DR. JOHN CLIFTON-BROWN (Orcid ID : 0000-0001-6477-5452)

PROF. ANTOINE HARFOUCHE (Orcid ID : 0000-0003-3998-0694)

DR. LAWRENCE B SMART (Orcid ID : 0000-0002-7812-7736)

DR. DANNY AWTY-CARROLL (Orcid ID : 0000-0001-5855-0775)

DR. LINDSAY CLARK (Orcid ID : 0000-0002-3881-9252)

DR. LIN SHIOW-FEN HUANG (Orcid ID : 0000-0003-3079-326X)

DR. JON MCCALMONT (Orcid ID : 0000-0002-5978-9574)

DR. PAUL ROBSON (Orcid ID : 0000-0003-1841-3594)

PROF. GAIL TAYLOR (Orcid ID : 0000-0001-8470-6390)

Article type   : Platform

PART OF A SPECIAL ISSUE ON BIOBASED VALUE CHAINS FOR A GROWING BIOECONOMY

Breeding progress and preparedness for mass-scale deployment of perennial lignocellulosic biomass crops switchgrass, miscanthus, willow, and poplar

Running title: Breeding Biomass Crops for the Bioeconomy

John Clifton-Brown[1,*], Antoine Harfouche[2,*], Michael D. Casler[3], Huw Dylan Jones[1], William J. Macalpine[4], Donal Murphy-Bokern[5], Lawrence B. Smart[6], Anneli Adler[7, 8], Chris Ashman[1], Danny Awty-Carroll[1], Catherine Bastien[9], Sebastian Bopper[10], Vasile Botnari[11], Maryse Brancourt-Hulmel[12], Zhiyong Chen[13], Lindsay V. Clark[14], Salvatore Cosentino[15], Sue Dalton[1], Chris Davey[1], Oene Dolstra[16], Iain Donnison[1], Richard Flavell[17], Joerg Greef[18], Steve Hanley[4], Astley Hastings[19], Magnus Hertzberg[7], Tsai-Wen Hsu[20], Lin Huang[1], Antonella Iurato[1], Elaine Jensen[1], Xiaoli Jin[21], Uffe Jørgensen[22], Andreas Kiesel[23], Do-Soon Kim[24], Jianxiu Liu[25], Jon P. McCalmont[1], Bernard G. McMahon[26], Michal Mos[27], Paul Robson[1], Erik J. Sacks[14], Anatolii Sandu[11], Giovanni Scalici[15], Kai Schwarz[18], Danilo Scordia[15], Reza Shafiei[28], Ian Shield[4], Gancho Slavov[4], Brian J. Stanton[29], Kankshita Swaminathan[30], Gail Taylor[31], Andres F. Torres[16], Luisa M. Trindade[16], Timothy Tschaplinski[32], Jerry Tuskan[32], Toshihiko Yamada[33], Chang Yeon Yu[34], Ron -Fs Zalesny [35], Junqin Zong[25], Iris Lewandowski[23]

[1]Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, SY23 3EE, UK

[2]Department for Innovation in Biological, Agrofood and Forest systems, University of Tuscia, Via S. Camillo de Lellis, Viterbo 01100, Italy

[3]USDA-ARS, U.S. Dairy Forage Research Center, 1925 Linden Dr. Madison, WI 53706-1108;
[4]Rothamsted Research, Harpenden, AL5 2JQ, UK

[5]Lindenweg 12, Kroge-Ehrendorf, Lohne, D-49393, Germany

[6]Horticulture Section, School of Integrative Plant Science, Cornell University, Cornell AgriTech, Geneva, NY 14456, USA

[7]SweTree Technologies AB, Box 4095, SE-90403, Umeå, Sweden

[8]Institute of Crop Production Ecology, Swedish University of Agricultural Sciences, P.O Box 7016, S-75007, Uppsala, Sweden

[9]INRA-BIOFORA, 45075 Orléans, France

[10]University of Hohenheim, Institute of Plant Breeding, Seed Science and Population Genetics Department of Seed Science and Technology, Stuttgart 70593, Germany

[11]Institute of Genetics, Physiology and Plant Protection (IGFPP) of Academy of Sciences of Moldova, 20 Padurii St., Chisinau, MD-2002, Republic of Moldova

[12]INRA-AgroImpact, Site d'Estrées-Mons, CS 50136, 80203 Péronne cedex, France

[13]Insitute of Miscanthus, Hunan Agricultural University, Hunan Changsha 410128, China [14]

[14]University of Illinois, Department of Crop Sciences & Center for Advanced Bioenergy and Bioproducts Innovation, 279 Edward R Madigan Laboratory, MC-051, 1201 W. Gregory Dr., Urbana, IL 61801, USA

[15]Dipartimento di Agricolture Alimentazione e Ambiente, Università degli Studi di Catania, via S. Sofia 100, Catania, Italy

[16]Wageningen University & Research, Plant Breeding, P.O. Box 386, 6700 AJ Wageningen, The Netherlands

[17]Battersea, London, UK

[18]Julius Kuhn-Institut (JKI), Bundesforschungsinstitut fur Kulturpflanzen, Braunschweig, Germany

[19]University of Aberdeen, Institute of Biological and Environmental Science, 23 St Machar Drive, Aberdeen AB24 3UU, UK

[20]Taiwan Endemic Species Research Institute (TESRI), Nantou County, 552, Taiwan

[21]Department of Agronomy & The Key Laboratory of Crop Germplasm Resouce of Zhejiang Province, Zhejiang University, Hangzhou, Zhejiang 310058, China

[22]Aarhus University Centre for Circular Bioeconomy, Department of Agroecology, Blichers Allé 20, 8830 Tjele, Denmark

[23]University of Hohenheim, Institute of Crop Science, Department of Biobased Products and Energy Crops, Stuttgart 70593, Germany

[24]Department of Plant Sciences, Research Institute of Agriculture & Life Sciences, CALS, Seoul National University, Seoul, Korea

[25]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, 210014, China

[26]University of Minnesota – Duluth, Natural Resources Research Institute, 5013 Miller Trunk Highway,Duluth MN 55811, USA

[27]Energene sp. z o.o., Plac Solny 15, 50-062 Wrocław, Poland

[28]University of Dundee at James Hutton Institute, Errol Road, Invergowrie, Dundee, DD2 5DA, UK

[29]GreenWood Resources, Inc. 1500 SW 1st Avenue, Suite 1150, Portland, Oregon 9720, USA

[30]601 Genome Way, Huntsville, AL 35806, USA

[31]Biological Sciences, University of Southampton, B85/6041, Life Sciences Building, Highfield Campus, Southampton SO17 1BJ, UK

[32]The Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN, 37831, USA

[33]Field Science Centre for the Northern Biosphere, Hokkaido University, Sapporo, Japan

[34]103-1, College of Agriculture and Life Sciences 2, Kangwon National University, Chuncheon, 200-701, South Korea

[35]Forest Service, Northern Research Station, Institute for Applied Ecosystem Studies, 5985 Highway K, Rhinelander, WI 54501, USA

[*]These authors contributed equally to this work.

Correspondence: John Clifton-Brown, tel. 00441970823191, email: jhc@aber.ac.uk

Type of paper: Research Platform

Abstract

Genetic improvement through breeding is one of the key approaches to increasing biomass supply. This paper documents the breeding progress to date for four perennial biomass crops (PBCs) that have high output-input energy ratios: namely *Panicum virgatum* (switchgrass), species of the genera *Miscanthus* (miscanthus), *Salix* (willow) and *Populus* (poplar). For each crop, we report on the size of

germplasm collections, the efforts to date to phenotype and genotype the diversity available for breeding, and on the scale of breeding work as indicated by number of attempted deliberate crosses. We also report on the development of faster and more precise breeding using molecular breeding techniques. Poplar is the model tree for genetic studies and is furthest ahead in terms of biological knowledge and genetic resources. Linkage maps, transgenesis, and genome editing methods are now being used in commercially focused poplar breeding. These are in development in switchgrass, miscanthus and willow generating large genetic and phenotypic datasets requiring concomitant efforts in informatics to create summaries that can be accessed and used by practical breeders. Cultivars of switchgrass and miscanthus can be seed-based synthetic populations, semi-hybrids or clones. Willow and poplar cultivars are commercially deployed as clones. At local and regional level, the most advanced cultivars in each crop are at technology readiness levels which could be scaled to planting rates of thousands of hectares per year in about 5 years with existing commercial developers. Investment in further development of better cultivars is subject to current market failure and the long breeding cycles. We conclude that sustained public investment in breeding plays a key role in delivering future mass-scale deployment of PBCs.

Introduction

Increasing sustainable biomass production is an important component of the transition from a fossil fuel-based economy to renewables. Taking the UK as an example, Lovett *et al.* (2014) suggested that 1.4 million ha of marginal agricultural land could be used for biomass production without compromising food production. Assuming a biomass dry matter (DM) yield of 10 Mg ha^{-1}

and a calorific value of 18 GJ Mg^{-1} DM, 1.4 million ha would deliver around 28 TWh of electricity (with 40% biomass conversion efficiency) which would be ~8% of primary UK electricity generation (336 TWh in 2017 (DUKES, 2017)). To achieve this by 2050, planting rates of ~35,000 ha y^{-1} would be needed from 2022, in line with calculations by Evans (2017). The current annual planting rates in the UK are orders of magnitude short of these levels at only several hundred hectares per year. Similar scenarios have been generated for other countries (BMU, 2009; Scarlat *et al.*, 2015).

If perennial biomass crops (PBCs) are to make a real contribution to sustainable development they should be grown on agricultural land which is less suitable for food crops (Lewandowski, 2015). This economically 'marginal' land is typically characterised by abiotic stresses (drought, flooding, stoniness, steep slope, exposure to wind, and sub-optimal aspect), low nutrients, and /or contaminated soils (Tóth *et al.*, 2016). In these challenging environments, PBCs need resilience traits. They also need high output:input ratios for energy (typically 20 to 50) to deliver large carbon savings. Land may also be marginal due to environmental vulnerability. Much of the value for society from the genetic improvement of these crops depends on positive effects arising from highly productive perennial systems. In addition to producing biomass as a carbon source to replace fossil carbon, these crops reduce nitrate leaching (Pugesgaard *et al.*, 2015), making them good candidates to help fulfil Water Framework Directive (2000/60/EC) and can increase soil carbon storage during their production (McCalmont *et al.*, 2017).

The objective of this paper is to report on the preparedness for wide deployment by summarising the technical state-of-the-art in breeding of four important PBCs: namely switchgrass, miscanthus, willow and poplar. These four crops are the most promising and advanced PBCs for temperate regions and have therefore the focus here. Switchgrass and miscanthus are both rhizomatous grasses with C_4 photosynthesis while willow and poplar are trees with C_3 photosynthesis. Specifically, this paper 1) reviews available crop trait genetic diversity information; 2) assesses the progress of conventional breeding technologies for yield resilience and biomass quality; 3) reports on progress

with new molecular based breeding technologies to increase speed and precision of selection; and 4) discusses the requirements and next steps for breeding of PBCs, including commercial considerations in order to sustainably meet the biomass requirements of a growing worldwide bioeconomy.

We summarise the crop specific attributes, the location of breeding programmes, the current availability of commercial cultivars and yield expectations in selected environments (Table 1), and the generalised breeding targets for all PBCs (Table 2). Economic information relating to the current market value of the biomass and the investment in breeding are presented for different countries / regions in Table 3. We also present a comparison of the pre-breeding and conventional breeding efforts step-by-step, starting with wild germplasm collection and evaluation before wide crossing of wild relatives (Table 4). Hybridisation is followed by at least 6 years of selection and evaluation before commercial upscaling can begin (Figure 1). Recurrent selection, often over decades, is used within parent populations as part of an ongoing long-term process to produce hybrid vigour (Brummer, 1999). In the following sections the state-of-the-art and new opportunities of breeding switchgrass, miscanthus, willow and poplar are described. The application of modern breeding technologies is compared for the four crops in Table 5. It is most advanced in poplar, and is therefore described in most detail.

Switchgrass

Switchgrass is indigenous to the North American prairies. It is grown from seed and harvested annually using technology similar to that used for pastures. Based on collections from thousands of wild prairie remnants, the genetic resources are roughly divided into lowland and upland ecotypes and there are distinct clades within each ecotype which occur along both latitudinal and longitudinal gradients (Zhang *et al.*, 2011; Lu *et al.*, 2013; Evans *et al.*, 2018). Genotype-by-environment interactions (G × E) are strong and must be considered in breeding (Casler, 2012; Casler *et al.*, 2012). Adaptation to environment is regulated principally by responses to day-length and temperature.

There are also strong genotype × environment interactions between the drier western regions and the wetter eastern regions (Casler *et al.*, 2017).

The growing regions of North America are divided into four adaptation zones for switchgrass, each roughly corresponding to two official hardiness zones. The lowland ecotypes are generally late flowering, high yielding and adapted to warmer climates, but have lower drought and cold resistance than upland ecotypes (Casler, 2012; Casler *et al.*, 2012).

In 2015, the United States Department of Agriculture (USDA) National Plant Germplasm System, GRIN (https://www.ars-grin.gov/npgs/), had 181 switchgrass accessions, of which only 96 were available for distribution due to limitations associated with seed multiplication (Casler *et al.*, 2015). There are well over 2,000 additional uncatalogued accessions (Table 1) held by various universities but the USDA access to these is also constrained by the effort needed in seed multiplication. Switchgrass is a model herbaceous species for conducting scientific research on biomass (Sanderson *et al.*, 2006), but little funding is available for the critical pre-breeding work that is necessary to link this biological research to commercial breeding. More than a million genotypes from ~2,000 accessions (seed accessions contain many genotypes) have been phenotypically screened in spaced plant nurseries and ten thousand of the most useful have been genotyped with different technologies, depending on the technology available at the time when these were performed. From these characterised genotypes, parents are selected for exploratory pairwise crosses to produce synthetic populations within ecotypes. Switchgrass, like many grasses, is outcrossing due to a strong genetically controlled self-incompatibly (akin to the S-Z-locus system of other grasses; (Martinez-Reyna & Vogel, 2002)). Thus, the normal breeding approaches used are F_1 wide crosses and recurrent selection cycles within synthetic populations.

The scale of these programmes varies from small-scale conventional breeding, based solely on phenotypic selection (e.g. REAP Canada, Montreal, Quebec), to large programmes incorporating modern molecular breeding methods (e.g., USDA-ARS, Madison, Wisconsin). Early agronomic

research and biomass production efforts were focused on the seed-based multiplication of promising wild accessions from natural prairies. Cultivars Alamo, Kanlow, and Cave-in-Rock were popular due to high yield and moderate-to-wide adaptation. Conventional breeding approaches focussed on biomass production traits and have led to the development of five cultivars particularly suited to biomass production: Cimarron, EG1101, EG1102, EG2101, and Liberty. The first four of these represent the lowland ecotype and were developed either in Oklahoma or Georgia. Liberty is a derivative of lowland × upland hybrids developed in Nebraska following selection for late flowering, the high yield of the lowland ecotype and cold tolerance of the upland ecotype (Vogel *et al.*, 2014). These five cultivars were all approximately 25-30 years in the making, counting from the initiation of these breeding programmes. Many more biomass-type cultivars are expected within the next few years as these and other breeding programmes mature. The average rate of gain for biomass yield in long-term switchgrass breeding programmes has been 1-4% per year, depending on ecotype, population, and location of the breeding programme (Casler & Vogel, 2014; Casler *et al.*, 2018). The hybrid derivative Liberty has a biomass yield 43% higher than the better of its two parents (Casler & Vogel, 2014; Vogel *et al.*, 2014). The development of cold-tolerant and late-flowering lowland-ecotype populations for the northern USA has increased biomass yields by 27% (Casler *et al.*, 2018).

Currently, more than twenty recurrent selection populations are being managed in the USA to select parents for improved yield, yield resilience and compositional quality of the biomass. For the agronomic development and upscaling, high seed multiplication rates need to be combined with lower seed dormancy to both reduce crop establishment costs and risks. Expresso is the first cultivar with significantly reduced seed dormancy which is the first step toward development of domesticated populations (Casler *et al.*, 2015). Most phenotypic traits of interest to breeders require a minimum of 2 years to be fully expressed which results in a breeding cycle that is at least two years. More complicated breeding programmes, or traits that require more time to evaluate, can extend the breeding cycle to 4 to 8 years per generation, e.g., progeny testing for biomass yield or

field-based selection for cold tolerance. Breeding for a range of traits with such long cycles calls for the development of molecular methods to reduce time scales and improve breeding efficiency.

Two association panels of switchgrass have been phenotypically and genotypically characterized to identify quantitative trait loci (QTLs) that control important biomass traits. The northern panel consists of 60 populations, approximately 65% from the upland ecotype. The southern panel consists of 48 populations, approximately 65% from the lowland ecotype. Numerous QTLs have been identified within the northern panel to date (Grabowski *et al.*, 2017). Both panels are the subject of additional studies focused on biomass quality, flowering and phenology, and cold tolerance. Additionally, numerous linkage maps have been created by the pairwise crossing of individuals with divergent characteristics, often to generate four-way crosses that are analysed as pseudo-F_2 crosses (Okada *et al.*, 2010; Liu *et al.*, 2012; Serba *et al.*, 2013; Tornqvist *et al.*, 2018). Individual markers and QTLs identified can be used to design marker-assisted selection (MAS) programmes to accelerate breeding and increase its efficiency. Genomic prediction and selection (GS) holds even more promise with the potential to double or triple the rate of gain for biomass yield and other highly complex quantitative traits of switchgrass (Ramstein *et al.*, 2016; Casler & Ramstein, 2018). The genome of switchgrass has recently been made public through the Joint Genome Institute (https://phytozome.jgi.doe.gov/).

Transgenic approaches have been heavily relied upon to generate unique genetic variants, principally for traits related to biomass quality (Merrick & Fei, 2015). Switchgrass is highly transformable using either *Agrobacterium* mediated transformation or biolistics bombardment, but regeneration of plants is the bottleneck to these systems. Traditionally, plants from the cultivar Alamo were the only re-generable genotypes, but recent efforts have begun to identify more genotypes from different populations that are capable of both transformation and subsequent regeneration (Li & Qu, 2011; King *et al.*, 2014; Ogawa *et al.*, 2014; Ogawa *et al.*, 2016). Cell-wall recalcitrance and improved sugar release are the most common targets for modification (Fu *et al.*,

2011; Biswal *et al.*, 2018). Transgenic approaches have the potential to provide traits that cannot be bred using natural genetic variability. However, they will still require about 10-15 years and will cost $70-100 million for cultivar development and deployment (Harfouche *et al.*, 2011). In addition, there is commercial uncertainty due to the significant costs and unpredictable timescales and outcomes of the regulatory approval process in the countries targeted for seed sales. As seen in maize, one advantage of transgenic approaches is that they can easily be incorporated into F_1 hybrid cultivars (Casler, 2012; Casler *et al.*, 2012), but this does not decrease the time required for cultivar development due to field evaluation and seed multiplication requirements.

