Tansley review

Genome-wide identification of *EMBRYO-DEFECTIVE* (*EMB*) genes required for growth and development in Arabidopsis

Author for correspondence:
David W. Meinke
Tel: +1 405 744 5559
Email: meinke@okstate.edu

Received: *2 May 2019*
Accepted: *10 July 2019*

David W. Meinke

Department of Plant Biology, Ecology, and Evolution, Oklahoma State University, Stillwater, OK 74078, USA

Contents

New Phytologist (2020) **226:** 306–325
doi: 10.1111/nph.16071

Key words: Arabidopsis, chloroplast translation, embryo-defective mutants, essential genes, insertional mutagenesis, plant auxotrophs, seed development.

Summary

With the emergence of high-throughput methods in plant biology, the importance of long-term projects characterized by incremental advances involving multiple laboratories can sometimes be overlooked. Here, I highlight my 40-year effort to isolate and characterize the most common class of mutants encountered in Arabidopsis (*Arabidopsis thaliana*): those defective in embryo development. I present an updated dataset of 510 *EMBRYO-DEFECTIVE* (*EMB*) genes identified throughout the Arabidopsis community; include important details on 2200 *emb* mutants and 241 *pigment-defective embryo* (*pde*) mutants analyzed in my laboratory; provide curated datasets with key features and publication links for each *EMB* gene identified; revisit past estimates of 500–1000 total *EMB* genes in Arabidopsis; document 83 double mutant combinations reported to disrupt embryo development; emphasize the importance of following established nomenclature guidelines and acknowledging allele history in research publications; and consider how best to extend community-based curation and screening efforts to approach saturation for this diverse class of mutants in the future. Continued advances in identifying *EMB* genes and characterizing their loss-of-function mutant alleles are needed to understand genotype-to-phenotype relationships in Arabidopsis on a broad scale, and to document the contributions of large numbers of essential genes to plant growth and development.

I. Introduction

Mutant analysis has long been a cornerstone of Arabidopsis research (Meyerowitz & Somerville, 1994; Koornneef & Meinke, 2010). Initially, attention focused on small collections of mutants with distinctive phenotypes (late flowering; transparent seed coat; absence of leaf trichomes) identified through forward genetic screens (Rédei, 1970). Selected examples of these mutants were then used to construct the first genetic maps of Arabidopsis (Koornneef *et al.*, 1983). With one notable exception, the early work of Andreas Müller (1963), embryo-lethal mutants were typically overlooked in these screens, most likely because they were viewed as biologically uninformative, difficult to maintain, too numerous to manage, and not amenable to detailed analysis. In retrospect, each of these concerns was overstated. Here, I highlight dramatic advances made over the past 40 years in the isolation and characterization of Arabidopsis mutants defective in embryo development. I present a detailed, curated dataset of 510 *EMBRYO-DEFECTIVE* (*EMB*) genes identified in my laboratory and elsewhere; illustrate the diversity of protein functions, mutant phenotypes, and patterns of inheritance encountered; highlight examples of important questions that can be addressed with large datasets of *EMB* genes in hand; and provide a foundation for approaching saturation for *EMB* genes in the future. My objective is to facilitate the continued analysis of *EMB* genes in Arabidopsis by establishing a high-profile, curated resource of basic information, organized into updatable spreadsheets, on the extensive collection of genes and mutant alleles identified to date.

The defining feature of recessive embryo-defective mutants is that siliques of selfed heterozygotes segregate mutant seeds with abnormal embryos. Mutants with altered seed pigmentation but normal embryo morphology are typically excluded. The frequency of mutant seeds in heterozygous siliques is usually about 25% but can range from 5% to 50%, depending on the locus and allele involved. In some cases, both embryo and gametophyte development are disrupted. In other cases, mutant embryos can germinate to produce defective seedlings. By utilizing the term embryo defective in place of embryo lethal, a strategy I adopted 25 years ago, one avoids the pitfalls of defining embryo viability when continued growth in culture is possible, thereby enabling a broader spectrum of mutants to be included. Several thousand embryo-defective mutants have been isolated and characterized in my laboratory alone, with many additional mutants described elsewhere.

I devoted my research career to the analysis of embryo-defective mutants of Arabidopsis (Meinke & Sussex, 1979a,b; Meinke, 2013a). Because existing datasets (Muralla *et al.*, 2011) and databases (www.seedgenes.org) of *EMB* genes and their loss-of-function mutant alleles have become outdated, with resources unavailable for continued updates, I decided to undertake a final review of information and seed stocks that my laboratory have long maintained, and to make this information available to the community in a format that is readily accessible. I also wanted to underscore a number of concerns I have about recent publications on *EMB* genes. Most important is the widespread failure of authors to acknowledge existing *EMB* gene symbols, describe past work on mutant alleles, and cite relevant publications. This ongoing problem conflicts with longstanding standards for Arabidopsis nomenclature (Meinke & Koornneef, 1997). In most cases, one can check The Arabidopsis Information Resource (TAIR; www.arabidopsis.org) for a list of publications and gene symbols associated with a locus of interest, and determine whether a mutant defective in embryo development has already been described. Authors and reviewers need to be more vigilant in this regard. Other troubling issues encountered while reviewing the literature include erroneous classification of undeveloped ovules as aborted seeds; inconsistent descriptions of mutants altered in both embryo and gametophyte development; failure to evaluate seed phenotypes in siliques of heterozygotes when homozygous mutant plants are not encountered; and a reluctance to place new *EMB* genes and mutant alleles in the context of the entire collection.

Although my laboratory focused at times on specific mutants of interest (Schneider *et al.*, 1989; Meinke, 1992; Vernon & Meinke, 1994; Liu & Meinke, 1998; Nickle & Meinke, 1998; Muralla *et al.*, 2007), I also advocated casting a wide net to analyze large numbers of mutants defective in embryo development (Meinke, 1985; Castle *et al.*, 1993; McElver *et al.*, 2001; Tzafrir *et al.*, 2004; Meinke *et al.*, 2008; Muralla *et al.*, 2011). What is the value of maintaining a comprehensive, curated dataset of known *emb* mutants and cloned *EMB* genes? Because with such information available, it becomes possible to better understand the full spectrum of mutant phenotypes, altered protein functions, and patterns of inheritance associated with loss-of-function mutant alleles, and to identify common features of *EMB* gene products, including shared protein complexes, cellular processes, and targeted intracellular compartments. Such a dataset can also be used to clarify differences between embryo and gametophyte lethality, determine the nature and diversity of essential genes in a model plant, highlight essential gene products with unknown functions that merit further analysis, and resolve fundamental questions in plant biology, including the importance of chloroplast translation for seed development, the scarcity of auxotrophic mutants identified at the seedling stage, and the minimal gene set required for plant growth and development.

II. Historical perspective

The original screens for embryo-lethal mutants of Arabidopsis involved examining immature siliques of chimeric M_1 plants, derived from X-irradiation or chemical mutagenesis, for the presence of segregating mutant seeds (Meinke & Sussex, 1979a). The defining work of Müller (1963) in Gatersleben, Germany, described 60 mutants isolated in the Estland accession that exhibited a wide range of seed phenotypes, including embryo abortion, albino or pale green embryos, and inappropriate accumulation of anthocyanin in developing cotyledons, a distinctive class that Müller named the fusca phenotype. The Müller embryo test for detecting induced mutations was established and utilized for decades to assess the efficacy of mutagenic agents in plants (Rédei, 1982; Abramov *et al.*, 2006). Sadly, the only remnants of this early collection are the *fusca* mutants, which helped to define important modulators of plant development (Meinke *et al.*, 1994; Miséra *et al.*, 1994).

My original collection of embryo-lethal mutants was generated following ethyl methanesulfonate (EMS) seed mutagenesis in the Columbia accession (Meinke & Sussex, 1979b; Meinke, 1985). Two independent populations were established: one at Yale (1976–77) and the other at Oklahoma State University (1983–84). Thirty-three mutants were initially described; 23 are still available through ABRC (Supporting Information Dataset S1). Another 40 EMS and X-ray mutants were identified soon thereafter from M_2 populations generated by Joseph Ecker at the University of Pennsylvania (Meinke *et al.*, 1989). Most of these mutants also remain available to the community. Early studies with this original mutant collection included descriptions of interesting terminal phenotypes (Marsden & Meinke, 1985), defects in male gametophyte function (Meinke, 1982), storage product accumulation (Heath *et al.*, 1986), ultrastructure (Patton & Meinke, 1990) and growth of mutant embryos in culture (Baus *et al.*, 1986; Franzmann *et al.*, 1989; Schneider *et al.*, 1989).

When large-scale genetic screens of T-DNA insertion lines for mutant phenotypes were first organized by Ken Feldmann at DuPont (Wilmington, DE) and later at the University of Arizona (Feldmann, 1991), my laboratory screened siliques from about half of the lines and identified 180 *emb* mutants that were subsequently analyzed in detail (Errampalli *et al.*, 1991; Castle *et al.*, 1993). The remaining lines were screened and specific mutants of interest characterized further by a collaborative group overseen by Robert Goldberg at UCLA (Yadegari *et al.*, 1994; Apuya *et al.*, 2001). Elsewhere, Martine Devic and colleagues in France began to screen the INRA collection of T-DNA lines (Devic *et al.*, 1996; Ortega *et al.*, 2002) and analyze *emb* mutants of interest (Albert *et al.*, 1999; Lahmy *et al.*, 2004). Other mutants identified through forward genetic screens and later cloned were reported by multiple laboratories, including *hydra* (Souter *et al.*, 2002), *edd* (Uwer *et al.*, 1998), *pasticcino* (Faure *et al.*, 1998), *sse1* (Lin *et al.*, 1999), *ssr16* (Tsugeki *et al.*, 1996) and *twn2* (Zhang & Somerville, 1997).

Gerd Jürgens and colleagues at the University of Tübingen pursued a different approach by undertaking a comprehensive screen of EMS mutagenized populations at the seedling stage for developmental abnormalities indicative of a defect in embryonic pattern formation (Mayer *et al.*, 1991). Among the handful of mutants with intriguing phenotypes and multiple alleles chosen for detailed analysis, several turned out to be allelic to mutants recovered during earlier screens in my laboratory, including *emb30/gnom, emb22/gurke*, and *emb40/fass*. Ultimately, the Jürgens collection provided valuable insights into a variety of biological processes in plants (Lukowitz *et al.*, 1996; Schrick *et al.*, 2000; De Smet *et al.*, 2010). One intriguing result of these studies was the realization that pattern mutants in Arabidopsis often have primary defects in basic cellular functions, not specialized regulatory proteins.

Large-scale screens for mutants altered in gametophyte development also uncovered mutants with overlapping defects in embryogenesis. Typically, these screens focused on identifying mutants with reduced transmission of a linked selectable marker (Procissi *et al.*, 2001; Lalanne *et al.*, 2004). One extensive collection included 130 candidate genes required for female gametophyte development, most represented by a single mutant allele and lacking confirmed gene identities (Pagnussat *et al.*, 2005). Other screens identified male gametophyte mutants (Boavida *et al.*, 2009), including some in the *quartet* background (Johnson *et al.*, 2004), which enables pollen tetrad analysis (Copenhaver *et al.*, 2000). An unusual group of mutants that produced heterozygous siliques with 50% defective seeds regardless of pollen genotype was also identified and studied in detail (Chaudhury *et al.*, 1997; Grossniklaus *et al.*, 1998; Choi *et al.*, 2002). One example of this class (*emb173/mea*) was originally found among the Feldmann lines and briefly characterized in my laboratory (Castle *et al.*, 1993). Overall, laboratories interested primarily in gametophyte development tended to focus less on embryo defects in these large collections, whereas laboratories interested in embryo development, including my own, often avoided mutants with high levels of gametophyte lethality, particularly on the female side. One major challenge in establishing datasets of *EMB* genes has therefore been to disentangle these overlapping phenotypes, attempt to define what constitutes an *EMB* gene, and determine why the loss of essential genes alters gametophyte development in some cases and embryo development in others (Meinke *et al.*, 2008; Muralla *et al.*, 2011).

III. Syngenta/SeedGenes project

The most extensive screen for embryo-defective mutants among T-DNA insertion lines was initiated in the mid-1990s and involved a long-term collaboration between my laboratory and David Patton and colleagues at Syngenta (originally named Ciba-Geigy, then Novartis) in North Carolina. From a corporate perspective, the rationale for this project was that some of the essential genes identified might uncover promising targets for novel herbicides. But it also represented a unique opportunity to increase the number of known *EMB* genes in Arabidopsis. More than 120 000 insertion lines were generated and screened for seed phenotypes at Syngenta, with putative mutants sent to my laboratory for further analysis (McElver *et al.*, 2001). The same population was also evaluated, to a lesser extent, for seedling lethal phenotypes (Budziszewski *et al.*, 2001). The vast majority of these lines were produced in the Columbia accession using a T-DNA vector that conferred Basta resistance. The *quartet* mutant background was used for some lines to facilitate parallel genetic screens. All of these lines were separate from the SAIL population, also produced by Syngenta, which was designed for reverse genetics and distributed throughout the community (Sessions *et al.*, 2002).

A total of 1740 *emb* mutants from the forward genetic screen were confirmed to be segregating for defective seeds (Table 1). Attention focused initially on classifying terminal embryo phenotypes, determining segregation ratios, and resolving tagging status. Tagged mutants exhibited tight genetic linkage between the T-DNA insert and *EMB* locus. Because the resistance marker in these lines was rarely silenced, unlike the kanamycin marker in lines produced at the Salk Institute (Alonso *et al.*, 2003), tagged mutants with a low ratio of resistant to sensitive progeny seedlings indicative of a single functional insert could be readily identified by the absence of wild-type (WT) plants among resistant transplants. By contrast, resistant transplants from untagged mutants with a single insert, and other mutants (both tagged and untagged) with additional inserts, contained a mixture of segregating and WT

Table 1 Collections of embryo-defective mutants analyzed in the Meinke laboratory.[a]

EMB gene symbols[b]	Mutants analyzed; maintained[c]	Mapped mutants[d]	Cloned genes[e]	Nature of mutant collection	Initial characterization	Supporting Information Dataset
1–71	73	50	15	Original EMS; X-ray[g]	1977–1990	S1–S3
72–181	92	63	16	Feldmann (DuPont) T-DNA[g]	1990–1998	S1–S3
201–295	88	31	9	Feldmann (Tucson) T-DNA[g]	1990–1998	S1–S3
1000–2759 2937–2999 3201–3235	1740[f]	(10)	164	Syngenta T-DNA[g,h]	1998–2004	S3–S6
2761–2820	132	(1)	48	Reverse genetic screens[i]	2004–2010	S3, S4
3002–3013	23	(1)	16	French (INRA) T-DNA[g,j]	2002–2004	S3, S4
3101–3147	52	(1)	28	Reverse genetic screens[k]	2008–2014	S3, S4
Totals	2200	144/123	296/264	EMS; X-ray; T-DNA; Transposon	1977–2014	S1–S6

[a] Refer to Supporting Information Table S1 for additional details on the total numbers of mutants and genes analyzed.

[b] Mutants were often assigned a gene symbol before allelism with other mutants was tested. The number of symbols assigned (mutants analyzed) is therefore greater than the number of *EMB* genes represented.

[c] Excludes 241 *pigment-defective embryo* (*pde*) mutants with albino or pale green embryos analyzed in the Meinke lab (Dataset S7), six mapped *fusca* mutants from the DuPont collection, and eight mapped *emb* mutants identified elsewhere.

[d] The number of mapped *EMB* genes is 123 after adjusting for 21 allelic mutants. Numbers in parentheses represent cloned (but not mapped) mutants allelic to mapped (but not cloned) mutants.

[e] The number of distinct *EMB* genes is 264 after adjusting for overlaps between collections.

[f] Includes 232 putatively tagged *emb* mutants with unknown gene identities (Dataset S5) and 1272 untagged and unresolved *emb* mutants not subjected to DNA isolation (Dataset S6). Refer to Table S3 for Syngenta T-DNA vector information.

