Accelerating Willow Breeding and Deployment (AWBD)

Conducted by Rothamsted Research

Funded by the Department for Energy Security & Net Zero

Executive summary

The Accelerating Willow Breeding and Deployment project was supported by the Department of Energy Security and Net Zero through the Net Zero Innovation Portfolio. The overall objective of this programme was to increase the UK's supply of sustainable biomass.

Rothamsted Research have progressed their willow (*Salix* spp.) biomass breeding programme by; expanding the range of environments they can give growing advice on, developing the ability to apply Genomic Selection to the breeding and exploiting Genomic Selection with rapid multiplication of breeding material.

A meta-analysis of previous field trial data was used to create a variety selection advisory leaflet for willow growers. Data were derived from a narrow set of environments. A network of five new field sites was established to address two objectives; precision deployment of optimal varieties for diverse growing environments and implementation of a Genomic Selection strategy that will accelerate the breeding of improved willow varieties.

A Training Population consisting of 560 distinct genotypes was designed and planted at four field sites. Intensive phenotyping generated data on phenology, pests, diseases and yield. Whole genome sequence data was generated for all genotypes and used with phenotype data to generate Genomic Estimated Breeding Values and test genomic prediction models. The fifth site contained 144 genotypes, 75 of which were common to the 560 planted elsewhere. That site was destined to form part of the independent testing of the Genomic Estimated Breeding Values.

A consequence of implementing Genomic Selection in the breeding programme will be shortened breeding cycles accompanied by a lesser quantity of planting material for commercial release. Micro-propagation could greatly accelerate multiplication of genotypes selected by Genomic Selection. The diversity of willow germplasm that can be reliably reproduced by micro-propagation has been greatly expanded. The project successfully propagated 38 genotypes for the first time.

The quantitative key performance metric attainable within the project was to increase the achievable yield from short rotation coppice willow from the current 10 t ha^{-1} dry matter per year to average 15 t ha^{-1} and be regularly recording up to 20 t ha^{-1}. Within the project, 15 t ha^{-1} was achieved by optimising the biomass variety to the growing environment. This improves the financial return to the grower and minimises the area of land required to contribute to net zero targets. Further improvements to yield by breeding, which will be significant, were beyond the timescale of the project.

It was expected that the emerging Bioenergy with Carbon Capture and Storage technology would drive new plantings. This has not happened. Market conditions currently preclude sales of planting material because alternative land uses and incentives are financially and logistically considerably more attractive. Rothamsted Research is conducting research that has potential to increase the value of the biomass produced from willow. However, that is at a low Technology Readiness Level and will take several years to come to fruition. The research conducted in this project will be of great value to that programme in future.

Table of Contents

List of Figures

List of Tables

Glossary

Term	Definition
Cultivar	Cultivated variety, normally named and protected by Plant Breeder's Rights.
Elites	A subsection of 33 of the Training Population genotypes that are contemporary varieties or near-market varieties and represent the most advanced breeding material
Genome editing	A process by which the genome of an organism is edited at a target location.
Genetic modification	A process by which DNA is inserted into the genome of a recipient organism
Genomic Selection	Predicts the genomic estimated breeding values in a population by associating their traits (e.g. resistance to pests) with their high-density genetic marker scores.
Genomic Estimated Breeding Values (GEBV)	A statistically generated number or score that estimates the total genetic potential of an individual with respect to a heritable trait
Genomic Best linear unbiased prediction (GBLUP)	A statistical method used to predict breeding values using single nucleotide polymorphisms for selection in animal and plant breeding.
Genotype	The genetic constitution of an individual organism. Can also be used to describe the process of determining the genetic constitution of an individual organism.
Green area duration	The time period between bud burst and senescence during which the tree has a leaf canopy intercepting light and photosynthesising.
Phenotyping	The assessment of observable characteristics (as influenced by genetic make-up and changes in the environment); a vital process in crop improvement programs

Abbreviations / Acronyms

AFBI	Agri-Food and Biosciences Institute
APHA	The Animal and Plant Health Agency
AWBD	Accelerating Willow Breeding and Deployment
BAP	6- Benzyl aminopurine
BBSRC	Biotechnology and Biological Sciences Research Council
BECCS	Bioenergy with Carbon Capture and Storage
BEIS	Business Energy and Industrial Strategy (UK Government Department)
BFI	Biomass Feedstocks Innovation Programme
BPS	Basic Payment Scheme
CCC	Climate Change Committee
DEFRA	Department for Environment, Food and Rural Affairs
DESNZ	Department of Energy Security and Net Zero
DAERA-PHI	Department of Agriculture, Environment and Rural Affairs – Plant Health Inspection
DM	Dry Matter
FarmPEP	Farm Performance Enhancement Platform
GEBV	Genomic Estimated Breeding Value
GS	Genomic Selection
GWAS	Genome-Wide Association Studies
IBA	Indole-3-butyric acid
LAT	Lawes Agricultural Trust
LSD	Least Significant Difference
LMM	Linear Mixed Model
MAS	Marker Assisted Selection
MS	Murashige and Skoog
NAA	1-Naphthalenacetic acid
NZIP	Net Zero Innovation Programme
ONT	Oxford Nanopore Technology

PGR	Plant Growth Regulator
QTL	Quantitative Trait Loci
REML	Restricted Maximum Likelihood
RRes	Rothamsted Research
SRCw	Short Rotation Coppice willow
SE	Standard Error (of the mean)
SFI	Sustainable Farming Initiative
SME	Small and Medium Sized Enterprise
SNP	Single Nucleotide Polymorphism
SRUC	Scotland’s Rural College
TIS	Temporary Immersion System
TP	Training Population
TRL	Technology Readiness Level
WPM	Woody Plant Media

1. Background

Rothamsted Research (RRes) is a world leading, nonprofit research Institute that focuses on strategic agricultural science to the benefit of farmers and society worldwide. The main site is in Harpenden, Hertfordshire. There are approximately 350 employees and an annual turnover of £35.7 M. Funding and other support for research is provided by the Biotechnology and Biological Sciences Research Council (BBSRC) of the UK, the Lawes Agricultural Trust (LAT), a wide range of other national and international funding bodies and industry.

RRes is also the longest running research organisation in the UK to have been engaged in research and development of willows for industrial and environmental purposes and has maintained an international reputation of excellence in willow throughout this time. RRes maintains a *Salix* germplasm collection of >1,500 accessions, this is the principal resource behind the Institute's work on genetic resources for bioenergy. The willow breeding programme at RRes began in 2003 building upon breeding and genetics work carried out at the sister institute at Long Ashton (and including staff transferred from that site). Willow breeding through the Bioenergy Genetic Improvement Network project was funded by DEFRA [NF0424]. To date five varieties have been registered for Plant Breeders Rights and licenced to nurseries for multiplication and sale. The breeding programme is closely allied to the genomics facility. Marker Assisted Selection (MAS) has already been applied to the programme to great effect.

To provide field sites in diverse environments, not previously used for detailed assessments of large numbers of Short Rotation Coppice willow (SRCw), RRes worked with four sub-contractors. Sub-contractors were The Agri-Food and Biosciences Institute (AFBI), Newcastle University, Scotland's Rural College (SRUC) and Somerset Willow Growers (Table 1).

2. Project Overview

The project built upon the advances made from past investments in the UK in willow genetic improvement and breeding to secure a sustainable supply of biomass for meeting government Net- Zero targets. Through the Department of Energy Security and Net Zero (DESNZ) Net Zero Innovation Portfolio (NZIP), Biomass Feedstocks Innovation (BFI) Programme Lot 1, Phase 1, planning project, RRes developed the plan and foundation materials needed to build on past advances to exploit the new advanced DNA technology of Genomic Selection (GS) in willow. In Phase 2, based on the successes of deploying GS in other crops, the aim was to significantly increase the speed and efficiency of breeding improved willows for the UK market by shortening the breeding cycle from 12 to 6 years, increasing biomass yield and supplying genetic diversity that ensures resilience against biotic and abiotic threats to sustaining yields in long-term perennial plantations. In so doing, breeding programme costs will be reduced, and the industry better supplied with planting material of greater potential sooner. This will occur in parallel with innovations aimed at rapid scaling of new breeding and generating even greater knowledge with which to optimise the use of past gains.

The project was divided into six Work Packages: 1 Project management, 2 Planting the TPs, 3 Phenotyping the TPs, 4 Genomics and molecular breeding, 5 Enabling Biotechnology and 6 Interactions and dissemination. The cost of the work was £2,298,800 and it was completed on budget.

2.1 Project Management

RRes were the project managers. The Heads of Agronomy and Genomics worked with the willow breeder and the bio-technology lead to direct the work. To address the objectives of this project, RRes worked with four sub-contractors. Sub-contractors provided field sites in environments not previously used for detailed assessments of large numbers of SRCw genotypes. The Agri-Food and Biosciences Institute (AFBI), Newcastle University and Scotland's Rural College (SRUC) are institutions with a strong reputation for agricultural field experimentation and therefore able to carry out all protocols independently. Somerset Willow Growers is a private company experienced in willow growing in seasonally flooded land, RRes staff conducted all phenotyping protocols at the site.

The project has resulted in three new full-time posts at Rothamsted. It has contributed to the recruitment of a new staff member at SRUC and AFBI. At Newcastle University it has contributed to securing the positions of two pre-existing staff. The project has contributed to safeguarding 6 existing positions at Rothamsted and one at AFBI. The sub-contract with Somerset Willow Growers contributes to the success of an SME and in a small way towards protecting their employee's jobs.

Only two issues impacted upon project management. Willow is considered a phytosanitary high-risk plant for import into the European Union. No specific threat is cited, the classification is based on the lack of a risk assessment having been approved by the EU. After much discussion with Animal and Plant Health Agency in England and DAERA-Plant Health Inspection Branch (Forest Service) in Northern Ireland, it became clear that the transfer of willows from England (GB) to Hillsborough (NI) would not be possible. The AFBI team collected all available willow germplasm (mostly, current and recent varieties or genotypes from historical yield trials), in Ireland and planted as a reduced population of 144 genotypes. There were 75 genotypes common to the TP planted at the four sites in Great Britain. The experimental design was as for the four larger sites.