The potential impacts of unintentional gene flow and establishment of non-native transgene sequences in native prairie species via cross-pollination are also major issues for the environmental risk assessment. These limit further the commercialization of varieties made using these technologies. Although there is active research into switchgrass sterility mechanisms to curb unintended pollen-mediated gene transfer, it is likely that the first transgenic cultivars proposed for release in the USA will be met with considerable opposition due to the potential for pollen flow to remaining wild prairie sites which account for less than 1% of the original prairie land area and are highly protected by various governmental and non-governmental organizations (Casler *et al.*, 2015). Evidence for landscape-level, pollen-mediated gene flow from genetically modified *Agrostis* seed multiplication fields (over a mountain range) to pollinate wild relatives (Watrud *et al.*, 2004) confirms the challenge of using transgenic approaches. Looking ahead, genome editing technologies hold considerable promise for creating targeted changes in phenotype (Burris *et al.*, 2016; Liu *et al.*, 2018) and at least in some jurisdictions it is likely that cultivars resulting from gene-editing will not need the same regulatory approval as GMOs (Jones, 2015a). However in July 2018 the European Court of Justice (ECJ) ruled that cultivars carrying mutations resulting from gene editing should be regulated in the same way as GMOs. The ECJ ruled that such cultivars be distinguished from those arising from untargeted mutation breeding which is exempted from regulation under Directive 2001/18/EC.

Miscanthus

Miscanthus is indigenous to eastern Asia and Oceania where it is traditionally used for forage, thatching and paper making (Xi, 2000; Xi & Jezowkski, 2004). In the 1960s, the high biomass potential of a Japanese genotype, introduced to Europe by Danish nurseryman Aksel Olsen in 1935, was first recognised in Denmark (Linde-Laursen, 1993). Later, this accession was characterised, described and named as '*M. × giganteus*' (Greef & Deuter, 1993; Hodkinson & Renvoize, 2001), commonly abbreviated as *Mxg.* It is a naturally occurring interspecies triploid hybrid between tetraploid *M. sacchariflorus* (2n = 4x) and diploid *M. sinensis* (2n = 2x). Despite its favourable agronomic characteristics and ability to produce high yields in a wide range of environments in Europe (Kalinina *et al.*, 2017), the risks of reliance on it as a single clone have been recognised. Miscanthus, like switchgrass, is outcrossing due to self-incompatibility (Jiang *et al.*, 2017). Thus seeded hybrids are an option for commercial breeding. Miscanthus can also be vegetatively propagated by rhizome or *invitro* culture, which allows the development of clones. The breeding approaches are usually based on F1 crosses and recurrent selection cycles within the synthetic populations. There are several breeding programmes that target improvement of miscanthus traits including stress resilience, targeted regional adaptation, agronomic 'scalability' through cheaper propagation, faster establishment, lower moisture and ash contents and greater usable yield (Clifton-Brown *et al.*, 2017).

Germplasm collections specifically to support breeding for biomass started in the late 1980s and early 90s in Denmark, Germany and the UK (Clifton-Brown *et al.*, 2015). These collections have continued with successive expeditions from European and US teams assembling diverse collections from a wide geographic range in eastern Asia, including from China, Japan, South Korea, Russia, and Taiwan (Stewart *et al.*, 2009; Hodkinson *et al.*, 2015). Three key miscanthus species for biomass production are *M. sinensis*, *M. floridulus* and *M. sacchariflorus*. *M. sinensis* is widely distributed throughout eastern Asia, with an adaptive range from the sub-tropics to southern Russia (Zhao *et al.*,

2013). This species has small rhizomes and produces many tightly packed shoots forming a 'tuft'. *M. floridulus* has a more southerly adaptive range with a rather similar morphology to *M. sinensis*, but grows taller with thicker stems and is evergreen and less cold tolerant than the other miscanthus species. *M. sacchariflorus* is the most northern-adapted species ranging to 50 °N in eastern Russia (Clark *et al.*, 2016). Populations of diploid and tetraploid *M. sacchariflorus* are found in China (Xi, 2000) and South Korea (Yook, 2016), and eastern Russia, but only tetraploids have been found in Japan (Clark *et al.*, 2018).

Germplasm has been assembled from multiple collections over the last century, though some early collections are poorly documented. This historical germplasm has been used to initiate breeding programmes largely based on phenotypic and genotypic characterisation. As many of the accessions from these collections are 'origin unknown', crucial environmental envelope data are not available. UK-led expeditions started in 2006 and continued until 2011 with European and Asian partners and have built up a comprehensive collection of 1,500 accessions from 500 sites across Eastern Asia, including China, Japan, South Korea and Taiwan. These collections were guided using spatial climatic data to identify variation in abiotic stress tolerance. Accessions from these recent collections were planted, following quarantine, in multi-location nursery trials at several locations in Europe to examine trait expression in different environments. Based on the resulting phenotypic and molecular marker data, several studies (1) characterised patterns of population genetic structure (Slavov *et al.*, 2013; Slavov *et al.*, 2014); (2) evaluated the statistical power of Genome-wide association study (GWAS) and identified preliminary marker-trait associations (Slavov *et al.*, 2013; Slavov *et al.*, 2014); and (3) assessed the potential of genomic prediction (Slavov *et al.*, 2014; Davey *et al.*, 2017; Slavov *et al.*, 2018b). Genomic index selection in particular offers the possibility of exploring scenarios for different locations or industrial markets (Slavov *et al.*, 2018a; Slavov *et al.*, 2018b).

Separately, US-led expeditions also collected about 1,500 accessions between 2010 and 2014 (Clark *et al.*, 2014; Clark *et al.*, 2015; Clark *et al.*, 2016; Clark *et al.*, 2018). A comprehensive genetic analysis of the population structure has been produced by RADseq for *M. sinensis* (Clark *et al.*, 2015; van der Weijde *et al.*, 2017c) and *M. sacchariflorus* (Clark *et al.*, 2018). Multi-location replicated field trials have also been conducted on these materials in North America and in Asia. GWAS has been conducted for both *M. sinensis* and a subset of *M. sacchariflorus* accessions (Clark *et al.*, 2016). To date, about 75% of these recent US-led collections are in nursery trials outside the USA. Due to lengthy USA quarantine procedures, these are not yet available for breeding in the USA. However, molecular analyses have allowed us to identify and prioritize sets of genotypes that best encompass the genetic variation in each species.

While most *M. sinensis* accessions flower in northern Europe, very few *M. sacchariflorus* accessions flower even in heated glasshouses. For this reason, the European programmes in the UK, Netherlands and France have performed mainly *M. sinensis* (intraspecies) hybridisations (Table 4). Selected progeny become the parents of later generations (recurrent selection, as in switchgrass). Seed sets of up to 400 seed per panicle occur in *M. sinensis*. In Aberystwyth and Illinois, significant efforts to induce synchronous flowering in *M. sacchariflorus* and *M. sinensis* have been made because interspecies hybrids have proven higher yield performance and wide adaptability (Kalinina *et al.*, 2017). In interspecies pairwise crosses in glasshouses, breathable bags and/or large crossing tubes or chambers in which 2 or more whole plants fit are used for pollination control. Seed sets are lower in bags than in the open air because bags restrict pollen movement whilst increasing temperatures and reducing humidity(Clifton-Brown *et al.*, 2018). About 30% of attempted crosses produced 10 to 60 seeds per bagged panicle. The seed (thousand seed mass ranges from 0.5 to 0.9 g) are threshed from the inflorescences and sown into modular trays to produce plug plants, which are then planted in field nurseries to identify key parental combinations.

A breeding programme of this scale must serve the needs of different environments, accepting the common purpose is to optimise the interception of solar radiation. An ideal hybrid for a given environment combines adaptation to date of emergence with optimisation of traits such as height, number of stems per plant, flowering and senescence time to optimise solar interception to produce a high biomass yield with low moisture content at harvest (Robson *et al.*, 2013a; Robson *et al.*, 2013b). By 2013/2014, conventional breeding in Europe had produced intra- and interspecific fertile seeded hybrids. When a cohort (typically about 5) of outstanding crosses have been identified, it is important to work on related upscaling matters in parallel. These are:

- Assessment of the yield and critical traits in selected hybrids using a network of field trials.

- Efficient cloning of the seed parents. While *in vitro* and macro cloning techniques are used some genotypes are amenable to neither technique.

- High seed production from field crossing trials conducted in locations where flowering in both seed and pollen parents are likely to happen synchronously.

- Scalable and adapted harvesting, threshing and seed processing methods for producing high seed quality.

The results of these parallel activities need to be combined to identify the up-scaling pathway for each hybrid; if this cannot be achieved the hybrid will likely not be commercially viable. The UK-led programme with partners in Italy and Germany show that seed-based multiplication rates of 1:2,000 are achievable for They provide potential routes several interspecific hybrids (Clifton-Brown *et al.*, 2017). The multiplication rate of *M. sinensis* is higher, probably 1:5,000-10,000. Conventional cloning from rhizome is limited to around 1:20, i.e. one ha of rhizome production could supply around 20 ha of new plantation.

Multi-location field testing of wild and novel miscanthus hybrids selected by breeding programmes in the Netherlands and the UK was performed as part of the project Optimizing

Miscanthus Biomass Production (OPTIMISC, 2012-2016). These trials showed that commercial yields and biomass qualities (Kiesel *et al.*, 2017; van der Weijde *et al.*, 2017a; Van der Weijde *et al.*, 2017d) could be produced in a wide range of climates and soil conditions from the temperate maritime climate of western Wales to the continental climate of eastern Russia and the Ukraine (Kalinina *et al.*, 2017). Extensive environmental measurements of soil and climate, combined with growth monitoring is being used to understand abiotic stresses (Nunn *et al.*, 2017; Van der Weijde *et al.*, 2017b) and develop genotype specific scenarios similar to those reported earlier in Hastings *et al.* (2009). Phenomics experiments on drought tolerance have been conducted on wild and improved germplasm (Malinowska *et al.*, 2017; Van der Weijde *et al.*, 2017b). Recently produced interspecific hybrids displaying exceptional yield under drought (~30% greater than control *Mxg*) in field trials in Poland and Moldova are being further studied in detail in the phenomics and genomics facility at Aberystwyth to better understand gene-trait associations which can be fed back into breeding.

Intraspecific seeded hybrids of *M. sinensis* produced in the Netherlands and interspecific *M. sacchariflorus x M. sinensis* hybrids produced by the UK-led breeding programme have entered yield testing in 2018 with the recently EU-funded project 'GRowing Advanced industrial Crops on marginal lands for biorEfineries (GRACE)' (https://www.grace-bbi.eu/). Substantial variation in biomass quality for saccharification efficiency (glucose release as % of dry matter), ash content, and melting point has already been generated in intraspecific *M. sinensis* hybrids (Van der Weijde *et al.*, 2017d) across environments (van der Weijde *et al.*, 2017a). GRACE aims to establish more than 20 hectares of new inter- and intraspecific seeded hybrids across 6 European countries. This project is building the know-how and agronomy needed to transition from small research plots to commercial scale field sites and linking biomass production directly to industrial applications. The biomass produced by hybrids in different locations will be supplied to innovative industrial end-users making a wide range of bio-based products, both for chemicals and energy. In the U.S., multi-location yield were initiated in 2018 to evaluate new triploid *M. $\times$ giganteus* genotypes developed at Illinois. Currently, infertile hybrids are favoured in the USA because this eliminates the risk of invasiveness from naturally

dispersed, viable seed. The precautionary principle is applied as fertile miscanthus has naturalised in several states (Quinn *et al.*, 2010). In North European multi-location field trials in the EMI and OPTIMISC projects have shown there is minimal risk of invasiveness even in years when fertile flowering hybrids produce viable seed. Naturalised stands have not established here due perhaps to low dormancy, poor overwintering, and low seedling competitive strength. In addition to breeding for non-shattering or sterile seeded hybrids, Quinn *et al.* (2010) suggest management strategies which can further minimize environmental opportunities to manage the risk of invasiveness.

Molecular breeding and biotechnology: In miscanthus, new plant breeding techniques (Table 5) have focussed on developing molecular markers for breeding in Europe, the USA, South Korea and Japan. There are several publications on QTL mapping populations for key traits such as flowering (Atienza *et al.*, 2003a) and compositional traits (Atienza *et al.*, 2003c). In the USA and UK, independent and interconnected bi-parental 'mapping' families have been studied (Gifford *et al.*, 2015; Dong *et al.*, 2018) alongside panels of diverse germplasm accessions for GWAS (Slavov *et al.*, 2014). Further developments calibrating GS with very large panels of parents and cross progeny are underway (Davey *et al.*, 2017). The recently completed first miscanthus reference genome sequence is expected to improve the efficiency of MAS strategies, and especially GWAS (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Msinensis_er). For example, without a reference genome sequence Clark *et al.* (2014) obtained 21,207 RAD-seq SNPs (single nucleotide polymorphisms) on a panel of 767 miscanthus genotypes (mostly *M. sinensis*), but subsequent reanalysis of the RAD-seq data using the new reference genome resulted in hundreds of thousands of SNPs being called.

Robust and effective *in vitro* regeneration systems have been developed for *Miscanthus sinensis*, *M. × giganteus* and *M. sacchariflorus* (Wang *et al.*, 2011; Zhang *et al.*, 2012; Dalton, 2013; Guo *et al.*, 2013; Rambaud *et al.*, 2013; Hwang *et al.*, 2014a; Ślusarkiewicz-Jarzina *et al.*, 2017). However, there is still significant genotype-specificity and these methods need 'in-house' optimisation and

development to be used routinely. They provide potential routes for rapid clonal propagation and also as a basis for genetic transformation. Stable transformation using both biolistics (Wang *et al.* 2011) and *Agrobacterium tumefaciens* DNA-delivery methods (Hwang *et al.* 2014) have been achieved in *M. sinensis*. Development of miscanthus transformation and gene editing to generate diplogametes for producing seed-propagated triploid hybrids are performed as part of the French project MISEDIT (miscanthus gene editing for seed propagated triploids). There are no reports of genome editing in any miscanthus species but new breeding innovations, including genome editing, are particularly relevant in this slow-to-breed, non-food, bioenergy crop (Table 4).

Willow

Willow (*Salix* spp.) is a very diverse group of catkin-bearing trees and shrubs. Willow belongs to the family *Salicaceae*, which also includes the *Populus* genus. There are approximately 350 willow species (Argus, 2007), found mostly in temperate and arctic zones in the northern hemisphere. A few are adapted to subtropical and tropical zones. The centre of diversity is believed to be in Asia, with over 200 species in China. Around 120 species are found in the former Soviet Union, over 100 in North America and around 65 species in Europe, and one species is native to South America (Dickmann & Kuzovkina, 2008; Karp *et al.*, 2011). Willows are dioecious, thus obligate outcrossers, and highly heterozygous. The haploid chromosome number is 19 (Hanley & Karp, 2014). Around 40% of species are polyploid (Suda & Argus, 1968), ranging from triploids to the atypical dodecaploid *S. maxxaliana* with 2n=190 (Zsuffa *et al.*, 1984).

Although almost exclusively native to the Northern Hemisphere, willow has been grown around the globe for many thousands of years to support a wide range of applications (Stott, 1992;

Kuzovkina & Quigley, 2005). However, it has been the focus of domestication for bioenergy purposes for only a relatively short period, since the 1970s in North America and Europe. For bioenergy, breeders have focused their efforts on the shrub willows (sub-genus *Vetix*) because of their rapid juvenile growth rates as a response to coppicing on a 2-4 year cycle that can be accomplished using farm machinery rather than forestry equipment (Smart & Cameron, 2012; Shield *et al.*, 2015).

Since shrub willow was not generally recognised as an agricultural crop until very recently, there has been little commitment to building and maintaining germplasm repositories of willow to support long-term breeding. One exception is the UK, where a large and well characterized *Salix* germplasm collection comprising over 1,500 accessions is held at Rothamsted Research (Stott, 1992; Trybush *et al.*, 2008). Originally initiated for use in basketry in 1923, accessions have been added ever since. In the USA, a germplasm collection of >350 accessions is located at Cornell University to support the breeding programme there. The UK and Cornell collections have a relatively small number of accessions in common (around 20). Taken together, they represent much of the species diversity, but only a small fraction of the overall genetic diversity within the genus. There are three active willow breeding programmes in Europe: Rothamsted Research (UK), Salixenergi Europa AB (SEE) and a programme at the University of Warmia and Mazury in Olsztyn (Poland) (abbreviations used in Table 1). There is one active USA programme based at Cornell University. Cultivars are still being marketed by the European Willow Breeding Programme (EWBP) (UK), which was actively breeding biomass varieties from 1996-2002. Cultivars are protected by plant breeders' rights (PBRs) in Europe and by plant patents in the USA. The sharing of genetic resources in the willow community is generally regulated by material transfer agreements (MTA) and tailored licensing agreements, although the import of cuttings into North America is prohibited except under special quarantine permit conditions.

Efforts to augment breeding germplasm collection from nature are continuing, with phenotypic screening of wild germplasm performed in field experiments with 177 *S. purpurea* genotypes in the

USA (at sites in Geneva and Portland, NY and Morgantown, WV) that have been genotyped using genotyping-by-sequencing (GBS) (Elshire *et al.*, 2011). In addition there are approximately 400 accessions of *S. viminalis* in Europe (near Pustnäs, Uppsala, Sweden and Woburn, UK (Berlin *et al.*, 2014; Hallingbäck *et al.*, 2016). The *S. viminalis* accessions were initially genotyped using 38 simple sequence repeats (SSR) markers to assess genetic diversity, and screened with ~1,600 SNPs in genes of potential interest for phenology and biomass traits. Genetics and genomics, combined with extensive phenotyping, have substantially improved the genetic basis of biomass related traits in willow and are now being developed in targeted breeding via MAS. This underpinning work has been conducted on large specifically developed bi-parental *Salix* mapping populations (Hanley & Karp, 2014; Zhou *et al.*, 2018), as well as GWAS panels (Hallingbäck *et al.*, 2016).

Once promising parental combinations are identified, crosses are usually performed using fresh pollen from material that has been subject to a phased removal from cold storage (-4°C) (Mosseler, 1990; Lindegaard & Barker, 1997; Macalpine *et al.*, 2008). Pollen storage is useful in certain interspecific combinations where flowering is not naturally synchronised. This can be overcome by using pollen collection and storage protocol which involves extracting pollen using toluene (Kopp *et al.*, 2002).