[g] Mutants were recovered through forward genetic screens for defective seeds segregating in siliques of heterozygotes.

[h] These Syngenta lines differ from the well-known SAIL (Garlic) lines also produced by Syngenta.

[i] Emphasis was placed on evaluating candidate essentials and identifying additional mutant alleles throughout the collection.

[j] These mutants were contributed by INRA/Genoplante, analyzed in more detail, and incorporated into the overall collection.

[k] Emphasis was placed on predicted chloroplast-localized proteins and insertion mutants that segregated aborted seeds.

plants. When more than one insert appeared to be present based on the initial ratio of resistant to sensitive seedlings, progeny subfamilies were examined to determine whether any had lost extraneous inserts through random assortment. Low-ratio subfamilies identified were then evaluated further. Other mutants remained unresolved with respect to tagging status because a low-ratio subfamily was not readily identified. Among the mutants with resolved tagging status, 35% appeared to contain an insert responsible for the mutant phenotype (McElver *et al.*, 2001). Tissue from these tagged mutants was returned to Syngenta, where genomic sequences flanking the T-DNA insert were recovered and analyzed. Overall, 440 tagged mutants were subjected to molecular characterization. Confidence that the altered gene had been identified improved when flanking sequence was obtained from both sides of the insert, and was confirmed when heterozygous plants from two independent mutants thought to be disrupted in the same gene were crossed and found to be allelic.

The SeedGenes database and website (www.seedgenes.org) arose from a desire to document the expanding list of genes with essential functions during seed development. Project objectives included facilitating the public release of internal Syngenta data on the identities of *EMB* genes uncovered, providing a centralized location for data on known *EMB* genes and their mutant alleles, and establishing a foundation for identifying every Arabidopsis gene with an essential function during growth and development. With support from the NSF 2010 Program and database expertise provided by Allan Dickerman and colleagues at the Virginia Bioinformatics Institute, the SeedGenes database was first released in March, 2002 (Tzafrir *et al.*, 2003), regularly updated for several years (Tzafrir *et al.*, 2004; Meinke *et al.*, 2008), and finalized in December, 2010 (Meinke, 2013a). No data have been added or removed since. The final release contained information on 481 genes and 888 mutants, 60% of which were studied in my laboratory. Cloned *emb* mutants analyzed throughout the community and *pigment-defective embryo* (*pde*) mutants from the Syngenta collection were also described. Key features of the original website included a query page to search for genes and mutants of interest, a profile page for each gene that summarized relevant gene and mutant information, flanking sequence files and supporting genetic and phenotypic data for the Syngenta collection, and a tutorial segment that outlined procedures for mutant analysis.

In the absence of continued support for the SeedGenes database, my attention soon turned to preparing reviews on genes required for embryo and/or gametophyte development (Muralla *et al.*, 2011) and genes with either a dominant (Meinke, 2013b) or recessive, loss-of-function (Lloyd & Meinke, 2012) phenotype of any kind. As the SeedGenes dataset became further outdated, I faced a dilemma familiar to other investigators: how to keep a specialized database current. Because producing each SeedGenes update was labor-intensive, with no successor identified to perform this task in the future, a decision was made to discontinue the host server and focus instead on preserving key features of SeedGenes in the detailed spreadsheets included with this review. Although the original query functions have been lost, most of the data from the final release are still available through the archival version maintained as static web pages at www.seedgenes.org.

IV. Meinke laboratory collection

My primary objective in this review was to assemble critical information on *EMB* genes and their mutant alleles and to establish a foundation, consistent with the original vision of Arabidopsis as a model genetic organism, to approach saturation for this class in the future. A logical first step was to organize information and seed stocks for the 2200 *emb* mutants isolated and characterized in my own laboratory. An overview of these collections is presented in Table 1. Additional information is provided in Supporting Information Tables S1–S3, with curated details in Datasets S1 through S6. The vast majority of these mutants are available through the Arabidopsis stock centers, with ABRC stock numbers noted in the datasets.

The initial collection of 253 EMS, X-ray, and T-DNA insertion mutants (Meinke, 1994; Datasets S1, S2) was the subject of a large-scale mapping project designed 30 years ago to enhance the classical genetic map of Arabidopsis and facilitate the identification of multiple alleles (Patton *et al.*, 1991; Franzmann *et al.*, 1995). Recombination data with linked visible markers were obtained for 169 *emb* mutants from this collection. Overall, 144 mutants were assigned a position on the classical genetic map. This total excludes six mapped *fusca* mutants from the Feldmann collection, six mapped *embRG* mutants (R. Goldberg T-DNA lines), and one *emb* mutant (*ssr16*) from a transposon collection that was mapped but analyzed elsewhere (Tsugeki *et al.*, 1996). To search for evidence of allelism within this collection, crosses were performed between mutants that mapped within a 5 cM window (Franzmann *et al.*, 1995). Closely linked *EMB* loci that remain on the map are therefore likely to represent distinct genes. Another 17 mutants were assigned to a chromosome but not placed on the map. In addition, eight mutants appeared from linkage data to contain a chromosomal translocation, consistent with results obtained elsewhere for frequent chromosomal aberrations in T-DNA insertion lines (Tax & Vernon, 2001; Clark & Krysan, 2010). Over a period of 8 years, 123 *EMB* genes were positioned on the genetic map, with 11 genes represented by multiple alleles (Table S2). The frequency of multiple alleles identified within this limited collection (Franzmann *et al.*, 1995) appeared to be consistent with a preliminary estimate of 500 total *EMB* genes in Arabidopsis based on the relative frequencies of embryo defectives and other types of visible mutations in T-DNA populations (Feldmann, 1991; Castle *et al.*, 1993).

Subsequent forward genetic screens of Syngenta insertion lines resulted in the identification of 1740 additional *emb* mutants, including 224 disrupted in known genes (Table S1; Dataset S4), 232 labeled as putatively tagged based on genetic data (Dataset S5) but lacking information on the disrupted gene, and 12 tagged mutants with problematic *EMB* gene identities (Table S1; Dataset S4). Another 1272 *emb* mutants from the Syngenta collection (Dataset S6) were isolated and characterized but not evaluated further because they appeared to be either not tagged (802 mutants), possibly tagged (23 mutants), or unresolved with respect to tagging status (447 mutants). In addition, my laboratory analyzed 18 mutants from the French collection of insertion lines with established gene identities (Dataset S4) and 178 insertion mutants identified through reverse genetic screens (Dataset S4) focused on additional alleles and candidate essentials (128 mutants; Tzafrir *et al.*, 2004) and genes encoding chloroplast-localized proteins (50 mutants; Bryant *et al.*, 2011). The remaining 23 mutants (12 from the Syngenta screen, five French lines, and six from reverse screens) were excluded from the *EMB* dataset because of questionable gene identities (Dataset S4). When combined with the initial collection of 253 mutants described earlier, the final collection contains a total of 2200 *emb* mutants: 2016 isolated through forward genetic screens and 184 through reverse. This excludes 29 mutants from the initial EMS, X-ray, and T-DNA populations that were analyzed and assigned *emb* numbers (Meinke, 1994) but were not maintained (Dataset S1). Another 241 *pde* mutants from the Syngenta collection (68 tagged and included in SeedGenes; 33 putatively tagged, and 140 not tagged or unresolved) were analyzed briefly in my laboratory (Dataset S7) and remain available for further studies.

In order to address questions about the number of distinct *EMB* genes represented in this mutant collection, I recognized a core set of 354 *EMB* genes defined by 566 mutant alleles analyzed in detail (Fig. 1; Table 2; Dataset S3). This dataset includes 264 cloned *EMB* genes (192 with a mutant allele first identified through forward genetics and 72 reverse) and 90 mapped but not cloned *EMB* genes from the initial collection. As noted before, the mapped genes likely represent distinct loci when viewed in isolation but may overlap with existing collections of cloned (but not mapped) *EMB* genes,

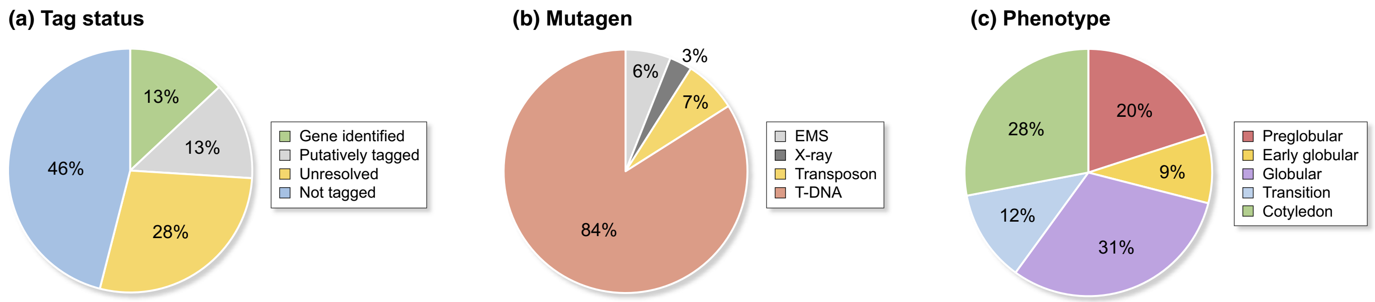

Fig. 1 Meinke laboratory collection of embryo-defective mutants of Arabidopsis. (a) Tagging status of 1740 mutants from the Syngenta forward genetic screen. (b) Mutagen used to generate the core collection of 566 mutants disrupted in 354 mapped and/or cloned *EMB* genes. (c) Terminal embryo phenotypes of the 566 mutants in the core collection.

Table 2 Origin of 566 mutant alleles of 354 *EMB* genes represented in the Meinke laboratory core collection.[a]

Mutant collection	Genetic screen	Mutants analyzed	Percent total
EMS; X-ray; Feldmann	Forward	147[b]	26.0
Syngenta	Forward	223[c]	39.4
French	Forward	18[d]	3.2
Additional alleles	Reverse	44	7.8
Candidate essentials	Reverse	84[d]	14.8
Chloroplast proteins	Reverse	50[d]	8.8

[a] Refer to Supporting Information Table S1 and Dataset S3 for additional details.

[b] Includes 46 cloned, mapped mutants; 98 mapped, not cloned mutants; *lec2* (cloned, not mapped); and two *embRG* alleles not found in Table 1. Excludes 79 mutants (not cloned or mapped) from the initial collection.

[c] Excludes *fus6-1S, emb2220* (gain-of-function allele), and 11 mutants with questionable gene identities.

[d] Excludes a total of 11 mutants with questionable gene identities (Dataset S4).

including those examined elsewhere. The extent of this overlap is relevant to ongoing questions about what percent saturation has already been reached. This issue was addressed through a pilot study focused on chromosomes 1 and 5 that involved integrating the physical and classical genetic maps of Arabidopsis and performing crosses between mutants from the two different subsets (mapped vs cloned) that localized to similar regions of the genome (Meinke *et al.*, 2009). The identities of 16 *EMB* genes from the mapped collection, including two of the original EMS mutants (Meinke & Sussex, 1979b), were revealed in this manner.

With respect to multiple alleles within the core set of 354 genes, 208 genes are represented by a single known allele, 104 by two alleles, 27 by three alleles, 11 by four alleles, and three (*BIO1, SUS2, EMB2016*) by five or six alleles (Table S2; Dataset S3). The remaining gene (*SUS1/DCL1*) has a total of nine mutant alleles, all isolated through forward genetic screens. Six of the 208 genes with a single allele within this collection have a weak (*pde*) allele that was also identified in the Syngenta screen. Forty-eight other genes have a different *emb* allele described in the literature, and eight more have a weak allele analyzed elsewhere. Thus, multiple alleles have been identified for a majority (59%) of the 354 genes in the core set. Many (55%) of the remaining genes represented by a single mutant allele are from the initial collection of mapped genes.

Characterizing the terminal phenotypes of defective embryos has been a challenge because of the sheer size of the collection, the number of years during which the analyses took place, and the inherent variability of mutant seeds produced by a single plant. One strategy involved assigning each mutant to a terminal phenotype class that illustrated the stage of development most often reached before seed desiccation. A second approach developed for the SeedGenes project involved examining 100 mutant seeds under a dissecting microscope, placing each embryo into a standardized phenotypic subclass (Meinke, 2013a), and saving images of representative embryos viewed with Nomarski optics. This information can still be accessed through the SeedGenes website, although it should be viewed with caution because of the different perspectives and levels of expertise of the dozens of investigators who performed the work.

V. SeedGenes review and literature curation

Several approaches were used here to generate an updated dataset of cloned *EMB* genes in Arabidopsis. The first involved determining what SeedGenes information was most important to maintain. Each published locus in SeedGenes (Tzafrir *et al.*, 2004; Meinke *et al.*, 2008; Bryant *et al.*, 2011; Muralla *et al.*, 2011; Lloyd & Meinke, 2012), including those removed from early database releases, was then used to identify relevant articles from TAIR locus pages. Attention focused on papers describing additional mutant alleles. Occasionally, publications appeared to conflict with existing *EMB* assignments. When the decision was made to exclude a questionable locus, selected information was maintained in a separate spreadsheet devoted to problematic genes.

PubMed queries were then performed to identify publications on *EMB* genes not included in SeedGenes. Emphasis was placed on articles published from 2010 to 2018, which followed the final SeedGenes update. No single method enabled detection of every locus or publication of interest. Querying for ‘Arabidopsis embryo lethal’ or simply ‘Arabidopsis embryo’ was quick but incomplete. A thorough query of PubMed abstracts was therefore performed using Arabidopsis and either embryo, embryos, embryogenesis, aborted, lethal, or lethality as key words. Relevant details on mutant alleles and protein function were then extracted from the full publication. Overall, several thousand abstracts were scanned and hundreds of publications evaluated. The final dataset includes more than 600 publications referenced with PubMed ID numbers.

Two recurrent questions needed to be addressed during literature curation: (1) how should an *EMB* gene and mutant phenotype be defined; and (2) what types of genetic and phenotypic data are required to support an *EMB* assignment? Both of these questions have been raised before but merit further consideration. Loss-of-function mutations in *EMB* genes disrupt embryo morphology and result in a recessive phenotype that can be detected visually, with or without a compound light microscope. Abnormal embryos segregate in selfed heterozygous siliques. Phenotypes revealed only through gene silencing are not included. Mutations that disrupt embryo pigmentation but not morphology, resulting in albino, pale green, or fusca phenotypes are also excluded. Likewise, mutations that alter embryo biochemistry or physiology (e.g. high chlorophyll fluorescence or modified levels of metabolites or storage products) without impacting morphology are excluded. Because a role for these excluded genes during embryo development has been documented, future datasets may need to include a wider spectrum of mutant phenotypes. By contrast, heritable changes in gene expression profiles; protein activity, accumulation or complex formation; and RNA modification without an associated morphological or biochemical defect are not considered mutant phenotypes (Lloyd & Meinke, 2012).

Problems often arose when dealing with mutations that resulted in subtle or transient phenotypes, especially those with reduced penetrance. Some of these genes were excluded from the current dataset, particularly when double mutants produced a more striking phenotype. Mutations that altered endosperm formation more than embryo development were included, although only a few examples were uncovered. Numerous mutations interfered with both embryo and gametophyte development (Muralla *et al.*, 2011). These were included in the dataset and noted in a column highlighting mutants with notable features, thereby facilitating future analyses. Mutations that primarily disrupted ovule or female gametophyte development rather than embryo development were generally excluded, even when the undeveloped ovules were incorrectly labeled as aborted seeds. Multiple examples were encountered of conflicting publications that reached different conclusions about whether a specific gene was required for embryo development. These were evaluated on a case-by-case basis. The most notable example involved two mutants thought to be deficient in a well-characterized auxin binding protein. Both the original allele (*abp1*), whose identity appeared to be confirmed through molecular complementation (Chen *et al.*, 2001), and a Syngenta mutant (*abp1-1S*) that my lab demonstrated was allelic to the original mutant, were later shown to be disrupted in an adjacent gene (*BSM*; Babiychuk *et al.*, 2011) that was responsible for embryo lethality (Dai *et al.*, 2015; Michalko *et al.*, 2015).