The intention was to repurpose the Hillsborough population as an independent test set. Models based on data from four sites would be used to predict outcomes on the fifth site. This would be additional to testing within the four sites conducted by dropping parts of the population and predicting outcomes from data for the remainder. However, in June 2024 it became apparent that the site had been differentially damaged by herbicide during a routine management operation. The data collection on site ended with early season rust assessment that year. Data collected prior to the incident was used. Managers at the other sites were alerted to the danger of using a powerful non-selective herbicide close to the willows and the possibility of rapidly changing weather conditions increasing the potential for damage to the crop.

Molecular genomics is a fast-moving field of science. The Genomics Team were able to incorporate some new techniques unavailable at the time of writing the project proposal. These were deployed within the overall budget and increased the amount of data captured.

3. Lot 1 Technical requirements

3.1 Introduction

The Committee on Climate Change 6th Carbon Budget estimated that expanding the growing of energy crops by ~23,000 hectares each year up to a predicted maximum area of 708,000 ha is needed to make a substantial contribution to the UK Net Zero Targets. The Business Energy and Industrial Strategy 2021 Biomass Policy Statement (2021 BEIS) emphasised the need to expand the land area under energy crops. Subsequently, the Biomass Strategy (2023) was less positive around domestic production of energy crops. The Clean Power 2030 Action Plan (2024) does not provide clarity on the future of role biomass, with various consultations being considered. Two dedicated perennial energy crop types are front runners in the UK agricultural landscape, miscanthus and SRCw, and are likely to form the greater part of any planting ambition.

Estimates of the potential contribution of SRCw biomass to the UK Net Zero Targets are based on current yields averaged from SRCw grown on less than 5,000 ha of mainly ex-arable land in England. Crop modelling shows that there is a considerable gap between potential and actual yield of SRCw in the UK, indicative of potential to increase biomass produced per unit of land. This would reduce total land use requirements and improve profitability of SRCw for growers. Such yield improvements would have to be reliably achievable under varied environmental conditions if the requirement for expansion onto suitable land areas, spatially planned in ways that harmonise with other land uses, is to be met. To meet financial and environmental targets for the crop, SRCw is not treated with many pesticides or given much fertiliser. During the latter growth stages of the harvest cycle, such applications are physically difficult due to the crop size and structure. Therefore, genetics is the main focus for improvements in crop performance, targeting of existing germplasm and, most importantly, breeding towards the right genotype in the right environment.

A limited knowledge of genotype response to the environment can be gained by a meta-analysis of previously collected data largely derived from trials rather than crop models. This was conducted in Phase 1 of the BFI programme. However, previous trials had been conducted in a limited range of environmental conditions, almost all arable fields. Quick gains in yield looked possible from growing in different environments. Beyond that, efficient, environment targeted breeding that results in quick genetic gains is urgently needed to provide a diversity of elite varieties tailored to suit the environments SRCw will need to be grown in. Application of GS to achieve breeding targets in a wide range of crops, trees and animals demonstrated GS to be the most highly effective technology to achieve this. Applying GS to the SRCw breeding programme should reduce the production time of a new variety by up to 6 years.

The multiple rounds of testing and selection in conventional breeding result in sufficient plants to base a multiplication stock upon. A consequence of rapid selection for market via GS would be a much-reduced stock of planting material for commercial deployment. Micro-propagation of plantlets from buds on a mother plant offered a route to rapid multiplication of a genotype. Previous work by Palomo-Rios *et al*. (2015) had been successful with a limited number of genotypes but shown some variability in repeatability. To avoid imposing a restriction on parent selection in the breeding programme, protocols for micro-propagation of a wider range of genotypes were required.

The programme of work addressed three objectives:

Objective 1. To generate the knowledge required to inform precision deployment of optimal varieties for different growing environments and maximize feedstock production.

Objective 2. To establish a Genomic Selection strategy that will accelerate the production, performance and security of UK SRC willow varieties for the bioenergy market.

Objective 3. To develop and optimise micropropagation technologies for rapid multiplication of optimal genotypes identified in GS-led breeding.

3.2 Method

3.2.1 Field-based activities

Field Sites

To obtain Genomic Estimated Breeding Values (GEBVs) a Training Population (TP) was selected from the extensive germplasm collection at RRes. This was informed by previous genomics work and, specifically, the intensive genotyping effort carried out in Phase 1 of the BFI Programme. Some constraints on selection were applied, for example, where there are known crossing barriers. The breeding programme is conducted by traditional crossing methodology; genetic modification and genome editing are not available in willow currently, so crossing compatibility is essential. A dendrogram was constructed showing genetic relatedness across the germplasm collection and 560 unique genotypes selected to form the TP, including 33 “Elites” (current varieties and near-market breeding material). Five environmentally contrasting sites were chosen to grow the TP that would both address GEBV discovery and extend knowledge of willow adaptation and its genetic basis (Table 1.).

Table 1. The sites where the Training Population was planted

Site	Co-ordinates	Host	Environment
Tertowie, **Aberdeen**	57.1°N, 2.3°W	Scotland's Rural College (SRUC)	Cooler temperatures, long daylight hours in growing season.
Newcastle University, Cockle Park Farm, **Northumberland**	55.2°N, 1.7°W	Newcastle University	A "control" site, similar to previous trial sites.
***Hillsborough**, Northern Ireland	54.5°N, 6.1°W	Agri-Food and Biosciences Institute	High disease pressure.
Woburn Experimental Farm, Bedford	52.0°N, 0.6°W	Rothamsted Research	Drought due to lower rainfall, higher temperatures and sandy soil.
Bussex Farm, Westonzoyland, **Somerset**	51.1°N, 2.9°W	Somerset Willow Growers	Flood inundation. Disease pressure.

**Hillsborough was planted with a different population due to phytosanitary regulations on planting material imports from Great Britain.*

Figure 1. Planting at Woburn, May 2023

Figure 2. Four weeks after planting at Somerset, June 2023

3.3 Phenotyping

The TP was assessed using standard protocols for traits considered most important for breeding high yielding, pest and disease resistant varieties suited to different target environments. Protocols used to generate this phenotype data can be viewed in Appendix A. Crop yield, level of infection with rust (*Melampsora* spp.), senescence timing and tip damage caused by terminalis midges of *Dasyneura* spp. were the primary traits of interest. Assessments involving other pests are more difficult to utilise in in genetic studies as the organisms are rarely well distributed across a site. The data collected are valuable indicators of pest resistance and provide useful baseline information but are not sufficiently robust for genomic analysis. The pests assessed were:

- The willow aphids *Tuberolachnus salignus* and *Pterocomma salicis*
- Chrysomelid beetles and their larvae, several species in the genera *Phratora* and *Galerucella* or *Crepidodera*
- Sawfly larvae *Nematus oligospilus*

All three were observed. Beetle and sawfly damage may also be recognised by the condition of the leaves; beetles leave the skeleton structure of the leaf (often referred to as ghosting) and sawflies leave the midrib only.

The mammals: deer, hares and rabbits can also damage willows, showing a feeding taste preference. All sites were fenced against mammals as the damage to their preferred genotypes can be so severe as to exclude collecting data on any other trait.

Figure 3. Aberdeen, September 2024

3.4 Statistical Design

At each site the planting was replicated in 4 statistical blocks, taking account of the most substantial source of within-site spatial variability, with each block containing 630 plots in a 30-by-21 arrangement at the Aberdeen, Newcastle, Somerset and Woburn sites and 162 plots in an 18-by-9 arrangement at Hillsborough. Within each block, the plots were grouped into sub-blocks, each containing 9 plots in a 3-by-3 arrangement, and each plot containing 6 individual trees planted in a 3-by-2 pattern. The TP allocated to the different plots were arranged so that each sub-block contained 8 different lines from the TP and one plot with a common control genotype (placed in the centre of the 3-by-3 array), and so that any pair of lines only appeared together in a sub-block a maximum of once per site. To further minimise the impact of spatial variability across what became a large field experiment, an additional blocking condition was imposed – within each column of 28 sub-blocks (12 sub-blocks at AFBI) any non-control genotype could only appear once. The implementation of these blocking constraints and inclusion of a regular pattern of controls across each site (them being present at the centre of every sub-block in the field) allowed spatial biases to be corrected during the analysis of phenotyping data. Guard rows were planted around the outside of the experiment to negate edge effects (*Figure 4*).

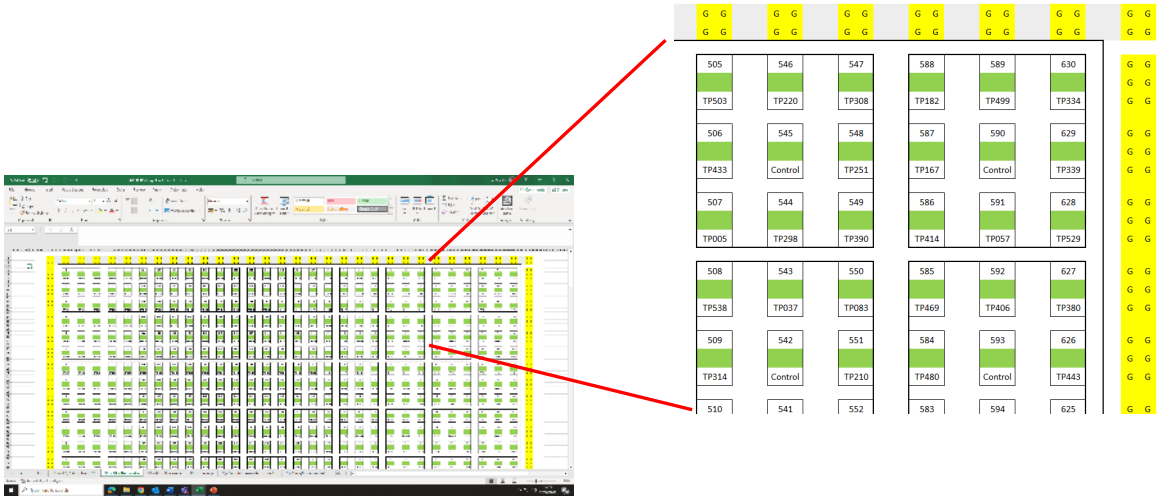


Figure 4. Plot arrangement and genotype allocation within sub-blocks

Figure 5. Somerset, September 2024, showing the arrangement of plots

3.5 Statistical Analysis

All phenotyping data were analysed by fitting a Linear Mixed Model (LMM) using the Restricted Maximum Likelihood (REML) algorithm, incorporating the blocking effects (random model) identified above, and using the TP lines as a treatment factor (fixed model). This facilitated estimates of a mean response for each genotype, free of spatial bias. Transformations were applied as appropriate to each response variable prior to analysis, to satisfy the assumptions of homogeneity of variance and normality needed for the LMM. Because of substantial differences in phenotypic responses between sites, comparisons of phenotypic performance between sites were primarily assessed through the relative rankings of the predicted means for

each TP lines, as well as by calculating both product-moment correlation coefficients and Spearman's rank correlation coefficients to quantify the consistency of the results. TP lines showing either consistency or substantial differences between sites were identified based on the variability of their ranks, with overall performance identified based on mean rank. Further comparisons were made with the AFBI phenotypic responses for those lines present in both this site and the four main sites. Analysis of the residuals from the linear models resulting from the analysis of each phenotyping variable provides information about the consistency of responses within a site, adding to the information about genotype-by-environment interactions provided by the comparisons of relative rankings between sites.