The main breeding approach to improve willow yields relies on species hybridization to capture hybrid vigour (Serapiglia *et al.*, 2014; Fabio *et al.*, 2017). In the absence of genotypic models for heterosis, breeders have extensively tested general and specific combining ability of parents to produce superior progeny. The UK breeding programmes (EWBP 1996-2002 and Rothamsted Research from 2003 on) have performed more than 1,500 exploratory cross pollinations. The Cornell programme has successfully completed about 550 crosses since 1998. Investment into the characterisation of genetic diversity combined with progeny tests from exploratory crosses has been used to produce hundreds of targeted intraspecies crosses in the UK and US respectively (see Table 1). To achieve long-term gains beyond F_1 hybrids, four intraspecific recurrent selection populations

have been created in the UK (for *S. dasyclados, S. viminalis* and *S. miyabeana*) and Cornell is pursuing recurrent selection of *S. purpurea*. Interspecific hybridisations with genotypes selected from the recurrent selection cycles are well advanced in willow, with such crosses to date totalling 420 in the UK and over 100 in the USA.

While species hybridisation is common in *Salix*, it is not universal. Of the crosses attempted, about 50 % hybridise and produce seed (Macalpine *et al.*, 2010). As the viability of seed from successful crosses is short (a matter of days at ambient temperatures), proper seed rearing and storage protocols are essential (Maroder *et al.*, 2000).

Progeny from crosses are treated in different ways among the breeding programmes at the seedling stage. In the USA, seedlings are planted into an irrigated field where plants are screened for two seasons before being progressed to further field trials. In the UK, seedlings are planted into trays of compost where they remain containerised in an irrigated nursery for the remainder of year one. In the UK, seedlings are subject to two rounds of selection in the nursery year. The first round takes place in September to select against susceptibility to rust infection (*Melampsora* spp.). A second round of selection in winter assesses tip damage from frost and giant willow aphid infestation. In the USA where the rust pressure is lower, screening for *Melampsora* spp. cannot be performed at the nursery stage. Both programmes monitor *Melampsora* spp., pest susceptibility, yield and architecture over multiple years in field trials. Selected material is subject to two rounds of field trials followed by a final multi-location yield trial to identify varieties for commercialisation.

Promising selections (i.e., potential cultivars) need to be clonally propagated. A rapid, *in vitro* tissue-culture propagation method has been developed (Palomo-Ríos *et al.*, 2015). This method can generate about 5,000 viable, transplantable clones from a single plant in just 24 weeks. An *in vitro* system can also accommodate early selection via molecular or biochemical markers to increase selection speed. Conventional breeding systems take 13 years via four rounds of selection from

crossing to selecting a variety (Figure 1), but this has the potential to be reduced to 7 years if micro-propagation and MAS selection is adopted (Hanley & Karp, 2014; Palomo-Ríos *et al.*, 2015).

Willows are currently propagated commercially by planting winter-dormant stem cuttings in spring. Commercial planting systems for willow use mechanical planters that cut and insert stem sections from whips into a well prepared soil. One hectare of stock plants grown in specific multiplication beds planted at 40,000 plants per ha produces planting material for 80 hectares of commercial short rotation coppice willow annually (planted at 15,000 cuttings per hectare) (Whittaker *et al.*, 2016). When commercial plantations are established, the industry standard is to plant intimate mixtures of ~5 diverse rust (*Melampsora spp.*) resistant varieties (McCracken & Dawson, 1997; van den Broek *et al.*, 2001).

The foundations for using new plant breeding techniques have been established with funding from both the public and private sectors. To establish QTL maps, 16 mapping populations from bi-parental crosses are under study in the UK. Nine are under study in the USA. The average number of individuals in these families ranges from 150 to 947 (Hanley & Karp, 2014). GS is also being evaluated in *S. viminalis*, and preliminary results indicate that multiomic approaches combining genomic and metabolomic data have great potential (Slavov & Davey, 2017). For both QTL and GS approaches, the field phenotyping demands are large as several thousand individuals need to be phenotyped for a wide range of traits. These include: dates of bud burst and growth succession, stem height, stem density, wood density and disease resistance. The greater the number of individuals the more precise are the QTL marker maps and GS models. However, the logistical and financial challenges of phenotyping large numbers of individuals are considerable, because the willow crop is >5 m tall in the second year. There is tremendous potential to improve the throughput of phenotyping using unmanned aerial systems, which is being tested in the USDA National Institute of Food and Agriculture (NIFA) Willow SkyCAP project at Cornell. Further, investment in these

approaches needs to be sustained over many years fully realise the potential of a marker assisted selection programme for willow.

To date, despite considerable efforts in Europe and the USA to establish a routine transformation system, there has not been a breakthrough in willow, but attempts are ongoing. As some form of transformation is typically a pre-requisite for genome editing techniques, these have not yet been applied to willow.

In Europe there are 53 short rotation coppice (SRC) biomass willow cultivars registered with the Community Plant Variety Office (CPVO) for PBRs, of which ~25 are available commercially in the UK. There are 8 patented cultivars commercially available in the USA. In Sweden there are 9 commercial cultivars registered in Europe and two others which are unregistered (http://salixenergi.se/planting-material/). Furthermore, there are about 20 pre-commercial hybrids in final yield trials both the USA and UK. It has been estimated that it would take two years to produce the stock required to plant 50 ha commercially from the plant stock in the final yield trials. Breeding programmes have already delivered rust resistant varieties and increases in yield to the market. The adoption of advanced breeding technologies will likely lead to a step change in improving traits of interest.

Poplar

Poplar, a fast growing tree from the northern hemisphere with a small genome size, has been adopted for commercial forestry and scientific purposes. The genus *Populus* consists of about 29 species, classified in six different sections: *Populus* (formerly *Leuce*), *Tacamahaca*, *Aigeiros*, *Abaso*, *Turanga* and *Leucoides* (Eckenwalder, 1996). The *Populus* species of most interest for breeding and testing in the USA and Europe are *P. nigra, P. deltoides*, *P. maximowiczii* and *P. trichocarpa* (Stanton, 2014). *Populus* clones for biomass production are being developed by intra- and interspecies hybridization (van der Schoot *et al.*, 2000; Richardson *et al.*, 2014; DeWoody *et al.*, 2015). Recurrent selection approaches are used for gradual population improvement and to create elite clonal lines

for commercialization (Neale & Kremer, 2011; Berguson *et al.*, 2017). Currently, poplar breeding in the USA occurs in industrial and academic programmes located in the Southeast, the Midwest, and the Pacific Northwest. These use six species and five inter-specific taxa (Stanton, 2014).

The southeastern programme historically focused on recurrent selection of *P. deltoides* from accessions made in the lower Mississippi River alluvial plain (Robison *et al.*, 2006). More recently the genetic base has been broadened to produce interspecific hybrids with resistance to the fungal infection *Septoria musiva,* which causes cankers.

In the Midwest of the USA, population improvement efforts are focused on *P. deltoides* selections from native provenances and hybrid crosses with accessions introduced from Europe. Inter-specific, intercontinental (Europe and America) hybrid crosses between *P. nigra* and *P. deltoides* (*P. × canadensis*) are behind many of the leading commercial hybrids which are the most advanced breeding materials for many applications and regions. In Minnesota, previous breeding experience and efforts utilizing *P. maximowiczii* and *P. trichocarpa* have been discontinued due to *Septoria* susceptibility and a lack of cold hardiness (Berguson *et al.,* 2017). Traits targeted for improvement include yield/growth rate, cold hardiness, adventitious rooting, resistance to *Septoria* and *Melampsora* leaf rust, and stem form. The Upper Midwest programme also carries out wide hybridizations within the section Populus. The *P. × wettsteinii (P. tremula × P. tremuloides)* taxon is bred for gains in growth rate, wood quality, and resistance to the fungus *Entoleuca mammata* which causes hypoxylon canker (David & Anderson, 2002).

In the Pacific Northwest, GreenWood Resources Inc. leads poplar breeding that emphasizes inter-specific hybrid improvement of *P. × generosa* (*P. deltoides × P. trichocarpa* and reciprocal) and *P. deltoides × P. maximowiczii* taxa for coastal regions, and the *P. × canadensis* taxon for the drier, continental regions. Intra-specific improvement of second-

generation breeding populations of *P. deltoides, P. nigra, P. maximowiczii* and *P. trichocarpa* are also involved (Stanton *et al.*, 2010). The present focus of GreenWood Resources's hybridization is bioenergy feedstock improvement concentrating on coppice yield, wood specific gravity and rate of sugar release.

Industrial interest in poplar in the USA has historically come from the pulp and paper sector, although veneer and dimensional lumber markets have been pursued at times. Currently, the biomass market for liquid transportation fuels is being emphasized, along with the use of traditional and improved poplar genotypes for ecosystem services such as phytoremediation (Tuskan & Walsh, 2001; Zalesny *et al.*, 2016).

In Europe there are breeding programmes in France, Germany, Italy and Sweden. These include: (i) Alasia Franco Vivai (AFV) programme in northern Italy; (ii) the French programme led by the poplar Scientific Interest Group (GIS Peuplier) and carried out collaboratively between the National Institute for Agricultural Research (INRA), the National Research Unit of Science and Technology for Environment and Agriculture (IRSTEA) and the Forest, Cellulose, Wood, Construction and Furniture Technology Institute (FCBA); (iii) the German programme at Northwest German Forest Research Station (NW-FVA) at Hannoversch Münden; and (iv) the Swedish programme at the Swedish University of Agricultural Sciences and SweTree Technologies AB (Table 1).

AFV leads an Italian poplar breeding programme using extensive field-grown germplasm collections of *P. alba*, *P. deltoides, P. nigra,* and *P. trichocarpa*. While inter-specific hybridization uses several taxa, the focus is on *P. × canadensis*. The breeding program addresses disease resistance (*Marssonina brunnea, Melampsora larici-populina, Discosporium populeum*, and poplar mosaic virus), growth rate, and photoperiod

adaptation. AFV and GreenWood Resources collaborate in poplar improvement in Europe through the exchange of frozen pollen and seed for reciprocal breeding projects. Plantations in Poland and Romania are currently the focus of the collaboration.

The ongoing French GIS Peuplier is developing a long term breeding programme based on intraspecific recurrent selection for the four parental species (*P. deltoides*, *P. trichocarpa*, *P. nigra* and *P. maximowiczii)* designed to better benefit from hybrid vigour demonstrated by the interspecific crosses *P. canadensis , P. deltoides* × *P. trichocarpa,* and *P. trichocarpa* × *P. maximowiczii*. Current selection priorities are targeting adaptation to soil and climate conditions, resistance and tolerance to the most economically important diseases and pests, high volume production under SRC and traditional poplar cultivation regimes as well as wood quality of interest by different markets. Currently, genomic selection is under exploration to increase selection accuracy and selection intensity while maintaining genetic diversity over generations.

The German NW-FVA programme is breeding inter-sectional Aigeiros – Tacamahaca hybrids with a focus on resistance to *Pollaccia elegans, Xanthomonas populi, Dothichiza* spp.*, Marssonina brunnea,* and *Melampsora* spp. (Stanton, 2014). Various cross combinations of *P. maximowiczii*, *P. trichocarpa, P. nigra,* and *P. deltoides* have led to new cultivars suitable for deployment in varietal mixtures of five to ten genotypes of complementary stature, high productivity, and phenotypic stability (Weisgerber, 1993). The current priority is the selection of cultivars for high-yield, short rotation biomass production. Six hundred *P. nigra* genotypes are maintained in an *ex situ* conservation programme. An *in situ P. nigra* conservation effort involves an inventory of native stands which have been molecular fingerprinted for identity and diversity.

The Swedish programme is concentrating on locally adapted genotypes used for short rotation forestry (SRF) because these meet the needs of the current pulping markets. Several field trials have

shown that commercialpoplar clones tested and deployed in Southern and Central Europe are not well adapted to photoperiods and low temperatures in Sweden and in the Baltics. Consequently, Swedish University of Agricultural Sciences and SweTree Technologies AB started breeding in Sweden in 1990's to produce poplar clones better adapted to local climates and markets.

Molecular breeding technologies: Poplar genetic improvement cannot be rapidly achieved through traditional methods alone because of the long breeding cycles, outcrossing breeding systems, and high heterozygosity. Integrating modern genetic, genomic, and phenomic techniques with conventional breeding has the potential to expedite poplar improvement.

The genome of poplar has been sequenced (Tuskan *et al.*, 2006). It has an estimated genome size of 485 $\pm$ 10 Mbp divided into 19 chromosomes. This is smaller than other PBCs and makes poplar more amenable to genetic engineering (transgenesis), GS, and genome editing. Poplar has seen major investment in both the USA and Europe, being the model system for woody perennial plant genetics and genomics research.

Targets for genetic modification: Traits targeted include wood properties (lignin content and composition), early/late flowering, male sterility to address biosafety regulation issues, enhanced yield traits, and herbicide tolerance. These extensive transgenic experiments have shown differences in recalcitrance to *in vitro* regeneration and genetic transformation in some of the most important commercial hybrid poplars (Alburquerque *et al.*, 2016). Further, transgene expression stability is being studied. So far, China is the only country known to have commercially used transgenic, insect-resistant poplar. A pre-commercial herbicide-tolerant poplar was trialed for 8 years in the USA (Li *et al.*, 2008) but could not be released due to stringent environmental risk assessments required for regulatory approval. This increases translation costs and delays reducing investor confidence for commercial deployment (Harfouche *et al.*, 2011).

The first field trials of transgenic poplar were performed in France in 1987 (Fillatti *et al.,* 2017) and in Belgium in 1988 (Deblock, 1990). Although there has been a total of 28 research-scale GM poplar field trials approved in the European Union under Council Directive 90/220/EEC since October 1991 (in Poland, Belgium, Finland, France, Germany, Spain, Sweden, and in the UK (Pilate *et al.*, 2016), only authorisations in Poland and Belgium are in place today. In the USA, regulatory notifications and permits for nearly 20,000 transgenic poplar trees derived from approximately 600 different constructs have been issued since 1995 by the USDA's Animal and Plant Health Inspection Service (APHIS) (Strauss *et al.*, 2016).

Genome-editing CRISPR technologies: Clustered regularly interspaced palindromic repeats (CRISPR) and the CRISPR-associated (CRISPR-Cas) nucleases are a groundbreaking genome-engineering tool that complements classical plant breeding and transgenic methods (Moreno-Mateos *et al.*, 2017). Only two published studies in poplar have applied the CRISPR/Cas9 technology. One is in *P. tomentosa*, in which an endogenous phytoene desaturase gene (*PtoPDS*) was successfully disrupted site-specifically in the first generation of transgenic plants resulting in an albino and dwarf phenotype (Fan *et al.*, 2015). The second was in *P. tremula* × *alba*, in which high CRISPR-Cas9 mutational efficiency was achieved for three 4-coumarate:CoA ligase (4CL) genes, 4CL1, 4CL2 and 4CL5, associated with lignin and flavonoid biosynthesis (Zhou *et al.*, 2015). Due to its low cost, precision and rapidness, it is very probable that cultivars or clones produced using CRISPR technology will be ready for marketing in the near future (Yin *et al.*, 2017). Recently, a CRISPR with a smaller associated endonuclease has been discovered from *Prevotella* and *Francisella* 1 (Cpf1) which may have advantages over Cas9. In addition, there are reports of DNA-free editing in plants, using both CRISPR Cpf1 and CRISPR Cas9 e.g. (Kim *et al.*, 2017; Mahfouz, 2017; Zaidi *et al.*, 2017).

It remains unresolved whether plants modified by genome editing will be regulated as genetically modified organisms (GMOs) by the relevant authorities in different countries (Lozano-Juste & Cutler, 2014). Regulations to cover these new breeding techniques are still evolving but those countries who have published specific guidance (including USA, Argentina and Chile) are indicating that plants

possessing simple genome edits will not be regulated as conventional transgenesis (Jones, 2015b). The first generation of genome-edited crops will likely be phenocopy gene knockouts that already exist to produce 'nature identical' traits, that is, traits that could also be derived by conventional breeding. Despite this, confidence in applying these new powerful breeding tools remains limited owing to the uncertain regulatory environment in many parts of the world (Gao, 2018) including the recent ECJ 2018 rulings mentioned earlier.

Genomics-based breeding technologies: Poplar breeding programs are becoming well equipped with useful genomics tools and resources that are critical to explore genome-wide variability and make use of the variation for enhancing genetic gains. Deep transcriptome sequencing, re-sequencing of alternate genomes and GBS technology for genome-wide marker detection using next-generation sequencing (NGS) are yielding valuable genomics tools. GWAS with NGS-based markers facilitate marker identification for MAS, breeding by design and GS.

GWAS approaches have provided a deeper understanding of genome function as well as allelic architectures of complex traits (Huang *et al.*, 2010) and have been widely implemented in poplar for wood characteristics (Porth *et al.*, 2013), stomatal patterning, carbon gain versus disease resistance (McKown *et al.*, 2014), height and phenology (Evans *et al.*, 2014), cell wall chemistry (Muchero *et al.*, 2015), growth and cell walls traits (Fahrenkrog *et al.*, 2017), bark roughness (Bdeir *et al.*, 2017), and height and diameter growth (Liu *et al.*, 2018). Using high-throughput sequencing and genotyping platforms, an enormous amount of SNP markers have been used to characterize the linkage disequilibrium (LD) in poplar (e.g., Slavov *et al.* 2012, discussed below).

The genetic architecture of photoperiodic traits in perennial trees is complex involving many loci. However, it shows high levels of conservation during evolution (Maurya & Bhalerao, 2017). These genomics tools can therefore be used to address adaptation issues and fine tune the movement of elite lines into new environments. For example, poor timing of spring bud burst and autumn bud set can result in frost damage resulting in yield losses (Ilstedt, 1996). These have been studied in *P.*

tremula genotypes along a latitudinal cline in Sweden (~56-66°N) and have revealed high nucleotide polymorphism in two nonsynonymous SNPs within and around the *phytochrome B2* locus (Ingvarsson *et al.*, 2006; Ingvarsson *et al.*, 2008). Resequencing 94 of these *P. tremula* genotypes for GWAS showed that non-coding variation of a single genomic region containing the *PtFT2* gene, described 65% of observed genetic variation in bud set along the latitudinal cline (Tan, 2018).

Re-sequencing genomes is currently the most rapid and effective method detecting genetic differences between variants and for linking loci to complex and important agronomical and biomass traits, thus addressing breeding challenges associated with long-lived plants like poplars.

To date, whole genome re-sequencing initiatives have been launched for several poplar species and genotypes. In *Populus*, LD studies based on genome re-sequencing suggested the feasibility of GWAS in undomesticated populations (Slavov *et al.*, 2012). This plant population is being used to inform breeding for bioenergy development. For example, the detection of reliable phenotype/genotype associations and molecular signatures of selection requires a detailed knowledge about genome-wide patterns of allele frequency variation, LD and recombination, suggesting that GWAS and GS in undomesticated populations may be more feasible in *Populus* than previously assumed. Slavov *et al.* (2012) have re-sequenced 16 genomes of *P. trichocarpa* and genotyped 120 trees from 10 subpopulations using 29,213 SNPs (Geraldes *et al.*, 2013). The largest-ever SNP dataset of genetic variations in poplar has recently been released, providing useful information for breeding https://www.bioenergycenter.org/besc/gwas/index.cfm. Also, deep sequencing of transcriptomes using RNA-Seq has been used for identification of functional genes and molecular markers, i.e., polymorphism markers and SSRs. A multi-tissue and multiple experimental dataset for *P. trichocarpa* RNA-Seq is publicly available (https://jgi.doe.gov/doe-jgi-plant-flagship-gene-atlas/).

