wide Association Studies as a Tool to Identify Virulence Factors in *Parastagonospora nodorum*', *Phytopathology*, 106(10), pp. 1177–1185. doi: 10.1094/PHYTO-02-16-0113-FI.
Gardiner, L. J. *et al.* (2015) 'A genome-wide survey of DNA methylation in hexaploid wheat', *Genome Biology*. Genome Biology, 16(1), pp. 1–15. doi: 10.1186/s13059-015-0838-3.
Ghislain, M. *et al.* (2019) 'Stacking three late blight resistance genes from wild species directly into African highland potato varieties confers complete field resistance to local blight races', *Plant Biotechnology Journal*. Blackwell Publishing Ltd, 17(6), pp. 1119–1129. doi: 10.1111/pbi.13042.
Ghosh, S. *et al.* (2018) 'Speed breeding in growth chambers and glasshouses for crop breeding and model plant research', *Nature Protocols*. doi: 10.1038/s41596-018-0072-z.
Gibson, D. G. *et al.* (2008) 'Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome', *Science*. American Association for the Advancement of Science, 319(5867), pp. 1215–1220. doi: 10.1126/science.1151721.
Gilly, A. *et al.* (2019) 'Very low-depth whole-genome sequencing in complex trait association studies', *Bioinformatics*. doi: 10.1093/bioinformatics/bty1032.
Guttieri, M. J. (2020) '*Ms3* dominant genetic male sterility for wheat improvement with molecular breeding', *Crop Science*. doi: 10.1002/csc2.20091.
Harlan, J. . R. . and de Wet, J. . M. . J. . (1971) 'Toward a Rational Classification of Cultivated Plants', *Taxon*, 20(4), pp. 509–517.
Harper, A. L. *et al.* (2012) 'Associative transcriptomics of traits in the polyploid crop species Brassica napus', *Nature Biotechnology*, 30(8), pp. 798–802. doi: 10.1038/nbt.2302.
Harris, C. J. *et al.* (2013) 'Stepwise artificial evolution of a plant disease resistance gene', *Proceedings of the National Academy of Sciences of the United States of America*, 110(52), pp. 21189–21194. doi: 10.1073/pnas.1311134110.
Hatta, M. A. M. *et al.* (2020) 'Extensive genetic variation at the sr22 wheat stem rust resistance gene locus in the grasses revealed through evolutionary genomics and functional analyses', *Molecular Plant-Microbe Interactions*, 33(11), pp. 1286–1298. doi: 10.1094/MPMI-01-20-0018-R.
Hayta, S. *et al.* (2019) 'An efficient and reproducible Agrobacterium-mediated transformation

method for hexaploid wheat (Triticum aestivum L.)', *Plant Methods*. BioMed Central Ltd., 15(1). doi: 10.1186/s13007-019-0503-z.
He, X. *et al.* (2020) 'Disease resistance evaluation of elite CIMMYT wheat lines containing the coupled *Fhb1* and *Sr2* genes', *Plant Disease*. Scientific Societies. doi: 10.1094/pdis-02-20-0369-re.
Hickey, L. T. *et al.* (2019) 'Breeding crops to feed 10 billion', *Nature Biotechnology*. doi: 10.1038/s41587-019-0152-9.
Hiebert, C. W. *et al.* (2020) 'Stem rust resistance in wheat is suppressed by a subunit of the mediator complex', *Nature Communications*. Nature Research, 11(1), pp. 1–10. doi: 10.1038/s41467-020-14937-2.
Hillocks, R. J. (2012) 'Farming with fewer pesticides: EU pesticide review and resulting challenges for UK agriculture', *Crop Protection*. Elsevier Ltd, 31(1), pp. 85–93. doi: 10.1016/j.cropro.2011.08.008.
Hovmøller, M. S., Walter, S. and Justesen, A. F. (2010) 'Escalating threat of wheat rusts.', *Science (New York, N.Y.)*. American Association for the Advancement of Science, 329(5990), p. 369. doi: 10.1126/science.1194925.
Huang, X. *et al.* (2009) 'High-throughput genotyping by whole-genome resequencing', *Genome Research*, 19(6), pp. 1068–1076. doi: 10.1101/gr.089516.108.
Hubbard, A. *et al.* (2015) 'Field pathogenomics reveals the emergence of a diverse wheat yellow rust population', *Genome Biology*, 16(1), p. 23. doi: 10.1186/s13059-015-0590-8.
Hurni, S. *et al.* (2014) 'The powdery mildew resistance gene Pm8 derived from rye is suppressed by its wheat ortholog Pm3', *Plant Journal*, 79(6), pp. 904–913. doi: 10.1111/tpj.12593.
Innes, R. L. and Kerber, E. R. (1994) 'Resistance to wheat leaf rust and stem rust in *Triticum tauschii* and inheritance in hexaploid wheat of resistance transferred from *T. tauschii*', *Genome*. Canadian Science Publishing, 37(5), pp. 813–822. doi: 10.1139/g94-116.
International Wheat Genome Sequencing Consortium (IWGSC), T. I. W. G. S. C. *et al.* (2018) 'Shifting the limits in wheat research and breeding using a fully annotated reference genome.', *Science (New York, N.Y.)*. American Association for the Advancement of Science, 361(6403), p. eaar7191. doi: 10.1126/science.aar7191.
Islam, M. T. *et al.* (2016) 'Emergence of wheat blast in Bangladesh was caused by a South American lineage of Magnaporthe oryzae.', *BMC biology*, 14(1), p. 84. doi: 10.1186/s12915-016-0309-7.
Jones, J. D. G. and Dangl, J. L. (2006) 'The plant immune system', *Nature*. Nature Publishing Group, pp. 323–329. doi: 10.1038/nature05286.
Jupe, F. *et al.* (2013) 'Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations', *Plant Journal*, 76(3), pp. 530–544. doi: 10.1111/tpj.12307.
Kangara, N. *et al.* (2020) 'Mutagenesis of *Puccinia graminis* f. sp. *tritici* and Selection of Gain-of-Virulence Mutants', *Frontiers in Plant Science*. Frontiers, 11, p. 1. doi: 10.3389/fpls.2020.570180.
Kawakatsu, T. *et al.* (2016) 'Epigenomic Diversity in a Global Collection of Arabidopsis thaliana Accessions', *Cell*. Cell Press, 166(2), pp. 492–505. doi: 10.1016/j.cell.2016.06.044.
Khan, A. W. *et al.* (2020) 'Super-Pangenome by Integrating the Wild Side of a Species for Accelerated Crop Improvement', *Trends in Plant Science*. Elsevier Inc., 25(2), pp. 148–158. doi: 10.1016/j.tplants.2019.10.012.
Kourelis, J. and Van Der Hoorn, R. A. L. (2018) 'Defended to the nines: 25 years of resistance gene cloning identifies nine mechanisms for R protein function', *Plant Cell*. American Society of Plant Biologists, pp. 285–299. doi: 10.1105/tpc.17.00579.
Krattinger, S. G. *et al.* (2009) 'A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat', *Science*. American Association for the Advancement of Science, 323(5919), pp. 1360–1363. doi: 10.1126/science.1166453.
Kuang, H. *et al.* (2004) 'Multiple genetic processes result in heterogeneous rates of evolution within the major cluster disease resistance genes in lettuce', *Plant Cell*. doi: 10.1105/tpc.104.025502.