Analysis of pairwise correlation coefficients and principal component analysis of the combined sets of phenotypic response variables for each site identifies the associations between the different response variables, with biplots allowing the identification of TP lines that have particularly strong responses to particular phenotypic traits or combinations of traits. Comparisons of the biplots for the different sites provides further information about the potential for genotype-by-environment interactions, including the identification of different TP lines suited to the different environments. Clustering of TP lines across all response variables can identify groups of TP lines with similar sets of phenotypic responses.

3.6 Genomics-based activities

3.6.1 Building required genome resources for target species

Background: A critical step in implementing the AWBD GS strategy was the requirement to genotype all individuals of the TP. This involves identifying points of variation in the genomes of the different individuals which can be used in conjunction with phenotype data to develop models for predicting phenotypes and subsequent selection of potential 'winners' in newly-bred (non-phenotyped) progenies as part of any future breeding programme. Different genotyping technologies exist and differ in the cost of use, and the amount and scope of data produced. The diversity within the material to be genotyped also needs to be considered – any genotyping assay developed must work well across the breadth of diversity present in the sample, i.e. it must produce data even for individuals at the extremes of the diversity range. Largely, genotyping approaches fall into three categories: 1) *targeted genotyping approaches* focus on generating data for particular, pre-selected genomic regions, 2) *non-targeted genotyping approaches* assay regions of the genome that are not preselected and semi-random (e.g. restriction enzyme-based Genotyping by Sequencing (GBS)), 3) *whole genome sequencing (WGS)* involves assaying the entire genome of all individuals in a study.

At the onset of the project proposal, we envisaged using a targeted genotyping approach due to the prohibitive cost of whole-genome sequencing at this time. To ensure any targeted genotyping assay would work across the diversity represented within the TP, some initial and novel work to understand the variation within the species and hybrids to be genotyped was initiated. As the project evolved and NGS technology evolved and costs decreased, it became evident that the preferred WGS approach would be possible for TP genotyping, shifting focus from ensuring a

targeted assay would work for mass TP genotyping to using the genome resources to improve analysis methods.

A selection of genotypes representing the species diversity relevant to the TP and wider willow breeding program were selected to form a reference set for detailed genome analysis. Suitable data were not available from public sources so were generated within project. Both short-read Illumina data (highly accurate raw data) and long read Oxford Nanopore Technology (ONT) sequencing data (lower raw read accuracy but better for assembly and understanding structural variation) were generated for all of these samples. As a particularly important species for conferring high biomass yield, multiple *S. viminalis* individuals representing different sub populations were included. More short-read data was generated for samples with expected polyploid genomes.

Table 2. A reference set of genotypes and species selected for more in-depth genome analysis. Pre-analysis ploidy estimates and quantities of sequence data per technology are provided. N50 values are indicative of data quality in terms of read length

Genotype	Species	Assumed Ploidy	Illumina (Gbp)	ONT (Gb : N50 kb))
NWC0099	*S. triandra*	2x	17.17	19.53 : 21.56
NWC0446	*S. aegyptiaca*	4x	24.43	22.34 : 32.72
NWC0473	*S. cinerea*	2x/4x	61.01	62.99 : 16.80
NWC0485	*S. scouleriana*	2x	25.24	43.00 : 22.36
NWC0488	*S. sitchensis*	2x	39.74	84.15 : 27.56
NWC0577	*S. dasyclados*	6x	83.25	58.38 : 13.67
NWC0607	*S. rehdriana*	4x	61.87	97.56 : 16.50
NWC0610	*S. udensis*	2x	21.86	72.75 : 16.21
NWC0615	*S. schwerinii*	2x	21.86	43.98 : 22.57
NWC0665	*S. viminalis*	2x	17.32	36.17 : 40.29
NWC0673	*S. viminalis*	2x	17.55	43.96 : 9.46
NWC0681	*S. viminalis*	2x	18.56	64.90 : 19.68
NWC0696	*S. viminalis*	2x	23.68	65.87 : 24.61
NWC0703	*S. viminalis*	2x	16.66	54.78 : 19.13
NWC0941	*S. miyabeana*	4x	86.50	63.27 : 23.50
NWC0954	*S. rigida*	6x	109.93	70.00 : 9.76
NWC1126	*S. caprea*	2x	29.87	18.17 : 16.10
NWC1153	*S. viminalis*	2x	16.59	90.30 : 26.20

Development of long-read ONT methodologies for willow

At the onset of the project, nanopore sequencing was a relatively new technology and yet to be routinely applied to willow. During the project, significant effort was put into developing protocols and sequencing workflows to ensure stable and optimal results. Several protocols were tested for obtaining high molecular weight (HMW) DNA of the required quality. Once achieved, different methods for size selection (removing smaller DNA molecules from the extract) were tested and library

preparation (ONT Ligation Sequencing Kits LSK109, LSK110 or LSK114) methods optimised. Finally, once stable performance on the nanopore flow cells was achieved, all samples were run on a PromethION 24 sequencer. The amount of sequence generated per sample is provided in Table 2 alongside N50 statistics to give an indication of the read length characteristics.

Short read Illumina data generation and subsequent analysis
Short read Illumina data (2 x 150bp paired end reads) was generated primarily as a basis for estimating ploidy and genome size in the reference set. Also, at the time of generation, these data sets were commonly used to 'polish' the more error-prone, longer nanopore reads in hybrid (both long & short read) *de* novo assembly approaches. Sequencing libraries (Illumina DNA PCR-free) were constructed using DNA available from the same HMW extracts used for ONT sequencing. Amounts of DNA sequence produced per sample are provided in Table 2.

Kmer analysis to estimate genome characteristics
To estimate the genome size, Illumina sequence reads were split into short sub-sequences called 'kmers' which will be represented in the genome multiple times. The software program 'jellyfish' was used to count and generate a histogram of 'kmer' frequencies. The R program 'GenomeScope' was then used to fit a model to the histogram which produces an estimate of genome size and genome heterozygosity level and an indication of ploidy level (an important consideration in subsequent variant calling analysis).

3.6.2 High-density TP genotyping required for GS models

Background: In this part of the project, the datasets required to identify points of difference in the genomes of the different TP individuals were generated. Alongside the phenotype data, these variants are used as the basis for development of genomic prediction models.

Whole genome sequencing using Illumina 2 x 150bp short-read sequencing was selected as the preferred method for TP genotyping, largely due to the improved variant calling accuracy expected to result from the high raw read accuracy associated with this technology. Furthermore, associated sequencing costs were possible within budget. WGS also avoids issues associated with ensuring any genotyping assay can capture all of the diversity present in the samples, and (unlike random methods such as GBS) provides reliably comparable datasets across different genomes. Leaf samples were collected from the Woburn field trial in Spring 2024 and DNA extracted using Qiagen Plant DNeasy spin columns. DNA quality and quantity was assessed using a Nanodrop spectrophotometer. Quantities were also assessed by the more accurate Qubit Broad Range DNA assay. Any sample extractions failing QC at this stage were repeated until a useful extraction was achieved. For the majority of samples, sequencing libraries were prepared using Illumina DNA PCR-free library kits. For samples with low DNA yields, the Illumina DNA Prep library kit was used. Libraries were sequenced on an Illumina Nextseq2000 system to a target depth of ~18X estimated genome coverage. DNA sequencing output quality was assessed using the fastQC software package.

Sequence reads were aligned to a *Salix viminalis* reference genome using the bwa-mem command of the sequence alignment software, Burrows Wheeler Aligner (BWA). Variants were called at gene locations of the reference genome in all sequenced lines of the TP using the variant sites processing software, BCFtools. Variants from all the lines were merged into a single variant file and biallelic single nucleotide polymorphic sites were extracted and filtered for missingness and allele frequencies. The resulting filtered genotype dataset was then used, in combination with the phenotyping data, to conduct genomic prediction and genome wide association mapping (GWAS).

3.6.3 GS data analysis

Background: Here data from TP phenotyping and high-density genotyping activities will be combined to generate prediction models that will be used demonstrate proof-of-concept and test different deployment approaches of GS technology in a complex system such as that which will be required in willow. Beyond proof-of-concept, this work will test and inform approaches for use of GS technology when considering the different environments in which the willow crop may be grown.

To conduct genomic prediction accuracy tests, the genomic best linear unbiased prediction model was applied to our genotype and phenotype data in the R package GAPIT. To achieve this, we first we select a random set of the willow lines, corresponding to 10% of the population, for which the phenotype data are set as missing - this is regarded as a 'test set' and the remaining 90% with their phenotype values known are regarded as a 'training set'. Next, we apply our model to the genotype dataset of the full willow population and the phenotype data of the training set to create a scenario whereby we can predict the trait value of the test set. The predicted values are the genomic estimated breeding values (GEBV) which are then correlated with the known trait value. The values of the correlations, which ranges from 0 (no correlation) to 1 (perfectly correlated), are regarded as the measure of the prediction accuracy. The prediction accuracy is a proxy for how well we can predict the traits of new germplasm within a breeding program. We repeated the process 100 times to generate values of prediction accuracies for different combinations of training and test set in the willow collection.

3.6.4 Convert disease resistance info to markers for MAS

Background: This area of work aimed to provide additional information about the underlying genetic basis of rust resistance in different genotypes. For deployment of durable resistance, it is important to understand the basis of resistance in released varieties. Here, AWBD aims to build a catalogue of known resistance loci and associated markers to inform future selection or deployment decisions. This knowledge will help future de-risking of the willow crop in terms of rust disease.