The SeedGenes database recognized two different levels of confidence that the gene responsible for the mutant phenotype had been identified. For genes labeled 'not confirmed', extensive co-segregation data were obtained, but neither molecular complementation nor additional mutant alleles were described. With confirmed genes, further support of gene identity was provided through molecular complementation, multiple alleles that failed to complement when heterozygotes were crossed, sector analysis following transposon excision, or additional genetic and biochemical studies. A similar approach was used when community mutants described in the literature were evaluated. Ultimately, the distinction between excluded mutants and those labeled as 'not confirmed' was subjective and should be revisited in the future. Large collections of mutants defective in female gametophyte development (Pagnussat *et al.*, 2005) and chloroplast-localized proteins (Savage *et al.*, 2013), along with other isolated cases of potential *emb* mutants encountered in the literature, were excluded from the dataset because of limited genetic evidence that the correct gene had been identified. Many Syngenta mutants with unresolved tagging status also remain to be evaluated.

VI. Comprehensive dataset of 510 *EMB* genes

Manual curation efforts described in the preceding section resulted in the establishment of a comprehensive dataset of 510 cloned *EMB* genes of Arabidopsis (Dataset S8), a significant advance beyond the earlier list of 396 genes (Muralla *et al.*, 2011). General features of the dataset are illustrated in Fig. 2. Mutants altered in 52% of these *EMB* genes were isolated and characterized in my laboratory. Overall, 51% of the gene identities were confirmed through molecular complementation. Another 30% were confirmed through the analysis of multiple alleles without complementation. Initially, all mutants were identified through forward genetic screens. Later, reverse genetics became more common, and was the only method used to identify embryo-defective mutants for 40% of the genes in the dataset. Weak alleles or transgenic plants with reduced expression have been described for 38% of the genes, thereby facilitating analysis of post-embryonic gene functions. Several classical loci with weak phenotypes identified through forward genetics, including *SERRATE* (Rédei & Hirono, 1964) and *FY* (Koornneef *et al.*, 1991), are included in the dataset because null alleles were found to be embryo-defective (Grigg *et al.*, 2005; Henderson *et al.*, 2005).

With 510 rows and 30 columns of data, the Excel spreadsheet presented in Dataset S8 represents a valuable resource of information on *EMB* genes and mutant alleles that is designed to be analyzed and further curated in the future. Information presented includes AGI locus numbers, gene symbols, confirmation status, terminal phenotype class, Meinke lab contributions, PubMed numbers for representative publications, and protein functions. Each publication is associated with further details on whether the mutants analyzed were new or known alleles, identified using forward or reverse genetics, and generated with chemical or insertional mutagenesis or ionizing radiation. The existence of weak alleles and plants with reduced expression is also noted.

One important feature of the dataset is the 'Notes' column, which highlights the varied ways that embryo-defective mutants can deviate from standard expectations. One third of the 510 *EMB* genes have at least one entry in this column. Overall, 23 different types of variations are noted. Common examples include subtle

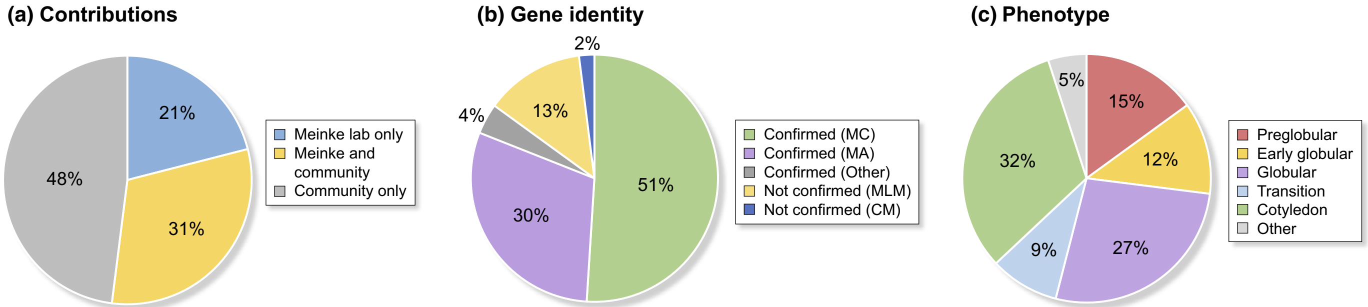

Fig. 2 Comprehensive dataset of 510 *EMB* genes of Arabidopsis. (a) Contributions of research groups to the analysis of embryo-defective mutants included in the dataset. (b) Confirmation of *EMB* gene identities. Genes were confirmed through molecular complementation (MC), the analysis of multiple alleles (MA) in the absence of molecular complementation, or some other experimental approach. Most of the genes lacking confirmation are represented by Meinke laboratory mutants (MLM) rather than community mutants (CM). Excludes 52 genes (Supporting Information Dataset S8E,F) originally designated as *EMB* but later removed from the main dataset because of uncertain gene identities or mutant phenotypes. (c) Terminal embryo phenotypes of strong mutant alleles.

embryo phenotypes that require a compound light microscope for detection (20 genes); deleterious effects on male or female gametophyte functions (49 and 31 genes, respectively), phenotypes initially described for seeds produced by homozygous mutant plants but also found or expected to segregate in heterozygous siliques (18 genes), and complex phenotypes or patterns of inheritance (14 genes). In eight cases, publications inconsistent with an *EMB* classification were identified but not viewed as sufficiently definitive to exclude the locus. Multiple examples are included of embryo rescue with a WT transgene driven by an embryo-specific promoter, an approach developed by Despres *et al.* (2001) to study post-embryonic functions of *EMB* genes.

The distribution of terminal embryo phenotypes observed for strong alleles of the 510 *EMB* genes (Fig. 2c) is similar to that noted for the core collection of 354 genes analyzed in my laboratory (Fig. 1c) and for previous collections of published *EMB* genes (Devic, 2008; Muralla *et al.*, 2011). Nevertheless, these phenotype ratios should be interpreted with caution because they are influenced by allele strength, functional redundancy, and in some instances, genetic background. Fortunately, terminal phenotypes are not usually influenced by growth conditions, silique location along the stem, or seed position within the silique. For mutants defective in gene products required throughout the life cycle, late terminal phenotypes may also reflect the ability of maternal tissues to partially rescue mutant embryos, and for heterozygous sporocytes to rescue mutant gametophytes (Muralla *et al.*, 2011). Because mutant embryos arrested before the cotyledon stage of development do not often survive seed desiccation, the 317 genes with terminal phenotypes at these early stages represent a logical focal point for studies on essential genes in the traditional sense.

VII. Diversity of EMB protein functions

Consistent with past reviews, the 510 *EMB* genes presented here perform a variety of essential cellular functions. Thirteen categories employed with earlier datasets (Bryant *et al.*, 2011; Muralla *et al.*, 2011; Lloyd & Meinke, 2012) were used to facilitate the sorting of *EMB* loci into functional groups. Several classes of proteins are particularly well represented (Table 3), including pentatricopeptide repeat (PPR) proteins, plastid-localized ribosomal proteins, proteins involved in RNA binding and modification, and chloroplast-localized proteins (Hsu *et al.*, 2010; Bryant *et al.*, 2011). In several cases, *EMB* genes encoding multiple components of essential protein complexes or cellular pathways have been identified. These include chloroplast protein import complexes, iron–sulfur clusters, chromosomal protein complexes (cohesin and condensin), and nuclear pore complexes. Proteins involved with microtubule and peroxisome assembly and histidine biosynthesis are also well-represented. The robustness of these collections is sufficient to begin addressing questions about what remaining members represent promising candidates for undiscovered *EMB* genes, and what factors enable knockouts of other members to complete embryo development. In the case of histidine biosynthesis, different members of a recent gene duplication (*HISN6A*; *HISN6B*) are either silenced or inactivated in different accessions, resulting in hybrid incompatibility when those accessions are crossed (Bikard *et al.*, 2009; Blevins *et al.*, 2017).

Table 3 Examples of *EMB* gene products with common, shared functions.

EMB genes identified[a]	Protein structure; function; complex	Representative publications
34	PPR protein	Lurin *et al*. (2004); Cushing *et al*. (2005); Barkan & Small (2014); Sun *et al*. (2018)
15	Chloroplast ribosomal protein	Romani *et al*. (2012); Parker *et al*. (2014)
14	Microtubule assembly and function	Tzafrir *et al*. (2002); Steinborn *et al*. (2002); Gillmor *et al*. (2016)
8	Iron–sulfur complex assembly	Bernard *et al*. (2013); Przybyla-Toscano *et al*. (2018); Lu (2018)
7	Peroxisome assembly and function	Rinaldi *et al*. (2017)
9	Cohesin; condensin complexes	Liu *et al*. (2002); Minina *et al*. (2017)
6	Nuclear pore complex	Parry (2014)
6	Minichromosome maintenance complex	Holding & Springer (2002); Herridge *et al*. (2014)
6	Chloroplast protein import motor (Heteromeric AAA-ATPase complex)	Kikuchi *et al*. (2018)
4	m^6A methyltransferase complex	Růžička *et al*. (2017)
5	Histidine biosynthesis	Muralla *et al*. (2007); Petersen *et al*. (2010)

[a]Refer to Supporting Information Dataset S8 for specific *EMB* genes and locus numbers.

Forty-six percent of the 510 *EMB* genes identified to date encode proteins assigned to one of three functional classes: metabolism, RNA binding or modification, and unknown. The relative frequencies of these and other classes varies with terminal embryo phenotype (Fig. 3) and protein localization within the cell. With increased emphasis on reverse genetic screens, the abundance of specific protein classes also reflects the number of investigators studying those functions. Among mutants with an early stage of arrest, proteins that function in metabolism, DNA synthesis and repair, or chromosome dynamics are relatively common. By contrast, transcription factors and proteins involved in signaling, cell structure, membrane function, and vesicle trafficking are more common among mutants that reach a cotyledon stage of development. About half of the mutants with a globular stage of arrest are defective in proteins that function in protein synthesis or RNA binding and modification, whereas only 8% of the EMB proteins associated with globular mutants function in chromosome dynamics, transcriptional regulation, signaling, cell structure, membrane function, or vesicle trafficking. The high frequency of proteins involved in translation or RNA binding among mutants arrested at the globular stage is a direct consequence of the requirement of chloroplast translation for embryo development in Arabidopsis (Parker *et al.*, 2014). By contrast, mutants defective in genes required for mitochondrial translation are often defective in gametophyte development, and null alleles of mutants defective in

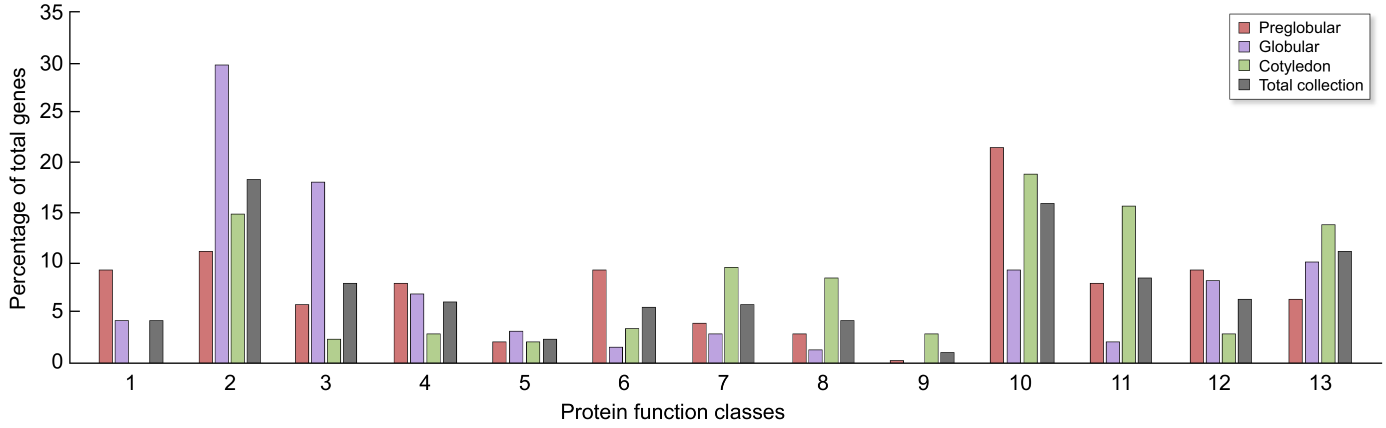

Fig. 3 Distribution of protein functions in the *EMB* dataset of Arabidopsis relative to terminal embryo phenotype. Protein classes: 1, DNA synthesis and repair; 2, RNA synthesis and modification; 3, Protein synthesis; 4, Protein modification and transport; 5, Protein degradation; 6, Chromosome dynamics; 7, Transcriptional regulation; 8, Signaling and regulatory proteins; 9, Energy and electron transport; 10, Metabolism; 11, Cell structure, membrane function, and vesicle trafficking; 12, Other; 13, Uncertain and unknown.

non-redundant proteins required for cytosolic translation are rarely encountered, most likely because they are lethal to both male and female gametophytes (Berg *et al.*, 2005). Understanding the relationship between protein function, genetic redundancy, allele strength, and mutant phenotype therefore remains a challenge for the future.

VIII. Atypical mutant alleles and patterns of inheritance

A number of fascinating variations were encountered while reviewing the literature on mutants defective in seed development. Some of these underscored the question of how *EMB* genes should be defined. On the paternal side, several cases are known where pollen genotype alone determines embryo phenotype. In other words, 50% defective seeds are produced when WT plants are crossed with pollen from heterozygotes. With knockouts of *CDC2A* (Iwakawa *et al.*, 2006; Nowack *et al.*, 2006), *FBL17* (Kim *et al.*, 2008; Gusti *et al.*, 2009), *DUO3* (Brownfield *et al.*, 2009), and *XRI1* (Dean *et al.*, 2009), the underlying cause is the formation of pollen with a single sperm cell. Aborted seeds are produced regardless of maternal genotype because double fertilization is compromised. These genes are included in the *EMB* dataset, despite the gametophytic nature of the initial defect, because limited fertilization is still possible. Knockouts of *KPL* exhibit preferential single fertilization (embryo or endosperm), even with two sperm cells present, highlighting a different underlying mechanism and a novel example of natural cis-antisense siRNA regulation involving an adjacent gene (Ron *et al.*, 2010).

With knockouts of *GEX2* (Mori *et al.*, 2014) and *GCS1/HAP2* (Mori *et al.*, 2006; Von Besser *et al.*, 2006), the primary defect involves sperm proteins required for cell recognition during fertilization. These genes are excluded from the *EMB* dataset because the protein functions are specific to fertilization, not embryo development. A third example (*DMP9*) is excluded because insertion mutants were not available and the phenotype was observed with RNAi alone (Takahashi *et al.*, 2018). A different mechanism of paternal inheritance is observed in knockouts of *SSP*, which encodes a receptor-associated protein kinase. In this case, transcripts produced in pollen are not translated until delivery to the zygote, where they function to activate the YODA protein kinase that regulates the first cell division (Bayer *et al.*, 2009). The *NIMNA* gene exhibits a weaker paternal effect in which heterozygous embryos with a paternal mutant allele are defective in embryo development, though to a lesser degree than homozygotes (Babu *et al.*, 2013).

The spectrum of mutants with gametophytic maternal effects is even more diverse (Ngo *et al.*, 2012; Armenta-Medina & Gillmor, 2019). The most striking examples are knockouts of five well-studied genes (*MEA*, *FIE1*, *FIS2*, *MSI1* and *DME*) required for Polycomb Repressive Complex 2 (PRC2) function during reproductive development. Heterozygous plants, regardless of pollen genotype, produce 50% defective seeds with a distinctive phenotype: endosperm over-proliferation and embryo arrest at the heart stage. A Syngenta/SeedGenes mutant (*emb2220; nf-ya1-D*) that was initially assigned to this class was later removed because it contains a novel gain-of-function mutation (Mu *et al.*, 2013). Paternal alleles of *MEA*, *FIS2*, and *FIE* were reported to be inactive early in seed development, consistent with the observed maternal effect (Kinoshita *et al.*, 1999; Luo *et al.*, 2000; Yadegari *et al.*, 2000). By contrast, the paternal allele of *MSI1* is expressed, but at levels insufficient to rescue mutant embryos lacking functional protein in female gametophytes (Leroy *et al.*, 2007). *DME* function is required for DNA hypomethylation in the central cell of the female gametophyte and maternal *MEA* function in the endosperm (Choi *et al.*, 2002).