915 Kuang, H. *et al.* (2006) 'The disease resistance gene *Dm3* is infrequent in natural populations of
916 *Lactuca serriola* due to deletions and frequent gene conversions at the RGC2 locus', *Plant*
917 *Journal*. Plant J, 47(1), pp. 38–48. doi: 10.1111/j.1365-313X.2006.02755.x.
918 De La Concepcion, J. C. *et al.* (2019) 'Protein engineering expands the effector recognition
919 profile of a rice NLR immune receptor', *eLife*, 8. doi: 10.7554/eLife.47713.
920 Liu, Y. *et al.* (2020) 'Pan-Genome of Wild and Cultivated Soybeans', *Cell*. Elsevier Inc., 182(1),
921 pp. 162-176.e13. doi: 10.1016/j.cell.2020.05.023.
922 Liu, Z. *et al.* (2012) 'The cysteine rich necrotrophic effector SnTox1 produced by *Stagonospora*
923 *nodorum* triggers susceptibility of wheat lines harboring *Snn1*', *PLoS Pathogens*. Edited by
924 B. Tyler. Public Library of Science, 8(1), p. e1002467. doi: 10.1371/journal.ppat.1002467.
925 Louwaars, N. P. (2018) 'Plant breeding and diversity: A troubled relationship?', *Euphytica*.
926 Springer Netherlands, p. 114. doi: 10.1007/s10681-018-2192-5.
927 Lowe, K. *et al.* (2016) 'Morphogenic regulators Baby boom and Wuschel improve monocot
928 transformation', *Plant Cell*. American Society of Plant Biologists, 28(9), pp. 1998–2015. doi:
929 10.1105/tpc.16.00124.
930 Lukaszewski, A. J. (2000) 'Manipulation of the 1RS.1BL translocation in wheat by induced
931 homoeologous recombination', *Crop Science*. doi: 10.2135/cropsci2000.401216x.
932 Luo, M. *et al.* (2021) 'A five-transgene cassette confers broad-spectrum resistance to a fungal
933 rust pathogen in wheat', *Nature Biotechnology*. doi: 10.1038/s41587-020-00770-x.
934 Ben M'Barek, S. *et al.* (2015) 'FPLC and liquid-chromatography mass spectrometry identify
935 candidate necrosis-inducing proteins from culture filtrates of the fungal wheat pathogen
936 Zymoseptoria tritici', *Fungal Genetics and Biology*. Elsevier Inc., 79, pp. 54–62. doi:
937 10.1016/j.fgb.2015.03.015.
938 Ben M'Barek, S. *et al.* (2020) 'Improved control of septoria tritici blotch in durum wheat using
939 cultivar mixtures', *Plant Pathology*. Blackwell Publishing Ltd, p. ppa.13247. doi:
940 10.1111/ppa.13247.
941 Marchal, C. *et al.* (2018) 'BED-domain containing immune receptors confer diverse resistance
942 spectra to yellow rust', *Nature Plants*. doi: 10.1101/299651.
943 McDonald, B. A. (2014) 'Using dynamic diversity to achieve durable disease resistance in
944 agricultural ecosystems', *Tropical Plant Pathology*, 39(3), pp. 191–196. doi: 10.1590/S1982-
945 56762014000300001.
946 McDonald, B. A. and Linde, C. (2002) 'Pathogen Population Genetics, Evolutionary Potential,
947 and Durable Resistance', *Annual Review of Phytopathology*, 40(1), pp. 349–379. doi:
948 10.1146/annurev.phyto.40.120501.101443.
949 McIntosh, R. A. *et al.* (2011) 'Rye-derived powdery mildew resistance gene Pm8 in wheat is
950 suppressed by the Pm3 locus', *Theoretical and Applied Genetics*. Springer Verlag, 123(3),
951 pp. 359–367. doi: 10.1007/s00122-011-1589-5.
952 Mizuno, N. *et al.* (2010) 'Hypersensitive Response-Like Reaction Is Associated with Hybrid
953 Necrosis in Interspecific Crosses between Tetraploid Wheat and *Aegilops tauschii* Coss',
954 *PLoS ONE*. Edited by E. Newbigin. Public Library of Science, 5(6), p. e11326. doi:
955 10.1371/journal.pone.0011326.
956 Moore, J. W. *et al.* (2015) 'A recently evolved hexose transporter variant confers resistance to
957 multiple pathogens in wheat', *Nature Genetics*. Nature Publishing Group, 47(12), pp. 1494–
958 1498. doi: 10.1038/ng.3439.
959 Niu, Z. *et al.* (2014) 'Development and characterization of wheat lines carrying stem rust
960 resistance gene *Sr43* derived from *Thinopyrum ponticum*', *Theoretical and Applied Genetics*.
961 Springer Verlag, 127(4), pp. 969–980. doi: 10.1007/s00122-014-2272-4.
962 Olson, E. L. *et al.* (2010) 'Development of wheat lines having a small introgressed segment
963 carrying stem rust resistance gene Sr22', *Crop Science*, 50(5), pp. 1823–1830. doi:
964 10.2135/cropsci2009.11.0652.
965 Omrane, S. *et al.* (2017) 'Plasticity of the MFS1 Promoter Leads to Multidrug Resistance in the
966 Wheat Pathogen Zymoseptoria tritici', *mSphere*. Edited by A. P. Mitchell. American Society
967 for Microbiology Journals, 2(5), pp. e00393-17. doi: 10.1128/mSphere.00393-17.
968 Pan, C. L. (1940) ' A Genetic Study of Mature Plant Resistance in Spring Wheat to Black Stem
969 Rust, Puccinia Graminis Tritici, and Reaction to Black Chaff, Bacterium Translucens, Var.