Development of markers for known QTL on willow chromosome XI

Additional microsatellite markers were developed in the genomic region around a known QTL for quantitative resistance to rust. Sequence data (from 4.1) was used to identify microsatellite motifs in the QTL region across different genotypes/species. Primers were designed in areas flanking polymorphic microsatellites for use in

screening. Following testing in a range of samples, four markers were deemed suitable for use in screening. This panel of markers was used to screen a subset of the TP to generate a data set that could be used in future tests for marker-trait associations. As the project developed and it became evident that whole genome sequencing of the entire TP would be possible within budget, work in this area was deemphasised as GWAS studies would now be possible using a greater density of marker information in the QTL region.

Identification of new markers for rust resistance in the AWBD TP.

Rust scores from in-project assessments were used to scan for marker-trait associations in Genome-Wide Association Studies (GWAS). To achieve this, a mixed linear model was applied to the genotype data in the R package GAPIT. To avoid identifying false positive associations, the mixed linear model accounts for the population structure and relatedness within the willow population using a combination of a principal components analysis and an additive genetic relationship matrix. From GWAS, we generate the probability values – probability that a SNP site is not associated with the trait of interest - for all the variant SNP sites. These values were used to create a Manhattan plot (figures shown), which is a plot of the logarithm of the probability values (y- axis) of all SNP sites by chromosomes (x-axis).

3.6.5 Refine markers for MAS based on known major effect loci

Phenotypes from in-project assessments of yield, maximum diameter (correlated with yield) and senescence were used to scan for marker-trait associations using Genome-Wide Association Studies (GWAS), with a view to identifying markers for use in future selections and to inform on the genetic architecture of key biomass-related traits. Equivalent approaches to those described above for GWAS analysis of rust scores were used.

3.6.6 Develop methodologies for identifying and confirming genotypes in the Hillsborough trial

Background: This work was performed to confirm that genotypes planted at the Hillsborough site in Northern Ireland were as expected. This was required as cuttings planted there were not provided by RRes due to import issues resulting from 'Brexit'.

Staff from RRes visited the Hillsborough trial in Summer 2024 and collected leaf samples for DNA extraction. DNA was extracted using the Qiagen DNeasy Plant 96 method. To confirm sample IDs, eight microsatellite markers were chosen for DNA profiling. Choice of markers was based on several round of preliminary testing to identify markers that were highly informative (detected large numbers of different alleles), discriminatory, and worked consistently across the diversity present in the sample. Multiplex screening protocols were optimised before running all samples on a SeqStudio Flex capillary sequencer. DNA samples from reference material at RRes was run alongside the Hillsborough test samples. Resulting profiles were analysed using Genemapper software and microsatellite profiles of test and reference samples compared.

3.7 Micropropagation methods

Background: Willow propagation (*Salix* spp.) is primarily achieved through vegetative methods, particularly the use of cuttings. These are usually taken when plants are winter dormant (January and February) and stored frozen before planting in spring. *In vitro* culture methods, offer an alternative, and potentially more rapid, approach. The conventional Rothamsted breeding scheme relies on vegetative propagation to multiply material from single seedlings to sufficient numbers of woody cuttings for field-based yield trials. This is a bottleneck in new variety production and takes several years. In a GS-led approach, *in vitro* micropropagation of promising genotypes from the seedling stage (following selection using GS) could circumvent this issue and greatly accelerate variety production.

The objective of micropropagation is to produce large numbers of plants that can survive under natural environmental conditions. It is a multi-stage process involving five different steps to produce transplantable propagules (Appendix C, Section 1): Stage 0: Establishment of donor plant/s; Stage I: Aseptic establishment and initiation of cultures; Stage II: Shoot multiplication; Stage III: Rooting and conditioning of shoots; Stage IV: Hardening, acclimatisation and transplantation in soil. These stages are universally applicable in the large-scale multiplication of plants.

Detailed methodologies of micropropagation activities carried out can be found in *Appendix C*, but are summarised here.

3.7.1 Establishment of material to in vitro conditions.

To establish material from diverse willow genotypes in aseptic *in vitro* conditions, clean and healthy stock material was required. Cuttings from clones growing in managed fields were collected and established successfully in the glasshouse. Material from plants showing sufficient and healthy growth was then introduced to *in vitro* conditions in the laboratory following methodology previously established at RRes (Palomo-Rios et al., 2015). Following sterilisation, stem segments were cultured upright in micropropagation stock media.

After six weeks of growth under controlled conditions, measurements were taken from explants to determine the establishment response of each genotype, and the multiplication rate was calculated. New shoots were then transferred to fresh standard media every six weeks for the duration of the project, increasing in number with each multiplication round. Once sufficient numbers were reached, this material was used in various optimisation experiments to determine optimal micropropagation conditions for each genotype.

3.7.2 Optimisation of micropropagation conditions.

For optimisation of micropropagation, several key parameters were considered, such as the optimisation of culture media and the environmental conditions. For tissue culture optimisation, two main parameters were studied: the effect of basal formulation and the effect of plant growth regulators (PGR).

Two different basal formulations, MS (Murashige and Skoog, 1962) and WPM (Lloyd and McCown, 1980), were assessed for their effect on the multiplication rate. Both formulations are widely used to study willow micropropagation in different species

and hybrids. Material from six genotypes was introduced to both media and micropropagation parameters were recorded after six weeks of growth. To determine the effect of plant growth hormones on micropropagation rate, nine treatment conditions were designed, comprising MS media containing either cytokinin alone or in combination with auxin, at various concentrations. Individual stem nodes including two auxiliary nodes were cultured in conical tissue culture tubes. Shooting and rooting response, new shoot length and number of nodes were recorded

To optimise the environmental conditions, the effect of vessel type was studied. Two different vessel types, ECO2BOX and OS140BOX, were assessed for their effect on multiplication rate. Material from seven genotypes was introduced to both vessel types and the micropropagation rate was measured in standard conditions after six weeks.

3.7.3 Demonstrating rapid scale-up of elite material.

The RITA® Temporary Immersion System (TIS) was used to demonstrate the rapid up-scaling of material, using liquid medium. The material was subcultured after 5 weeks in the TIS. Various factors were tested in preliminary experiments, including concentration of cytokinin in the media and number of explants per vessel, to determine the optimal conditions for multiplication. Rapid up-scaling of material was demonstrated with seven genotypes, with measurements of shooting response and shoot length taken after five weeks. Following this, explants required rooting in agar-based stock media for six weeks prior to cutting production.

To demonstrate the production of cuttings for mass deployment to yield trials, healthy *in vitro* explants from ten genotypes were acclimatised in the glasshouse and transferred to the nursery. Twenty-four plants per genotype were allowed to grow as a source of cuttings for future assessments of performance in a Biomass Connect field trial.

3.8 Results

Weather data collected close to each site shows that the growing seasons 2023 and 2024 followed the expected pattern of conditions for which the sites had been selected (Appendix B, Table B1). Aberdeen was cooler than other sites, with reasonable rainfall and sunshine hours. The maximum daylength was almost 18 hours at the summer solstice compared to 16.5 hours at the southernmost site Somerset. Woburn was warm, dry and sunny, leading to the predicted period of drought stress (Appendix B Figure B1). this was greatly exacerbated by the very low water holding capacity of the sandy soil. Catt *et*. *al*. reported 71% sand (63 – 2000 µm) in the topsoil rising to 91% at 800 mm depth with 100 mm of total available water between 0 - 800 mm rooting depth. Evidence from other crops suggested that drought stress would increase between approximately 50 - 100 mm soil moisture deficit, becoming severe above 100 mm (Penman 1970, French & Legg 1979). Willows may root more deeply than annual crops, but it is known that the soil continues to have a very high sand content to 1,800 mm depth.

In Somerset, surface water was first noted in early November 2023 following Storm Ciaran. On 4th November, it was 450 mm deep, by 21st November the water had receded to a maximum of 300 mm depth. First year cutback was possible in February with only small areas of surface water, no more than 100 mm deep. In January 2025, the water was again 400 mm deep (*Figure 6*).

Figure 6. The Somerset site, 30th January 2025

3.8.1 Phenotyping.

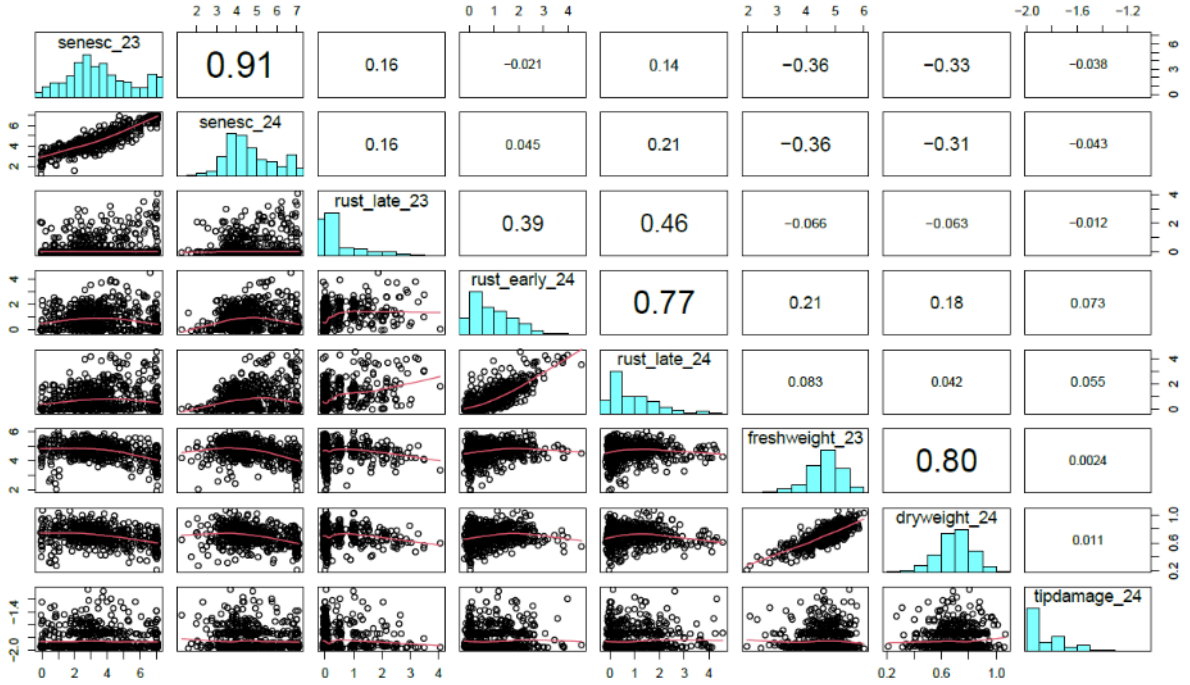
Data from assessments made in different years strongly correlated as did assessments made at the same time at different sites (*Figure 7*. and Appendix B, Table B2 and B3 and Principal Component Analysis, output not shown). This indicates that operators were consistent in their interpretation of the protocols.