A detailed consideration of the full spectrum of female gametophytic mutants with overlapping defects in embryo development is beyond the scope of this review. In some of these mutants, the underlying causes of abnormal development have been analyzed in detail (see Ngo *et al.*, 2012). In many cases, however, information is limited to the frequencies of aborted seeds and undeveloped ovules produced with different types of crosses. Frequently, this mixture appears to reflect incomplete penetrance in the loss of a gene product required for both embryo and female gametophyte development. However, the situation is complicated by differences in genetic redundancy, paternal allele contributions, spatial-temporal patterns of gene expression, and requirements for gene function in male gametophytes. Inconsistent descriptions of mutant phenotypes also are encountered. Despite these limitations, an initial collection of *EMB* genes assigned to this class can be obtained by referring to Note 11 in Dataset S8.

Several examples of *EMB* genes that exhibit haploinsufficiency, with a weak phenotype detected in heterozygotes, have been described in the literature (Du *et al.*, 2010; Mansilla *et al.*, 2015). In addition to providing insights into gene function at later stages of development, these examples raise the broader question of whether plants heterozygous for null alleles of other *EMB* genes exhibit a subtle phenotype that has escaped detection. Although haploinsufficiency is a frequent cause of disease phenotypes in humans (Bartha *et al.*, 2018), relatively few examples have been described in Arabidopsis (Meinke, 2013b). Several other strategies have been employed to study *EMB* gene functions at later stages of development (Candela *et al.*, 2011). The most common involve the analysis of weak mutant alleles, transgenic plants with reduced gene function, and embryo rescue with limited expression of a WT transgene. Multiple examples of these approaches can be found in the dataset. Embryo rescue in culture has also proven to be effective, particularly with embryos that reach a cotyledon stage. One important outcome of this approach was the identification of the leafy cotyledon phenotype in Arabidopsis (Meinke *et al.*, 1994). Mosaic analysis using marked genetic sectors has also been explored, although not fully utilized (Latvala-Kilby & Kilby, 2006; Muñoz-Nortes *et al.*, 2017). One of the most intriguing examples of post-embryonic development involves a mutant defective in a cell wall glycoprotein (EXTENSIN3) in which 'self-rescue' generates viable homozygotes (phenotypic revertants) that activate alternative gene expression programs (Saha *et al.*, 2013).

In some cases, novel mutant alleles have enabled the analysis of defects in embryo development that might otherwise have escaped detection. One class involves weak alleles of genes required for male gametophyte development, where null alleles are not transmitted through the pollen (see Note 6). Additional examples of such genes await the discovery of mutant alleles with suitable levels of residual gene function. Male gametophyte mutants have been described where defects in embryo development are observed when pollen lethality is circumvented by pollen-specific expression of a WT transgene (Note 7), or when RNAi or antisense methods are used to reduce gene function in embryos (Note 19). Limited pollination strategies, which reduce pollen competition, have also been used to obtain mutant embryos when pollen tube guidance is disrupted (H.J. Li *et al.*, 2011). One atypical allele involves an insertion mutant that produces a truncated protein (GEX1) of sufficient length to rescue the gametophytic lethality observed in true null alleles (Alandete-Saez *et al.*, 2011). In another case, a truncated protein encoded by a mutant allele exhibits a dominant negative effect on embryo development not found with other null alleles (Yu *et al.*, 2016). A different mechanism of allelic variation has been described for a single locus that produces two transcripts (short and long) encoding different isoforms of a geranylgeranyl diphosphate synthase required for isoprenoid biosynthesis (Ruiz-Sola *et al.*, 2016). Null mutations that disrupt only the product of the long transcript, which is targeted to plastids, exhibit defects in seedling pigmentation, whereas those that disrupt the product of the short transcript, which performs a different function in the cytosol, exhibit embryo lethality. Mutant phenotypes can also vary depending on genetic background (Parker *et al.*, 2014; Blevins *et al.*, 2017), a phenomenon frequently encountered in maize genetics.

When null alleles are not available, alternative methods can be used to determine whether gene function is required for embryo development. One promising approach involves TALEN and/or CRISPR-Cas9-targeted gene inactivation (Moussu *et al.*, 2017; Shin *et al.*, 2018). This strategy should become even more widespread in the future. Antagonistic peptide technology is more limited in scope but has been used to study a peptide signaling pathway that functions during embryo development (Song *et al.*, 2013). Another strategy involves combining an artificial miRNA construct with a weak insertion allele to generate 'two allele' embryos with an enhanced mutant phenotype (Park *et al.*, 2018). With further advances in these and other technologies, it should be possible to generate the types of mutant alleles required for detailed analysis.

IX. Double mutants defective in embryo development

Some protein functions required for embryo development will be missed in genetic screens for mutant phenotypes because they are encoded by multiple genes with related functions. Such examples of synthetic lethality have been widely studied in yeast and humans (Ooi *et al.*, 2006; Zhan & Boutros, 2016). Related efforts in Arabidopsis have been more limited in scope (Bolle *et al.*, 2013). Because earlier attempts to catalog double mutants with defects in embryo development were incomplete (Tzafrir *et al.*, 2004; Lloyd & Meinke, 2012), an expanded dataset was established here through literature curation (Dataset S9). Additional complexities were encountered, including differing genotypes of parental plants (double heterozygotes vs homozygous for one locus and heterozygous for the other); the presence or absence of single mutant phenotypes, which reflected differences in the degree of functional overlap; inclusion of genes in the single mutant *EMB* dataset; and whether sesquimutant embryos homozygous for one mutation and heterozygous for the other were also defective. Occasionally, embryo defects were identified in one direction (*aaBb* parent) and gametophyte defects in the other (*Aabb* parent). Close genetic linkage also complicated the production of some double mutants and altered segregation ratios.

Because some descriptions of double mutants were more definitive than others, each pair was assigned a status rank: (A) straightforward example with consistent seed phenotype, expected segregation ratios, and minimal defects in gametophyte development; (B) more complex or less definitive example; and (C) questionable example with limited analysis, ambiguous phenotype descriptions, confusing segregation ratios, or problematic defects in both embryo and gametophyte development. Once again, a notes column with additional details was included to facilitate subsequent analyses. Reciprocal BLASTP scores were also noted to differentiate between duplicated genes with a high degree of sequence similarity and pairs of dissimilar genes with overlapping cellular functions.

Eighty-three well-supported (Ranks A and B) double mutant combinations that produce defective embryos were identified. Among the protein functions involved, transcription factors and components of signaling pathways are well represented, in part because of widespread interest in studying their role in plant development. With 68 of the 83 pairs, neither gene is included in the single mutant dataset. These 131 distinct genes can thus be added to the 510 genes in Dataset S8 when tracking protein functions for which a loss-of-function phenotype indicates an essential role during embryo development. In most (66%) of these cases, the two genes are reciprocal BLASTP top matches, indicating a high degree of sequence similarity that likely reflects a recent gene duplication. In 22 of those pairs, neither single mutant was reported to have an obvious phenotype, consistent with significant functional overlap.

With the remaining 15 gene pairs at the bottom of Dataset S9, which includes eight genes not listed elsewhere, at least one member of the pair is also found in the single mutant *EMB* dataset. The double mutant phenotype is either more severe or exhibits a higher level of penetrance. Genetic screens for enhancers of embryo defects in *pinoid* single mutants were particularly successful at identifying double mutants with more striking phenotypes (Ito *et al.*, 2011). In some cases, the phenotypes of single mutant embryos are too subtle or infrequent to merit inclusion in the single mutant dataset. Examples include the slightly enlarged lipid droplets found in *seipin1* and *seipin2* mutants (Taurino *et al.*, 2018) and the low frequency of cup-shaped cotyledons described for *cuc1* and *cuc2* mutants (Aida *et al.*, 1997). In each of these cases, mutant phenotypes are consistent with the inability of one gene to compensate fully for disruption of the other during embryo development.

Fourteen questionable (Rank C) double mutant combinations requiring further analysis are listed in a separate tab of the dataset. Some of these have primary defects in gametophyte development or additional complexities that remain unresolved. Other double mutants are excluded because the embryonic defects are maternal in origin (Kunieda *et al.*, 2008) or less severe than single mutant phenotypes (Collinge *et al.*, 2004), or because lethality was noted without evidence of what developmental stage was disrupted. Promising examples of embryo lethals among the DUPLO collection of double knockouts of paralogous genes (Bolle *et al.*, 2013) are also excluded because of the limited analyses described.

X. Higher-order mutants defective in embryo development

When three or more genes with related functions are required for embryo development, higher-order mutants may be needed to observe a phenotype. The challenges of dealing with such multiple mutants on a broad scale have been noted previously (Lloyd & Meinke, 2012). Well-documented examples with alterations in embryo development include triple mutants defective in auxin influx carriers (Robert *et al.*, 2015) and a *bypass*-mediated signaling pathway (Lee *et al.*, 2012), and a quintuple mutant deficient in serine acetyltransferases (Watanabe *et al.*, 2008). An intriguing octuple mutant deficient in ethylene production and constructed using a combination of T-DNA insertions and artificial micro RNAs (Tsuchisaka *et al.*, 2009) was reported to exhibit embryo lethality but appears to be defective in female gametophyte development instead. Other multiple mutants are maternal in nature (Chen *et al.*, 2015) or exhibit defects that are simply more pronounced than those of lower-order mutants (Prigge *et al.*, 2005; Taurino *et al.*, 2018). The strength of individual mutant alleles can also determine the number of gene disruptions required to detect a phenotype. In one case involving a family of cytokinin response factors, the quadruple mutant (*crf1 crf2 crf5 crf6*) is defective in embryo development when a weak *crf5* allele is used, whereas the *crf5 crf6* double mutant is embryo-defective when a strong *crf5* allele is chosen (Rashotte *et al.*, 2006; Raines *et al.*, 2016). Thus, including weak alleles in multiple mutant combinations can sometimes enable additional genes with overlapping functions to be identified. Overall, these examples revealed 15 additional genes with redundant functions required for embryo development.

XI. Significance of a comprehensive dataset of *EMB* genes

The comprehensive dataset of *EMB* genes presented here provides a framework for determining which gene functions and mutant phenotypes are frequently encountered in existing collections, and by contrast, which atypical examples merit further analysis. It also documents the broad spectrum of essential genes and loss-of-function mutant alleles discovered to date. This perspective needs to be considered in future publications on embryo-defective mutants. The high frequency of mutant embryos arrested at the globular stage reflects both the abundance of EMB proteins targeted to plastids and the requirement of chloroplast gene expression for embryo development (Bryant *et al.*, 2011; Parker *et al.*, 2014). Unusual mutant phenotypes often reflect novel functions of a small number of genes or unique alleles that perturb common functions in novel ways. The failure of mutant embryos to complete a developmental pathway of interest does not necessarily indicate that the gene in question regulates that pathway, but rather that its cellular function is required for normal growth and development.

Beyond these general considerations, large datasets of Arabidopsis *EMB* genes are a valuable resource for addressing important questions in plant biology. For example, several recent studies on the maternal to zygotic transition in gene expression during reproductive development made use of existing datasets of *EMB* genes, either to evaluate transcriptional profiles of developing embryos for the presence of transcripts thought to be required at early stages of development (Nodine & Bartel, 2012), or to examine young seeds from segregating siliques to determine whether some heterozygous embryos are delayed in development, consistent with delayed expression of the paternal allele (Del Toro-De León *et al.*, 2014). Datasets of essential genes in Arabidopsis have also been used to identify candidate genes of interest in other plant species (Lloyd *et al.*, 2015) and to focus attention on essential genes with unknown cellular functions. They also offer promise in evaluating the embryonic transcriptome of Arabidopsis (Gao *et al.*, 2019; Hofmann *et al.*, 2019).

XII. Embryo defectives and the search for plant auxotrophs

One informative example of the value of large collections of embryo-defective mutants involves the longstanding search for plant auxotrophs. Nutritional mutants disrupted in seedling development were among the first Arabidopsis mutants to receive widespread attention (Langridge, 1955). The initial expectation was that a robust collection of auxotrophs could be established in Arabidopsis and used in a manner similar to that pursued with microorganisms. But despite extensive screens for rescuable mutants at the seedling stage, the vast majority of auxotrophs recovered in Arabidopsis were defective in a single pathway: thiamine biosynthesis (Li & Rédei, 1969). This outcome was unexpected and not readily explained. As a result, attention gradually shifted from seedling screens for auxotrophic mutants in Arabidopsis (Rédei, 1975) to the establishment of auxotrophic cell cultures in species more amenable to regeneration in culture (Blonstein, 1986).

After expanding the search for plant auxotrophs to a broader spectrum of forward and reverse genetic screens, the question of missing plant auxotrophs in Arabidopsis has finally been resolved. Elimination of biosynthetic pathways for an essential amino acid or vitamin in Arabidopsis frequently results in embryo lethality (Muralla *et al.*, 2007, 2008). Information on 19 *EMB* gene products required for the biosynthesis of an essential nutrient is found in Table 4. Embryo rescue with nutritional supplementation has been demonstrated for knockouts of 12 genes required for the biosynthesis of two amino acids (histidine and proline) and four vitamins (biotin, folate, pyridoxine and pantothenate). In two other cases (lysine and arginine), embryo rescue failed for unknown reasons. In five additional cases involving folate, serine, purine, phenylalanine, and aromatic amino acids, embryo rescue experiments have not been reported. Embryo lethality in the *cyt1* mutant defective in vitamin C biosynthesis is not rescuable because more than one cellular process is disrupted (Lukowitz *et al.*, 2001). Strong mutant alleles of some of the *EMB* genes exhibit defects in gametophyte development in addition to embryo lethality (Muralla *et al.*, 2007; Huang *et al.*, 2017). Double mutants defective in redundant genes required for two additional steps in folate and histidine biosynthesis also exhibit both embryo and gametophyte defects (Muralla *et al.*, 2007; Mehrshahi *et al.*, 2010). Conversely, seedling defects have been noted for weak alleles of several *EMB* genes listed in Table 4 (Conklin *et al.*, 2000; Song *et al.*, 2004; Noutoshi *et al.*, 2005; Mo *et al.*, 2006). Seedling mutant phenotypes have also been documented for Arabidopsis auxotrophs altered in eight additional genes required for the biosynthesis of thiamine (Komeda *et al.*, 1988; Papini-Terzi *et al.*, 2003; Raschke *et al.*, 2007; Mimura *et al.*, 2016), tryptophan (Last & Fink, 1988), riboflavin (Ouyang *et al.*, 2010; Hedtke *et al.*, 2012), and isoleucine (Yu *et al.*, 2013). Seedling defects in the *trp1* auxotroph are not rescuable by tryptophan (Last & Fink, 1988) because other intermediates in tryptophan biosynthesis are required for auxin-mediated plant growth and development.

Taken together, these reports and related studies indicate that the mutant phenotype expected from disruption of a gene involved in amino acid, vitamin, or purine biosynthesis in Arabidopsis depends on: (1) strength of the mutant allele being studied; (2) normal cellular function of the nutrient in question; (3) requirement for the nutrient at different stages of reproductive development; (4) redundancy of the biosynthetic pathway at both the genomic and biochemical levels; (5) expression patterns of duplicated genes with overlapping functions; and (6) ability of surrounding heterozygous tissue to rescue mutant embryos and gametophytes deficient in the essential nutrient. The defects observed in aminoacyl-tRNA synthetase mutants unable to incorporate specific amino acids into elongating polypeptide chains (Berg *et al.*, 2005) is consistent with the conclusion that without partial rescue from surrounding maternal or paternal tissues, the default state for amino acid auxotrophs (null alleles) is gametophyte lethality. But in practice, the most common phenotype observed for both amino acid and vitamin auxotrophs of Arabidopsis is embryo lethality.