970 Undulosum 1 ', *Agronomy Journal*. doi: 10.2134/agronj1940.00021962003200020003x.
971 Periyannan, S. *et al.* (2017) 'An overview of genetic rust resistance: From broad to specific
972 mechanisms', *PLoS Pathogens*, 13(7), pp. 1–6. doi: 10.1371/journal.ppat.1006380.
973 Petersen, L. K. and Stowers, R. S. (2011) 'A Gateway Multisite recombination cloning toolkit',
974 *PLoS ONE*. PLoS One, 6(9). doi: 10.1371/journal.pone.0024531.
975 Pont, C. *et al.* (2019) 'Tracing the ancestry of modern bread wheats', *Nature Genetics*. Springer
976 US, 51, pp. 905–911. doi: 10.1038/s41588-019-0393-z.
977 Poore, J. and Nemecek, T. (2018) 'Reducing food ' s environmental impacts through producers
978 and consumers', 992(June), pp. 987–992.
979 Praz, C. R. *et al.* (2017) 'AvrPm2 encodes an RNase-like avirulence effector which is conserved
980 in the two different specialized forms of wheat and rye powdery mildew fungus', *New*
981 *Phytologist*. Blackwell Publishing Ltd, 213(3), pp. 1301–1314. doi: 10.1111/nph.14372.
982 Radhakrishnan, G. V. *et al.* (2019) 'MARPLE, a point-of-care, strain-level disease diagnostics
983 and surveillance tool for complex fungal pathogens', *BMC Biology*. BioMed Central Ltd.,
984 17(1), pp. 65–65. doi: 10.1186/s12915-019-0684-y.
985 Rahman, A. *et al.* (2018) 'Association mapping from sequencing reads using k-mers', *eLife*, 7.
986 doi: 10.7554/eLife.32920.
987 Randhawa, M. S. *et al.* (2018) 'Interactions among genes *Sr2*/*Yr30*, *Lr34*/*Yr18*/*Sr57* and *Lr68*
988 confer enhanced adult plant resistance to rust diseases in common wheat (*Triticum aestivum*
989 L.) line "Arula"', *Australian Journal of Crop Science*, 12(06), pp. 1023–1033. doi:
990 10.21475/ajcs.18.12.06.PNE1305.
991 Richardson, T. *et al.* (2014) 'Efficient Agrobacterium transformation of elite wheat germplasm
992 without selection', *Plant Cell, Tissue and Organ Culture*. Kluwer Academic Publishers,
993 119(3), pp. 647–659. doi: 10.1007/s11240-014-0564-7.
994 Saari, E. E. and Wilcoxson, R. D. (1974) 'Plant Disease Situation of High-Yielding Dwarf
995 Wheats in Asia and Africa', *Annual Review of Phytopathology*, (12), pp. 49–68.
996 Saintenac, C. *et al.* (2018) 'Wheat receptor-kinase-like protein Stb6 controls gene-for-gene
997 resistance to fungal pathogen *Zymoseptoria tritici*', *Nature Genetics*. Nature Publishing
998 Group, 50(3), pp. 368–374. doi: 10.1038/s41588-018-0051-x.
999 Salcedo, A. *et al.* (2017) 'Variation in the *AvrSr35* gene determines *Sr35* resistance against
1000 wheat stem rust race Ug99', *Science*, 358(6370), pp. 1604–1606. doi:
1001 10.1126/science.aao7294.
1002 Sánchez-Vallet, A. *et al.* (2018) 'Nature's genetic screens: Using genome-wide association
1003 studies for effector discovery', *Molecular Plant Pathology*, 19(1), pp. 3–6. doi:
1004 10.1111/mpp.12592.
1005 Saur, I. M. L. *et al.* (2019) 'A cell death assay in barley and wheat protoplasts for identification
1006 and validation of matching pathogen AVR effector and plant NLR immune receptors', *Plant*
1007 *Methods*. BioMed Central Ltd., 15(1), p. 118. doi: 10.1186/s13007-019-0502-0.
1008 Savary, S. *et al.* (2019) 'The global burden of pathogens and pests on major food crops', *Nature*
1009 *Ecology and Evolution*, 3(3), pp. 430–439. doi: 10.1038/s41559-018-0793-y.
1010 Segretin, M. E. *et al.* (2014) 'Single amino acid mutations in the potato immune receptor R3a
1011 expand response to Phytophthora effectors', *Molecular Plant-Microbe Interactions*. doi:
1012 10.1094/MPMI-02-14-0040-R.
1013 Seifbarghi, S. *et al.* (2009) 'Studies on the host range of Septoria species on cereals and some
1014 wild grasses in Iran', *Phytopathologia Mediterranea*, 48(3), pp. 422–429. doi:
1015 10.14601/PHYTOPATHOL_MEDITERR-2940.
1016 Song, W. Y. *et al.* (1995) 'A receptor kinase-like protein encoded by the rice disease resistance
1017 gene, Xa21', *Science*, 270(5243), p. 1804. doi: 10.1126/science.270.5243.1804.
1018 Stam, R. and McDonald, B. A. (2018) 'When resistance gene pyramids are not durable-the role
1019 of pathogen diversity', *Molecular Plant Pathology*, 19(3), pp. 521–524. doi:
1020 10.1111/mpp.12636.
1021 Steuernagel, B. *et al.* (2016) 'Rapid cloning of disease-resistance genes in plants using
1022 mutagenesis and sequence capture', *Nature Biotechnology*. Nature Publishing Group, 34(6),
1023 pp. 652–655. doi: 10.1038/nbt.3543.
1024 Stewart, E. L. *et al.* (2016) 'An improved method for measuring quantitative resistance to the