3.8.2 Rust

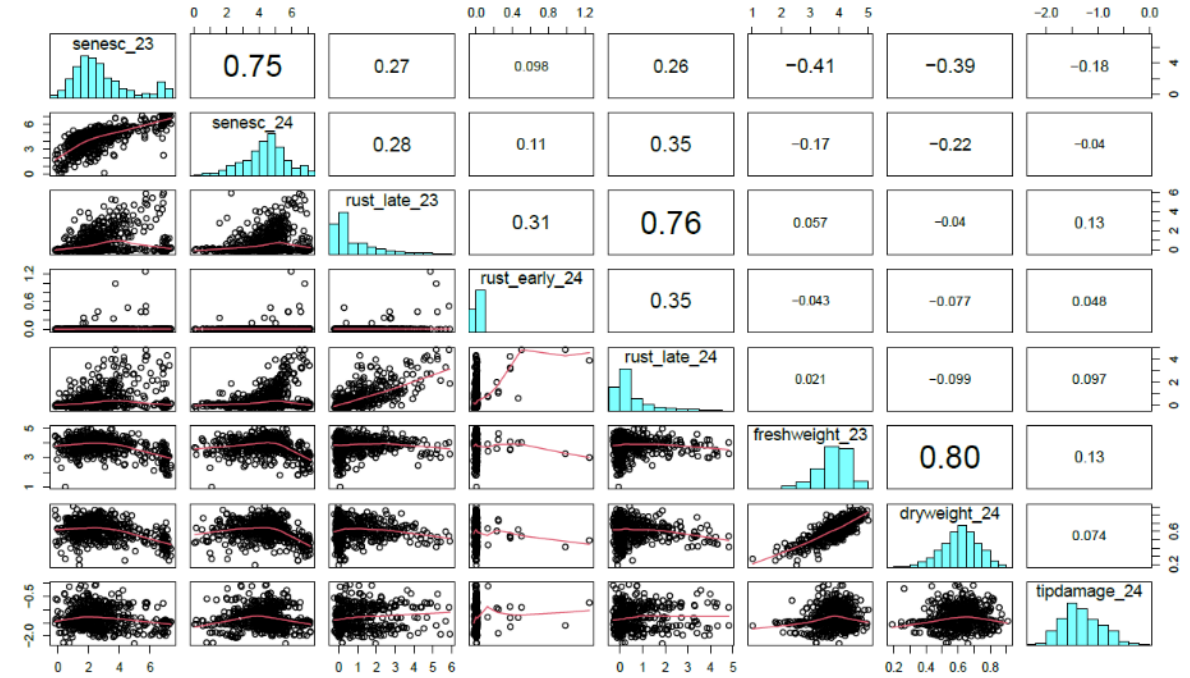
The scores for the control genotype (susceptible to rust infection) are shown in the 3-D plots in Appendix B, Figure B2, where rust was observed across all sites. Correlation coefficients between rust scores for each genotype were strongest between assessment occasions within sites but also significant across sites (*Figure 7*). The exceptions were occasions when rust was not observed to any great extent (late summer 2023 Aberdeen and early summer 2024 Newcastle).

The rust scores did not correlate with yield at any assessment date on any of the sites. However, rust is a very important pathogen to consider in willow breeding and variety selection, as there have been numerous instances of catastrophic collapse in resistance in some varieties. Breeding with new sources of resistance and stacking multiple sources of resistance in one variety are essential to prevent this breakdown. In addition, it is recommended that commercial crops are planted as mixtures of at least 6 varieties selected from different resistance sources. This has been shown to slow the spread of a rust pathotype that any one variety becomes susceptible to. The “Grower’s guide to SRCw varieties for biomass” leaflet, produced as an output of AWBD Phase 1 of the BFI programme, groups the current varieties by resistance source. Phase 2 offers the opportunity to add to the knowledge of the genetic basis of those sources.

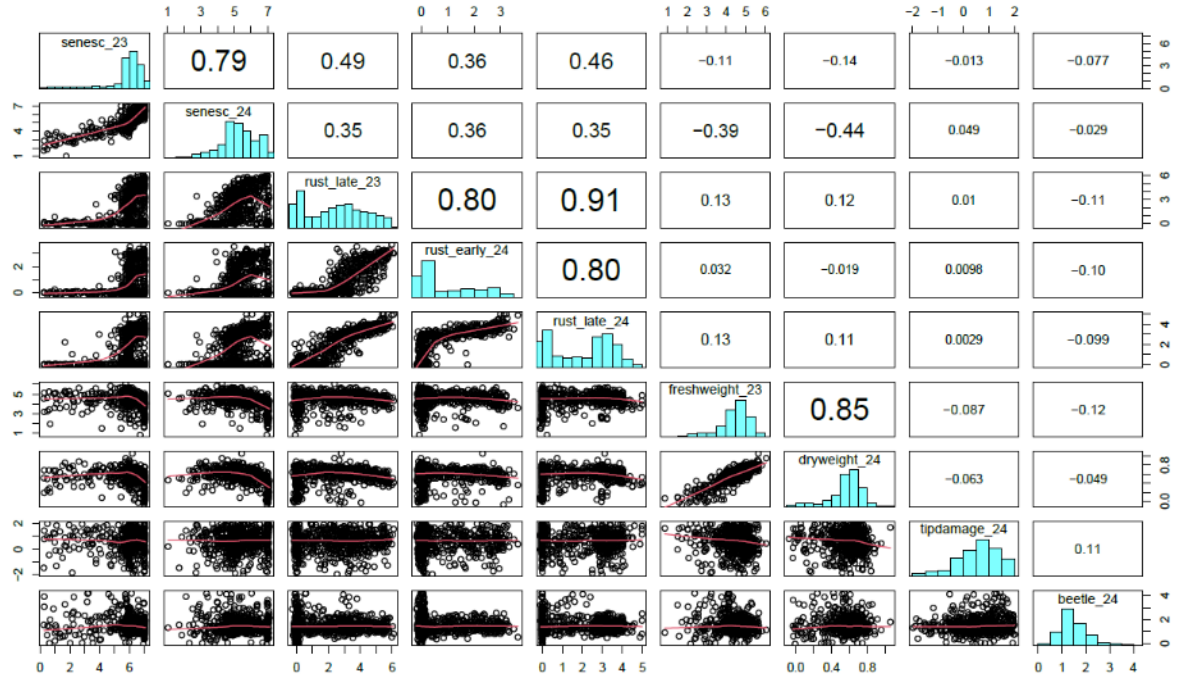
a) Aberdeen



b) Newcastle



c) Somerset



d) Woburn

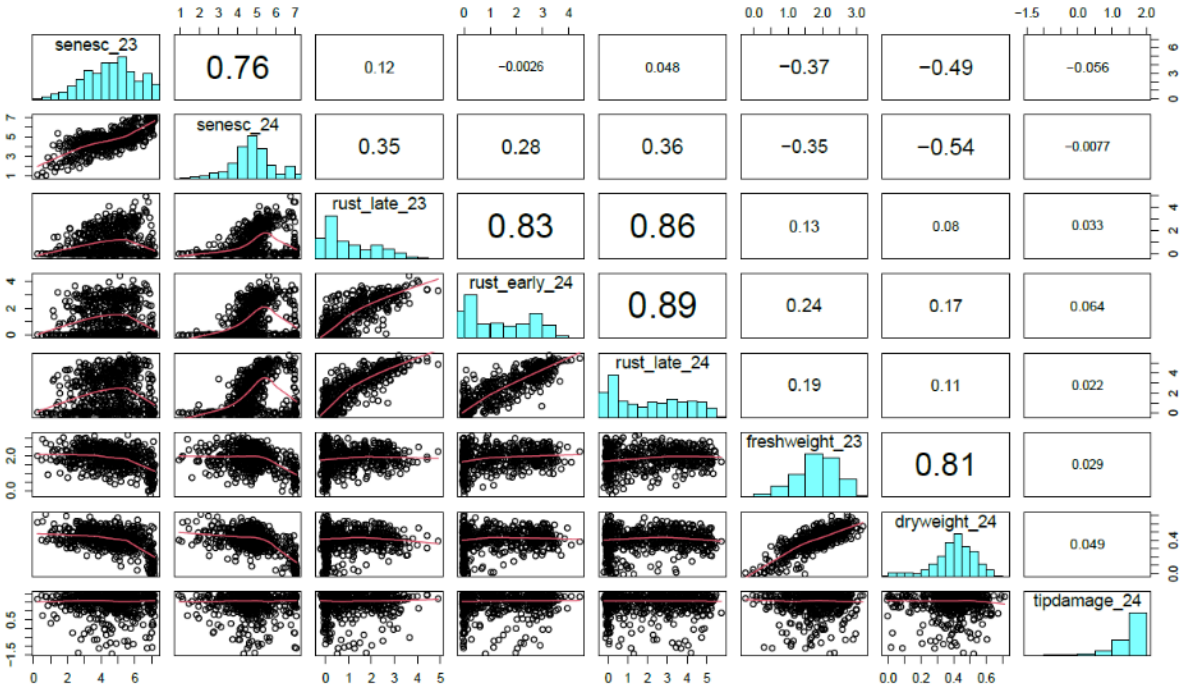


Figure 7. Frequency distributions of assessments and correlation coefficients comparing senescence, rust, yield and tip damage at the four sites. Somerset also included beetle damage scores

Table 3. Summary of rust scores

	Late Rust Year 1 (Sep 2023)	Late Rust Year 1 (Sept 2023)	Early Rust Year 2 (June/July 2024)	Early Rust Year 2 (June/July 2024)	Late Rust Year 2 (Sep 2024)	Late Rust Year 2 (Sep 2024)
	Mean score	No. mean score <1	Mean score	No. mean score <1	Mean score	No. mean score <1
4 GB sites	1.15	[a]102	0.76	[a]31	1.34	[a]18
Aberdeen	0.37	395	0.91	[b]47	0.89	147
Newcastle	0.81	[b]0	No rust	No rust	0.55	119
Somerset	2.42	122	0.87	262	1.93	139
Woburn	1.00	207	1.25	172	2.04	110

[a] Common to all 4 sites

[b] Larger residual variance indicative of variable scores across replicates leading to fewer genotypes with certainty over value of predicted mean <1 (larger confidence interval).