XIII. Chloroplast translation and embryo development

A second example of the value of large mutant collections involves the question of why a loss of chloroplast translation in Arabidopsis consistently results in embryo lethality and seed abortion, whereas in maize, dependent upon genetic background, it leads to embryo lethality with normal endosperm development, or to albino seedlings or variegated leaf sectors (Parker *et al.*, 2014). The importance of chloroplast translation for embryo development in Arabidopsis is evident from the multiple *EMB* genes that encode chloroplast ribosomal proteins. The question is therefore which chloroplast genes, beyond those needed for protein synthesis, are required to support embryo development. Two candidates (*ycf1* and *ycf2*) known to be essential for plant growth (Drescher *et al.*, 2000) have recently been shown to encode components of chloroplast protein import complexes (Kikuchi *et al.*, 2013, 2018). The importance of these proteins for embryo development is confirmed by the knockout phenotypes of *EMB* genes encoding other components of the same complexes (Dataset S8). However, these complexes are also needed to import other proteins required for chloroplast translation, which leaves open the original question.

Three observations in my laboratory helped to establish that a single chloroplast gene (*accD*) encoding one subunit of the heteromeric acetyl-CoA carboxylase (ACCase) required for fatty acid biosynthesis in plastids is central to the importance of chloroplast gene function for embryo development in Arabidopsis. First, mutations in a nuclear gene that encodes another subunit of heteromeric ACCase also result in embryo lethality (X. Li *et al.*, 2011). Second, the null phenotype of embryos disrupted in chloroplast-localized ribosomal proteins differs by accession; mutant embryos in the 'Nossen' background (Riken insertion mutants) arrest at the preglobular stage, whereas those in the Columbia accession (Syngenta and Salk mutants) arrest at the globular stage. Finally, the 'Nossen' accession is unusually sensitive to a loss of chloroplast translation, as measured by seedling growth on spectinomycin.

Based on these observations, a genetic screen was performed to search for factors in a spectinomycin-tolerant accession that enabled further growth of arrested embryos from mutants defective in chloroplast translation in a sensitive accession (Parker *et al.*, 2014). The initial locus identified in this screen was *ACC2*, a tandemly-duplicated nuclear gene that encodes a plastid-localized version of homomeric ACCase (Babiychuk *et al.*, 2011). The other homomeric ACCase gene (*ACC1/EMB22/GURKE*) is highly expressed and functions in cytosolic long-chain fatty acid biosynthesis. Hypersensitive accessions harbor a variety of deleterious mutations in *ACC2* that prevent further development in the absence of chloroplast translation (Parker *et al.*, 2016). By contrast, functional ACC2 in tolerant accessions results in partial rescue of embryos and seedlings defective in chloroplast translation because it supports a low level of fatty acid biosynthesis in plastids. The globular stage of embryo arrest characteristic of mutants disrupted in a wide range of chloroplast functions in the Columbia accession therefore reflects a shared defect in fatty acid biosynthesis. The availability of large collections of ribosomal protein mutants in multiple accessions enabled this relationship to be established. This work also contributed to recent advances in chloroplast transformation in Arabidopsis (Yu *et al.*, 2017; Ruf *et al.*, 2019). A different situation is encountered in maize, where chloroplasts lack *accD, ycf1*, and *ycf2*, and instead utilize a plastid-targeted, homomeric ACCase for fatty acid biosynthesis. Embryo-specific lethality in maize kernels defective in plastid-localized ribosomal proteins in selected genetic backgrounds remains to be explained, but may reflect alterations in a retrograde signal (Woodson & Chory, 2012) that activates expression of nuclear genes required for embryogenesis but not endosperm formation or leaf development.

XIV. Approaching saturation for *EMB* genes in Arabidopsis

Despite past concerns about feasibility and effort required, research on embryo-defective mutants has progressed to where questions about saturation merit consideration. The total number of *EMB* genes in Arabidopsis remains unknown. Initial projections of 4000 loci based on a forward genetic screen of EMS mutagenized lines for abnormal seedling phenotypes (Jürgens *et al.*, 1991) were likely to be an overestimate. Spontaneously aborted seeds are frequently encountered in such screens, particularly when large numbers of plants are grown under suboptimal conditions, and putative *emb* mutants in this case were not analyzed in detail or confirmed in subsequent generations. Estimates from my laboratory based on the frequency of confirmed *emb* mutations relative to other types of

Table 4 Examples of auxotrophic *emb* mutants identified through forward and reverse genetic screens.

Nutrient class	Biosynthetic pathway	Genetic screen	Mutants analyzed[a]	Rescue experiments Embryos[b]	Plants[c]	References
Vitamin	Biotin	Forward	*bio1*	C, S	V, F	Schneider *et al*. (1989)
	Biotin	Both	*bio2*	C, S	V[f]	Patton *et al*. (1998)
				S	V, F	Arnal *et al*. (2006)
	Biotin	Reverse	*bio3 (bio1)*[d]	C, S	V, F	Muralla *et al*. (2008)
	Folate	Forward	*gla1 (emb9)*	C	ND	Ishikawa *et al*. (2003)
	Pyridoxine	Both	*pdx2 (emb2407)*	C	V	Tambasco-Studart *et al*. (2007)
	Pantothenate	Reverse	*pts*	S	V, F	Jonczyk *et al*. (2008)
Amino acid	Histidine	Both	*hisn6a*	S	V	Muralla *et al*. (2007)
	Histidine	Reverse	*hisn2; hisn3; hisn4; hisn8*[e]	S	V	Muralla *et al*. (2007)
	Histidine	Reverse	*hisn7/impl2*	S	ND	Petersen *et al*. (2010)
				S	ND	Sato *et al*. (2011)
	Proline	Reverse	*p5cs2*	C	(V)	Székely *et al*. (2008)
				S	ND	Mattioli *et al*. (2009)
	Lysine	Both	*agd2*	Failed	ND	Song *et al*. (2004)
				ND	ND	Hudson *et al*. (2006)
	Arginine	Forward	*nagk*	Failed	ND	Huang *et al*. (2017)
Vitamin	Folate	Reverse	*emb3127*	ND	ND	Bryant *et al*. (2011)
Amino acid	Serine	Reverse	*pgdh1*	ND	ND	Toujani *et al*. (2013)
				ND	ND	Benstein *et al*. (2013)
	Phenylalanine	Reverse	*adt2*	ND	ND	El-Azaz *et al*. (2018)
	Aromatic	Forward	*emb1144*	ND	ND	Tzafrir *et al*. (2004)
Other	Purine	Reverse	*emb2818*	ND	ND	Muralla *et al*. (2011)

[a] Allelic *emb* mutants are listed in parentheses. Alias symbols are separated by a slash.
[b] C, cultured: mutant embryos are rescued (reach a later stage of development or produce callus) when cultured on media containing the required nutrient; S, supplemented: mutant embryos are rescued when heterozygous plants are watered or sprayed with the required nutrient; ND, not documented.
[c] V, viable: homozygous mutant plants are viable when supplemented with the required nutrient; F, fertile: homozyogus mutant plants are fertile and produce mature seeds when supplemented with the required nutrient; ND, not documented; (V), incomplete rescue of rosettes.
[d] The *BIO3-BIO1* locus encodes a bifunctional enzyme that catalyzes two sequential steps in biotin biosynthesis.
[e] The *hisn8* mutant exhibits an ovule abortion/zygotic embryo phenotype that is rescued by supplemental histidine.
[f] The original *bio2* mutant allele *(emb49)* contains a deletion that removes a linked gene (*FPA*) required for flowering.

mutations in the Feldmann T-DNA collection, and the frequency of allelic *emb* mutants obtained from several forward genetic screens, ranged from 500 to 1000 genes (Franzmann *et al.*, 1995; Meinke *et al.*, 2008; Muralla *et al.*, 2011). These estimates also have limitations given that some genes mutate more readily than others. After reviewing past work and taking into account the number of known *EMB* genes already identified, I postulate that 50–70% of the estimated 750–1000 *EMB* genes with a single mutant phenotype in Arabidopsis have likely been found.

Strategies for approaching saturation have been discussed before (Meinke *et al.*, 2008), with emphasis placed on reverse genetic approaches involving homologs of essential genes in other organisms, shared processes, pathways and protein interactors with known EMB proteins, and genes expressed in embryos or excluded from collections of knockout homozygotes. Recently, machine-learning methods were utilized to identify additional candidates for essential genes in Arabidopsis (Lloyd *et al.*, 2015). Future datasets might need to include chloroplast genes required for the synthesis of essential proteins (*c.* 30 genes) and nuclear genes encoding miRNAs required for embryo development (Armenta-Medina *et al.*, 2017). Broadening the definition of *EMBs* to include redundant genes with multiple mutant phenotypes would further increase the number of genes involved and the challenges of reaching saturation.

Reverse genetic studies not initially focused on embryogenesis should continue to reveal additional *EMB* genes in the future. But progress with this approach will be slow, with some classes of proteins excluded, particularly those with unknown functions. Gene editing methods could simplify the production of null alleles for candidate genes not represented in existing collections. But whether this approach can be scaled to the analysis of hundreds of candidate genes remains to be seen. Large-scale forward genetic screens for embryo-defective mutants are labor-intensive and will become less efficient as saturation is approached, with a higher percentage of mutants representing new alleles of known genes. Forward screens focused on mutant phenotypes of special interest will continue to be explored (Gillmor *et al.*, 2016). But another broad screen of insertion lines similar in scale to that undertaken with Syngenta seems unlikely.

With these limitations in mind, the best strategy forward may be to combine targeted reverse genetic screens of candidate essentials with further analysis of existing collections of mutants not found in the updated dataset. These collections include mutants with overlapping defects in female gametophyte development (Pagnussat *et al.*, 2005), mutants disrupted in predicted chloroplast-localized proteins (Savage *et al.*, 2013), and community mutants with ambiguous phenotypes or unresolved gene identities. Extensive collections from my laboratory also merit further analysis, including mutants that were initially mapped but not cloned (Dataset S2) and Syngenta mutants that are putatively tagged, unresolved, or not tagged (Datasets S5, S6). These mutants are well-curated, with seed stocks available to the community without restrictions. With mapped mutants, the original focus on chromosomes 1 and 5 (Meinke *et al.*, 2009) could be updated with additional cloned *EMB* genes and expanded genome-wide. This approach would not reveal new *EMB* genes, but it could uncover additional alleles of known genes, thereby focusing attention on loci that remain to be identified. These loci have the advantage of an established map location, which could facilitate gene identification using whole-genome sequencing methods similar to those pursued with essential genes in *Chlamydomonas reinhardtii* (Breker *et al.*, 2018).

XV. Conclusions

Advances in genome sequencing over the past 20 years have yielded invaluable information on the molecular landscape of a broad spectrum of plant species. Mutant analysis in model genetic organisms such as Arabidopsis, however, remains a powerful tool for identifying genes with essential functions in plant growth and development. One obstacle to further advances in mutant analysis of seed development continues to be the sheer number of mutants and genes involved. This comprehensive review was designed to maximize the utilization of existing collections of embryo-defective mutants of Arabidopsis, and to focus attention not on what has already been learned about plant development from the analysis of selected mutants of interest, but rather on the value of analyzing and expanding the mutant collection as a whole. Looking ahead, further community-based curation efforts will clearly be required to update the datasets of *EMB* genes and mutant alleles presented here, and to track continued progress towards saturation. Ideally, someone might volunteer to oversee this work early in their career to assure long-term progress. Multiple groups could also coordinate different types of genetic screens to expand the collection. With appropriate training, teams of undergraduate students could make significant contributions, particularly with literature curation. Further insights could be obtained from longstanding efforts to uncover Mendelian inheritance genes in humans (Amberger *et al.*, 2019) and essential genes in other model organisms (Zhan & Boutros, 2016). The challenges ahead remain daunting. But the potential benefits of a coordinated effort to approach saturation for embryo-defective mutants of Arabidopsis are substantial, and a prerequisite to determining the function of every gene in a model angiosperm.

Acknowledgements

This review is dedicated to the 90 members of my laboratory at Oklahoma State University, from students to research associates, who contributed to the isolation and characterization of embryo-defective mutants over a period of 36 years. Contributions of external collaborators, especially David Patton (Syngenta) and Allan Dickerman (Biocomplexity Institute, University of Virginia), and individuals throughout the Arabidopsis community, are gratefully acknowledged. Research in my laboratory was supported by multiple grants from the National Science Foundation, with additional funding from USDA, Syngenta, the S.R. Noble Foundation, and the Oklahoma Center for the Advancement of Science and Technology. Nicole Parker assisted with preparation of this article and with submission of residual seed stocks to the Arabidopsis Biological Resource Center.

ORCID

David W. Meinke https://orcid.org/0000-0002-7292-6318

References

Abramov VI, Rubanovich AV, Shevchenko VA, Shevchenko VV, Grinikh LI. 2006. Genetic effects in plant populations in the zone of the Chernobyl accident. *Radiatsionnaia Biologiia Radioecologiia* **46**: 259–267.

Aida M, Ishida T, Fukaki H, Fujisawa H, Tasaka M. 1997. Genes involved in organ separation in Arabidopsis: an analysis of the cup-shaped cotyledon mutant. *Plant Cell* **9**: 841–857.

Alandete-Saez M, Ron M, Leiboff S, McCormick S. 2011. *Arabidopsis thaliana* GEX1 has dual functions in gametophyte development and early embryogenesis. *The Plant Journal* **68**: 620–632.

Albert S, Després B, Guilleminot J, Bechtold N, Pelletier G, Delseny M, Devic M. 1999. The *EMB 506* gene encodes a novel ankyrin repeat containing protein that is essential for the normal development of Arabidopsis embryos. *The Plant Journal* **17**: 169–179.

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R *et al.* 2003. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* **301**: 653–657.

Amberger JS, Bocchini CA, Scott AF, Hamosh A. 2019. OMIM.org: leveraging knowledge across phenotype-gene relationships. *Nucleic Acids Research* **47**: D1038–D1043.

Apuya NR, Yadegari R, Fischer RL, Harada JJ, Zimmerman JL, Goldberg RB. 2001. The Arabidopsis embryo mutant *schlepperless* has a defect in the *chaperonin-60α* gene. *Plant Physiology* **126**: 717–730.

Armenta-Medina A, Gillmor CS. 2019. Genetic, molecular and parent-of-origin regulation of early embryogenesis in flowering plants. *Current Topics in Developmental Biology* **131**: 497–543.

Armenta-Medina A, Lepe-Soltero D, Xiang D, Datla R, Abreu-Goodger C, Gillmor CS. 2017. *Arabidopsis thaliana* miRNAs promote embryo pattern formation beginning in the zygote. *Developmental Biology* **431**: 145–151.

Arnal N, Alban C, Quadrado M, Grandjean O, Mireau H. 2006. The *Arabidopsis* Bio2 protein requires mitochondrial targeting for activity. *Plant Molecular Biology* **62**: 471–479.

Babiychuk E, Vandepoele K, Wissing J, Garcia-Diaz M, De Rycke R, Akbari H, Joubès J, Beeckman T, Jänsch L, Frentzen M *et al.* 2011. Plastid gene expression and plant development require a plastidic protein of the mitochondrial transcription termination factor family. *Proceedings of the National Academy of Sciences, USA* **108**: 6674–6679.

Babu Y, Musielak T, Henschen A, Bayer M. 2013. Suspensor length determines developmental progression of the embryo in Arabidopsis. *Plant Physiology* **162**: 1448–1458.

Barkan A, Small I. 2014. Pentatricopeptide repeat proteins in plants. *Annual Review of Plant Biology* **65**: 415–442.