The availability of genomic information of DNA-containing cell organelles (nucleus, chloroplast and mitochondria) will also allow a holistic approach in poplar molecular breeding in the near future

(Kersten *et al.*, 2016). Complete *Populus* genome sequences are available for nucleus (*P. trichocarpa*; section Tacamahaca) and chloroplasts (seven species, and two clones from *P. tremula* W52 and *P. tremula* × *P. alba* 717-1B4). A comparative approach revealed structural and functional information, broadening the knowledge base of *Populus* cpDNA and stimulating future diagnostic marker development. The availability of whole genome sequences of these cellular compartments of *P. tremula* holds promise for boosting marker-assisted poplar breeding. Other nuclear genome sequences from additional *Populus* species are now available (e.g., *P. deltoides* (https://phytozome.jgi.doe.gov/pz/) and will become available in the forthcoming years (e.g. *P. tremula* and *P. tremuloides* – PopGenIE (Sjodin *et al.*, 2009)). Recently, the characterization of the poplar pan-genome by genome-wide identification of structural variation in three crossable poplar species: *P. nigra*, *P. deltoides*, and *P. trichocarpa,* revealed a deeper understanding of the role of inter- and intraspecific structural variants in poplar phenotype and may have important implications for breeding, particularly, interspecific hybrids (Pinosio *et al.,* 2016).

GS has been proposed as an alternative to MAS in crop improvement (Bernardo & Yu, 2007; Heffner *et al.*, 2009). GS is particularly well suited for species with long generation times, for characteristics that display moderate to low heritability, for traits that are expensive to measure and for selection of traits expressed late in the life cycle, as is the case for most traits of commercial value in forestry (Harfouche *et al.*, 2012). Current joint genome sequencing efforts to implement GS in poplar using genomic-estimated breeding values for bioenergy conversion traits from 49 *P. trichocarpa* families and 20 full-sib progeny are taking place at the Oak Ridge National Laboratory and GreenWood Resources (Brian Stanton, personal communication https://cbi.ornl.gov/). These data together with the re-sequenced GWAS population data will be the basis for developing GS algorithms. Genomic breeding tools have been developed for the intra-specific program targeting yield, resistance to *Venturia* shoot blight, *Melampsora* leaf rust, resistance to *Cryptorhynchus lapathi*, stem form, wood specific gravity, and wind firmness (Evans *et al.*, 2014; Guerra *et al.*, 2016). A newly developed "breeding with rare defective alleles" (BRDA) technology has been developed to

exploit natural variation of *P. nigra* and identify defective variants of genes predicted by prior transgenic research to impact lignin properties. Individual trees carrying naturally defective alleles can then be incorporated directly into breeding programs, thereby bypassing the need for transgenics (Vanholme *et al.*, 2013). This novel breeding technology offers a reverse genetics complement to emerging GS for targeted improvement of quantitative traits (Tsai, 2013).

Phenomics-assisted breeding technology: Phenomics involves the characterization of phenomes – the full set of phenotypes of given individual plants (Houle *et al.*, 2010). Traditional phenotyping tools, which inefficiently measure a limited set of phenotypes, have become a bottleneck in plant breeding studies. High-throughput plant phenotyping facilitates to provide accurate screening of thousands of plant breeding lines, clones or populations over time (Fu, 2015) are critical for accelerating genomics-based breeding. Automated image collection and analysis, phenomics technologies allow accurate and non-destructive measurements of a diversity of phenotypic traits in large breeding populations (Goggin *et al.*, 2015; Doonan *et al.*, 2016; Ludovisi *et al.*, 2017; Shakoor *et al.*, 2017). One important consideration is the identification of relevant and quantifiable target traits that are early diagnostic indicators of biomass yield. Good progress has been made in elucidating these underpinning morpho-physiological traits that are amenable to remote sensing in *Populus* (Rae *et al.*, 2004; Harfouche *et al.*, 2014). More recently, Ludovisi *et al.,* (2017) developed a novel methodology for field phenomics of drought stress in a *P. nigra* F$_2$ partially inbred population using thermal infrared images recorded from an unmanned aerial vehicle-based platform.

Energy is the current main market for poplar biomass but the market return provided is not sufficient to support production expansion even with added demand for environmental and land management 'ecosystem services' such as the treatment of effluent, phytoremediation, riparian buffer zones and agro-forestry plantings. Aviation fuel is a significant target market (Crawford *et al.*, 2016). To serve this market and to reduce current carbon costs of production (Budsberg *et al.*, 2016), key improvement traits in addition to yield (e.g., coppice regeneration, pest /disease resistance,

water- and nutrient-use efficiencies) will be trace greenhouse gas (GHG) emissions (e.g., isoprene volatiles), site adaptability and biomass conversion efficiency. Efforts are underway to have national environmental protection agencies approval for poplar hybrids qualifying for renewable energy credits.

Reflections on the commercialisation challenge

The research and innovation activities reviewed in this paper aim to advance the genetic improvement of crop species that are viable candidate feedstocks for bioenergy applications should those markets eventually develop. These markets need to generate sufficient revenue and adequately distribute it to the actors along the value chain. The work on all four crops shares one thing in common: long-term efforts to integrate fundamental knowledge into breeding and crop development along a research and development (R&D) pipeline. The development of miscanthus led by Aberystwyth University exemplifies the concerted research effort that has integrated the R&D activities from eight projects over 14 years with background core research funding along an emerging innovation chain (Figure 2). This programme has produced a first range of conventionally bred seeded interspecies hybrids, which are now in upscaling trials (Table 3). The application of molecular approaches (Table 5) with further conventional breeding (Table 4) offers the prospect of a second range of improved seeded hybrids. This example shows that research-based support of the development of new crops or crop types requires a long-term commitment that goes beyond that normally available from project-based funding (Figure 1). Innovation in this sector requires continuous resourcing of conventional breeding operations and capability to minimise time and investment losses caused by funding discontinuities.

This challenge is increased further by the well-known market failure in the breeding of many agricultural crop species. The UK Department for Environment, Food and Rural Affairs (Defra) examined the role of genetic improvement in relation to non-market outcomes, such as environmental protection, and concluded that public investment in breeding was required if

profound market failure is to be addressed (Defra, 2002). With the exception of widely-grown hybrid crops, such as maize, and some high-value horticultural crops, royalties arising from plant breeders rights or other returns to breeders fail to adequately compensate for the full cost for research-based plant breeding. The result, even for major crops such as wheat, is sub-optimal investment and sub-optimal returns for society. This market failure is especially acute for perennial crops developed for improved sustainability, rather than consumer appeal (Tracy *et al.*, 2016). Figure 3 illustrates the underlying challenge of capturing value for the breeding effort. The 'valley of death' that results from the low and delayed returns to investment applies generally to the research-to-product innovation pipeline (Beard *et al.*, 2009) and certainly to most agricultural crop species. However, this schematic is particularly relevant to PBCs. Most of the value for society from the improved breeding of these crops comes from changes in how agricultural land is used, i.e., it depends on the increased production of these crops. The value for society includes many ecosystems benefits: the effects of a return to semi-natural perennial crop cover that protects soils, the increase of soil carbon storage, the protection of vulnerable land or the cultivation of polluted soils and the reductions in GHG emissions (Lewandowski, 2016). By its very nature, the production of biomass on agricultural land marginal for food production challenges farm-level profitability. The costs of planting material and one-time nature of crop establishment are major early stage costs, and therefore the opportunities for conventional royalty capture by breeders that are manifold for annual crops are limited for PBCs (Hastings *et al.,* 2017). Public investment in developing PBCs for the non-food bio-based sector needs to provide more long-term support for this critical foundation to a sustainable bio-economy.

Conclusions

This paper provides an overview of research-based plant breeding in four leading PBCs. For all four PBC genera, significant progress has been made in genetic improvement through collaboration between research scientists and those operating on-going breeding programmes. Compared with the main food crops most PBC breeding programmes date back only a few decades (Table 1). This

breeding effort has thus co-evolved with molecular biology and the resulting -omics technologies that can support breeding. The development of all four PBCs has depended strongly on public investment in research and innovation. The nature and driver of the investment varied. In close association with public research organisations or universities, all these programmes started with germplasm collection and characterisation, which underpin the selection of parents for exploratory wide crosses for progeny testing (Figure 1, Table 4).

Public support for switchgrass in North America was explicitly linked to plant breeding with 12 breeding programmes supported in the USA and Canada. Switchgrass breeding efforts to date, using conventional breeding, have resulted in over 36 registered cultivars in the USA (Table 1), with the development of dedicated biomass-type cultivars coming within the past few years. While -omics technologies have been incorporated into several of these breeding programmes, they have not yet led to commercial deployment in either conventional or hybrid cultivars.

Willow genetic improvement was led by the research community closely linked to plant breeding programmes. Willow and poplar have the longest record of public investment in genetic improvement that can be traced back to 1920s in the UK and USA respectively. Like switchgrass, breeding programmes for willow are connected to public research efforts. The UK, in partnership with the programme based at Cornell University, remains the European leader in willow improvement with a long term breeding effort closely linked to supporting biological research at Rothamsted. In willow, F_1 hybrids have produced impressive yield gains over parental germplasm by capturing hybrid vigour. Over 30 willow clones are commercially available in the USA and Europe, and a further ~90 are under pre-commercial testing (Table 1).

Compared with willow and poplar, miscanthus is a relative newcomer with all the current breeding programmes starting in the 2000's. Clonal *M. × giganteus* propagated by rhizomes is expected to be replaced by more readily scalable seeded hybrids from intra- (*M. sinensis*) and inter-

(*M. sacchariflorus* × *M. sinensis*) species crosses with high seed multiplication rates (of >2,000). The first group of hybrid cultivars is expected to be market-ready around 2022.

Of the four genera used as PBCs, *Populus* is the most advanced in terms of achievements in biological research as a result of its use as a model for basic research of trees. Much of this biological research is not directly connected to plant breeding. Nevertheless, reflecting the fact that poplar is widely grown as a single-stem tree in SRF, there are about 60 commercially available clones and an additional 80 clones in commercial pipelines (Table 1). Transgenic poplar hybrids have moved beyond proof-of-concept to commercial reality in China.

Many PBC programmes have initiated long-term conventional recurrent selection breeding cycles for population improvement, which is a key process in increasing yield through hybrid vigour. As this approach requires many years, most programmes are experimenting with molecular breeding methods as these have the potential to accelerate precision breeding. For all four PBCs, investments in basic genetic and genomic resources, including the development of mapping populations for QTLs and whole genome sequences, are available to support long-term advances. More recently, association genetics with panels of diverse germplasm are being used as training populations for GS models (Table 5). These efforts are benefitting from publicly-available DNA sequences and whole genome assemblies in crop databases. Key to these accelerated breeding technologies are developments in novel phenomic technologies to bridge the genotype/phenotype gap. In poplar, novel remote sensing field phenotyping is now being deployed to assist breeders. These advances are being combined with *in-vitro* and *in-planta* modern molecular breeding techniques such as CRISPR (Table 5). CRISPR technology for genome editing has been proven in poplar. This technology is also being applied in switchgrass and miscanthus (Table 5), but the future of CRISPR in commercial breeding for the European market is uncertain in the light of recent ECJ 2018 rulings.

There is integration of research and plant breeding itself in all four PBCs. Therefore, estimating the ongoing costs of maintaining these breeding programmes is difficult. Investment in research also

seeks wider benefits associated with technological advances in plant science rather than cultivar development *per se*. However, in all cases, the conventional breeding cycle shown in Figure 1 is the basic 'engine' with molecular technologies (-omics) serving to accelerate this engine. The history of the development shows that the existence of these breeding programmes is essential to gain benefits from the biological research. Despite this, it is this essential step that is at most risk from reductions in investment. A conventional breeding programme typically requires a breeder and several technicians who are supported over the long term (20-30 years, Figure 1) at costs of about 0.5 to 1.0 million Euro per year (as of 2018). The analysis reported here shows that the time needed to perform one cycle of conventional breeding, bringing germplasm from the wild to a commercial hybrid ranged from 11 years in switchgrass to 26 years in poplar (Figure 1). In a mature crop grown on over 100,000 ha, with effective cultivar protection and a suitable business model, this level of revenue could come from royalties. Until such levels are reached, PBCs lie in the innovation valley of death (Figure 3) and need public support.

Applying industrial 'technology readiness levels' (TRL) originally developed for aerospace (Héder, 2017) to our plant breeding efforts we estimate many promising hybrids cultivars are at TRL levels of 3-4. In Table 3 experts in each crop estimate that it would take 3 years from now to upscale planting material from leading cultivars in plot trials to 100 ha.

Taking the UK example mentioned in the introduction, planting rates of ~35,000 ha per year from 2022 onwards are needed to reach over 1 m ha by 2050. Ongoing work in the UK funded Miscanthus Upscaling Technology (MUST) project shows that ramping annual hybrid seed production from the current level of sufficient seed for 10 ha in 2018 to 35,000 ha would take about 5 years, assuming no setbacks. If current hybrids of any of the four PBCs in the upscaling pipeline fail on any step e.g. lower than expected multiplication rates or unforeseen agronomic barriers, then further selections from ongoing breeding are needed to replace earlier candidates.

In conclusion, the breeding foundations have been laid well for switchgrass, miscanthus, willow and poplar owing to public funding over the long time periods necessary. Improved cultivars or genotypes are available that could be scaled up over a few years if real, sustained market opportunities emerged in response to sustained favourable policies and industrial market pull. The potential contributions of growing and using these PBCs for socioeconomic and environmental benefits are clear but how farmers and others in commercial value chains are rewarded for mass scale deployment, as is necessary, is not obvious at present. Therefore mass scale deployment of these lignocellulose crops needs developments outside the breeding arenas to drive breeding activities more rapidly and extensively.

Acknowledgments

Conflict of interest

The authors declare that progress reported in this paper, which includes input from industrial partners, is not biased by their business interests.

(Atienza *et al.*, 2003c; Atienza *et al.*, 2003b; Rayburn *et al.*, 2009; Jannink *et al.*, 2010; Hwang *et al.*, 2014b; Lazarus *et al.*, 2015; Wheeler *et al.*, 2015; Zhou *et al.*, 2015; Evans, 2016; Fabio *et al.*, 2017; Kalinina *et al.*, 2017; Krzyżak *et al.*, 2017)

References

Alburquerque N, Baldacci-Cresp F, Baucher M *et al.* (2016) New Transformation Technologies for Trees. In: *Biosafety of Forest Transgenic Trees.* pp 31-66. Springer.

Argus GW (2007) Salix (Salicaceae) distribution maps and a synopsis of their classification in North America, north of Mexico. *Harvard Papers in Botany,* **12**, 335-368.

Atienza SG, Ramirez MC, Martin A (2003a) Mapping-QTLs controlling flowering date in *Miscanthus sinensis* Anderss. *Cereal Research Communications,* **31**, 265-271.

Atienza SG, Satovic Z, Petersen KK, Dolstra O, Martin A (2003b) Identification of QTLs influencing agronomic traits in *Miscanthus sinensis* Anderss. I. Total height, flag-leaf height and stem diameter. *Theoretical and Applied Genetics,* **107**, 123-129.

Atienza SG, Satovic Z, Petersen KK, Dolstra O, Martin A (2003c) Influencing combustion quality in *Miscanthus sinensis* Anderss.: identification of QTLs for calcium, phosphorus and sulphur content. *Plant Breeding,* **122**, 141-145.

Bdeir R, Muchero W, Yordanov Y, Tuskan GA, Busov V, Gailing O (2017) Quantitative trait locus mapping of Populus bark features and stem diameter. *BMC plant biology,* **17**, 224.

Beard TR, Ford GS, Koutsky TM, Spiwak LJ (2009) A Valley of Death in the innovation sequence: an economic investigation. *Research Evaluation,* **18**, 343-356.

Berguson W, Mcmahon B, Riemenschneider D (2017) Additive and Non-Additive Genetic Variances for Tree Growth in Several Hybrid Poplar Populations and Implications Regarding Breeding Strategy. *Silvae Genetica*.

Berlin S, Trybush S, Fogelqvist J *et al.* (2014) Genetic diversity, population structure and phenotypic variation in European *Salix viminalis* L.(Salicaceae). *Tree genetics & genomes,* **10**, 1595-1610.

Bernardo R, Yu J (2007) Prospects for genomewide selection for quantitative traits in maize. *Crop Science,* **47**, 1082-1090.

Biswal AK, Atmodjo MA, Li M *et al.* (2018) Sugar release and growth of biofuel crops are improved by downregulation of pectin biosynthesis. *Nature Biotechnology,* **36**, 249-+.

Bmu (2009) National Biomass Action Plan for Germany pp 17. Bundesministerium für Umwelt Naturschutz und Reaktorsicherheit, Bundesministerium für Ernährung (BMU) & Landwirtschaft und Verbraucherschutz (BMELV), 11055 Berlin | Germany. http://www.bmel.de/SharedDocs/Downloads/EN/Publications/BiomassActionPlan.pdf?__blob=publicationFile (accessed 8 Sept 2018).

Brummer EC (1999) Capturing heterosis in forage crop cultivar development. *Crop Science,* **39**, 943-954.

Budsberg E, Crawford JT, Morgan H, Chin WS, Bura R, Gustafson R (2016) Hydrocarbon bio-jet fuel from bioconversion of poplar biomass: life cycle assessment. *Biotechnology for biofuels,* **9**, 170.

Burris KP, Dlugosz EM, Collins AG, Stewart CN, Lenaghan SC (2016) Development of a rapid, low-cost protoplast transfection system for switchgrass (*Panicum virgatum* L.). *Plant Cell Reports,* **35**, 693-704.

Casler M (2012) Switchgrass breeding, genetics, and genomics. In: *Switchgrass* (ed Monti A), pp 29-54. London, Springer.

Casler M, Mitchell R, Vogel K (2012) Switchgrass. In: *Handbook of bioenergy crop plants Vol. 2.* (eds Kole C, Joshi CP, Shonnard DR). New York, Taylor & Francis.

Casler M, Vogel K, Lee D *et al.* (2018) 30 Years of progress toward increasing biomass yield of switchgrass and big bluestem. *Crop Science,* **58**, (in press).

Casler MD, Ramstein GP (2018) Breeding for Biomass Yield in Switchgrass Using Surrogate Measures of Yield. *BioEnergy Research,* **11**, 6-12.

Casler MD, Sosa S, Hoffman L *et al.* (2017) Biomass Yield of Switchgrass Cultivars under High- versus Low-Input Conditions. *Crop Science,* **57**, 821-832.

Casler MD, Vogel KP (2014) Selection for Biomass Yield in Upland, Lowland, and Hybrid Switchgrass. *Crop Science,* **54**, 626-636.