wheat pathogen Zymoseptoria tritici using high throughput automated image analysis', *Phytopathology*, 106, pp. 782–788. doi: 10.1094/PHYTO-01-16-0018-R.
Stewart, E. L. and McDonald, B. A. (2014) 'Measuring Quantitative Virulence in the Wheat Pathogen Zymoseptoria tritici Using High-Throughput Automated Image Analysis', *Phytopathology*, 104(9), pp. 985–992. doi: 10.1094/PHYTO-11-13-0328-R.
Stotz, H. U. *et al.* (2014) 'Effector-triggered defence against apoplastic fungal pathogens', *Trends in Plant Science*. Elsevier Current Trends, pp. 491–500. doi: 10.1016/j.tplants.2014.04.009.
Talas, F. *et al.* (2016) 'Genome-Wide Association Study Identifies Novel Candidate Genes for Aggressiveness, Deoxynivalenol Production, and Azole Sensitivity in Natural Field Populations of Fusarium graminearum', *Molecular Plant-Microbe Interactions*. Molecular Plant-Microbe Interactions, 29(5), pp. 417–430. doi: 10.1094/MPMI-09-15-0218-R.
Togninalli, M. *et al.* (2018) 'The AraGWAS Catalog: A curated and standardized Arabidopsis thaliana GWAS catalog', *Nucleic Acids Research*. Oxford University Press, 46(D1), pp. D1150–D1156. doi: 10.1093/nar/gkx954.
Uauy, C., Wulff, B. B. H. and Dubcovsky, J. (2017) 'Combining Traditional Mutagenesis with New High-Throughput Sequencing and Genome Editing to Reveal Hidden Variation in Polyploid Wheat', *Annual Review of Genetics*. doi: 10.1146/annurev-genet-120116-024533.
Voichek, Y. and Weigel, D. (2020) 'Identifying genetic variants underlying phenotypic variation in plants without complete genomes', *Nature Genetics*. Springer US, 52(5), pp. 534–540. doi: 10.1038/s41588-020-0612-7.
Wang, Y. *et al.* (2014) 'Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew', *Nature Biotechnology*. Nature Publishing Group, 32(9), pp. 947–951. doi: 10.1038/nbt.2969.
Watson, A. *et al.* (2018) 'Speed breeding is a powerful tool to accelerate crop research and breeding', *Nature Plants*. Nature Publishing Group, 4(1), pp. 23–29. doi: 10.1038/s41477-017-0083-8.
Weber, E. *et al.* (2011) 'A Modular Cloning System for Standardized Assembly of Multigene Constructs', *PLoS ONE*. Edited by J. Peccoud. Public Library of Science, 6(2), p. e16765. doi: 10.1371/journal.pone.0016765.
Wellings, C. R. *et al.* (2009) 'The development and application of near-isogenic lines for monitoring cereal rust pathogens', in R., M. (ed.) *Proceedings of Oral Papers and Posters, 2009 Technical Workshop,*. Cd. Obregón, Sonora, Mexico: BGRI, p. pp 77–87.
Van de Weyer, A. L. *et al.* (2019) 'A Species-Wide Inventory of NLR Genes and Alleles in Arabidopsis thaliana', *Cell*. Cell Press, 178(5), pp. 1260-1272.e14. doi: 10.1016/j.cell.2019.07.038.
Wingen, L. U. *et al.* (2014) 'Establishing the A. E. Watkins landrace cultivar collection as a resource for systematic gene discovery in bread wheat', *Theor Appl Genet*, 127, pp. 1831–1842. doi: 10.1007/s00122-014-2344-5.
Wolfe, M. S. *et al.* (1992) 'Barley mildew in Europe: population biology and host resistance', *Euphytica*, 63(1–2), pp. 125–139. doi: 10.1007/BF00023918.
Wu, J. Q. *et al.* (2017) 'Comparative Genomics Integrated with Association Analysis Identifies Candidate Effector Genes Corresponding to Lr20 in Phenotype-Paired Puccinia triticina Isolates from Australia', *Frontiers in Plant Science*. Frontiers, 8, p. 148. doi: 10.3389/fpls.2017.00148.
Wulft, B. B. H. and Moscou, M. J. (2014) 'Strategies for transferring resistance into wheat: From wide crosses to GM cassettes', *Frontiers in Plant Science*. Frontiers Media SA, p. 692. doi: 10.3389/fpls.2014.00692.
Xing, L. *et al.* (2018) '*Pm21* from *Haynaldia villosa* Encodes a CC-NBS-LRR Protein Conferring Powdery Mildew Resistance in Wheat', *Molecular Plant*. Cell Press, pp. 874–878. doi: 10.1016/j.molp.2018.02.013.
Xu, P. *et al.* (2017) 'Automatic Wheat Leaf Rust Detection and Grading Diagnosis via Embedded Image Processing System', in *Procedia Computer Science*. doi: 10.1016/j.procs.2017.03.177.
Yates, S. *et al.* (2019) 'Precision Phenotyping Reveals Novel Loci for Quantitative Resistance to

Septoria Tritici Blotch', *Plant Phenomics*. American Association for the Advancement of Science (AAAS), 2019, pp. 1–11. doi: 10.34133/2019/3285904.
Yoshida, K. *et al.* (2009) 'Association Genetics Reveals Three Novel Avirulence Genes from the Rice Blast Fungal Pathogen *Magnaporthe oryzae*', *The Plant Cell*, 21, pp. 1573–1591. doi: 10.1105/tpc.109.066324.
Zaidi, S. S.-A. *et al.* (2019) 'New plant breeding technologies for food security', *Science*. doi: 10.1126/science.aav6316.
Zhang, D. *et al.* (2019) 'Wheat powdery mildew resistance gene *Pm64* derived from wild emmer (*Triticum turgidum* var. *dicoccoides*) is tightly linked in repulsion with stripe rust resistance gene *Yr5*', *Crop Journal*. Crop Science Society of China/ Institute of Crop Sciences, 7(6), pp. 761–770. doi: 10.1016/j.cj.2019.03.003.
Zhang, J. *et al.* (2020) 'How Target-Sequence Enrichment and Sequencing (TEnSeq) Pipelines Have Catalyzed Resistance Gene Cloning in the Wheat-Rust Pathosystem', *Frontiers in Plant Science*, 11(May), pp. 1–13. doi: 10.3389/fpls.2020.00678.
Zhang, X. *et al.* (2016) 'Development and verification of wheat germplasm containing both Sr2 and Fhb1', *Molecular Breeding*. Springer Netherlands, 36(7), pp. 1–14. doi: 10.1007/s11032-016-0502-y.
Zhang, Z. *et al.* (2010) 'Mixed linear model approach adapted for genome-wide association studies', *Nature Genetics*, 42(4), pp. 355–360. doi: 10.1038/ng.546.
Zhong, Z. *et al.* (2017) 'A small secreted protein in *Zymoseptoria tritici* is responsible for avirulence on wheat cultivars carrying the *Stb6* resistance gene', *New Phytologist*, 214(2), pp. 619–631. doi: 10.1111/nph.14434.
Zhou, Y. *et al.* (2020) 'Convergence within divergence : insights of wheat adaptation from Triticum population sequencing'.
Zurayk, R. and Khalidi, R. (2011) *Food, farming, and freedom : sowing the Arab spring*. Just World Books.