Bartha I, Di Lulio J, Venter JC, Telenti A. 2018. Human gene essentiality. *Nature Reviews Genetics* **19**: 51–62.

Baus AD, Franzmann L, Meinke DW. 1986. Growth *in vitro* of arrested embryos from lethal mutants of *Arabidopsis thaliana*. *Theoretical and Applied Genetics* **72**: 577–586.

Bayer M, Nawy T, Giglione C, Galli M, Meinnel T, Lukowitz W. 2009. Paternal control of embryonic patterning in *Arabidopsis thaliana*. *Science* **323**: 1485–1488.

Benstein RM, Ludewig K, Wulfert S, Wittek S, Gigolashvili T, Frerigmann H, Gierth M, Flügge UI, Krueger S. 2013. *Arabidopsis* phosphoglycerate dehydrogenase1 of the phosphoserine pathway is essential for development and required for ammonium assimilation and tryptophan biosynthesis. *Plant Cell* **25**: 5011–5029.

Berg M, Rogers R, Muralla R, Meinke D. 2005. Requirement of aminoacyl-tRNA synthetases for gametogenesis and embryo development in Arabidopsis. *The Plant Journal* **44**: 866–878.

Bernard DG, Netz DJA, Lagny TJ, Pierik AJ, Balk J. 2013. Requirements of the cytosolic iron-sulfur cluster assembly pathway in Arabidopsis. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **368**: 20120259.

Bikard D, Patel D, Le Metté C, Giorgi V, Camilleri C, Bennett MJ, Loudet O. 2009. Divergent evolution of duplicate genes leads to genetic incompatibilities within *A. thaliana*. *Science* **323**: 623–626.

Blevins T, Wang J, Pflieger D, Pontvianne F, Pikaard CS. 2017. Hybrid incompatibility caused by an epiallele. *Proceedings of the National Academy of Sciences, USA* **114**: 3702–3707.

Blonstein AD. 1986. Auxotroph isolation *in vitro*. In: Blonstein AD, King PJ, eds. *Plant gene research: a genetic approach to plant biochemistry*. New York, NY, USA: Springer, 259–279.

Boavida LC, Shuai B, Yu HJ, Pagnussat GC, Sundaresan V, McCormick S. 2009. A collection of Ds insertional mutants associated with defects in male gametophyte development and function in *Arabidopsis thaliana*. *Genetics* **181**: 1369–1385.

Bolle C, Huep G, Kleinbölting N, Haberer G, Mayer K, Leister D, Weisshaar B. 2013. GABI-DUPLO: a collection of double mutants to overcome genetic redundancy in *Arabidopsis thaliana*. *The Plant Journal* **75**: 157–171.

Breker M, Lieberman K, Cross FR. 2018. Comprehensive discovery of cell-cycle-essential pathways in C*hlamydomonas reinhardtii*. *Plant Cell* **30**: 1178–1198.

Brownfield L, Hafidh S, Durbarry A, Khatab H, Sidorova A, Doerner P, Twell D. 2009. *Arabidopsis* DUO POLLEN3 is a key regulator of male germline development and embryogenesis. *Plant Cell* **21**: 1940–1956.

Bryant N, Lloyd J, Sweeney C, Myouga F, Meinke D. 2011. Identification of nuclear genes encoding chloroplast-localized proteins required for embryo development in Arabidopsis. *Plant Physiology* **155**: 1678–1689.

Budziszewski GJ, Lewis SP, Glover LW, Reineke J, Jones G, Ziemnik LS, Lonowski J, Nyfeler B, Aux G, Zhou Q *et al.* 2001. Arabidopsis genes essential for seedling viability: isolation of insertional mutants and molecular cloning. *Genetics* **159**: 1765–1778.

Candela H, Pérez-Pérez JM, Micol JL. 2011. Uncovering the post-embryonic functions of gametophytic- and embryonic-lethal genes. *Trends in Plant Science* **16**: 336–345.

Castle LA, Errampalli D, Atherton TL, Franzmann LH, Yoon EH, Meinke DW. 1993. Genetic and molecular characterization of embryonic mutants identified following seed transformation in *Arabidopsis*. *Molecular and General Genetics* **241**: 504–514.

Chaudhury AM, Ming L, Miller C, Craig S, Dennis ES, Peacock WJ. 1997. Fertilization-independent seed development in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences, USA* **94**: 4223–4228.

Chen JG, Ullah H, Young JC, Sussman MR, Jones AM. 2001. ABP1 is required for organized cell elongation and division in Arabidopsis embryogenesis. *Genes & Development* **15**: 902–911.

Chen LQ, Lin IW, Qu XQ, Sosso D, McFarlane HE, Londoño A, Samuels AL, Frommer WB. 2015. A cascade of sequentially expressed sucrose transporters in the seed coat and endosperm provides nutrition for the Arabidopsis embryo. *Plant Cell* **27**: 607–619.

Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, Jacobsen SE, Fischer RL. 2002. DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. *Cell* **110**: 33–42.

Clark KA, Krysan PJ. 2010. Chromosomal translocations are a common phenomenon in *Arabidopsis thaliana* T-DNA insertion lines. *The Plant Journal* **64**: 990–1001.

Collinge MA, Spillane C, Köhler C, Gheyselinck J, Grossniklaus U. 2004. Genetic interaction of an origin recognition complex subunit and the *Polycomb* group gene *MEDEA* during seed development. *Plant Cell* **16**: 1035–1046.

Conklin PL, Saracco SA, Norris SR, Last RL. 2000. Identification of ascorbic acid-deficient *Arabidopsis thaliana* mutants. *Genetics* **154**: 847–856.

Copenhaver GP, Keith KC, Preuss D. 2000. Tetrad analysis in higher plants. A budding technology. *Plant Physiology* **124**: 7–16.

Cushing DA, Forsthoefel NR, Gestaut DR, Vernon DM. 2005. *Arabidopsis emb175* and other *ppr* knockout mutants reveal essential roles for pentatricopeptide repeat (PPR) proteins in plant embryogenesis. *Planta* **221**: 424–436.

Dai X, Zhang Y, Zhang D, Chen J, Gao X, Estelle M, Zhao Y. 2015. Embryonic lethality of *Arabidopsis abp1-1* is caused by deletion of the adjacent *BSM* gene. *Nature Plants* **1**: 15183.

De Smet I, Lau S, Mayer U, Jürgens G. 2010. Embryogenesis – the humble beginnings of plant life. *The Plant Journal* **61**: 959–970.

Dean PJ, Siwiec T, Waterworth WM, Schlögelhofer P, Armstrong SJ, West CE. 2009. A novel ATM-dependent X-ray-inducible gene is essential for both plant meiosis and gametogenesis. *The Plant Journal* **58**: 791–802.

Del Toro-De León G, García-Aguilar M, Gillmor CS. 2014. Non-equivalent contributions of maternal and paternal genomes to early plant embryogenesis. *Nature* **514**: 624–627.

Despres B, Delseny M, Devic M. 2001. Partial complementation of embryo defective mutations: a general strategy to elucidate gene function. *The Plant Journal* **27**: 149–159.

Devic M. 2008. The importance of being essential: *EMBRYO-DEFECTIVE* genes in Arabidopsis. *Comptes Rendus Biologies* **331**: 726–736.

Devic M, Albert S, Delseny M. 1996. Induction and expression of seed-specific promoters in Arabidopsis embryo-defective mutants. *The Plant Journal* **9**: 205–215.

Drescher A, Ruf S, Calsa T Jr, Carrer H, Bock R. 2000. The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. *The Plant Journal* **22**: 97–104.

Du Y, Cui M, Qian D, Zhu L, Wei C, Yuan M, Zhang Z, Li Y. 2010. AtTFC B is involved in control of cell division. *Frontiers in Bioscience (Elite Edition)* **2**: 752–763.

El-Azaz J, Cánovas FM, Ávila C, de la Torre F. 2018. The arogenate dehydratase ADT2 is essential for seed development in *Arabidopsis*. *Plant Cell Physiology* **59**: 2409–2420.

Errampalli D, Patton D, Castle L, Mickelson L, Hansen K, Schnall J, Feldmann K, Meinke D. 1991. Embryonic lethals and T-DNA insertional mutagenesis in *Arabidopsis*. *Plant Cell* **3**: 149–157.

Faure JD, Vittorioso P, Santoni V, Fraisier V, Prinsen E, Barlier I, Van Onckelen H, Caboche M, Bellini C. 1998. The *PASTICCINO* genes of *Arabidopsis thaliana* are involved in the control of cell division and differentiation. *Development* **125**: 909–918.

Feldmann KA. 1991. T-DNA insertion mutagenesis in *Arabidopsis*: mutational spectrum. *The Plant Journal* **1**: 71–82.

Franzmann L, Patton DA, Meinke DW. 1989. *In vitro* morphogenesis of arrested embryos from lethal mutants of *Arabidopsis thaliana*. *Theoretical and Applied Genetics* **77**: 609–616.

Franzmann LH, Yoon ES, Meinke DW. 1995. Saturating the genetic map of *Arabidopsis thaliana* with embryonic mutations. *The Plant Journal* **7**: 341–350.

Gao P, Xiang D, Quilichini TD, Venglat P, Pandey PK, Wang E, Gillmor CS, Datla R. 2019. Gene expression atlas of embryo development in *Arabidopsis*. *Plant Reproduction* **32**: 93–104.

Gillmor CS, Roeder AH, Sieber P, Somerville C, Lukowitz W. 2016. A genetic screen for mutations affecting cell division in the *Arabidopsis thaliana* embryo identifies seven loci required for cytokinesis. *PLoS ONE* **11**: e0146492.

Grigg SP, Canales C, Hay A, Tsiantis M. 2005. SERRATE coordinates shoot meristem function and leaf axial patterning in Arabidopsis. *Nature* **437**: 1022–1026.

Grossniklaus U, Vielle-Calzada JP, Hoeppner MA, Gagliano WB. 1998. Maternal control of embryogenesis by *MEDEA*, a *Polycomb* group gene in *Arabidopsis*. *Science* **280**: 446–450.

Gusti A, Baumberger N, Nowack M, Pusch S, Eisler H, Potuschak T, De Veylder L, Schnittger A, Genschik P. 2009. The *Arabidopsis thaliana* F-box protein FBL17 is essential for progression through the second mitosis during pollen development. *PLoS ONE* **4**: e4780.

Heath JD, Weldon R, Monnot C, Meinke DW. 1986. Analysis of storage proteins in normal and aborted seeds from embryo-lethal mutants of *Arabidopsis thaliana*. *Planta* **169**: 304–312.

Hedtke B, Alawady A, Albacete A, Kobayashi K, Melzer M, Roitsch T, Masuda T, Grimm B. 2012. Deficiency in riboflavin biosynthesis affects tetrapyrrole biosynthesis in etiolated Arabidopsis tissue. *Plant Molecular Biology* **78**: 77–93.

Henderson IR, Liu F, Drea S, Simpson GG, Dean C. 2005. An allelic series reveals essential roles for FY in plant development in addition to flowering-time control. *Development* **132**: 3597–3607.

Herridge RP, Day RC, Macknight RC. 2014. The role of MCM2-7 helicase complex during *Arabidopsis* seed development. *Plant Molecular Biology* **86**: 69–84.

Hofmann F, Schon MA, Nodine MD. 2019. The embryonic transcriptome of *Arabidopsis thaliana*. *Plant Reproduction* **32**: 77–91.

Holding DR, Springer PS. 2002. The Arabidopsis gene *PROLIFERA* is required for proper cytokinesis during seed development. *Planta* **214**: 373–382.

Hsu SC, Belmonte MF, Harada JJ, Inoue K. 2010. Indispensable roles of plastids in *Arabidopsis thaliana* embryogenesis. *Current Genomics* **11**: 338–349.

Huang J, Chen D, Yan H, Xie F, Yu Y, Zhang L, Sun M, Peng X. 2017. Acetylglutamate kinase is required for both gametophyte function and embryo development in *Arabidopsis thaliana*. *Journal of Integrative Plant Biology* **59**: 642–656.

Hudson AO, Singh BK, Leustek T, Gilvarg C. 2006. An LL-diaminopimelate aminotransferase defines a novel variant of the lysine biosynthesis pathway in plants. *Plant Physiology* **140**: 292–301.

Ishikawa T, Machida C, Yoshioka Y, Kitano H, Machida Y. 2003. The *GLOBULAR ARREST1* gene, which is involved in the biosynthesis of folates, is essential for embryogenesis in *Arabidopsis thaliana*. *The Plant Journal* **33**: 235–244.

Ito J, Sono T, Tasaka M, Furutani M. 2011. *MACCHI-BOU 2* is required for early embryo patterning and cotyledon organogenesis in Arabidopsis. *Plant Physiology* **52**: 539–552.

Iwakawa H, Shinmyo A, Sekine M. 2006. Arabidopsis CDKA;1, a cdc2 homologue, controls proliferation of generative cells in male gametogenesis. *The Plant Journal* **45**: 819–831.

Johnson MA, von Besser K, Zhou Q, Smith E, Aux G, Patton D, Levin JZ, Preuss D. 2004. Arabidopsis *hapless* mutations define essential gametophytic functions. *Genetics* **168**: 971–982.

Jonczyk R, Ronconi S, Rychlik M, Genschel U. 2008. Pantothenate synthetase is essential but not limiting for pantothenate biosynthesis in Arabidopsis. *Plant Molecular Biology* **66**: 1–14.

Jürgens G, Mayer U, Torres Ruiz RA, Berleth T, Miséra S. 1991. Genetic analysis of pattern formation in the *Arabidopsis* embryo. *Development* **113**(Supplement 1): 27–38.

Kikuchi S, Asakura Y, Imai M, Nakahira Y, Kotani Y, Hashiguchi Y, Nakai Y, Takafuji K, Bédard J, Hirabayashi-Ishioka Y *et al.* 2018. A Ycf2-FtsHi heteromeric AAA-ATPase complex is required for chloroplast protein import. *Plant Cell* **30**: 2677–2703.

Kikuchi S, Bédard J, Hirano M, Hirabayashi Y, Oishi M, Imai M, Takase M, Ide T, Nakai M. 2013. Uncovering the protein translocon at the chloroplast inner envelope membrane. *Science* **339**: 571–574.

Kim HJ, Oh SA, Brownfield L, Hong SH, Ryu H, Hwang I, Twell D, Nam HG. 2008. Control of plant germline proliferation by $\mathrm{SCF}^{\mathrm{FBL17}}$ degradation of cell cycle inhibitors. *Nature* **455**: 1134–1137.

Kinoshita T, Yadegari R, Harada JJ, Goldberg RB, Fischer RL. 1999. Imprinting of the *MEDEA* polycomb gene in the Arabidopsis endosperm. *Plant Cell* **11**: 1945–1952.

Komeda Y, Tanaka M, Nishimune T. 1988. A *th-1* mutant of *Arabidopsis thaliana* is defective for a thiamin-phosphate-synthesizing enzyme: thiamin phosphate pyrophosphorylase. *Plant Physiology* **88**: 248–250.

Koornneef M, Hanhart CJ, van der Veen JH. 1991. A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. *Molecular and General Genetics* **229**: 57–66.

Koornneef M, Meinke D. 2010. The development of Arabidopsis as a model plant. *The Plant Journal* **61**: 909–921.

Koornneef M, Van Eden J, Hanhart CJ, Stam P, Braaksma FJ, Feenstra WJ. 1983. Linkage map of *Arabidopsis thaliana*. *Journal of Heredity* **74**: 265–272.

Kunieda T, Mitsuda N, Ohme-Takagi M, Takeda S, Aida M, Tasaka M, Kondo M, Nishimura M, Hara-Nishimura I. 2008. NAC family proteins NARS1/NAC2 and NARS2/NAM in the outer integument regulate embryogenesis in *Arabidopsis*. *Plant Cell* **20**: 2631–2642.

Lahmy S, Guilleminot J, Cheng CM, Bechtold N, Albert S, Pelletier G, Delseny M, Devic M. 2004. *DOMINO1*, a member of a small plant-specific gene family, encodes a protein essential for nuclear and nucleolar functions. *The Plant Journal* **39**: 809–820.