Following planting (spring 2023) it was not possible to score the severity of rust infection until late summer (Late Rust Year 1). At that time the crop had a very open canopy as it was planted on wide row spacing in anticipation of the coppiced form to be created by the first-year cutback. The scores recorded as Early Rust Year 2 were from the regrowth after cutback in winter 2023-24 and so although coppice form were as short new stems.

In late summer 2024, the crop canopy was dense therefore, Late Rust Year 2 was the point of greatest disease pressure on the TP. At that time the two northern sites had notably lower mean predicted score, but the genotypes correlated well with their counterparts at the two southern sites. Aberdeen and Newcastle also demonstrated greater variability in mean predicted score across the replicates, especially in Late Rust Year 1 (Newcastle) and Early Rust Year 2 (Aberdeen).

The assessments have identified useful potential parents for the breeding programme showing strong resistance to infection by the rust fungus (Table 4.). Species such as *S. dasyclados*, *S. miyabeana* and *S. rehderiana* were already known to show good resistance. Six genotypes in Table 4. are the result of crosses already made within the programme (RR prefix to the “Name”). *S. eriocephela* Michx. flowers in September, whereas the majority of willows flower in early spring, so will require implementation of pollen storage protocols.

Generally, it appears from these results that rust infections are less severe in the northern parts of the UK. However, this is based upon two northern sites and only one assessment in the highest disease pressure crop condition. An additional angle to the rust infection severity is the potential for genetic variation in the fungus to exist within the different environments. The reasonably strong correlations between scores across sites suggest that this was not a big component of the results. Samples of rust were collected and stored to observe the genetic variation in the rust pathotypes in future work.

Table 4. The 18 genotypes recorded as having rust scores <1 at all 4 sites in late summer 2024

Genotype	Name
Tordis × RR08155 (Tora × *S. dasyclados* 'Grandis')	RR10127
S. sachalinensis F. Schmidt	-
S. glaucophyloides Fern.	-
S. hookeriana Barratt	-
S. aegyptiaca × *S. miyabeana*	RR10121
S. rehderiana × *S. miyabeana*	RR10210
S. schwerinii E.Wolf.	-
S. rehderiana Schneid. x *S. dasyclados* '77056'	RR06070
S. miyabeana Seem.	III
S. gracilistyla x *S. bakko* (syn. *S. caprea*)	-
RR08402 (SW930812 × *S. dasyclados* 'Loden') × RR09233 ('Tordis' × 'Sven')	RR13106
S. miyabeana Seem.	I
S. viminalis	I
S. schwerinii Wolf	-
S. alba	Corvinus
S. integra × *S. vulpina*	-
S. dasyclados 'Aud' × *S. dasyclados* 'Loden'	RR07130
S.eriocephala Michx.	-

3.8.3 Beetles

Leaf damage caused by beetles (ghosting or skeletal leaf structure) was observed at all sites. Adults and larvae were seen occasionally. There was only one occasion when a site experienced a relatively even distribution of beetles and the damage caused could be fully assessed: Somerset late summer 2024 (Appendix B, figure B3). The prevalence of other willows in the immediate environment to the site likely contributing to this opportunity. Only one genotype, a *Salix miyabeana*, had a score of <1, indicating <5% of leaf area damaged. An additional 107 genotypes had scores <2, indicating <10% of leaf area damaged. When <10% of leaf area is damaged it is possible that beetles landed on the plant, began feeding, but were deterred by aspects of leaf chemistry or structure and moved on. This constitutes a useful pool of potential parents for future crosses.

No other site had a population size and distribution that would allow collection of a useful data set. Beetle damage is known to be sporadic and largely unpredictable. The absence of damage at other sites during 2023 and 2024 does not indicate that beetles may not be a problem for willows growing there in any other year.

3.8.4 Terminalis midge

The terminalis midge (*Dasyneura* spp.) causes tip damage to the growing point of willow stems resulting in further branching from buds below the tip. Of the insect pests of interest to the breeding programme it was the most widely distributed. However, as Figure B4 in Appendix B shows, the distribution was variable, from almost absent at Aberdeen to almost all plants affected at Woburn.

3.8.5 Senescence

The senescence scores for the control genotype were evenly distributed (Appendix B). The environment is generally assumed to act relatively uniformly on the plots in terms of driving senescence. Senescence 2023 was scored at Aberdeen, Newcastle and Somerset on the exact same dates, with Woburn a week later. The Somerset assessment was delayed because of Storm Ciaran and so more genotypes were senesced or nearing senescence compared with other sites (*Figure 8*). However, the data is still distributed across a range of scores and, therefore, of use in GS modelling. In 2024, the distribution of senescence scores was more evenly distributed (*Figure 8*).

A leaf damaging fungus such as willow rust may have caused defoliation that could have been confused with early senescence. In contrast, for leaves damaged by insect pests evidence remains, preventing such an error; For beetle damage - skeletonised leaves, for sawfly larvae - leaf mid-rib only and for aphids stems are blackened. Correlations between rust and senescence data were weak, only exceeding r = 0.4 in Somerset (*Figure 7*). This suggests that the data has captured the genetic basis of senescence.

Ultimately, senescence data will be used in conjunction with bud burst timing in spring to determine the green area duration and therefore potential for photosynthesis and growth. Pest and pathogen resistance is then required to protect that green area from damage. It has not been possible to score bud burst so far as the TP was cutback in winter 2023-24.

Figure 7 shows that there was a yield loss from early senescence. The effect was stronger in the south (Somerset and Woburn) with correlation coefficients between senescence and yield in 2024 being less strong in Newcastle and Aberdeen. The cooler autumn weather at the northern sites (Appendix B, Table B1.) may result in less potential to accumulate further yield in autumn when compared to the southern sites. Senescence scores >5 showed a sharp decline in yield, and this was less pronounced at Aberdeen.

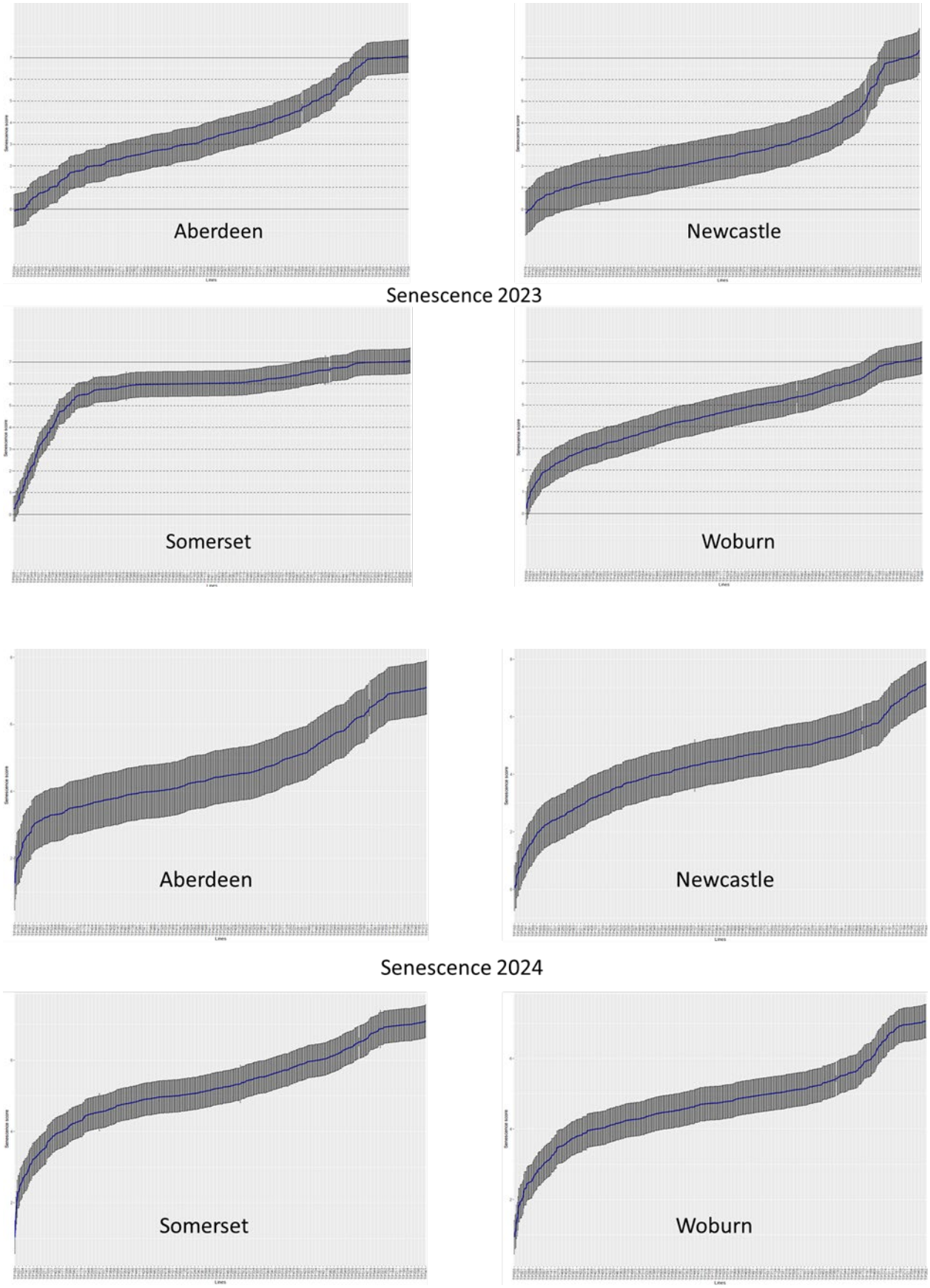


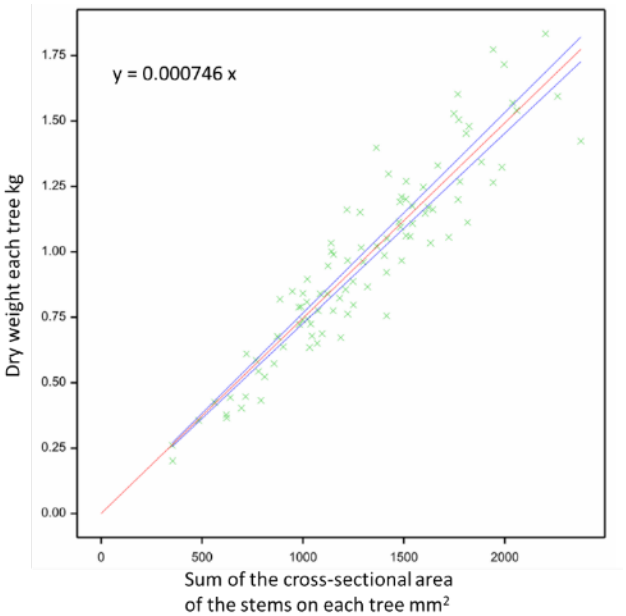
Figure 8. The range of senescence scores with confidence intervals from the Training Population, in 2023 and 2024

Accepting this yield loss may be advantageous when autumn harvesting is desirable e.g. in fields that become too wet to harvest in winter. Early leaf loss maintains nutrient cycling in the field and avoids those nutrients that are coincidentally deleterious elements in combustion systems from entering the biomass supply chain. Leaves also ultimately contribute to increasing soil carbon stocks. In Canada it has been observed that willows that senesce early show greater cold resistance over the following winter (Richard Krygier, Canadian Forest Service, pers. comm.).