Table S1. The 534 native and introgressed *R* genes designated in wheat.
Table S2. Cloned *R* genes from wheat and wild wheat progenitors.
Table S3. Cloned effectors corresponding to cloned wheat *R* genes.
Table S4. Estimated budget for *R* gene atlas.
Table S5. Details of patents associated with cloned *R* genes.

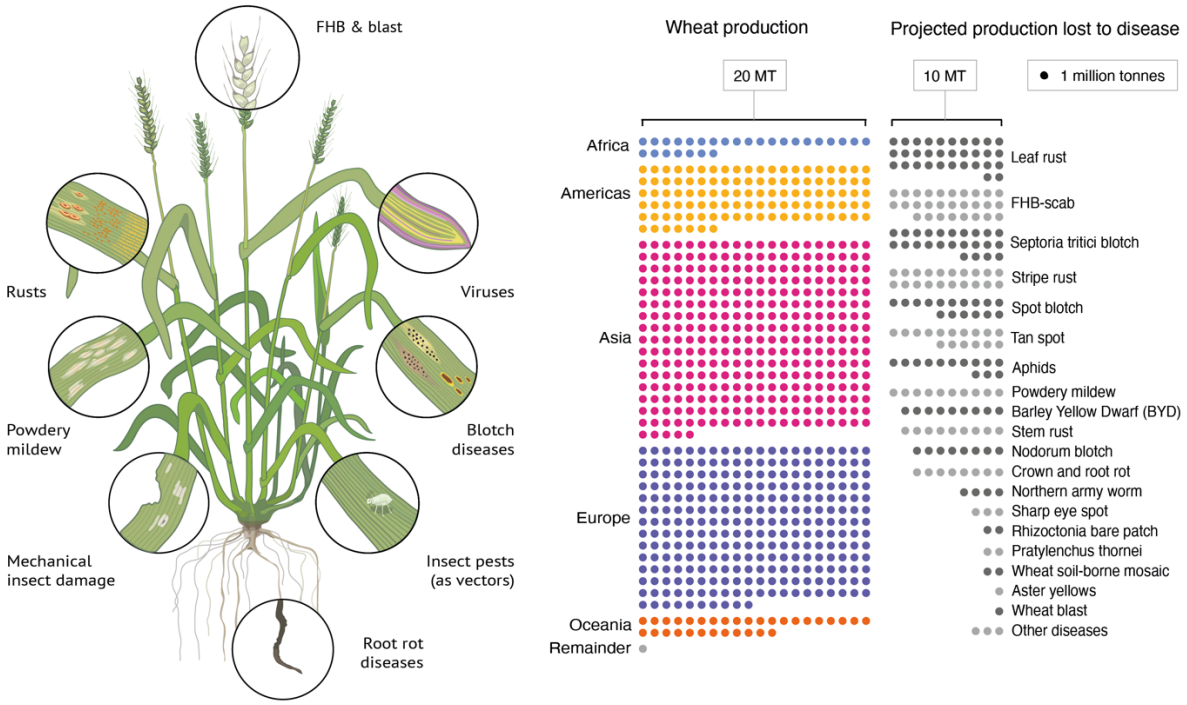

Figure 1: Major wheat diseases and their global impacts on wheat production. Wheat production statistics were taken from FAO (http://www.fao.org/faostat/en/#data/QC; 2017 data). Estimates of yield loss to pathogens were adapted from an expert-based assessment of crop health (Savary *et al.*, 2019).