Lalanne E, Michaelidis C, Moore JM, Gagliano W, Johnson A, Patel R, Howden R, Vielle-Calzada JP, Grossniklaus U, Twell D. 2004. Analysis of transposon insertion mutants highlights the diversity of mechanisms underlying male progamic development in Arabidopsis. *Genetics* **167**: 1975–1986.

Langridge J. 1955. Biochemical mutations in the crucifer *Arabidopsis thaliana* (L.) Heynh. *Nature* **176**: 260–261.

Last RL, Fink GR. 1988. Tryptophan-requiring mutants of the plant *Arabidopsis thaliana*. *Science* **240**: 305–310.

Latvala-Kilby SM, Kilby NJ. 2006. Uncovering the post-embryonic role of embryo essential genes in *Arabidopsis* using the controlled induction of visibly marked genetic mosaics: *EMB506*, an illustration. *Plant Molecular Biology* **61**: 179–194.

Lee DK, Van Norman JM, Murphy C, Adhikari E, Reed JW, Sieburth LE. 2012. In the absence of BYPASS1-related gene function, the *bps* signal disrupts embryogenesis by an auxin-independent mechanism. *Development* **139**: 805–815.

Leroy O, Hennig L, Breuninger H, Laux T, Köhler C. 2007. Polycomb group proteins function in the female gametophyte to determine seed development in plants. *Development* **134**: 3639–3648.

Li HJ, Xue Y, Jia DJ, Wang T, Hi DQ, Liu J, Cui F, Xie Q, Ye D, Yang WC. 2011. POD1 regulates pollen tube guidance in response to micropylar female signaling and acts in early embryo patterning in *Arabidopsis*. *Plant Cell* **23**: 3288–3302.

Li SL, Rédei GP. 1969. Thiamine mutants of the crucifer, *Arabidopsis*. *Biochemical Genetics* **3**: 163–170.

Li X, Ilarslan H, Brachova L, Qian HR, Li L, Che P, Wurtele ES, Nikolau BJ. 2011. Reverse-genetic analysis of the two biotin-containing subunit genes of the heteromeric acetyl-coenzyme A carboxylase in Arabidopsis indicates a unidirectional functional redundancy. *Plant Physiology* **155**: 293–314.

Lin Y, Sun L, Nguyen LV, Rachubinski RA, Goodman HM. 1999. The Pex16p homolog SSE1 and storage organelle formation in *Arabidopsis* seeds. *Science* **284**: 328–330.

Liu CM, McElver J, Tzafrir I, Joosen R, Wittich P, Patton D, Van Lammeren AAM, Meinke D. 2002. Condensin and cohesin knockouts in *Arabidopsis* exhibit a *titan* seed phenotype. *The Plant Journal* **29**: 405–415.

Liu CM, Meinke DW. 1998. The *titan* mutants of *Arabidopsis* are disrupted in mitosis and cell cycle control during seed development. *The Plant Journal* **16**: 21–31.

Lloyd J, Meinke D. 2012. A comprehensive dataset of genes with a loss-of-function mutant phenotype in Arabidopsis. *Plant Physiology* **158**: 1115–1129.

Lloyd JP, Seddon AE, Moghe GD, Simenc MC, Shiu SH. 2015. Characteristics of plant essential genes allow for within- and between- species prediction of lethal mutant phenotypes. *Plant Cell* **27**: 2133–2147.

Lu Y. 2018. Assembly and transfer of iron-sulfur clusters in the plastid. *Frontiers in Plant Science* **9**: 336.

Lukowitz W, Mayer U, Jürgens G. 1996. Cytokinesis in the Arabidopsis embryo involves the syntaxin-related KNOLLE gene product. *Cell* **84**: 61–71.

Lukowitz W, Nickle TC, Meinke DW, Last RL, Conklin PL, Somerville CR. 2001. *Arabidopsis cyt1* mutants are deficient in a mannose-1-phosphate guanylyltransferase and point to a requirement of N-linked glycosylation for cellulose biosynthesis. *Proceedings of the National Academy of Sciences, USA* **98**: 2262–2267.

Luo M, Bilodeau P, Dennis ES, Peacock WJ, Chaudhury A. 2000. Expression and parent-of-origin effects for *FIS2*, *MEA*, and *FIE* in the endosperm and embryo of developing *Arabidopsis* seeds. *Proceedings of the National Academy of Sciences, USA* **97**: 10637–10642.

Lurin C, Andrés C, Aubourg S, Bellaoui M, Bitton F, Bruyère C, Caboche M, Debast C, Gualberto J, Hoffmann B *et al.* 2004. Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* **16**: 2089–2103.

Mansilla N, Garcia L, Gonzalez DH, Welchen E. 2015. AtCOX10, a protein involved in haem *o* synthesis during cytochrome *c* oxidase biogenesis, is essential for plant embryogenesis and modulates the progression of senescence. *Journal of Experimental Botany* **66**: 6761–6775.

Marsden MPF, Meinke DW. 1985. Abnormal development of the suspensor in an embryo-lethal mutant of *Arabidopsis thaliana*. *American Journal of Botany* **72**: 1801–1812.

Mattioli R, Falasca G, Sabatini S, Altamura MM, Costantino P, Trovato M. 2009. The proline biosynthetic genes *P5CS1* and *P5CS2* play overlapping roles in *Arabidopsis* flower transition but not in embryo development. *Physiologia Plantarum* **137**: 72–85.

Mayer U, Torres Ruiz RA, Berleth T, Miséra S, Jürgens G. 1991. Mutations affecting body organization in the *Arabidopsis* embryo. *Nature* **353**: 402–407.

McElver J, Tzafrir I, Aux G, Rogers R, Ashby C, Smith K, Thomas C, Schetter A, Zhou Q, Cushman MA *et al.* 2001. Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana*. *Genetics* **159**: 1751–1763.

Mehrshahi P, Gonzalez-Jorge S, Akhtar TA, Ward JL, Santoyo-Castelazo A, Marcus SE, Lara-Núñez A, Ravanel S, Hawkins ND, Beale MH *et al.* 2010. Functional analysis of folate polyglutamylation and its essential role in plant metabolism and development. *The Plant Journal* **64**: 267–279.

Meinke D, Koornneef M. 1997. Community standards for *Arabidopsis* genetics. *The Plant Journal* **12**: 247–253.

Meinke D, Muralla R, Sweeney C, Dickerman A. 2008. Identifying essential genes in *Arabidopsis thaliana*. *Trends in Plant Science* **13**: 483–491.

Meinke D, Patton D, Shellhammer J, Reynolds-Duffer A, Franzmann L, Schneider T, Robinson K. 1989. Developmental and molecular genetics of embryogenesis in *Arabidopsis thaliana*. In: Goldberg R, ed. *UCLA Symposia on Molecular and Cellular Biology, Vol. 92, The molecular basis of plant development*. New York, NY, USA: Alan Liss, 121–132.

Meinke D, Sweeney C, Muralla R. 2009. Integrating the genetic and physical maps of *Arabidopsis thaliana*: identification of mapped alleles of cloned essential (*EMB*) genes. *PLoS ONE* **4**: e7386.

Meinke DW. 1982. Embryo-lethal mutants of *Arabidopsis thaliana*: evidence for gametophytic expression of the mutant genes. *Theoretical and Applied Genetics* **63**: 381–386.

Meinke DW. 1985. Embryo-lethal mutants of *Arabidopsis thaliana*: analysis of mutants with a wide range of lethal phases. *Theoretical and Applied Genetics* **69**: 543–552.

Meinke DW. 1992. A homoeotic mutant of *Arabidopsis thaliana* with leafy cotyledons. *Science* **258**: 1647–1650.

Meinke DW. 1994. Seed development in *Arabidopsis*. In: Meyerowitz EM, Somerville CR, eds. *Arabidopsis thaliana*. New York, NY, USA: Cold Spring Harbor Laboratory Press, 253–295.

Meinke DW. 2013a. Large-scale mutant analysis of seed development in *Arabidopsis*. In: Becraft PW, ed. *Seed genomics*. Ames, Iowa, USA: Wiley-Blackwell, 5–20.

Meinke DW. 2013b. A survey of dominant mutations in *Arabidopsis thaliana*. *Trends in Plant Science* **18**: 84–91.

Meinke DW, Franzmann LH, Nickle TC, Yeung EC. 1994. *Leafy cotyledon* mutants of *Arabidopsis*. *Plant Cell* **6**: 1049–1064.

Meinke DW, Sussex IM. 1979a. Embryo-lethal mutants of *Arabidopsis thaliana*: a model system for genetic analysis of plant embryo development. *Developmental Biology* **72**: 50–61.

Meinke DW, Sussex IM. 1979b. Isolation and characterization of embryo-lethal mutants of *Arabidopsis thaliana*. *Developmental Biology* **72**: 62–72.

Meyerowitz EM, Somerville CR. 1994. *Arabidopsis*. New York, NY, USA: Cold Spring Harbor Laboratory Press.

Michalko J, Dravecká M, Bollenbach T, Friml J. 2015. Embryo-lethal phenotypes in early *abp1* mutants are due to disruption of the neighboring *BSM* gene. *F1000Research* **4**: 1104.

Mimura M, Zallot R, Niehaus TD, Hasnain G, Gidda SK, Nguyen TN, Anderson EM, Mullen RT, Brown G, Yakunin AF *et al.* 2016. Arabidopsis *TH2* encodes the orphan enzyme thiamin monophosphate phosphatase. *Plant Cell* **28**: 2683–2696.

Minina EA, Reza SH, Gutierrez-Beltran E, Elander PH, Bozhkov PV, Moschou PN. 2017. The *Arabidopsis* homolog of Scc4/MAU2 is essential for embryogenesis. *Journal of Cell Science* **130**: 1051–1063.

Miséra S, Müller AJ, Weiland-Heidecker U, Jürgens G. 1994. The *FUSCA* genes of Arabidopsis: negative regulators of light responses. *Molecular and General Genetics* **244**: 242–252.

Mo X, Zhu Q, Li X, Li J, Zeng Q, Rong H, Zhang H, Wu P. 2006. The *hpa1* mutant of Arabidopsis reveals a crucial role of histidine homeostasis in root meristem maintenance. *Plant Physiology* **141**: 1425–1435.

Mori T, Igawa T, Tamiya G, Miyagishima SY, Berger F. 2014. Gamete attachment requires GEX2 for successful fertilization in *Arabidopsis*. *Current Biology* **24**: 170–175.

Mori T, Kuroiwa H, Higashiyama T, Kuroiwa T. 2006. GENERATIVE CELL SPECIFIC 1 is essential for angiosperm fertilization. *Nature Cell Biology* **8**: 64–71.

Moussu S, Doll NM, Chamot S, Brocard L, Creff A, Fourquin C, Widiez T, Nimchuk ZL, Ingram G. 2017. ZHOUPI and KERBEROS mediate embryo/endosperm separation by promoting the formation of an extracuticular sheath at the embryo surface. *Plant Cell* 29: 1642–1656.

Mu J, Tan H, Hong S, Liang Y, Zuo J. 2013. *Arabidopsis* transcription factor genes *NF-YA1, 5, 6*, and *9* play redundant roles in male gametogenesis, embryogenesis, and seed development. *Molecular Plant* 6: 188–201.

Müller AJ. 1963. Embryonentest zum Nachweis rezessiver Letalfaktoren bei *Arabidopsis thaliana. Biologisches Zentralblatt* 82: 133–163.

Muñoz-Nortes T, Candela H, Micol JL. 2017. Suitability of two distinct approaches for the high-throughput study of the post-embryonic effects of embryo-lethal mutations in Arabidopsis. *Scientific Reports* 7: 17010.

Muralla R, Chen E, Sweeney C, Gray JA, Dickerman A, Nikolau BJ, Meinke D. 2008. A bifunctional locus (*BIO3-BIO1*) required for biotin biosynthesis in Arabidopsis. *Plant Physiology* 146: 60–73.

Muralla R, Lloyd J, Meinke D. 2011. Molecular foundations of reproductive lethality in *Arabidopsis thaliana. PLoS ONE* 6: e28398.

Muralla R, Sweeney C, Stepansky A, Leustek T, Meinke D. 2007. Genetic dissection of histidine biosynthesis in Arabidopsis. *Plant Physiology* 144: 890–903.

Ngo QA, Baroux C, Guthörl D, Mozerov P, Collinge MA, Sundaresan V, Grossniklaus U. 2012. The Armadillo repeat gene *ZAK IXIK* promotes *Arabidopsis* early embryo and endosperm development through a distinctive gametophytic maternal effect. *Plant Cell* 24: 4026–4043.

Nickle TC, Meinke DW. 1998. A cytokinesis-defective mutant of *Arabidopsis (cyt1)* characterized by embryonic lethality, incomplete cell walls, and excessive callose. *The Plant Journal* 15: 321–332.

Nodine MD, Bartel DP. 2012. Maternal and paternal genomes contribute equally to the transcriptome of early plant embryos. *Nature* 482: 94–97.

Noutoshi Y, Ito T, Shinozaki K. 2005. *ALBINO AND PALE GREEN 10* encodes BBMII isomerase involved in histidine biosynthesis in *Arabidopsis thaliana. Plant and Cell Physiology* 46: 1165–1172.

Nowack MK, Grini PE, Jakoby MJ, Lafos M, Koncz C, Schnittger A. 2006. A positive signal from the fertilization of the egg cell sets off endosperm proliferation in angiosperm embryogenesis. *Nature Genetics* 38: 63–67.

Ooi SL, Pan X, Peyser BD, Ye P, Meluh PB, Yuan DS, Irizarry RA, Bader JS, Spencer FA, Boeke JD. 2006. Global synthetic-lethality analysis and yeast functional profiling. *Trends in Genetics* 22: 56–63.

Ortega D, Raynal M, Laudié M, Llauro C, Cooke R, Devic M, Genestier S, Picard G, Abad P, Contard P *et al.* 2002. Flanking sequence tags in *Arabidopsis thaliana* T-DNA insertion lines: a pilot study. *Comptes Rendus Biologies* **325**: 773–780.

Ouyang M, Ma J, Zou M, Guo J, Wang L, Lu C, Zhang L. 2010. The photosensitive *phs1* mutant is impaired in the riboflavin biogenesis pathway. *Journal of Plant Physiology* 167: 1466–1476.

Pagnussat GC, Yu HJ, Ngo QA, Rajani S, Mayalagu S, Johnson CS, Capron A, Xie LF, Ye D, Sundaresan V. 2005. Genetic and molecular identification of genes required for female gametophyte development and function in Arabidopsis. *Development* 132: 603–614.

Papini-Terzi FS, Galhardo RS, Farias LP, Menck CF, Van Sluys MA. 2003. Point mutation is responsible for Arabidopsis *tz-201* mutant phenotype affecting thiamin biosynthesis. *Plant and Cell Physiology* 44: 856–860.

Park M, Krause C, Karnahl M, Reichardt I, El Kasmi F, Mayer U, Stierhof YD, Hiller U, Strompen G, Bayer M *et al.* 2018. Concerted action of evolutionarily ancient and novel SNARE complexes in flowering-plant cytokinesis. *Developmental Cell* 44: 500–511.

Parker N, Wang Y, Meinke D. 2014. Natural variation in sensitivity to a loss of chloroplast translation in Arabidopsis. *Plant Physiology* 166: 2013–2027.

Parker N, Wang Y, Meinke D. 2016. Analysis of Arabidopsis accessions hypersensitive to a loss of chloroplast translation. *Plant Physiology* 172: 1862–1875.

Parry G. 2014. Components of the Arabidopsis nuclear pore complex play multiple diverse roles in control of plant growth. *Journal of Experimental Botany* 65: 6057–6067.

Patton DA, Franzmann LH, Meinke DW. 1991. Mapping genes essential for embryo development in *Arabidopsis thaliana. Molecular and General Genetics* 227: 337–347.

Patton DA, Meinke DW. 1990. Ultrastructure of arrested embryos from lethal mutants of *Arabidopsis thaliana. American Journal of Botany* 77: 653–661.