Casler MD, Vogel KP, Harrison M (2015) Switchgrass Germplasm Resources. *Crop Science,* **55**, 2463-2478.

Clark LV, Brummer JE, Głowacka K *et al.* (2014) A footprint of past climate change on the diversity and population structure of *Miscanthus sinensis*. *Annals of Botany,* **114**, 97-107.

Clark LV, Dzyubenko E, Dzyubenko N *et al.* (2016) Ecological characteristics and in situ genetic associations for yield-component traits of wild Miscanthus from eastern Russia. *Annals of Botany,* **118**, 941-955.

Clark LV, Jin X, Petersen KK *et al.* (2018) Population structure of *Miscanthus sacchariflorus* reveals two major polyploidization events, tetraploid-mediated unidirectional introgression from diploid *M. sinensis*, and diversity centred around the Yellow Sea. *Annals of Botany*, doi: 10.1093/aob/mcy1161.

Clark LV, Stewart JR, Nishiwaki A *et al.* (2015) Genetic structure of *Miscanthus sinensis* and *Miscanthus sacchariflorus* in Japan indicates a gradient of bidirectional but asymmetric introgression. *Journal of Experimental Botany,* **66**, 4213-4225.

Clifton-Brown J, Hastings A, Mos M *et al.* (2017) Progress in upscaling Miscanthus biomass production for the European bio-economy with seed-based hybrids. *Global Change Biology Bioenergy,* **9**, 6-17.

Clifton-Brown J, Schwarz KU, Hastings A (2015) History of the Development of Miscanthus as a Bioenergy Crop: From Small Beginnings to Potential Realisation. *Biology and Environment-Proceedings of the Royal Irish Academy,* **115B**, 45-57.

Clifton-Brown J, Senior H, Purdy S *et al.* (2018) Investigating the potential of novel non-woven fabrics to increase pollination efficiency in plant breeding. *PloS one*, accepted subject to revisions.

Crawford JT, Shan CW, Budsberg E, Morgan H, Bura R, Gustafson R (2016) Hydrocarbon bio-jet fuel from bioconversion of poplar biomass: techno-economic assessment. *Biotechnology for biofuels,* **9**, 141.

Dalton S (2013) Biotechnology of Miscanthus. In: *Biotechnology of Neglected and Underutilized Crops.* pp 243-294. Springer.

Davey CL, Robson P, Hawkins S, Farrar K, Clifton-Brown JC, Donnison IS, Slavov GT (2017) Genetic relationships between spring emergence, canopy phenology, and biomass yield increase the accuracy of genomic prediction in Miscanthus. *Journal of Experimental Botany,* **68**, 5093-5102.

David X, Anderson X (2002) Aspen and Larch Genetics Cooperative Annual Report 13. Department of Forest Resources, University of Minnesota, St. Paul, Minnesota, USA.

Deblock M (1990) Factors Influencing the Tissue-Culture and the Agrobacterium-Tumefaciens-Mediated Transformation of Hybrid Aspen and Poplar Clones. *Plant Physiology,* **93**, 1110-1116.

Defra (2002) The role of future public research investment in the genetic improvement of UK-grown crops. pp 220, sciencesearch.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=None&Completed=220&ProjectID=10412 (accessed on 10418 Sept 12018). London.

Dewoody J, Trewin H, Taylor G (2015) Genetic and morphological differentiation in Populus nigra L.: isolation by colonization or isolation by adaptation? *Molecular Ecology,* **24**, 2641-2655.

Dickmann D, Kuzovkina J (2008) Poplars and willows of the world, with emphasis on silviculturally important species. In: *Poplars and Willows in the World: Meeting the needs of society and the environment.* (eds Isebrands J, Richardson J), pp 125 pp. Food and Agriculture Organization of the United Nations, Rome, Italy.

Dong HX, Liu SY, Clark LV *et al.* (2018) Genetic mapping of biomass yield in three interconnected Miscanthus populations. *Global Change Biology Bioenergy,* **10**, 165-185.

Doonan JH, Camargo A, Gay A, Gegas VC (2016) Challenges of Crop Phenomics in the Post-genomic Era. In: *Phenomics.* pp 146-175. CRC Press.

Dukes (2017) Digest of UK Energy Statistics (DUKES). pp PP 264, www.gov.uk/government/collections/digest-of-uk-energy-statistics-dukes (accessed on 268th Sept 2018). London, Department for Business, Energy & Industrial Strategy.

Eckenwalder J (1996) Systematics and evolution of Populus. In: *Biology of Populus and its implications for management and conservation.* (eds Stettler R, Bradshaw H, Heilman J, Hinckley T), pp 7-32

Ottawa, Ontario, Canada, National Research Council of Canada.

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS one,* **6**, e19379.

Evans H (2016) Bioenergy crops in the UK: Case studies of successful whole farm integration. pp 23, https://www.eti.co.uk/library/bioenergy-crops-in-the-uk-case-studies-on-successful-whole-farm-integration-evidence-pack (accessed on 28 Sept. 2018). Loughborough, Energy Technologies Institute.

Evans H (2017) Increasing UK biomass production through more productive use of land. pp 7, http://www.eti.co.uk/library/an-eti-perspective-increasing-uk-biomass-production-through-more-productive-use-of-land (accessed on the 8 Sept 2018). Loughborough, Energy Technologies Institute.

Evans J, Sanciangco MD, Lau KH *et al.* (2018) Extensive Genetic Diversity is Present within North American Switchgrass Germplasm. *Plant Genome,* **11**.

Evans LM, Slavov GT, Rodgers-Melnick E *et al.* (2014) Population genomics of Populus trichocarpa identifies signatures of selection and adaptive trait associations. *Nature Genetics,* **46**, 1089-1096.

Fabio ES, Volk TA, Miller RO *et al.* (2017) Genotype× environment interaction analysis of North American shrub willow yield trials confirms superior performance of triploid hybrids. *Global Change Biology Bioenergy,* **9**, 445-459.

Fahrenkrog AM, Neves LG, Resende MF *et al.* (2017) Genome-wide association study reveals putative regulators of bioenergy traits in Populus deltoides. *New Phytologist,* **213**, 799-811.

Fan D, Liu TT, Li CF, Jiao B, Li S, Hou YS, Luo KM (2015) Efficient CRISPR/Cas9-mediated Targeted Mutagenesis in Populus in the First Generation. *Scientific Reports,* **5**.

Fu CX, Mielenz JR, Xiao XR *et al.* (2011) Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proceedings of the National Academy of Sciences of the United States of America,* **108**, 3803-3808.

Fu YB (2015) Understanding crop genetic diversity under modern plant breeding. *Theoretical and Applied Genetics,* **128**, 2131-2142.

Gao C (2018) The future of CRISPR technologies in agriculture. *Nature reviews. Molecular cell biology*.

Geraldes A, Difazio SP, Slavov GT *et al.* (2013) A 34K SNP genotyping array for Populus trichocarpa: Design, application to the study of natural populations and transferability to other Populus species. *Molecular Ecology Resources,* **13**, 306-323.

Gifford JM, Chae WB, Swaminathan K, Moose SP, Juvik JA (2015) Mapping the genome of Miscanthus sinensis for QTL associated with biomass productivity. *Global Change Biology Bioenergy,* **7**, 797-810.

Goggin FL, Lorence A, Topp CN (2015) Applying high-throughput phenotyping to plant-insect interactions: picturing more resistant crops. *Current Opinion in Insect Science,* **9**, 69-76.

Grabowski PP, Evans J, Daum C *et al.* (2017) Genome-wide associations with flowering time in switchgrass using exome-capture sequencing data. *New Phytologist,* **213**, 154-169.

Greef JM, Deuter M (1993) Syntaxonomy of *Miscanthus x giganteus* GREEF et DEU. *Angewandte Botanik,* **67**, 87-90.

Guerra FP, Richards JH, Fiehn O *et al.* (2016) Analysis of the genetic variation in growth, ecophysiology, and chemical and metabolomic composition of wood of Populus trichocarpa provenances. *Tree genetics & genomes,* **12**, 6.

Guo HP, Shao RX, Hong CT, Hu HK, Zheng BS, Zhang QX (2013) Rapid in vitro propagation of bioenergy crop Miscanthus sacchariflorus. In: *Applied Mechanics and Materials.* pp 181-186, Trans Tech Publ.

Hallingbäck HR, Fogelqvist J, Powers SJ *et al.* (2016) Association mapping in *Salix viminalis* L. (Salicaceae)–identification of candidate genes associated with growth and phenology. *Global Change Biology Bioenergy,* **8**, 670-685.

Hanley SJ, Karp A (2014) Genetic strategies for dissecting complex traits in biomass willows (*Salix* spp.). *Tree physiology,* **34**, 1167-1180.

Harfouche A, Meilan R, Altman A (2011) Tree genetic engineering and applications to sustainable forestry and biomass production. *Trends in biotechnology,* **29**, 9-17.

Harfouche A, Meilan R, Altman A (2014) Molecular and physiological responses to abiotic stress in forest trees and their relevance to tree improvement. *Tree physiology,* **34**, 1181-1198.

Harfouche A, Meilan R, Kirst M, Morgante M, Boerjan W, Sabatti M, Mugnozza GS (2012) Accelerating the domestication of forest trees in a changing world. *Trends in Plant Science,* **17**, 64-72.

Hastings A, Clifton-Brown J, Wattenbach M, Mitchell CP, Stampfl P, Smith P (2009) Future energy potential of Miscanthus in Europe. *Global Change Biology Bioenergy,* **1**, 180-196.

Héder M (2017) From NASA to EU: The evolution of the TRL scale in Public Sector Innovation. *The Innovation Journal,* **22**, 1–23.

Heffner EL, Sorrells ME, Jannink JL (2009) Genomic Selection for Crop Improvement. *Crop Science,* **49**, 1-12.

Hodkinson TR, Klaas M, Jones MB, Prickett R, Barth S (2015) Miscanthus: a case study for the utilization of natural genetic variation. *Plant Genetic Resources-Characterization and Utilization,* **13**, 219-237.

Hodkinson TR, Renvoize S (2001) Nomenclature of *Miscanthus xgiganteus (Poaceae)*. *Kew Bulletin,* **56**, 759-760.

Houle D, Govindaraju DR, Omholt S (2010) Phenomics: the next challenge. *Nature Reviews Genetics,* **11**, 855-866.

Huang LS, Flavell R, Donnison IS (2019) Collecting wild Miscanthus germplasm in Asia for bioenergy crop improvement in Europe: a practical implementation of the United Nations. *Annals of Botany*, in review.

Huang XH, Wei XH, Sang T *et al.* (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. *Nature Genetics,* **42**, 961-U976.

Hwang O-J, Cho M-A, Han Y-J *et al.* (2014a) Agrobacterium-mediated genetic transformation of *Miscanthus sinensis*. *Plant Cell, Tissue and Organ Culture (PCTOC),* **117**, 51-63.

Hwang O-J, Lim S-H, Han Y-J, Shin A-Y, Kim D-S, Kim J-I (2014b) Phenotypic characterization of transgenic *Miscanthus sinensis* plants overexpressing *Arabidopsis* phytochrome B. *International Journal of Photoenergy,* **2014**.

Ilstedt B (1996) Genetics and performance of Belgian poplar clones tested in Sweden. *International Journal of Forest Genetics*.

Ingvarsson PK, Garcia MV, Hall D, Luquez V, Jansson S (2006) Clinal variation in phyB2, a candidate gene for day-length-induced growth cessation and bud set, across a latitudinal gradient in European aspen (*Populus tremula*). *Genetics,* **172**, 1845-1853.

Ingvarsson PK, Garcia MV, Luquez V, Hall D, Jansson S (2008) Nucleotide polymorphism and phenotypic associations within and around the phytochrome B2 locus in European aspen (*Populus tremula*, Salicaceae). *Genetics,* **178**, 2217-2226.

Jannink J-L, Lorenz AJ, Iwata H (2010) Genomic selection in plant breeding: from theory to practice. *Briefings in functional genomics,* **9**, 166-177.

Jiang J, Guan Y, Mccormick S, Juvik J, Lubberstedt T, Fei S-Z (2017) Gametophytic Self-Incompatibility Is Operative in *Miscanthus sinensis* (Poaceae) and Is Affected by Pistil Age. *Crop Science,* **57**, 1948-1956.

Jones HD (2015a) Future of breeding by genome editing is in the hands of regulators. *GM crops & food,* **6**, 223-232.

Jones HD (2015b) Regulatory uncertainty over genome editing. *Nature plants,* **1**, 14011.

Kalinina O, Nunn C, Sanderson R *et al.* (2017) Extending Miscanthus Cultivation with Novel Germplasm at Six Contrasting Sites. *Frontiers in plant science,* **8**.

Karp A, Hanley SJ, Trybush SO, Macalpine W, Pei M, Shield I (2011) Genetic Improvement of Willow for Bioenergy and BiofuelsFree Access. *Journal of Integrative Plant Biology,* **53**, 151-165.

Kersten B, Rampant PF, Mader M *et al.* (2016) Genome Sequences of *Populus tremula* Chloroplast and Mitochondrion: Implications for Holistic Poplar Breeding. *PloS one,* **11**.

Kiesel A, Nunn C, Iqbal Y *et al.* (2017) Site-specific management of miscanthus genotypes for combustion and anaerobic digestion: a comparison of energy yields. *Frontiers in plant science,* **8**, 347.

Kim H, Kim ST, Ryu J, Kang BC, Kim JS, Kim SG (2017) CRISPR/Cpf1-mediated DNA-free plant genome editing. *Nature Communications,* **8**.

King ZR, Bray AL, Lafayette PR, Parrott WA (2014) Biolistic transformation of elite genotypes of switchgrass (*Panicum virgatum* L.). *Plant Cell Reports,* **33**, 313-322.

Kopp RF, Maynard CA, De Niella PR, Smart LB, Abrahamson LP (2002) Collection and storage of pollen from Salix (Salicaceae). *American Journal of Botany,* **89**, 248-252.

Krzyżak J, Pogrzeba M, Rusinowski S *et al.* (2017) Heavy metal uptake by novel miscanthus seed-based hybrids cultivated in heavy metal contaminated soil. *Civil and Environmental Engineering Reports,* **26**, 121-132.

Kuzovkina YA, Quigley MF (2005) Willows beyond wetlands: uses of Salix L. species for environmental projects. *Water, Air, and Soil Pollution,* **162**, 183-204.

Lazarus W, Headlee WL, Zalesny RS (2015) Impacts of supplyshed-level differences in productivity and land costs on the economics of hybrid poplar production in Minnesota, USA. *BioEnergy Research,* **8**, 231-248.

Lewandowski I (2015) Securing a sustainable biomass supply in a growing bioeconomy. *Global Food Security,* **6**, 34-42.

Lewandowski I (2016) The role of perennial biomass crops in a growing bioeconomy. In: *Perennial Biomass Crops for a Resource-Constrained World.* (eds Barth S, Murphy-Bokern D, Kalinina O, Taylor G, Jones M), pp 3-13. Springer.

Li JL, Meilan R, Ma C, Barish M, Strauss SH (2008) Stability of herbicide resistance over 8 years of coppice in field-grown, genetically engineered poplars. *Western Journal of Applied Forestry,* **23**, 89-93.

Li RY, Qu RD (2011) High throughput Agrobacterium-mediated switchgrass transformation. *Biomass & Bioenergy,* **35**, 1046-1054.

Linde-Laursen IB (1993) Cytogenetic analysis of *Miscanthus* 'Giganteus', an interspecific hybrid. *Hereditas,* **119**, 297-300.

Lindegaard KN, Barker JHA (1997) Breeding willows for biomass. *Aspects of Applied Biology*, 155-162.

Liu LL, Wu YQ, Wang YW, Samuels T (2012) A High-Density Simple Sequence Repeat-Based Genetic Linkage Map of Switchgrass. *G3-Genes Genomes Genetics,* **2**, 357-370.

Liu Y, Merrick P, Zhang ZZ, Ji CH, Yang B, Fei SZ (2018) Targeted mutagenesis in tetraploid switchgrass (*Panicum virgatum* L.) using CRISPR/Cas9. *Plant Biotechnology Journal,* **16**, 381-393.

Lovett A, Sünnenberg G, Dockerty T (2014) The availability of land for perennial energy crops in Great Britain. *Global Change Biology Bioenergy,* **6**, 99-107.

Lozano-Juste J, Cutler SR (2014) Plant genome engineering in full bloom. *Trends in Plant Science,* **19**, 284-287.

Lu F, Lipka AE, Glaubitz J *et al.* (2013) Switchgrass Genomic Diversity, Ploidy, and Evolution: Novel Insights from a Network-Based SNP Discovery Protocol. *Plos Genetics,* **9**.

Ludovisi R, Tauro F, Salvati R, Khoury S, Mugnozza GS, Harfouche A (2017) UAV-Based Thermal Imaging for High-Throughput Field Phenotyping of Black Poplar Response to Drought. *Frontiers in plant science, 8*.

Macalpine W, Shield I, Karp A (2010) Seed to near market variety; the BEGIN willow breeding pipeline 2003–2010 and beyond. In: *Bioten conference proceedings, Birmingham.* pp 21-23.

Macalpine WJ, Shield IF, Trybush SO, Hayes CM, Karp A (2008) Overcoming barriers to crossing in willow (Salix spp.) breeding. *Aspects of Applied Biology*, 173-180.

Mahfouz MM (2017) The efficient tool CRISPR-Cpf1. *Nature Plants, 3*.

Malinowska M, Donnison IS, Robson PR (2017) Phenomics analysis of drought responses in Miscanthus collected from different geographical locations. *Global Change Biology Bioenergy, 9*, 78-91.

Maroder HL, Prego IA, Facciuto GR, Maldonado SB (2000) Storage Behaviour of *Salix alba* and *Salix matsudana* Seeds. *Annals of Botany, 86*, 1017-1021.

Martinez-Reyna J, Vogel K (2002) Incompatiblity systems in switchgrass. *Crop Science, 42*, 1800-1805.

Maurya JP, Bhalerao RP (2017) Photoperiod-and temperature-mediated control of growth cessation and dormancy in trees: a molecular perspective. *Annals of Botany, 120*, 351-360.

Mccalmont JP, Hastings A, Mcnamara NP, Richter GM, Robson P, Donnison IS, Clifton-Brown J (2017) Environmental costs and benefits of growing Miscanthus for bioenergy in the UK. *Global Change Biology Bioenergy, 9*, 489-507.

Mccracken AR, Dawson WM (1997) Using mixtures of willow clones as a means of controlling rust disease. *Aspects of Applied Biology*, 97-103.

Mckown AD, Klapste J, Guy RD *et al.* (2014) Genome-wide association implicates numerous genes underlying ecological trait variation in natural populations of *Populus trichocarpa*. *New Phytologist, 203*, 535-553.

Merrick P, Fei SZ (2015) Plant regeneration and genetic transformation in switchgrass - A review. *Journal of Integrative Agriculture, 14*, 483-493.