In terms of selecting parents for future crosses, it may be desirable to pick from the range of senescence timings dependent on a target market. Breeding for high value extractives (see Commercialisation below) may require specific senescence timing dependent on whether the leaves being in the harvested crop will have positive or negative effect on extraction.

3.8.6 Yield

Yield integrates many factors influencing the plant and therefore can be variable. Appendix B, Figures B6 and B7, show the yield of the control plots. The fresh weight yield of SRCw in the first growing season after planting is very low (Table 5) and not often measured. In this case, fresh weight was measured at first year cutback as the scale of the work precluded drying all cut stems. The dry matter content of SRCw is relatively constant at 45 - 48%, especially within the core majority genotypes in the TP.



To preserve the opportunity for future work with the TP it was decided to make non-destructive measurements indicative of yield. A relationship between the sum of the cross-sectional area of the stems and the dry weight of the tree was derived from data collected previously (*Figure 9*). The sites were planted at 16,667 trees per ha. Assuming a 5% loss during establishment, a figure of 15,800 trees per ha was used as a multiplier to generate a crude estimate of yield in tonnes of dry matter per ha.

Figure 9. The relationship between the sum of the cross-sectional area of all stems at 55 cm above the soil surface (mm2) and yield (kg per tree). Data from previous work

The two yield measurements were strongly correlated at all 4 sites (*Figure 7* and Appendix B, Table B3). This is valuable evidence that first year cutback yield can be an indicator of future performance and is therefore valuable as a selection criterion in the breeding programme.

The sandy soils at Woburn not only hold very little water but are also depleted in nutrients resulting in lower yields. First year establishment was weak and yield very low, but the crop was better able to grow in 2024, once established (Table 5). It was notable that the northern sites produced the greater mean yield once cutback and in coppiced form.

Table 5. The mean yield at first year cutback (fresh weight), estimated mean and modelled maximum yield after one year of coppiced growth (dry weight) of the Training Population and Control at the four sites

Site	Mean Yield g / tree FW, 2023	Estimated mean g / tree DM, 2024	Modelled maximum t ha^{-1} DM, 2024
Aberdeen	125	698	16.85
Newcastle	50	611	14.37
Somerset	108	564	16.72
Woburn	7	402	11.26

In terms of the yields of the coppiced trees in 2024, the highest yielding 20 genotypes in Newcastle and Somerset were not that different except in that the highest yielding genotype produced an exceptional yield at Somerset (Table 6). The 20 highest yielding genotypes at Aberdeen produced considerably greater yield than the other sites. Table 6. shows a mixture of germplasm collected from the wild, newer breeding outputs (RR and LA prefixes), older varieties and Elites. This is encouraging for current deployment of Elites, medium term improvements using existing breeding outputs and longer-term improvements using new parental genotypes.

The average ranking across the four sites reveals some consistent high yielding genotypes as indicated by the low residual variance on the mean rank. Endurance is known to be an excellent variety and features across the sites. A second well regarded variety, Resolution, does not appear. Stott10 had been considered an outclassed variety as it became extremely susceptible to rust, however, it has performed strongly at the sites.

There was no clear steer to select genotypes for the long growing season daylight hours at Aberdeen as many genotypes performed particularly well. The cultivar Aurora had a large residual variance on the average rank across all four sites. It was 24th highest yielding at Somerset and averaged 309th across the other three sites, indicating some potential traits valuable in flooded conditions. The yields of 21_CZ and R326 were considerably greater than all other genotypes at Somerset. However, they performed well across all sites and so it cannot be inferred that they were particularly high performing in flooded conditions. Woburn showed the fewest genotypes in common with other sites in the 20 highest yielding genotypes (6). Three of the four highest yielding genotypes came from the breeding programme. RR10051 and RR10088 both being part of a programme of genetic diversification in the breeding involving *S. rehderiana*, *S. aegyptiaca* and *S. miyabeana*, species identified above as valuable in introducing rust resistance. RR10051 did not produce a high

yield at other sites, suggesting that traits may be present conferring drought tolerance. Similarly, RR10127 and RR10028 showed the greatest and 3rd greatest variance on the mean ranking. They were 11th and 14th highest yielding at Woburn but averaged 434th and 306th respectively across the other three sites. This indicated that further investigation of their ability to grow under drought conditions should be undertaken.

It must be stressed that to date the TP has only been assessed as a first year after cutback crop and that yields at that point have been estimated from a non-destructive assessment of stem diameter using a relationship obtained from previous data. It is reasonable to expect changes in yield to occur in the subsequent growing seasons before the next true yield is measured when cut. In addition, second and subsequent harvest rotations have been observed to produce greater yields than the first rotation.

3.8.7 Black Spot

During 2024, there were reports of some SRCw crops suffering leaf damage. The symptoms were dark brown or black spots on the leaf and in the worst cases leaf margins blackening. One specific cultivar in a field crop in Berkshire was defoliated. The Biomass Connect BFI demonstrator reported several project sites showing symptoms on two-year old SRCw plants. Samples were collected for analysis by plant pathologists at RRes and an attempt was made to capture some basic data from the TPs.

At both Biomass Connect and AWBD sites, symptoms were more severe in the south of the UK. At AWBD Woburn 71/631 and at AWBD Somerset 208/631 genotypes showed no symptoms. The cultivar badly affected in Berkshire was moderately affected at Woburn and Somerset. The overall situation was similar to an incident in Ireland in 2012, in that case the infection was not repeated the following year or since.

Figure 10. Black spot symptoms on willow leaves

The pathologists isolated two broad fungi groups from the infected leaves. Attempts will be made to recreate the symptoms on healthy trees by reinfecting with cultures of each as one or both could have been saprophytic and not the causal agent.

Table 6. The estimated yield of the 20 highest yielding genotypes at each of the 4 sites and the average ranking across all 4 sites. Those in blue font appearing in top 20 at 2 sites, those in red at 3 sites. Pale green highlighting "Elites"

Aberdeen	t ha^{-1} DM
S. viminalis XIII	14.66
S. viminalis XIV	14.68
S. viminalis XX	14.69
S. viminalis I	14.71
Stott10	14.74
Astrid	14.75
RR09296	14.85
Endurance	14.99
RR04125	15.03
S. dasyclados Skv. II	15.15
RR10308	15.26
RR05281	15.30
S. viminalis XVIII	15.52
S. viminalis X	15.59
S. viminalis VIII	15.73
RR06269	15.97
Beagle	16.08
S. viminalis VI	16.26
S. viminalis V	16.73
S. viminalis XXIII	16.85

Newcastle	t ha^{-1} DM
Beagle	12.79
RR10308	12.82
S. viminalis II	12.88
S. viminalis X	12.88
Sven	12.89
S. viminalis XXIV	12.92
RR05281	12.95
S. viminalis L. x *schwerinii* Wolf I	12.96
Bowles Hybrid	12.97
RR10100	12.98
RR06269	13.19
S. viminalis VI	13.25
LA970126	13.30
S. viminalis L. x *schwerinii* Wolf II	13.46
S. viminalis VIII	13.51
S. viminalis XXII	13.52
Ester	13.65
Endurance	13.66
S. viminalis XVI	13.78
RR04125	14.37

Somerset	t ha^{-1} DM
RR07160	12.64
S. dasyclados Skv. I	12.67
S. viminalis XI	12.71
RR09453	12.77
Emma	12.77
RR04261	12.81
S. viminalis XV	12.85
S. viminalis L. x *caprea* L.	12.93
S. viminalis XVII	12.98
RR09296	13.05
RR10100	13.07
S. viminalis VIII	13.31
Linnea	13.41
S. dasyclados Skv. II	13.47
Roth Hambleton	13.76
RR04125	13.92
RR10088	13.95
S. viminalis XXI	14.43
S. viminalis XVI	15.09
S. viminalis VI	16.72

Woburn	t ha^{-1} DM
S. viminalis XXIII	9.51
S. viminalis IV	9.52
Roth Cheviot	9.55
S. viminalis XIX	9.56
S. viminalis VII	9.60
S. viminalis XXII	9.63
RR10028	9.68
RR07054	9.73
Discovery	9.82
RR10127	9.91
RR10121	9.92
S. viminalis XII	9.94
S. viminalis IX	9.96
Emma	9.98
RR05154	9.98
Quest	10.21
RR09296	10.63
RR10088	10.80
Endurance	10.94
RR10051	11.26

	Rank	ResVar
Emma	61.25	4040
S. viminalis XXIII	60.25	3601
S. viminalis V	57	2450
S. viminalis XI	57	2034
S. viminalis IX	52.5	1343
RR09296	49.75	6444
RR10100	49.5	3408
S. viminalis XXII	48.5	4159
S. viminalis X	48.25	1782
RR06269	47.5	3075
Linnea	47.25	2460
S. viminalis XXV	37	109
S. viminalis XVI	36.5	3123
S. viminalis L. x *schwerinii* Wolf I	34	437
Stott10	30	163
S. viminalis VI	25.25	1762
Endurance	24	1321
RR05281	21	173
L810203	20	678
RR04125	15.25	363

3.9. Genome Based Activities

3.9.1 Building required genome resources for target species

Genomes sequencing for the reference genome sample set

Following method optimisation for willow, both long and short read sequencing approaches yielded satisfactory data sets, both in terms of quality and sequence yield) for all samples. Details of data output from the long (ONT) and short (Illumina) read sequencing is provided in Table 2 (Method section). As the first full genome sequencing data sets for some willow species, these data represent an important and novel resource for use within this project but also for the wider tree research community.