Patton DA, Schetter AL, Franzmann LH, Nelson K, Ward ER, Meinke DW. 1998. An embryo-defective mutant of Arabidopsis disrupted in the final step of biotin synthesis. *Plant Physiology* 116: 935–946.

Petersen LN, Marineo S, Mandalà S, Davids F, Sewell BT, Ingle RA. 2010. The missing link in plant histidine biosynthesis: Arabidopsis *myoinositol monophosphatase-like2* encodes a functional histidinol-phosphate phosphatase. *Plant Physiology* 152: 1186–1196.

Prigge MJ, Otsuga D, Alonso JM, Ecker JR, Drews GN, Clark SE. 2005. Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in Arabidopsis development. *Plant Cell* 17: 61–76.

Procissi A, de Laissardière S, Férault M, Vezon D, Pelletier G, Bonhomme S. 2001. Five gametophytic mutations affecting pollen development and pollen tube growth in *Arabidopsis thaliana. Genetics* 158: 1773–1783.

Przybyla-Toscano J, Roland M, Gaymard F, Couturier J, Rouhier N. 2018. Roles and maturation of iron-sulfur proteins in plastids. *Journal of Biological Inorganic Chemistry* 23: 545–566.

Raines T, Shanks C, Cheng CY, McPherson D, Argueso CT, Kim HJ, Franco-Zorrilla JM, López-Vidriero I, Solano R, Vaňková R *et al.* 2016. The cytokinin response factors modulate root and shoot growth and promote leaf senescence in Arabidopsis. *The Plant Journal* 85: 134–147.

Raschke M, Bürkle L, Müller N, Nunes-Nesi A, Fernie AR, Arigoni D, Amrhein N, Fitzpatrick TB. 2007. Vitamin B1 biosynthesis in plants requires the essential iron sulfur cluster protein, THIC. *Proceedings of the National Academy of Sciences, USA* 104: 19637–19642.

Rashotte AM, Mason MG, Hutchison CE, Ferreira FJ, Schaller GE, Kieber JJ. 2006. A subset of *Arabidopsis* AP2 transcription factors mediates cytokinin responses in concert with a two-component pathway. *Proceedings of the National Academy of Sciences, USA* 103: 11081–11085.

Rédei GP. 1970. *Arabidopsis thaliana* (L.) Heynh. A review of the genetics and biology. *Bibliographia Genetica* 20: 1–151.

Rédei GP. 1975. Induction of auxotrophic mutations in plants. In: Ledoux L, ed. *Genetic manipulations with plant material.* New York, NY, USA: Plenum Press, 329–350.

Rédei GP. 1982. Mutagen assay with Arabidopsis. A report of the U.S. Environmental Protection Agency Gene-Tox Program. *Mutation Research* 99: 243–255.

Rédei GP, Hirono Y. 1964. Linkage studies. *Arabidopsis Information Service* 1: 9–10.

Rinaldi MA, Fleming WA, Gonzalez KL, Park J, Ventura MJ, Patel AB, Bartel B. 2017. The PEX1 ATPase stabilizes PEX6 and plays essential roles in peroxisome biology. *Plant Physiology* 174: 2231–2247.

Robert HS, Grunewald W, Sauer M, Cannoot B, Soriano M, Swarup R, Weijers D, Bennett M, Boutilier K, Friml J. 2015. Plant embryogenesis requires AUX/LAX-mediated auxin influx. *Development* 142: 702–711.

Romani I, Tadini L, Rossi F, Masiero S, Pribil M, Jahns P, Kater M, Leister D, Pesaresi P. 2012. Versatile roles of Arabidopsis plastid ribosomal proteins in plant growth and development. *The Plant Journal* 72: 922–934.

Ron M, Alandete Saez M, Eshed Williams L, Fletcher JC, McCormick S. 2010. Proper regulation of a sperm-specific cis-nat-siRNA is essential for double fertilization in *Arabidopsis. Genes & Development* 24: 1010–1021.

Ruf S, Forner J, Hasse C, Kroop X, Seeger S, Schollbach L, Schadach A, Bock R. 2019. High-efficiency generation of fertile transplastomic *Arabidopsis* plants. *Nature Plants* 5: 282–289.

Ruiz-Sola MÁ, Barja MV, Manzano D, Llorente B, Schipper B, Beekwilder J, Rodriguez-Concepcion M. 2016. A single Arabidopsis gene encodes two differentially targeted geranylgeranyl diphosphate synthase isoforms. *Plant Physiology* 172: 1393–1402.

Růžička K, Zhang M, Campilho A, Bodi Z, Kashif M, Saleh M, Eeckhout D, El-Showk S, Li H, Zhong S *et al.* 2017. Identification of factors required for m^6A mRNA methylation in *Arabidopsis* reveals a role for the conserved E3 ubiquitin ligase HAKAI. *New Phytologist* 215: 157–172.

Saha P, Ray T, Tang Y, Dutta I, Evangelous NR, Kieliszewski MJ, Chen Y, Cannon MC. 2013. Self-rescue of an EXTENSIN mutant reveals alternative gene expression programs and candidate proteins for new cell wall assembly in *Arabidopsis*. *The Plant Journal* **75**: 104–116.

Sato Y, Yazawa K, Yoshida S, Tamaoki M, Nakajima N, Iwai H, Ishii T, Satoh S. 2011. Expression and functions of *myo*-inositol monophosphatase family genes in seed development of Arabidopsis. *Journal of Plant Research* **124**: 385–394.

Savage LJ, Imre KM, Hall DA, Last RL. 2013. Analysis of essential Arabidopsis nuclear genes encoding plastid-targeted proteins. *PLoS ONE* **8**: e73291.

Schneider T, Dinkins R, Robinson K, Shellhammer J, Meinke DW. 1989. An embryo-lethal mutant of *Arabidopsis thaliana* is a biotin auxotroph. *Developmental Biology* **131**: 161–167.

Schrick K, Mayer U, Horrichs A, Kuhnt C, Bellini C, Dangl J, Schmidt J, Jürgens G. 2000. FACKEL is a sterol C-14 reductase required for organized cell division and expansion in *Arabidopsis* embryogenesis. *Genes & Development* **14**: 1471–1484.

Sessions A, Burke E, Presting G, Aux G, McElver J, Patton D, Dietrich B, Ho P, Bacwaden J, Ko C *et al.* 2002. A high-throughput Arabidopsis reverse genetics system. *Plant Cell* **14**: 2985–2994.

Shin J, Jeong G, Park JY, Kim H, Lee I. 2018. *MUN* (*MERISTEM UNSTRUCTURED*), encoding a SPC24 homolog of NDC80 kinetochore complex, affects development through cell division in *Arabidopsis thaliana*. *The Plant Journal* **93**: 977–991.

Song JT, Lu H, Greenberg JT. 2004. Divergent roles in *Arabidopsis thaliana* development and defense of two homologous genes, *ABERRANT GROWTH AND DEATH2* and *AGD2-LIKE DEFENSE RESPONSE PROTEIN1*, encoding novel aminotransferases. *Plant Cell* **16**: 353–366.

Song XF, Guo P, Ren SC, Xu TT, Liu CM. 2013. Antagonistic peptide technology for functional dissection of *CLV3/ESR* genes in Arabidopsis. *Plant Physiology* **161**: 1076–1085.

Souter M, Topping J, Pullen M, Friml J, Palme K, Hackett R, Grierson D, Lindsey K. 2002. *hydra* mutants of Arabidopsis are defective in sterol profiles and auxin and ethylene signaling. *Plant Cell* **14**: 1017–1031.

Steinborn K, Maulbetsch C, Priester B, Trautmann S, Pacher T, Geiges B, Küttner F, Lepiniec L, Stierhof YD, Schwarz H *et al.* 2002. The *Arabidopsis PILZ* group genes encode tubulin-folding cofactor orthologs required for cell division but not cell growth. *Genes & Development* **16**: 959–971.

Sun YK, Gutmann B, Yap A, Kindgren P, Small I. 2018. Editing of chloroplast *rps14* by PPR editing factor EMB2261 is essential for *Arabidopsis* development. *Frontiers in Plant Science* **9**: 841.

Székely G, Ábrahám E, Cséplo Á, Rigó G, Zsigmond L, Csiszár J, Ayaydin F, Strizhov N, Jásik J, Schmelzer E *et al.* 2008. Duplicated *P5CS* genes of Arabidopsis play distinct roles in stress regulation and developmental control of proline biosynthesis. *The Plant Journal* **53**: 11–28.

Takahashi T, Mori T, Ueda K, Yamada L, Nagahara S, Higashiyama T, Sawada H, Igawa T. 2018. The male gamete membrane protein DMP9/DAU2 is required for double fertilization in flowering plants. *Development* **145**: doi: 10.1242/dev.170076.

Tambasco-Studart M, Tews I, Amrhein N, Fitzpatrick TB. 2007. Functional analysis of PDX2 from Arabidopsis, a glutaminase involved in vitamin B6 biosynthesis. *Plant Physiology* **144**: 915–925.

Taurino M, Costantini S, Domenico SD, Stefanelli F, Ruano G, Delgadillo MO, Sánchez-Serrano JJ, Sanmartín M, Santino A, Rojo E. 2018. SEIPIN proteins mediate lipid droplet biogenesis to promote pollen transmission and reduce seed dormancy. *Plant Physiology* **176**: 1531–1546.

Tax FE, Vernon DM. 2001. T-DNA-associated duplication/translocations in Arabidopsis. Implications for mutant analysis and functional genomics. *Plant Physiology* **126**: 1527–1538.

Toujani W, Muñoz-Bertomeu J, Flores-Tornero M, Rosa-Téllez S, Anoman AD, Alseekh S, Fernie AR, Ros R. 2013. Functional characterization of the plastidial 3-phosphoglycerate dehydrogenase family in Arabidopsis. *Plant Physiology* **163**: 1164–1178.

Tsuchisaka A, Yu G, Jin H, Alonso JM, Ecker JR, Zhang X, Gao S, Theologis A. 2009. A combinatorial interplay among the 1-Aminocyclopropan-1-Carboxylate isoforms regulates ethylene biosynthesis in *Arabidopsis thaliana*. *Genetics* **183**: 979–1003.

Tsugeki R, Kochieva EZ, Fedoroff NV. 1996. A transposon insertion in the *Arabidopsis SSR16* gene causes an embryo-defective lethal mutation. *The Plant Journal* **10**: 479–489.

Tzafrir I, Dickerman A, Brazhnik O, Nguyen Q, McElver J, Frye C, Patton D, Meinke D. 2003. The *Arabidopsis* SeedGenes Project. *Nucleic Acids Research* **31**: 90–93.

Tzafrir I, McElver JA, Liu CM, Yang LJ, Wu JQ, Martinez A, Patton DA, Meinke DW. 2002. Diversity of TITAN functions in Arabidopsis seed development. *Plant Physiology* **128**: 38–51.

Tzafrir I, Pena-Muralla R, Dickerman A, Berg M, Rogers R, Hutchens S, Sweeney TC, McElver J, Aux G, Patton D *et al.* 2004. Identification of genes required for embryo development in Arabidopsis. *Plant Physiology* **135**: 1206–1220.

Uwer U, Willmitzer L, Altmann T. 1998. Inactivation of a glycyl-tRNA synthetase leads to an arrest in plant embryo development. *Plant Cell* **10**: 1277–1294.

Vernon DM, Meinke DW. 1994. Embryogenic transformation of the suspensor in *twin*, a polyembryonic mutant of *Arabidopsis*. *Developmental Biology* **165**: 566–573.

Von Besser K, Frank AC, Johnson MA, Preuss D. 2006. *Arabidopsis HAP2 (GCS1)* is a sperm-specific gene required for pollen tube guidance and fertilization. *Development* **133**: 4761–4769.

Watanabe M, Mochida K, Kato T, Tabata S, Yoshimoto N, Noji M, Saito K. 2008. Comparative genomics and reverse genetics analysis reveal indispensable functions of the serine acetyltransferase gene family in *Arabidopsis*. *Plant Cell* **20**: 2484–2496.

Woodson JD, Chory J. 2012. Organelle signaling: how stressed chloroplasts communicate with the nucleus. *Current Biology* **22**: R690–R692.

Yadegari R, Kinoshita T, Lotan O, Cohen G, Katz A, Choi Y, Katz A, Nakashima K, Harada JJ, Goldberg RB *et al.* 2000. Mutations in the *FIE* and *MEA* genes that encode interacting polycomb proteins cause parent-of-origin effects on seed development by distinct mechanisms. *Plant Cell* **12**: 2367–2382.

Yadegari R, Paiva G, Laux T, Koltunow AM, Apuya N, Zimmerman JL, Fischer RL, Harada JJ, Goldberg RB. 1994. Cell differentiation and morphogenesis are uncoupled in Arabidopsis *raspberry* embryos. *Plant Cell* **6**: 1713–1729.

Yu H, Zhang F, Wang G, Liu Y, Liu D. 2013. Partial deficiency of isoleucine impairs root development and alters transcript levels of the genes involved in branched-chain amino acid and glucosinolate metabolism in *Arabidopsis*. *Journal of Experimental Botany* **64**: 599–612.

Yu Q, Lutz KA, Maliga P. 2017. Efficient plastid transformation in Arabidopsis. *Plant Physiology* **175**: 186–193.

Yu TY, Shi DQ, Jia PF, Tang J, Li HJ, Liu J, Yang WC. 2016. The Arabidopsis receptor kinase ZAR1 is required for zygote asymmetric division and its daughter cell fate. *PLoS Genetics* **12**: e1005933.

Zhan T, Boutros M. 2016. Towards a compendium of essential genes – From model organisms to synthetic lethality in cancer cells. *Critical Reviews in Biochemistry and Molecular Biology* **51**: 74–85.

Zhang JZ, Somerville CR. 1997. Suspensor-derived polyembryony caused by altered expression of valyl-tRNA synthetase in the *twn2* mutant of *Arabidopsis*. *Proceedings of the National Academy of Sciences, USA* **94**: 7349–7355.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Dataset S1 Overview of initial collection of 253 *emb* mutants.

Dataset S2 Details of initial collection of 253 *emb* mutants.

Dataset S3 Core collection of 354 *EMB* genes (264 cloned; 90 mapped, not cloned) and mutant alleles.

Dataset S4 Details of Syngenta/SeedGenes collection of 443 tagged *emb* mutants.

Dataset S5 Supplemental collection of 232 Syngenta putatively tagged *emb* mutants.

Dataset S6 Collection of 1272 Syngenta untagged and unresolved *emb* mutants.

Dataset S7 *Pigment-defective embryo* (*pde*) mutants in the Meinke lab collection.

Dataset S8 Curated dataset of 510 *EMB* genes of Arabidopsis.

Dataset S9 Double mutants with an embryo-defective phenotype.

Table S1 Additional information on mutants and genes included in Table 1.

Table S2 152 *EMB* genes with multiple mutant alleles in the Meinke lab collection.

Table S3 T-DNA vector information for the Syngenta mutant collection.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

About *New Phytologist*

- *New Phytologist* is an electronic (online-only) journal owned by the New Phytologist Trust, a **not-for-profit organization** dedicated to the promotion of plant science, facilitating projects from symposia to free access for our Tansley reviews and Tansley insights.
- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as ready' via *Early View* – our average time to decision is <26 days. There are **no page or colour charges** and a PDF version will be provided for each article.
- The journal is available online at Wiley Online Library. Visit **www.newphytologist.com** to search the articles and register for table of contents email alerts.
- If you have any questions, do get in touch with Central Office (np-centraloffice@lancaster.ac.uk) or, if it is more convenient, our USA Office (np-usaoffice@lancaster.ac.uk)
- For submission instructions, subscription and all the latest information visit **www.newphytologist.com**

See also the Commentary on this article by Ponce & Micol, **226**: 289–291.

© 2019 The Author
New Phytologist © 2019 New Phytologist Trust

14698137, 2020, 2, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.16071, Wiley Online Library on [14/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License