Moreno-Mateos MA, Fernandez JP, Rouet R *et al.* (2017) CRISPR-Cpf1 mediates efficient homology-directed repair and temperature-controlled genome editing. *Nature communications, 8*, 2024.

Mosseler A (1990) Hybrid performance and species crossability relationships in willows (Salix). *Canadian Journal of Botany, 68*, 2329-2338.

Muchero W, Guo J, Difazio SP *et al.* (2015) High-resolution genetic mapping of allelic variants associated with cell wall chemistry in Populus. *BMC genomics, 16*.

Neale DB, Kremer A (2011) Forest tree genomics: growing resources and applications. *Nature Reviews Genetics, 12*, 111-122.

Nunn C, Hastings AFS, Kalinina O *et al.* (2017) Environmental Influences on the Growing Season Duration and Ripening of Diverse Miscanthus Germplasm Grown in Six Countries. *Frontiers in plant science,* **8**.

Ogawa Y, Honda M, Kondo Y, Hara-Nishimura I (2016) An efficient Agrobacterium-mediated transformation method for switchgrass genotypes using Type I callus. *Plant Biotechnology,* **33**, 19-+.

Ogawa Y, Shirakawa M, Koumoto Y, Honda M, Asami Y, Kondo Y, Hara-Nishimura I (2014) A simple and reliable multi-gene transformation method for switchgrass. *Plant Cell Reports,* **33**, 1161-1172.

Okada M, Lanzatella C, Saha MC, Bouton J, Wu RL, Tobias CM (2010) Complete Switchgrass Genetic Maps Reveal Subgenome Collinearity, Preferential Pairing and Multilocus Interactions. *Genetics,* **185**, 745-760.

Palomo-Ríos E, Macalpine W, Shield I *et al.* (2015) Efficient method for rapid multiplication of clean and healthy willow clones via in vitro propagation with broad genotype applicability. *Canadian Journal of Forest Research*, 1-6.

Pilate G, Allona I, Boerjan W *et al.* (2016) Lessons from 25 years of GM tree field trials in Europe and prospects for the future. In: *Biosafety of Forest Transgenic Trees.* pp 67-100. Springer.

Porth I, Klapste J, Skyba O *et al.* (2013) Network analysis reveals the relationship among wood properties, gene expression levels and genotypes of natural Populus trichocarpa accessions. *New Phytologist,* **200**, 727-742.

Pugesgaard S, Schelde K, Larsen SU, Lærke PE, Jørgensen U (2015) Comparing annual and perennial crops for bioenergy production–influence on nitrate leaching and energy balance. *Global Change Biology Bioenergy,* **7**, 1136-1149.

Quinn LD, Allen DJ, Stewart JR (2010) Invasiveness potential of *Miscanthus sinensis*: implications for bioenergy production in the United States. *Global Change Biology Bioenergy,* **2**, 310-320.

Rae AM, Robinson KM, Street NR, Taylor G (2004) Morphological and physiological traits influencing biomass productivity in short-rotation coppice poplar. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere,* **34**, 1488-1498.

Rambaud C, Arnoult S, Bluteau A, Mansard MC, Blassiau C, Brancourt-Hulmel M (2013) Shoot organogenesis in three Miscanthus species and evaluation for genetic uniformity using AFLP analysis. *Plant Cell, Tissue and Organ Culture (PCTOC),* **113**, 437-448.

Ramstein GP, Evans J, Kaeppler SM, Mitchell RB, Vogel KP, Buell CR, Casler MD (2016) Accuracy of Genomic Prediction in Switchgrass (*Panicum virgatum* L.) Improved by Accounting for Linkage Disequilibrium. *G3-Genes Genomes Genetics,* **6**, 1049-1062.

Rayburn AL, Crawford J, Rayburn CM, Juvik JA (2009) Genome size of three Miscanthus species. *Plant Molecular Biology Reporter,* **27**, 184.

Richardson J, Isebrands J, Ball J (2014) Ecology and physiology of poplars and willows. In: *Poplars and willows: trees for society and the environment.* (eds Isebrands J, Richardson J), pp 92-123. The Food and Agriculture Organization of United Nations and CABI.

Robison TL, Rousseau RJ, Zhang J (2006) Biomass productivity improvement for eastern cottonwood. *Biomass & Bioenergy,* **30**, 735-739.

Robson P, Jensen E, Hawkins S *et al.* (2013a) Accelerating the domestication of a bioenergy crop: identifying and modelling morphological targets for sustainable yield increase in Miscanthus. *Journal of Experimental Botany,* **64**, 4143-4155.

Robson PR, Farrar K, Gay AP, Jensen EF, Clifton-Brown JC, Donnison IS (2013b) Variation in canopy duration in the perennial biofuel crop Miscanthus reveals complex associations with yield. *Journal of Experimental Botany,* **64**, 2373-2383.

Sanderson MA, Adler PR, Boateng AA, Casler MD, Sarath G (2006) Switchgrass as a biofuels feedstock in the USA. *Canadian Journal of Plant Science,* **86**, 1315-1325.

Scarlat N, Dallemand JF, Monforti-Ferrario F, Nita V (2015) The role of biomass and bioenergy in a future bioeconomy: Policies and facts. *Environmental Development,* **15**, 3-34.

Serapiglia MJ, Gouker FE, Smart LB (2014) Early selection of novel triploid hybrids of shrub willow with improved biomass yield relative to diploids. *BMC plant biology,* **14**, 74.

Serba D, Wu LM, Daverdin G *et al.* (2013) Linkage Maps of Lowland and Upland Tetraploid Switchgrass Ecotypes. *BioEnergy Research,* **6**, 953-965.

Shakoor N, Lee S, Mockler TC (2017) High throughput phenotyping to accelerate crop breeding and monitoring of diseases in the field. *Current Opinion in Plant Biology,* **38**, 184-192.

Shield I, Macalpine W, Hanley S, Karp A (2015) Breeding willow for short rotation coppice energy cropping. In: *Industrial Crops.* pp 67-80. Springer.

Sjodin A, Street NR, Sandberg G, Gustafsson P, Jansson S (2009) The Populus Genome Integrative Explorer (PopGenIE): a new resource for exploring the Populus genome. *New Phytologist,* **182**, 1013-1025.

Slavov G, Davey C (2017) Integrated genomic prediction in bioenergy crops In: *Abstracts from the International Conference on 'Developing biomass crops for future climates'* pp 34, 24-27 September 2017, Oriel College Oxford, www.watbio.eu.

Slavov G, Davey C, Bosch M, Robson P, Donnison I, Mackay I (2018a) Genomic index selection provides a pragmatic framework for setting and refining multi-objective breeding targets in Miscanthus. *Annals of Botany*, (accepted with minor revisions).

Slavov G, Davey C, Robson P, Donnison I, Mackay I (2018b) Domestication of Miscanthus through genomic index selection. In: *Plant and Animal Genome Conference.* 13 January 2018, San Diego, CA, https://pag.confex.com/pag/xxvi/meetingapp.cgi/Paper/30622.

Slavov G, Robson P, Jensen E *et al.* (2013) Contrasting geographic patterns of genetic variation for molecular markers vs. phenotypic traits in the energy grass *Miscanthus sinensis*. *Global Change Biology Bioenergy,* **5**, 562-571.

Slavov GT, Difazio SP, Martin J *et al.* (2012) Genome resequencing reveals multiscale geographic structure and extensive linkage disequilibrium in the forest tree *Populus trichocarpa*. *New Phytologist,* **196**, 713-725.

Slavov GT, Nipper R, Robson P *et al.* (2014) Genome-wide association studies and prediction of 17 traits related to phenology, biomass and cell wall composition in the energy grass *Miscanthus sinensis*. *New Phytologist,* **201**, 1227-1239.

Ślusarkiewicz-Jarzina A, Ponitka A, Cerazy-Waliszewska J, Wojciechowicz MK, Sobańska K, Jeżowski S, Pniewski T (2017) Effective and simple in vitro regeneration system of *Miscanthus sinensis*, *M.× giganteus* and *M. sacchariflorus* for planting and biotechnology purposes. *Biomass and Bioenergy,* **107**, 219-226.

Smart L, Cameron K (2012) Shrub willow. In: *Handbook of Bioenergy Crop Plants.* (eds Kole C, Joshi S, Shonnard D), pp 687-708. Taylor and Francis Group, Boca Raton, FL.

Stanton BJ (2014) The domestication and conservation of *Populus* and *Salix* genetic resources (Chapter 4) In: *In Poplars and Willows: Trees for Society and the Environment. .* (eds Isebrands J, Richardson J) pp 124-200, Rome, Italy, CAB International, Food and Agricultural Organization of the United Nations.

Stanton BJ, Neale DB, Li S (2010) Populus breeding: from the classical to the genomic approach. In: *Genetics and genomics of populus.* pp 309-348. Springer.

Stewart JR, Toma Y, Fernandez FG, Nishiwaki A, Yamada T, Bollero G (2009) The ecology and agronomy of *Miscanthus sinensis*, a species important to bioenergy crop development, in its native range in Japan: a review. *Global Change Biology Bioenergy,* **1**, 126-153.

Stott KG (1992) Willows in the service of man. *Proceedings of the Royal Society of Edinburgh. Section B. Biological Sciences,* **98**, 169-182.

Strauss SH, Ma C, Ault K, Klocko AL (2016) Lessons from Two Decades of Field Trials with Genetically Modified Trees in the USA: Biology and Regulatory Compliance. In: *Biosafety of Forest Transgenic Trees.* pp 101-124. Springer.

Suda Y, Argus GW (1968) Chromosome numbers of some North American Salix. *Brittonia,* **20**, 191-197.

Tan B (2018) Genomic selection and genome-wide association studies to dissect quantitative traits in forest trees. Umeå University.

Tornqvist C, Taylor M, Jiang Y, Evans J, Buell C, Kaeppler S, Casler M (2018) Quantitative trait locus mapping for flowering time in a lowland x upland switchgrass pseudo-F2 population. *The Plant Genome* **11**, (in press).

Tóth G, Hermann T, Da Silva M, Montanarella L (2016) Heavy metals in agricultural soils of the European Union with implications for food safety. *Environment international,* **88**, 299-309.

Tracy W, Dawson J, Moore V, Fisch J (2016) Intellectual Property Rights and Public Plant Breeding: Recommendations, and proceedings of a conference on best practices for intellectual property protection of publically developed plant germplasm. pp 70, https://host.cals.wisc.edu/agronomy/wp-content/uploads/sites/16/2016/2005/Proceedings-IPR-Final.pdf (accessed 2018 Sept. 2018). University of Wisconsin-Madison.

Trybush S, Jahodová Š, Macalpine W, Karp A (2008) A genetic study of a Salix germplasm resource reveals new insights into relationships among subgenera, sections and species. *BioEnergy Research,* **1**, 67-79.

Tsai CJ (2013) Next-generation sequencing for next-generation breeding, and more. *New Phytologist,* **198**, 635-637.

Tuskan GA, Difazio S, Jansson S *et al.* (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science,* **313**, 1596-1604.

Tuskan GA, Walsh ME (2001) Short-rotation woody crop systems, atmospheric carbon dioxide and carbon management: a US case study. *The Forestry Chronicle,* **77**, 259-264.

Van Den Broek R, Treffers D-J, Meeusen M, Van Wijk A, Nieuwlaar E, Turkenburg W (2001) Green energy or organic food? A life-cycle assessment comparing two uses of set-aside land. *Journal of Industrial Ecology,* **5**, 65-87.

Van Der Schoot J, Pospiskova M, Vosman B, Smulders MJM (2000) Development and characterization of microsatellite markers in black poplar (Populus nigra L.). *Theoretical and Applied Genetics,* **101**, 317-322.

Van Der Weijde T, Dolstra O, Visser RG, Trindade LM (2017a) Stability of cell wall composition and saccharification efficiency in Miscanthus across diverse environments. *Frontiers in plant science,* **7**, 2004.

Van Der Weijde T, Huxley LM, Hawkins S *et al.* (2017b) Impact of drought stress on growth and quality of miscanthus for biofuel production. *Global Change Biology Bioenergy,* **9**, 770-782.

Van Der Weijde T, Kamei CLA, Severing EI *et al.* (2017c) Genetic complexity of miscanthus cell wall composition and biomass quality for biofuels. *BMC genomics,* **18**, 406.

Van Der Weijde T, Kiesel A, Iqbal Y *et al.* (2017d) Evaluation of *Miscanthus sinensis* biomass quality as feedstock for conversion into different bioenergy products. *Global Change Biology Bioenergy,* **9**, 176-190.

Vanholme B, Cesarino I, Goeminne G *et al.* (2013) Breeding with rare defective alleles (BRDA): a natural Populus nigra HCT mutant with modified lignin as a case study. *New Phytologist,* **198**, 765-776.

Vogel KP, Mitchell RB, Casler MD, Sarath G (2014) Registration of 'Liberty' Switchgrass. *Journal of Plant Registrations,* **8**, 242-247.

Wang X, Yamada T, Kong FJ *et al.* (2011) Establishment of an efficient in vitro culture and particle bombardment-mediated transformation systems in *Miscanthus sinensis* Anderss., a potential bioenergy crop. *Global Change Biology Bioenergy,* **3**, 322-332.

Watrud LS, Lee EH, Fairbrother A *et al.* (2004) Evidence for landscape-level, pollen-mediated gene flow from genetically modified creeping bentgrass with CP4 EPSPS as a marker. *Proceedings of the National Academy of Sciences of the United States of America,* **101**, 14533-14538.

Weisgerber H (1993) Poplar Breeding for the Purpose of Biomass Production in Short-Rotation Periods in Germany - Problems and First Findings. *Forestry Chronicle,* **69**, 727-729.

Wheeler NC, Steiner KC, Schlarbaum SE, Neale DB (2015) The evolution of forest genetics and tree improvement research in the United States. *Journal of Forestry,* **113**, 500-510.

Whittaker C, Macalpine WJ, Yates NE, Shield I (2016) Dry Matter Losses and Methane Emissions During Wood Chip Storage: the Impact on Full Life Cycle Greenhouse Gas Savings of Short Rotation Coppice Willow for Heat. *BioEnergy Research,* **9**, 820-835.

Xi Q (2000) *Investigation on the distribution and potential of giant grasses in China,* Goettingen, Cuvillier Verlag.

Xi Q, Jezowkski S (2004) Plant resources of *Triarrhena* and *Miscanthus* species in China and its meaning for Europe. *Plant Breeding and Seed Science,* **49**, 63-77.

Yin KQ, Gao CX, Qiu JL (2017) Progress and prospects in plant genome editing. *Nature Plants,* **3**.

Yook M (2016) Genetic diversity and transcriptome analysis for salt tolerance in *Miscanthus*. PhD thesis, Seoul National University, Korea.

Zaidi SSEA, Mahfouz MM, Mansoor S (2017) CRISPR-Cpf1: A New Tool for Plant Genome Editing. *Trends in Plant Science,* **22**, 550-553.

Zalesny RS, Stanturf JA, Gardiner ES *et al.* (2016) Ecosystem Services of Woody Crop Production Systems. *BioEnergy Research,* **9**, 465-491.

Zhang QX, Sun Y, Hu HK *et al.* (2012) Micropropagation and plant regeneration from embryogenic callus of Miscanthus sinensis. *In Vitro Cellular & Developmental Biology-Plant,* **48**, 50-57.

Zhang YW, Zalapa JE, Jakubowski AR *et al.* (2011) Post-glacial evolution of *Panicum virgatum*: centers of diversity and gene pools revealed by SSR markers and cpDNA sequences. *Genetica,* **139**, 933-948.

Zhao H, Wang B, He J, Yang J, Pan L, Sun D, Peng J (2013) Genetic diversity and population structure of *Miscanthus sinensis* germplasm in China. *PloS one,* **8**, e75672.

Zhou R, Macaya-Sanz D, Rodgers-Melnick E *et al.* (2018) Characterization of a large sex determination region in *Salix purpurea* L.(Salicaceae). *Molecular Genetics and Genomics*, 1-16.

Zhou XH, Jacobs TB, Xue LJ, Harding SA, Tsai CJ (2015) Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial Populus reveals 4-coumarate: CoA ligase specificity and redundancy. *New Phytologist,* **208**, 298-301.

Zsuffa L, Mosseler A, Raj Y (1984) Prospects for interspecific hybridization in willow for biomass production. *Rapport-Sveriges Lantbruksuniversitet, Institutionen foer Ekologi och Miljoevaard (Sweden)*.

Figure captions

Figure 1 Cumulative minimum years needed for the conventional breeding cycle through the steps from wild germplasm to the commercial hybrids in switchgrass, miscanthus, willow and poplar. Information links between the steps are indicated by dotted arrows, and highlight the importance of long-term informatics to maximise breeding gain.

Figure 2 A schematic development pathway for miscanthus in the UK related to the investment in R&D projects at Aberystwyth (top coloured areas for projects in the three categories: basic research, breeding, and commercial upscaling) leading to a projected cropping area of 350,000 ha by 2030 with clonal and successive ranges of improved seed based hybrids. Purple represents the Biotechnology and Biological Sciences Research Council (BBSRC) and brown the Department for Environment, Food and Rural Affairs (Defra) (UK National funding), blue bars represents EU funding, green private sector funding (Terravesta and CERES) and GIANT-LINK and Miscanthus Upscaling Technology (MUST) are public-private-initiatives (PPI).