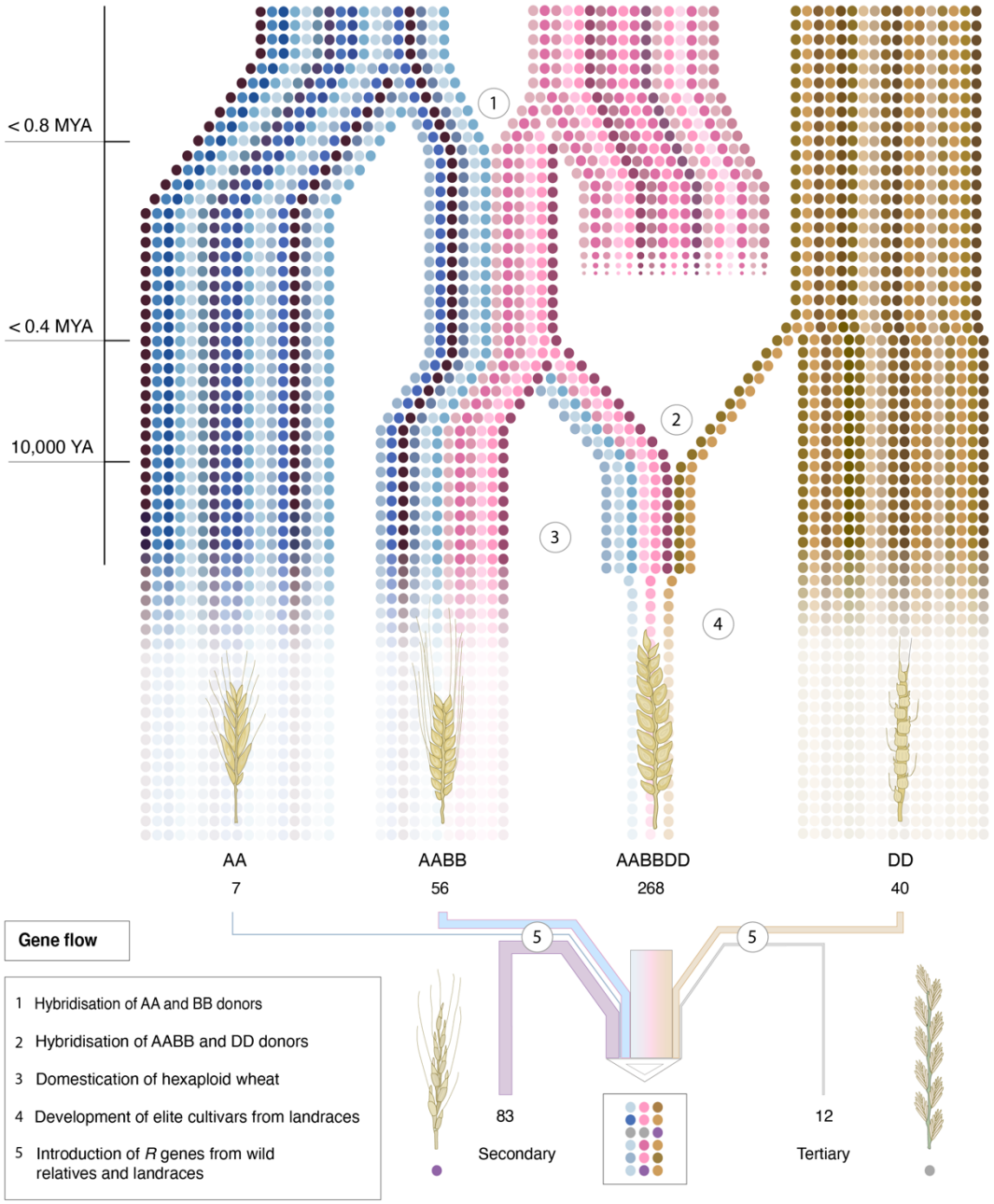

Figure 2. The diversity of *R* genes in hexaploid bread wheat has been influenced by genetic bottlenecks during its speciation and domestication, and through controlled inter-species crossing. (**1**) Hybridisation of *Triticum urartu* (AA) with a close relative of *Aegilops speltoides* (BB) to form allotetraploid emmer wheat, *Triticum turgidum* ssp. *diccocoides* (AABB) 0.8 million years ago (MYA). (**2**) Hybridisation of emmer wheat with *Aegilops tauschii* (DD) to produce hexaploid wheat (AABBDD), between 0.4 MYA (genetic evidence) and 10,000 years ago (archaeological evidence). (**3**) Domestication likely imposed a bottleneck, but there are no wild hexaploid wheat species so it is not possible to measure how much diversity was lost between bottlenecks 2 and 3. (**4**) Genetic bottleneck due to wheat breeding. Pangenome diversity is represented by coloured dots proportionate to the gene flow estimated by Zhou *et al*. (2020). The scope of this figure is the last ~1 million years of wheat's phylogenetic history. (**5**) In the last 100 years, inter-species crosses with the primary, secondary and tertiary gene pools have enriched the known native wheat resistance gene pool of 306 genes with 228 exogenous *R* genes (**Table S1**). The secondary gene pool defined here excludes *Aegilops tauschii.*

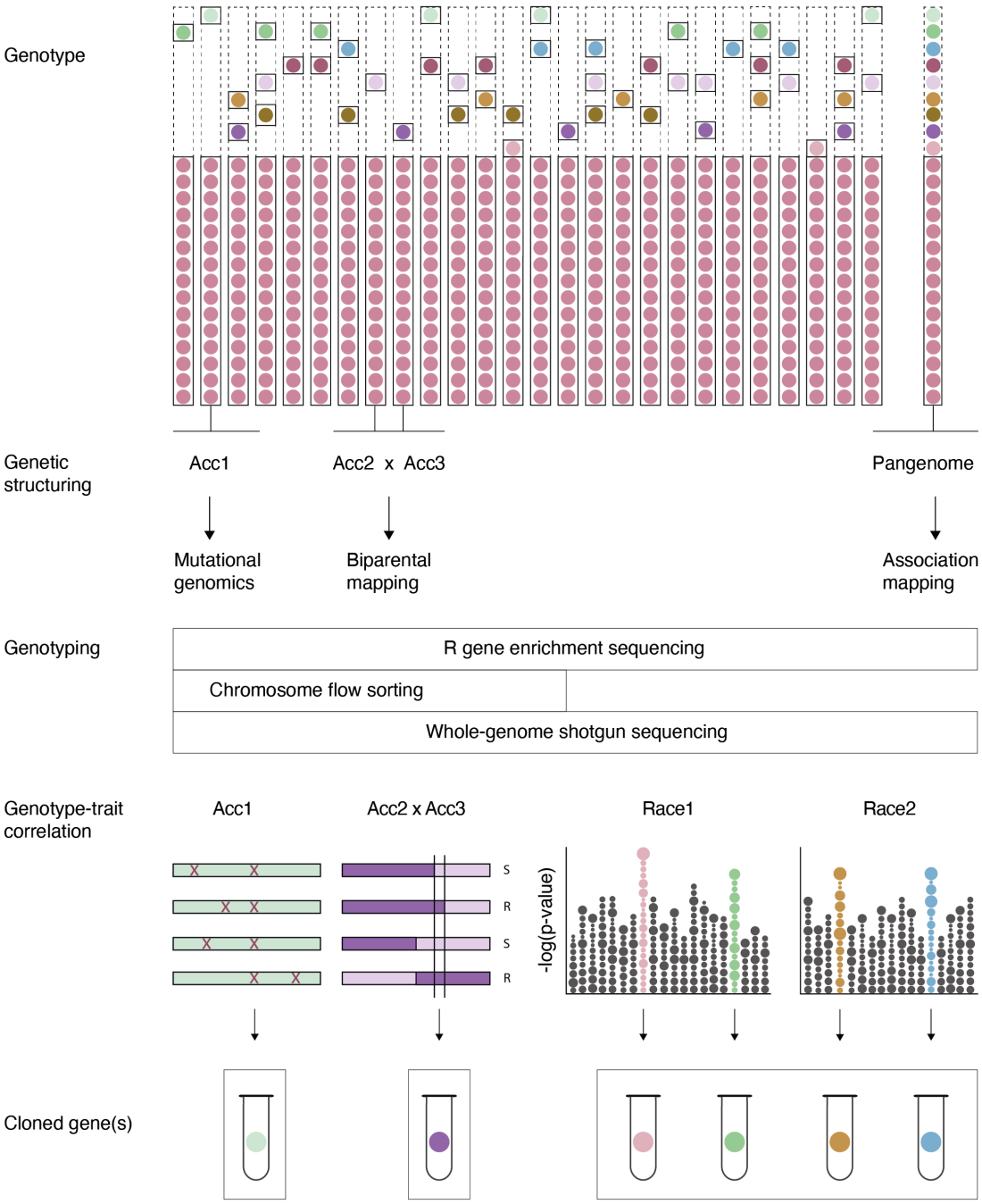

Figure 3. Map-based cloning and mutational genomics can be used to interrogate a narrow genetic base – that of the parent of the mutant population, or of two parents in a biparental mapping population. Association genetics, on the other hand, inspects a wide genetic base through the use of the pangenome variation available in a diversity panel. *R* gene enrichment or whole-genome shotgun sequencing are suitable genotyping approaches for all three types of genetically structured populations, but chromosome flow sorting may not be cost effective across diversity panels. In all three methods, correlations between the genotype and the trait are examined to identify candidate genes; however, only association mapping offers the possibility of cloning multiple genes from one structured population.