Figure 3 A schematic relating some of the steps in the innovation chain from relatively basic crop science research through to the deployment in commercial cropping systems and value chains. The shape of the funnel above the expanding development and deployment represents the availability of investment along the development chain from relatively basic research at the top to the upscaled

deployment at the bottom. Plant breeding links the research effort with the development of cropping systems. The constriction represents the constrained funding for breeding that links conventional public research investment and the potential returns from commercial development. The handover points between publicly funded work to develop the germplasm resources (often known as pre-breeding), the breeding, and the subsequent crop development are shown on the left. The constriction point is aggravated by the lack academic rewards for this essential breeding activity. The outcome is such that this innovation system is constrained by the precarious resourcing of plant breeding. The authors' assessment of development status of the four species is shown (poplar having two: one for short rotation coppice (SRC) poplar and one for the more traditional short rotation forestry (SRF)). The four new perennial biomass crops (PBCs) are now in the critical phase of depending of plant breeding progress without the income stream from a large crop production base.
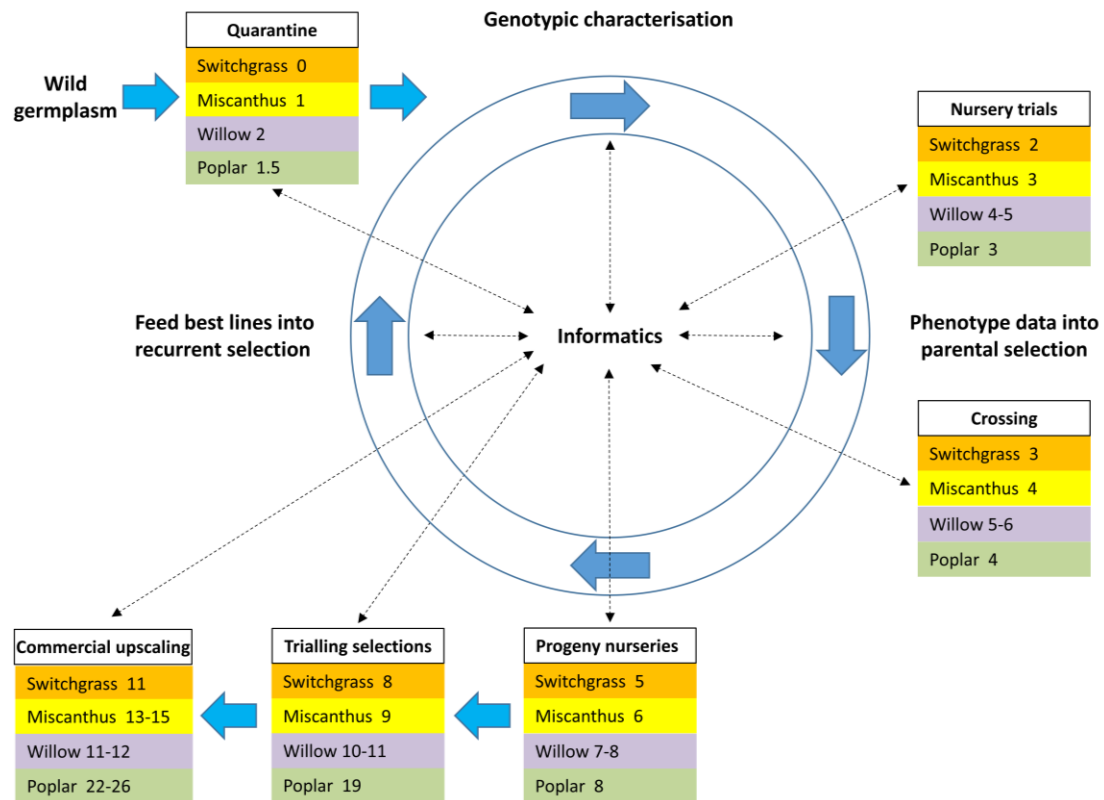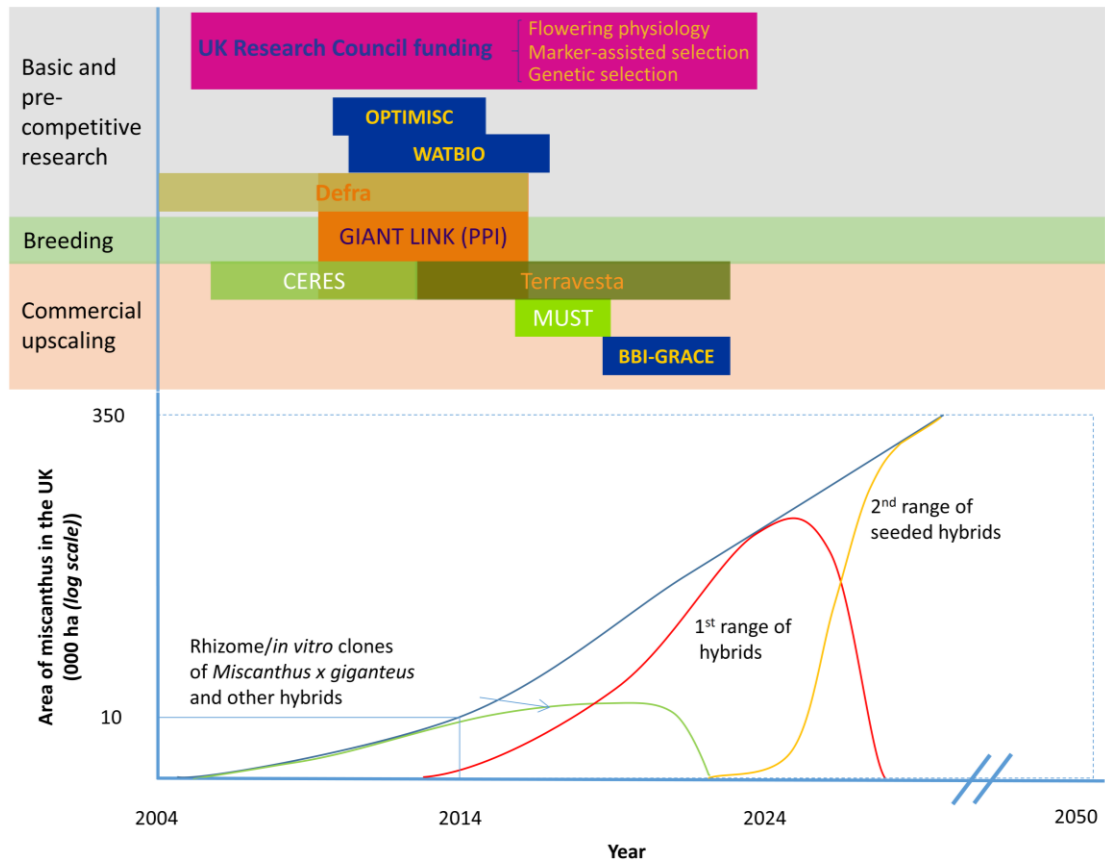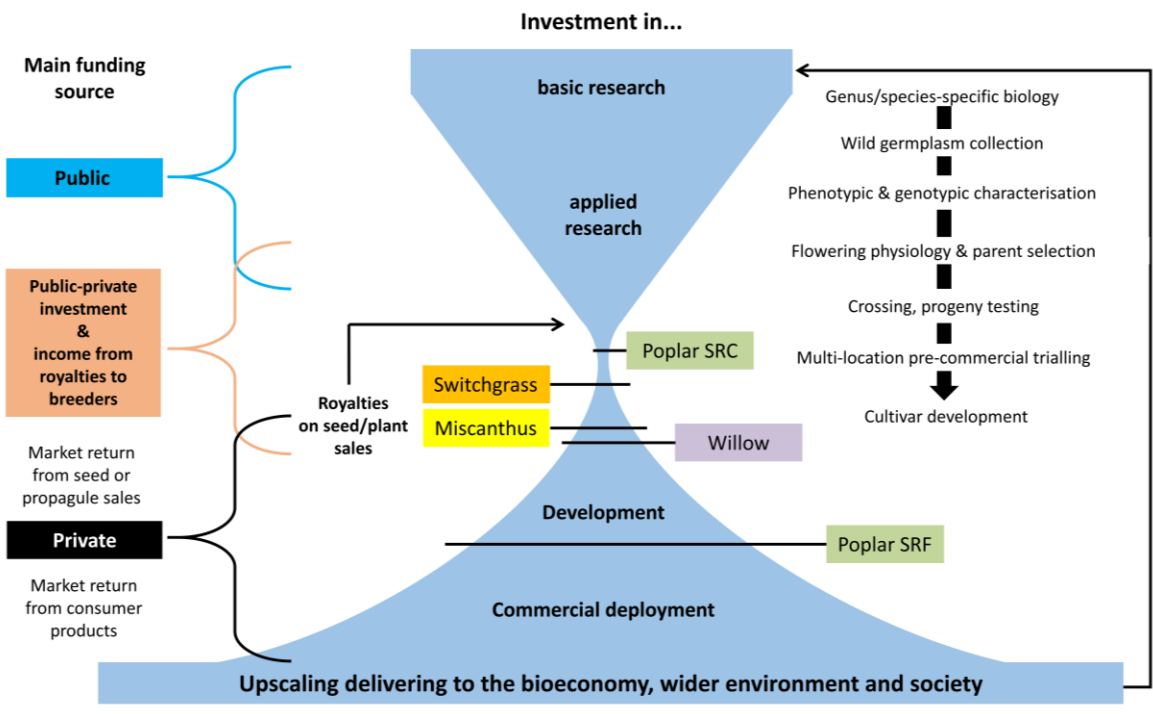

Figure 1

Figure 2

Figure 3

Table 1 Breeding related attributes for four leading perennial biomass crops (PBCs).

Species	Switchgrass	Miscanthus	Willow	Poplar
Type	C4 – Grass	C4 – Grass	C3 – SRC	C3 – SRC/SRF
Sources of indigenous germplasm	CA[§] to MX, east of Rocky Mountains.	Eastern Asia and Oceania	Predominantly Northern hemisphere	Northern hemisphere
Breeding system	Monoecious, outcrossing	Monoecious, outcrossing	Dioecious, outcrossing	Dioecious, outcrossing
Ploidy	4x, 8x	2x, 3x, 4x	2x-12x	2x
Species within *genus*	~450	~14	~400	~30-32
Types used mainly for breeding	**US:** Lowland ecotype (subtropical climates) and Upland ecotype (temperate climates)	**EU[$], JP, SK** and **US:** *M.sin.* and *M.sac.*	**EU:** *S. viminalis × schwerinii, S. dasyclados × rehderiana, S. dasyclados, S. viminalis* **US & UK:** *S. viminalis × miyabeana, S. miyabeana* **US:** *S. purpurea × miyabeana, S. purpurea*	*Populus trichocarpa, P. deltoides, P. nigra, P. suaveolens subsp. maximowiczii, P. balsamifera P. alba* and hybrids
Typical haploid genome size (Mbp)	~1500	*M.sin.* ~5400 and *M.sac.* ~4400 (Rayburn *et al.,* 2009)	~450	~485 ± 10
Breeding programmes:	CA: 2000 (REAP, Quebec) US: - 1992 (Nebraska) - 1992 (Oklahoma) - 1992 (Georgia) - 1996 (Wisconsin) - 1996 (South Dakota) - 2000 (Tennessee) - 2002 (Mississippi) - 2007 (Oklahoma) - 2008 (Rutgers, New Jersey) - 2012 (Cornell, New York) - 2012 (Urbana-Champaign, Illinois)	DE: 1990s (Klein-Wanzleben) NL: 2000s (Wageningen) UK: 2004 (Aberystwyth) CN: 2006 (Changsha) JP: 2006 (Hokkaido) US: - 2006 (California – CERES Inc.) - 2006 (California & Indiana – MBI) - 2008 (Urbana-Champaign) SK: 2009 (Suwon-SNU) & (Muan, NICS of RDA) FR: 2011 (Estrées-Mons)	UK: 1980s (Long Ashton, relocated to Rothamsted Research in 2002) SE: 1980s (Svalöf Weibull / Salixenergi Europa AB) US: 1990s (Cornell, New York) PL: 2000s (Olsztyn, University of Warmia and Mazury)	SE: - 1939 (Mykinge, Ekebo Research Institute) - 1990's (Upsala, SLU) - 2010 (Uppsala, STT) US: - 1927 (New York, Oxford Paper Company & New York Botanical Garden, Wheeler *et al.,* 2015) - 1979 (Washington, UW & Oregon, GWR) - 1980-1995 (Mississippi, MSState & GWR) - 1996 (Minnesota, UMD NRRI & GWR) IT: 1983 (Piedmont, AFV) FR: 2001 (Orléans & Nancy, INRA; Charrey-sur-Saône & Pierroton, FCBA; Nogent-sur-Vernisson, IRSTEA) DE: 2008 (Göttingen, NW-FVA)
Current commercial varieties on the market	US: No commercial hybrids	CA: 2 EU: 1 (*M×g* from different origins) + selected *M.sin.* for thatching in DK US: 3 (but 2 are genetically identical)	UK: 25 US: 8	EU: DE: <10, FR: 44, IT: 10-15, SE: ~14, US: 8-12 Southeast, 8-14 Upper Midwest, 10 Pacific Northwest
Pre-commercial cultivars expected to be on the market in 3 years	US: 36 registered cultivars (half are random seed increases from natural prairies, half are bred varieties); Most are public releases; few are protected, patented, or licensed.	NL: 8, seeded hybrids, van Dinter Semo, MTA. UK: 4, seeded hybrids, CERES (Land O'Lakes) & Terravesta Ltd., MTA US: None FR: None	EU: 53 registered with CPVO for PBR UK: 20 registered with CPVO for PBR US: 18 clones from Cornell in multisite trials	CA: unquantified, UAlberta & Quebec FR: 8-12 SRF clones, license-based IT: 5 SRC & 6 SRF, AFV, MTA or license SE: 14, STT, license-based & MTA US: ~19 Southeast; 5-7 Upper Midwest from UMD NRRI, MTA; 6-12 North Central from GWR, MTA; Pacific Northwest: 8 clones, GWR, MTA; 24 in multi-location yield trials, GWR.
Commercial yield (t DM ha⁻¹ y⁻¹)	US: 3-18 EU: 8-12	CN: 20 - 30 US: 10- 25 EU: 7 – 20	UK: 8-14 US: 8-14	EU: 5-20 (SE and Baltic Countries: 8-12) US: 10-22 (10-12 North Central, 12-16 Southeast, 15-22 Pacific Northeast)
Harvest rotation and commercial stand life span	Annual for 10-12 years	Annual for 10-25 years (*M .sac* has been used for ~30 years in China)	2-4 year cycle for 22-30 years	SRC: 3-7 year cycle for 20 years SRF: 10-12 year cycle for >50 years
Adaptive range	Open-pollinated and synthetic cultivars are limited in adaptation by temperature and precipitation (~8 breeding zones in US and CA)	Standard *M×g* is widely adapted in EU, but is limited in the US by insufficient winter hardiness for Northern Midwest and heat intolerance in the south. Novel *M.sac. × M.sin.* hybrids and *M.sin. × M.sin.* hybrids selected in continental DE have shown a wide adaptive range in EU (Kalinina *et al.,* 2017). Ongoing trials on heavy metal contaminated soils indicate tolerance by exclusion (Krzyżak *et al.,* 2017).	Different hybrids are needed for different zones Best hybrids show some G × E (US), some hybrids low G × E (Fabio *et al.,* 2017)	Different hybrids are needed for different climatic zones. Hybrids that are adapted for growing seasons of ~6 months & relatively short days in Southern Europe are maladapted to short growing seasons of ~4 months & relatively long days in Northern EU. The most broadly adapted varieties come from the *P. canadensis* taxon.

[§]ISO Alpha-2 letter country codes; [Φ]EU is used for Europe.

Abbreviations: AFV (Alasia Franco Vivai); CPVO (Community Plant Variety Office); Cornell (Cornell University); FCBA (Forest, Cellulose, Wood, Construction and Furniture Technology Institute); G × E (genotype-by-environment interaction); GWR (GreenWood Resources); INRA (French National Institute for Agricultural Research); IRSTEA (National Research Unit of Science and Technology for Environment and Agriculture); MBI (Mendel Biotechnology Inc.), Mbp (mega base pair); *M. sin* (*Miscanthus sinensis*); *M. sac* (*M. sacchariflorus and M×g* (*M. × giganteus*)); MSState (Mississippi State University); MTA (material transfer agreement); SRC (short rotation coppice); NICS (National Institute of Crop Science); NW-FVA (Northwest German Forest Research Institute); PBR (Plant breeder's right); RDA (Rural Development Administration); REAP (Resource Efficient Agriculture Production); SLU (Swedish University of Agricultural Sciences); SRF (short rotation forestry); SNU (Seoul National Uni.); STT (SweTree Tech); t DM ha⁻¹ y⁻¹ (tons of dry matter per hectare per year); University of Alberta (UAlberta); University of Minnesota Duluth's Natural Resources Research Institute (UMD NRRI); UMN (University of Minnesota); UWM = University of Warmia & Mazury; UW (University of Washington).

Table 2 Generalised improvement targets for perennial biomass crops (PBCs).

Net energy yield per hectare
• Increased yield
• Reduced moisture content at harvest
Physical and chemical composition for different end-use applications
• Increased lignin content and decreased corrosive elements for thermal conversion
• Reduced recalcitrance through decreased lignin content and/or modified lignin monomer composition to reduce pretreatment requirements for next generation biofuels by saccharification and fermentation
• Plant morphological differences which influence biomass harvest, transport and storage (e.g. stem thickness)
Propagation costs
• Improved cloning systems (trees and grasses)
• Seed systems (grasses)
• Optimising agronomy for each new cultivar
Resilience through enhanced
• Abiotic stress tolerance/resistance (e.g. drought, salinity, and high and low temperature)
• Biotic stress resistance (e.g. insects, fungal, bacterial, and viral diseases)
• Site adaptability especially to those of marginal / contaminated agricultural land

Table 3 Preparedness for mass upscaling: Current market value and research investment in four perennial biomass crops (PBCs).

Species	Switchgrass	Miscanthus	Willow	Poplar
Current commercial planting costs per ha	US[§]: 200 USD[¡] in the establishment year	DE: 3375 Euro[¡] (Xue et al., 2015) UK: 2153 GBP[¡] (Evans, 2016) reduced to 1169 GBP with *Mxg* rhizome where the farmer does the land preparation (www.terravesta.com) US: 1730-2225 USD	UK: 1500 – 1739 GBP plus land preparation (Evans, 2016) US: 1976 USD (= 800 USD / acre)	IT: 1100 Euro FR/SE: 1000-2000 Euro US: 863 USD / ha (Lazarus et al., 2015)
Current market value of the biomass per t DM	US: 80-100 USD	UK: ~80 GBP (bales) (Terravesta, personal communication) US: 94 USD (chipped)	UK: 49.41 GBP (chipped) (Evans, 2016) US: 55-70 USD	IT: 100 Euro US: 80-90 USD delivered based on 40 miles of haulage
Science for genetic improvement: projects in the last 10 years	US: > 50 projects funded by US-DOE and USDA-NIFA	CN: >30 projects funded by CN-NSFC and MBI EU[◊]: 6 projects (OPTIMISC, OPTIMA, WATBIO, SUNLIBB, FIBRA, GRACE) FR: BFF SK: 2 projects funded by IPET and PMBC. US: EBI, CABBI, multiple DOE Feedstock Genomics and USDA AFRI projects	UK: BSBEC, BEGIN (2003-2010) US: US DOE JGI Genome Sequencing (2009-2012; 2015-2018); USDA Northeast Sun Grant (2009-2012); USDA NIFA NEWBIO Consortium (2012-2017), USDA NIFA Willow SkyCAP (2018-2021)	EU: FP7 (Energy Poplar, NovelTree, WATBIO, Tree4Future) SE: Climate adapted poplars & Nanowood, SLU & STT US: Bioenergy Science (ORNL), USDA feedstocks genomics programme, BRDI, BRC-CBI
Major projects supporting crossing and selection cycles in the last 10 years	US and CA: 12 projects	NL: RUE miscanthus, PPP, 2015-2019, 50-100 K Euro / year UK: GIANT-LINK, PPI, 2011-2016, 1.3 M GBP / year US: EBI / CABBI, ~0.5 M USD / year	UK: BEGIN 2000-2010 US: USDA Northeast Sun Grant (2009-2012); USDA NIFA NEWBio Consortium (2012-2017), USDA NIFA Willow SkyCAP (2018-2021)	FR: 14 regional & national projects, 100-150k Euro /year SE: STT: one breeding effort as an internal project in 2010 with a resulting progeny trial US: DOE Sun Grant feedstock development partnership programme (including Willow)
Current annual investment in projects for translation into commercial hybrids	US: > 20M USD / year	UK: MUST, 2016-2019, 0.5 M GBP / year US: CABBI/USDA AFRI, 2017-2022 0.5 M USD / year	US: USDA NIFA NEWBio ~0.4 M USD / year	FR: Science for improvement: ~20k Euro/year US: USDA under AFRI Co-ordinated Ag. Producer projects. And several translational genomics projects without public funding.
Upscaling time (years) to produce sufficient propagules to plant > 100 ha	US: Using seed 2-3 years	UK: Rhizome for 1-2000 hectares can be ready in 6 months. UK/NL: 20 ha of seed hybrids planted in 2018. In 2019 sufficient seed for 50 – 100 ha is expected.	UK: 3 years using conventional cuttings, faster using micro-propagation	FR, IT: 3 years by vegetative propagation SE: >3 years by vegetative propagation (cuttings), 3 years by micro-propagation

[§]ISO Alpha-2 letter country codes; [◊]EU is used for Europe; [¡]Local currencies are used as at 2018: Euro, GBP (Great Britain Pound); USD (US Dollar).
Abbreviations: AFRI (Agriculture and Food Research Initiative in the US); BEGIN (Biomass for Energy Genetic Improvement Network); BFF (Biomass For the Future); BRC-CBI (Bioenergy Research Centre- Centre for Bioenergy Innovation); BRDI (Biomass research development initiative); BSBEC (BBSRC Sustainable Bioenergy Centre (BSBEC); CABBI (Center for advanced bioenergy and bioproducts innovation in the US); CN-NSFC (Natural Science Foundation of China); DOE (Department of Energy in the US); EBI (Energy Biosciences Institute); FIBRA (Fibre crops as sustainable source of biobased material for industrial products in Europe and China); FP7 (Seventh Framework Programme in the EU); GIANT-LINK (Genetic improvement of miscanthus as a sustainable feedstock for bioenergy in the UK); GRACE (GRowing Advanced industrial Crops on marginal lands for biorEfineries); IPET (Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry and Fisheries); JGI (Joint Genome Institute); MBI (Mendel Biotechnology Inc.); MUST (Miscanthus UpScaling Technology); NEWBio (Northeast Woody/Warm-season Biomass Consortium); NIFA (National Institute of Food and Agriculture in the US); NovelTree (Novel tree breeding strategies); OPTIMA (Optimization of perennial grasses for biomass production in the Mediterranean area); OPTIMISC (Optimizing bioenergy production from Miscanthus); ORNL (Oak Ridge National Lab); PMBC (Plant Molecular Breeding Center of the Next Generation Biogreen Research Centers of the Republic of Korea); PPI (Public-private investment); PPP (Public private partnership); RUE (radiation use efficiency); SkyCAP (Coordinated Agricultural Project); SLU (Swedish University of Agricultural Sciences); STT (SweTree Tech); SUNLIBB (Sustainable Liquid Biofuels from Biomass Biorefining); Tree4Future (Designing Trees for the future); USDA (United States Department of Agriculture); WATBIO (Development of improved perennial non-food biomass and bioproduct crops for water stressed environments).

Table 4 Pre-breeding research and the status of conventional breeding in four leading perennial biomass crops (PBCs).

Breeding technology step	Use	Pre-requisite steps	Limitations	Switchgrass	Miscanthus	Willow	Poplar
1. Collected wild accessions or secondary sources available for breeding.	Provide a broad base of useful traits	Respect for CBD and Nagoya protocol on collections after 2014	Not all indigenous genetic resources are accessible under CBD for political reasons	US[§]: ~2000 (181 in official GRIN gene bank of which 96 were available; others in various unofficial collections)	CN: Changsha, ~1000 and Nanjing, 2000 from China DK: ~120 from JP FR: working collection of ~ 100 (mainly *M sin*) JP: ~1500 from JP, SK and CN NL: working collection of ~300 *M.sin.* SK: ~700 from SK, CN, JP and RU UK: ~1500 accessions (from >500 sites) of which 1000 are in field nurseries in 2018 US: 14 in GRIN global US: Illinois, ~1500 collections made in Asia with 25% currently available in US for breeding	UK: 1500 with about 20 in common with the US US: 350 largely unique	FR : 3370 *Populus deltoides* (650), *P. nigra* (2000), *P. trichocarpa* (600), *P. maximowiczii* (120) IT: 30000, *P. deltoids, P. nigra, P. maximowiczii* and *P. trichocarpa* SE: 13000 accessions, mainly *P. trichocarpa*, by STT and SLU US: - GWR: 150 *P. trichocarpa* and 300 *P. maximowiczii* accessions for Pacific Northwest program; 535 *P. deltoides* and ~200 *P. nigra* accessions for Southeast program & 77 *P. simonii* from Mongolia. - UMD NRRI: 550+ clonal accessions (primarily *P. × canadnesis*, along with *P. deltoides* and *P. nigra* parents) for Upper Midwest program
2. Wild accessions which have undergone phenotypic screening in field trials	Selection of parental lines with useful traits	Strong partnerships to run multi-location field trials to phenotype consistently the accessions / genotypes in different environments & databases	Cost of running multi-location trials	US: >500,000 genotypes	CN: ~1250 in Changsha, ~1700 in Hunan, Jiangsu, Shandong, Hainan NL: ~ 250 genotypes JP: ~1200 SK: ~400 UK led: ~1000 genotypes in a replicated trial US led: ~1200 genotypes in multi-location replicated trials in Asia and North America: *M.sin.*: SK, CN, JP, CA (Onatario), US (Illinois & Colorado); *M.sac.*: SK (Kangwon), CN (Zhejiang), JP (Sapporo), CA (Onatario), US (Illinois), DK (Aarhus)	UK: >400 US: 180	FR : 2720 IT: 10000 SE: ~150 US: - GWR: ~1500 wild *P. trichocarpa* phenotypes and OP seedlots (total of 70) from the *P. maximowiczii* species complex (*suaveolens, cathayana, koreana, ussuriensis, maximowiczii*) and 2000-3000 *P. deltoides* collections (based on 104 2nd generation families) - UMD NRRI: screened seedlings from a *P. deltoides* (OP or CP) breeding program (1996-2016). >7200 OP seedlings of *P. nigra* under Minnesota climatic conditions (improve parental population)

				US	Country data[§]		
3. Wild germplasm genotyping	Construct phylogenetic trees and dissimilarity indices	A managed living collection of clonal types	The type of molecular analysis - AFLP, cpDNA, RADseq, whole genome sequencing	US: ~20,000 genotypes	CN: Changsha, ~1000; Nanjing, 37 FR: 44 by cpDNA (Feng et al., 2014) JP: ~1200 NL: 250 genotypes by RADseq UK: ~1000 by RADseq SK: ~300 US: Illinois: RADseq on all wild accessions.	UK: ~400 US: 225	FR: 2310 SE: ~150 US: 1500
4. Exploratory crossing and progeny tests	Discover good parental combinations - general combining ability	Geographic separation, species or phylogenetic trees	Costly long-term multi-location trials for progeny	US: >10,000	CN: ~10 FR: ~200 JP: ~30 NL: 3600 SK: ~20 UK: ~4000 (~1500 *M.sin.*) US: 500-1000	UK: - >700 since 2003 - >800 EWBP (1996-2002) US: 800	FR: Cloned 13x13 factorial mating design with 3 species [10 *P. deltoides*, 8 *P. trichocarpa*, 8 *P. nigra*) US: - GWR: 1000 exploratory crossings of *P. fremontii*, *P. simonii* for improved drought tolerance - UMD NRRI: crossed native *P. deltoides* with non-native *P. trichocarpa* and *P. maximowiczii*. Majority of progeny populations suffered from climatic and/or pathogenic susceptibility issues
5. Wide intraspecies hybridisation	Discover good parental combinations - general combining ability	1 to 4 above, informatics	Flowering synchronisation and low seed set	US: ~200	CN: ~10 NL: 1800 SK: ~10 UK: ~500	UK: 276 US: 400	IT: 500 SE: 120 US: 50-100
6. Within species recurrent selection	Concentration of positive traits	Identification of the right heterotic groups in steps 1-5	Difficult to introduce new germplasm without dilution of best traits	US: >40 populations undergoing recurrent selection	NL: 2 populations UK: 6 populations US: ~12 populations to be established by 2019	UK: 4 US: 150	FR: *P. deltoides* (>30 FS families – 900 clones), *P. nigra* (>40 FS families – 1600 clones), *P. trichocarpa* (15 FS Families -350 clones) IT: 80 SE: ca. 50 parent breeding populations within *P. trichocarpa* US: Goals is ~100 parent breeding populations within *P. trichocarpa*, *P. deltoides*, *P. nigra* and *P. maximowiczii*
7. Interspecies hybrid breeding	Combining complementary traits to produce good morphotypes and heterosis effects	All the steps 1-6	In early stage improvements are unpredictable. A wide base is costly to manage	None	CN: Changsha, 3; Nanjing, 120 *M.sin.* × *M.flor.* JP: ~5 with *Saccharum spontaneum* SK: 5	UK: 420 US: 250	FR : *P. canadensis* (1800), *P. deltoides* × *P. trichocarpa* (800), *P. trichocarpa* × *P. maximowiczii* (1200) IT: ~120 US: 525 genotypes in eastside hybrid program, 205 westside hybrid program and ~1200 in southeast *P. deltoides* program
8. Chromosome doubling	A route to triploid seeded types, doubling a diploid parent of known breeding value from recurrent selection	Needs 1-7 to help identify the right parental lines	Doubled plants are notorious for reverting to diploid	US: ~50 plants taken from tetraploid to octaploid, which are now in field trials	CN: ~10 *M.sac.* triploids DK: Several attempts in mid 90s FR: 5 genotypes of M sin UK: 2 *M.sin.*, 1 *M.flor.* and 1 *M.sac.* US: ~30	UK: Attempted but not routinely used US: Not part of the regular breeding because there existing natural polyploids	N/A
9. Double Haploids	A route to producing homogeneous progeny, and also for the introduction of transgenes or for genome editing	Needs 1-7 to help identify the right parental lines	Fully homozygous plants are weak and easily die, and may not flower synchronously	US: None yet attempted, due to poor vigour and viability of haploids	CN: 2 genotypes of *M.sin* and *M.flor* UK: Di-haploids of *M.sin*, *M.flor* and *M.sac.* and 2 transgenic *M.sin.* via anther culture US: 6 created and used to identify miss-identifying paralogue loci (caused by recent genome duplication in miscanthus). Used in creating the reference genome for *M.sin*. Very weak plants, and difficult to retain the lines	N/A	N/A
10. Embryo rescue	An attractive technique for recovering plants from sexual crosses		The majority of embryos cannot survive in vivo or become long-time dormant	None	UK: one 3x hybrid	UK: Embryos rescue protocol proved robust at 8 days post pollination	FR: An embryo rescue protocol proved robust and improved hybridization success for *P. deltoides* × *P. trichocarpa* crosses
11. Pollen storage	Flexibility to cross interesting parents without need of flowering synchronisation			None	No success to date, notoriously difficult with grasses including sugarcane	Fresh pollen commonly used in crossing. Pollen storage in some interspecific crosses	FR: Both stored (cryobank) and fresh individual pollen

[§]ISO Alpha-2 letter country codes.

Abbreviations: AFLP (amplified fragment length polymorphism); CBD (convention on biological diversity); CP (controlled pollination); CpDNA (chloroplast DNA); EWBP (European willow breeding programme); FS (full-sib); G × E (genotype-by-environment); GRIN (germplasm resources information network); GWR (GreenWood Resources); *M.flor.* (*Miscanthus floridulus*); *M.sac.* (*M. sacchariflorus*); *M.sin.* (*M. sinensis*); N/A (not applicable); NRRI (Natural Resources Research Institute); OP (open pollination); RADseq (restriction site-associated DNA sequencing); SLU (Swedish University of Agricultural Sciences); STT (SweTree Tech); UMD (University of Minnesota Duluth).

Table 5 Status of modern plant breeding techniques in four leading perennial biomass crops (PBCs).

Breeding technology	Use	Pre-requisite steps	Limitations	Switchgrass	Miscanthus	Willow	Poplar
MAS	Use of marker sequences that correlate with a trait allowing early progeny selection or rejection / Locating known genes for useful traits (such as height) from other species in your crop	Breeding programme relevant bi-parental crosses (Table 4) to create 'mapping populations' and QTL identification and estimation to identify robust markers; Alternatively, a comprehensive panel of unrelated genotypes for GWAS	Ineffective where traits are affected by many genes with small effects	**US:** Literally dozens of studies to identify SNPs and markers of interest. No attempts at MAS as yet, largely due to population specificity	**CN:** SSR, ISSR markers for *M.sin.*, *M.sac.* & *M.flor.* **FR:** 2 *M.sin* mapping populations (BFF project) **NL:** 2 *M.sin* mapping populations (Atienza *et al.*, 2003b, Atienza *et al.*, 2003a, van der Weijde *et al.*, 2017) **SK:** 2 GWAS panels (1 collection of *M. sin*, 1 diploid bi-parental F₁ population) in the pipeline **UK:** 12 families to find QTL in common in different families of the same and different species and hybrids **US:** 2 large GWAS panels, 4 diploid bi-parental F₁ populations, 1 F₂ population (additional in pipeline)	**UK:** 16 families for QTL discovery, 2 crosses MAS screened, 1 potential variety selected **US:** 9 families, genotyped by GBS (8 are F₁, one is F₂), a number of promising QTL, development of markers in progress.	**FR (INRA):** tested in large FS families and in factorial mating design, low efficiency due to family specificity **US:** 49 families
GS	Method to accelerate breeding through reducing the resources for cross attempts by predicting the performance of progeny of crosses (and in research to predict best parents to use in bi-parental crosses)	A 'training population' of individuals from stages above that have been both genotyped and phenotyped to develop a model that takes genotypic data from a 'candidate population' of untested individuals and produces GEBVs (Jannink *et al.*, 2010)	Risks of poor prediction. Progeny testing needs to be continued during the time the training set is being phenotyped and genotyped. In this time, next generation germplasm.	**US:** One program so far - USDA in Wisconsin. Three cycles of genomic selection completed and ready to begin the second round of training and recalibration of genomic prediction models	**CN:** ~1000 genotypes **NL:** preliminary models for *M.sin* developed **UK:** ~1000 genotypes **US:** Preliminary models for *M.sac* and *M.sin* developed. Work is underway to deal more effectively with polyploid genetics	Very suitable application, not attempted yet; Maybe challenging for inter-species hybrids	**FR (INRA):** 1200 genotypes *Populus nigra* are being used to develop intraspecific GS calibrations **US:** 1250 genotypes are being used to develop GS calibrations
Traditional transgenesis	Efficient introduction of 'foreign' traits (possibly from other genus / species) into an elite plant that needs a simple trait to be improved; Validate candidate genes from QTL studies	MAS and knowledge of the biology of the trait and source of genes to confer the relevant changes to phenotype. Working transformation protocol	IP issues / cost of regulatory approval / GMO labeling / marketing issues/ tricky to use transgenes in out-breeders because of complexity of transforming and gene flow risks	**US:** Many programs in US are creating transgenic plants, using Alamo as a source of transformable genotypes. Many traits of interest. Nothing commercial yet	**CN:** Changsha: *Bt* gene transformed into *M.sac.* in 2004 *M. sin* transformed with *MINAC2* gene and *M.sac.* transformed with *Cry2Aa* gene, both with a marker gene by *Agrobacterium*. **JP:** 1 *M.sin.* transgenic for low temperature tolerance with increased expression of fructans (unpublished) **NL:** Improving protocol for *M.sin* transformation **SK:** 2 *M.sin* with *Arabidopsis* and *Brachypodium Phytochrome B* genes (Hwang *et al.*, 2014) **UK:** 2 *M.sin* and 1 *M.flor.* genotypes variously transformed with four cell wall genes, *ipt* and *uidA* genes using two selection systems by biolistics and *Agrobacterium*. Transformed plants are being analysed. **US:** Preliminary work will be taken forward in 2018-2022 in CABBI	**UK:** Routine transformation not yet possible. Research to overcome recalcitrance ongoing. Currently trying different species and conditions. Poplar transformation used at present **US:** Frequently attempted with very limited success	**US:** 600 transgenic lines have been characterized; mostly performed in the aspen hybrids; reproductive sterility, drought tolerance, with Knockdowns
Genome editing CRISPR	Refinement of existing traits in useful parents or promising hybrids by generating targeted mutations in genes known to control the trait of interest. First a double stranded break is made in the DNA which is repaired by natural DNA repair machinery. Leads to frame-shift / SNP / or can use a 'repair template' or can be used to insert a transgene into a 'safe harbour locus'; It could be used to delete repressors (or transcription factors)	Identification, mapping and sequencing of target genes (from DNA sequence); Avoiding / screening out of unintended edits	Many regulatory authorities have not decided whether CRISPR and other genome editing technologies are GMOs or not. If GMOs see comment on stage 10 above. If not, edited crops will be regulated as conventional varieties	**CA:** Technology is still too new and switchgrass genome is very complex; Other labs are interested, but not yet moving on this **US:** One program so far - USDA in Albany	**FR:** Initiated in 2016 (MISEDIT project) **NL:** Initiating in 2018 **US:** Initiating in 2018, in CABBI **UK:** Not started	Not possible until transformation achieved	**US:** CRISPR-Cas9 and Cpf1 have been successfully developed in *Populus* and are highly efficient; Examples for Lignin biosynthesis

§ISO Alpha-2 letter country codes.

Abbreviations: BFF (Biomass For the Future); *Bt* (*Bacillus thuringiensis*); CABBI (Center for advanced bioenergy and bioproducts innovation); CRISPR (clustered regularly interspaced palindromic repeats; CRISPR-associated (CRISPR-Cas); CRISPR from *Prevotella* and *Francisella* 1 (Cpf1); Cry2Aa (crystal toxins 2Aa subfamily produced by *Bt*); FS (full-sib); GBS (genotyping-by-sequencing); GEBVs (genomic estimated breeding values); GS (genomic selection); GMO (genetically modified organism); GWAS (genome-wide association study); INRA (French National Institute for Agricultural Research); IP (intellectual property); IPT (isopentenyltransferase); ISSR (inter SSR); *M.flor.* (*Miscanthus floridulus*); *M.sac.* (*M. sacchariflorus*); *M.sin.* (*M. sinensis*); MAS (marker assisted selection); MISEDIT (miscanthus gene editing for seed propagated triploids); NAC (no apical meristem, ATAF1/2, and cup-shaped cotyledon2)-like); QTL (quantitative trait locus); SNP (single nucleotide polymorphism); SSR (simple sequence repeats); USDA (United States Department of Agriculture)