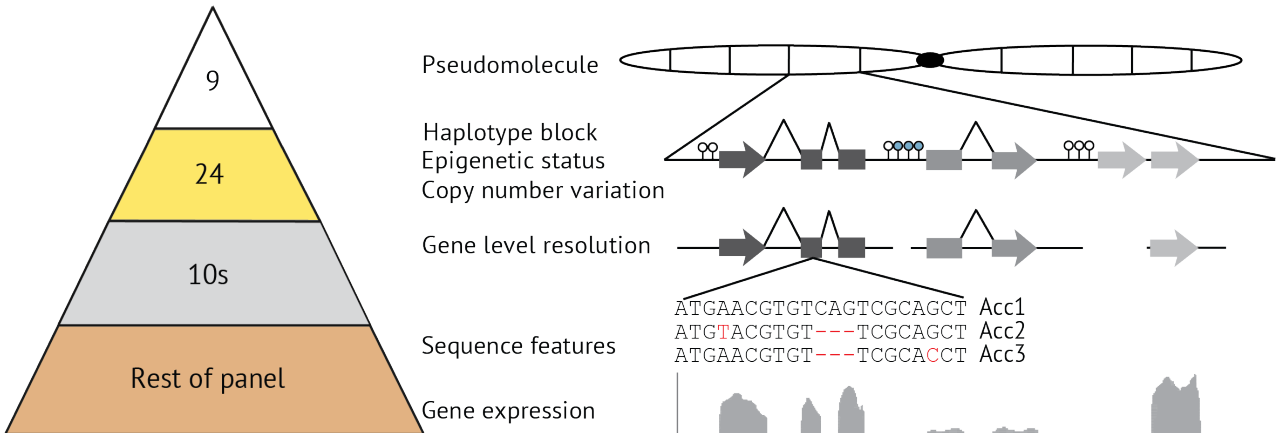

Figure 4. A tiered sequencing approach can be implemented to cost-effectively capture genetic diversity in panels. The tiers are gold+, gold, silver and bronze (top to bottom). Gold+-level assemblies are chromosome-scale pseudomolecules, providing the physical context of genes for key population sub-groups (nine in wheat; Wingen *et al.*, 2014). Gold-level assemblies should consist of scaffolds 100 kb to several Mb long, and can be used to define haplotype blocks in the genome. Silver-level assemblies require a scaffold size of only 5 to 15 kb to provide genes with their regulatory elements for a larger number of accessions. Finally, bronze-level assemblies can be generated from 5- to 10-fold Illumina reads and can enable the identification of SNPs and micro-indels across the majority of the diversity panel. The variation across these genotyping categories can be captured through mapping reads to higher-tier assemblies or by generating *k*-mers. Capturing the epigenetic variation (methylation of cytosines indicated by blue stalked circles, unmethylated sites by white stalked circles) of gold-level accessions along with RNA-seq data for the rest of the panel would provide insight into gene expression that could aid the identification of candidate genes underlying traits of interest.

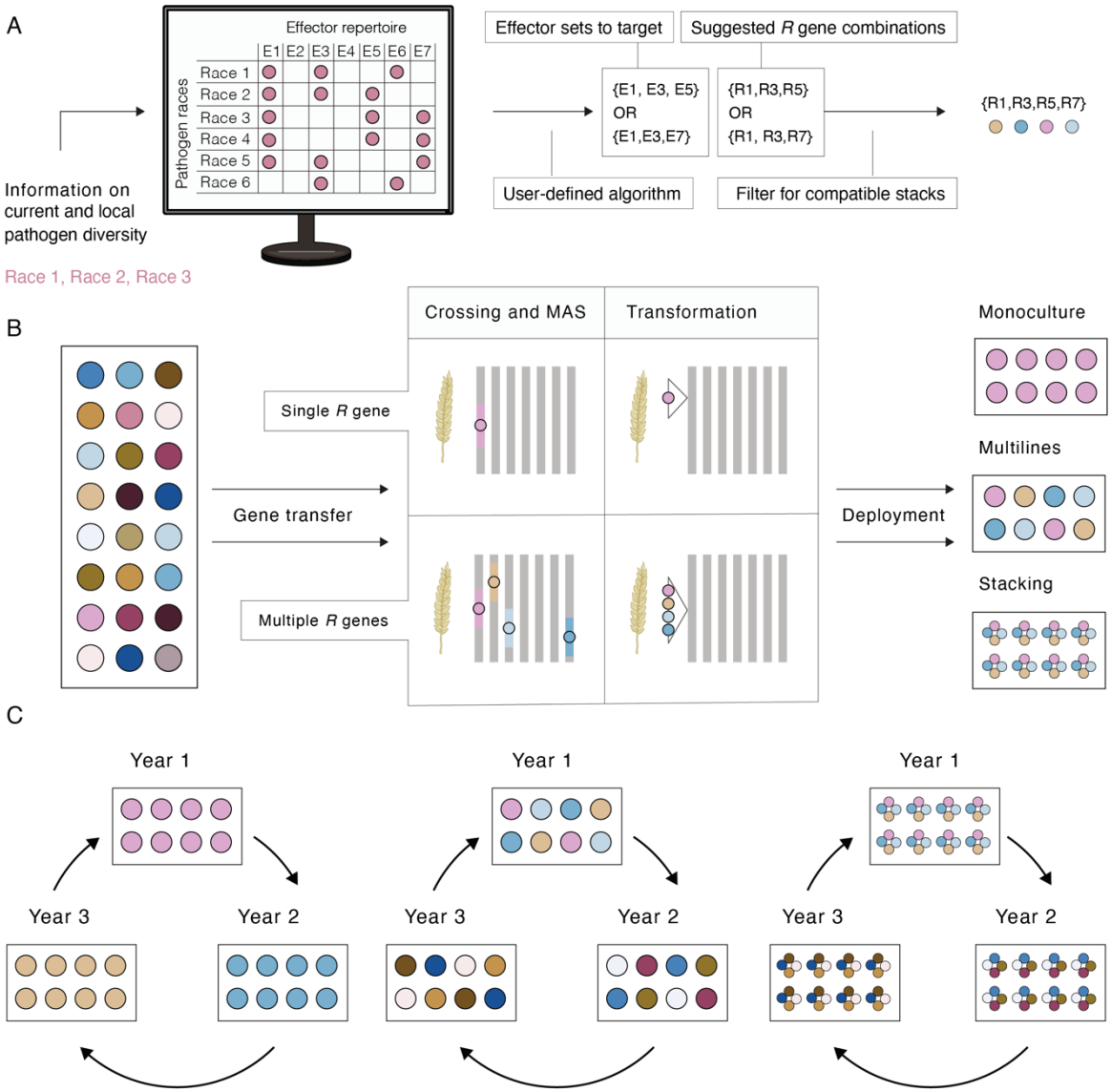

Figure 5. Using the *R* gene atlas and pathogen diversity information to determine appropriate stacks. **(A)** Information on current local pathogen diversity could be added to a database of races and their effector repertoires. A user-defined algorithm could then be used to calculate the largest number of effectors that could be targeted by the least number of *R* genes which maintain sufficient redundancy to maximise durable resistance; these *R* genes would be the best candidates for stacking. In the example shown, the suggested *R* gene stack targets 2-3 effectors per pathogen race. **(B)** Both traditional methods (crossing and marker-assisted selection (MAS)) and transformation can be employed to transfer single *R* genes or *R*-gene stacks from wild species or landraces into elite cultivars. Transformation has the advantage of speed, avoiding linkage drag and incorporating stacks at a single locus, while breeding can focus on combining *R* genes with partial resistance. Cultivars can then be deployed in monoculture or as multilines. **(C)** Whether cultivars are deployed with a single *R* gene or stack or as multilines, cultivars can be changed year on year to both combat the most destructive virulences in the pathogen and reduce the time that individual *R* genes are exposed to pathogens, reducing the chances of virulence to them evolving. This is a dynamic diversity approach (McDonald, 2014).

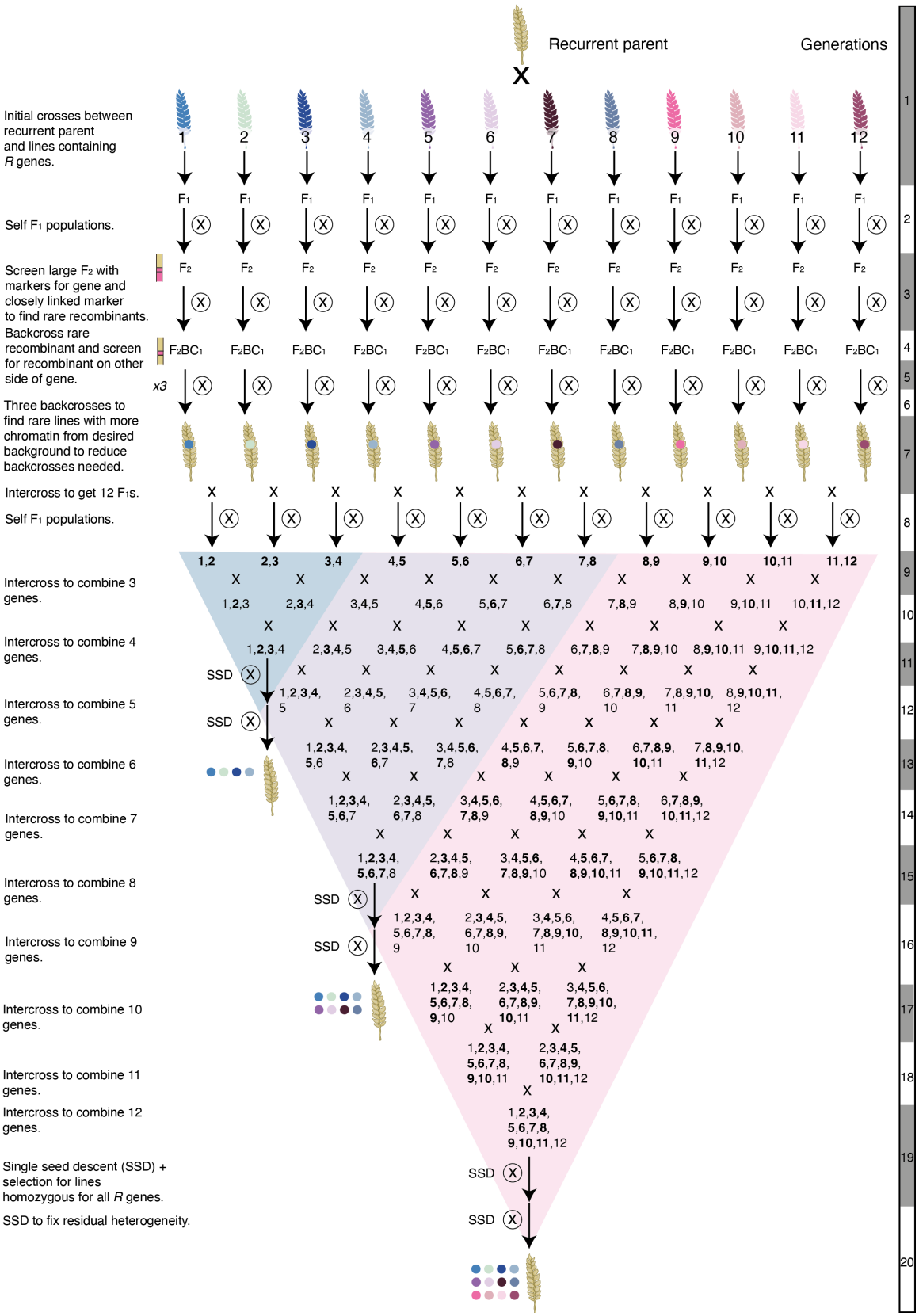

Figure 6. A scheme for mobilising 12 *R* genes into wheat through crossing and marker-assisted selection. Generations are indicated by the bar on the right. The blue, lilac and pink triangles indicate the crosses required to generate stacks of four, eight and 12 genes, respectively. Homozygous genes are emboldened. SSD = single seed descent.