Kmer analysis highlighted likely differences between different willow genomes in the reference set. Examples of the output of GenomeScope analysis are provided in *Figure 11*. This analysis was informative on the ploidy levels of the samples, including confirming the previously questioned ploidy of NWC1126 (*S. caprea*) as diploid. Different estimates of genome size resulted but these will require confirmation by additional analysis. However, this suggests that there may be significant structural genome variation in the TP set and analysis of these differences may be informative in association studies and GS-based prediction of phenotypes in any future willow breeding work.

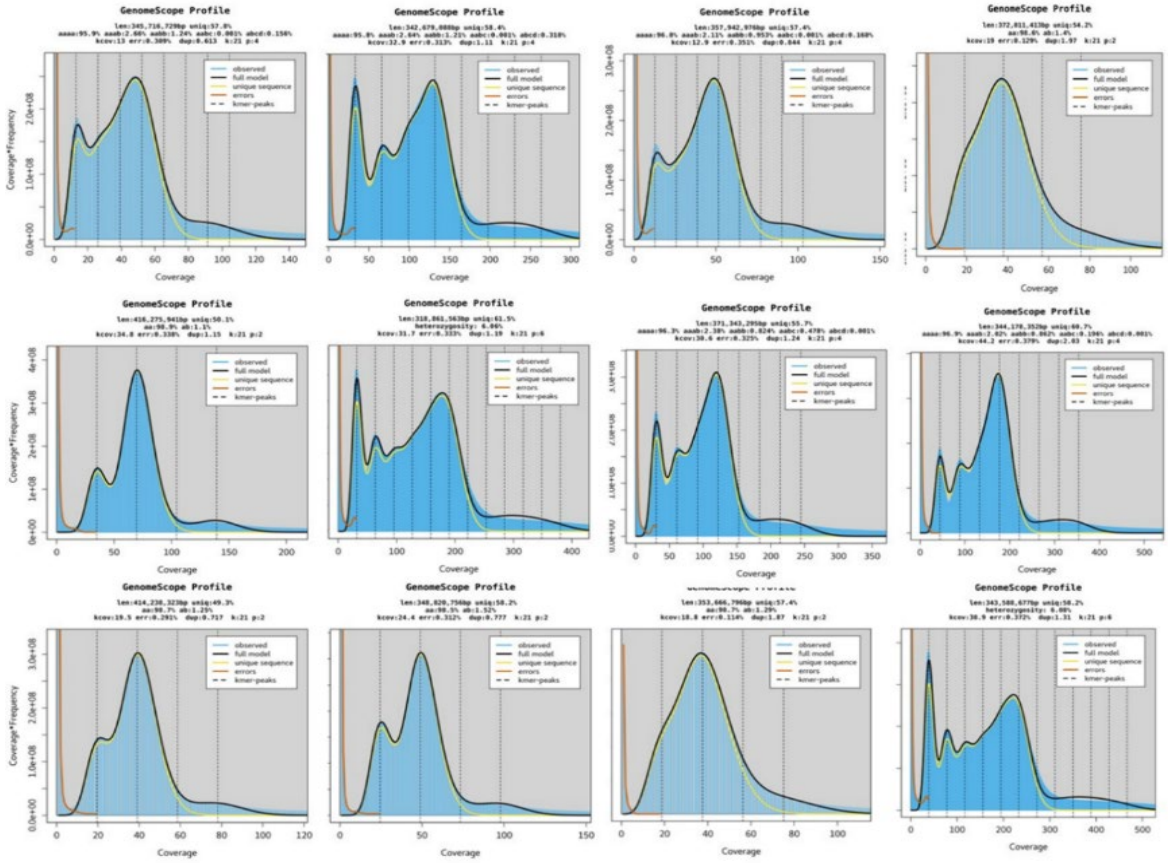


Figure 11. Examples of GenomeScope output for a subset of 12 of the reference set of willow genomes

3.9.2 High-density TP genotyping required for GS models

Whole genome-derived short read (Illumina) DNA sequence information was successfully generated for 554 of the 560 of the TP individuals. Summary statistics describing sequence quality and amount are provided for each sample in Appendix C Overall, data quality metrics were in line with those expected with the technology used and achieved sequence yields were largely as targeted. The data set was therefore considered suitable for downstream variant calling.

Failed samples were attributed to the presence of secondary compounds or contaminants in some extractions that caused library construction to fail, despite repeated attempts. Although only a small proportion of the total TP population, and therefore deemed unlikely to have significant impact on the overall project goals, new extractions using different methods may overcome this issue if required in future.

Results confirmed very high levels of sequence diversity within the TP. Over 20 million potential variant sites were identified in the gene space, which after merging and filtering, were reduced to approximately 280,000 biallelic variant sites. Table 7 summarises the number of sites on each of the 19 willow chromosomes that were used in downstream analysis.

Table 7. Number of gene-based SNP variants identified per willow chromosome

Chr	# SNPs						
1	16162	6	17227	11	14376	16	28854
2	16452	7	11783	12	11080	17	14675
3	14324	8	13636	13	14610	18	12575
4	13449	9	9237	14	12529	19	13963
5	18316	10	15986	15Z	13010	**Total**	**282244**

3.9.3 GS data analysis

Genomic prediction accuracy values for maximum stem diameter, late rust, senescence and yield at all locations ranged between 0.5 to 0.9 for all test and training set combinations, with mean accuracy values ranging between 0.6 – 0.8 for these traits (*Figure 12*). These results imply that the application of genomic selection is practicable and can be used to enhance future breeding efforts of willow.

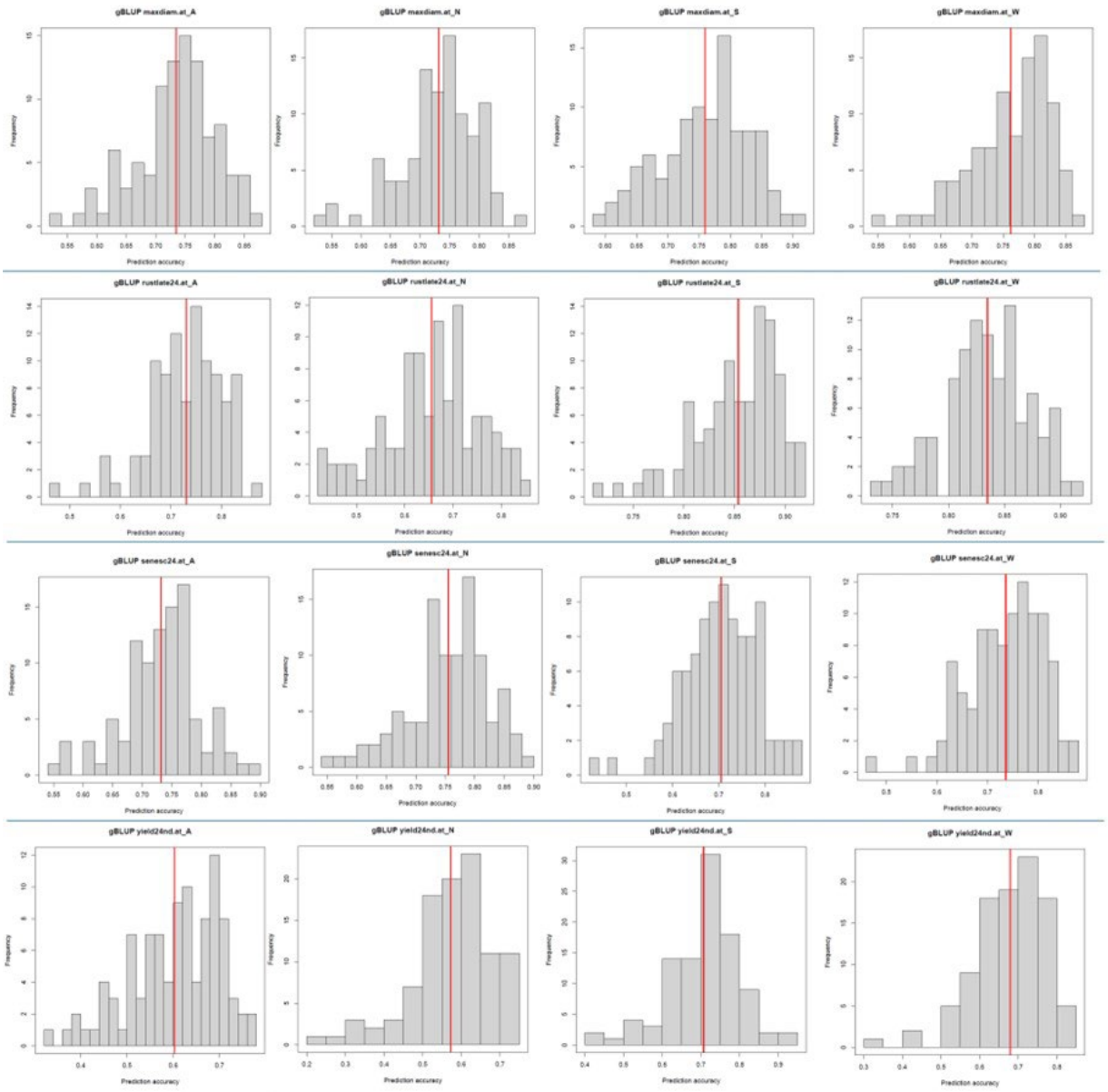


Figure 12. Prediction accuracy values derived from 100 cycles of 5-fold cross validations on maximum stem diameter, rust resistance, senescence and non-destructive yield data from (all from 2024) collected at Aberdeen (A), Newcastle (N), Somerset (S) and Woburn (W) TP sites. Red lines show the mean prediction accuracy. Prediction accuracy is calculated as the correlation of the predicted genomic estimated breeding value with the observed phenotype value of a trait. Higher correlation values indicate high prediction accuracy

3.9.4 Convert disease resistance information into markers for MAS

Scans for marker-trait associations using GWAS analysis identified several putative SNPs worthy of further investigation in future projects. Manhattan plots summarising these results are provided in *Figure 13*. Some associations appeared consistent across sites (e,g, on chr02 at Woburn and Aberdeen) whilst others were site specific.

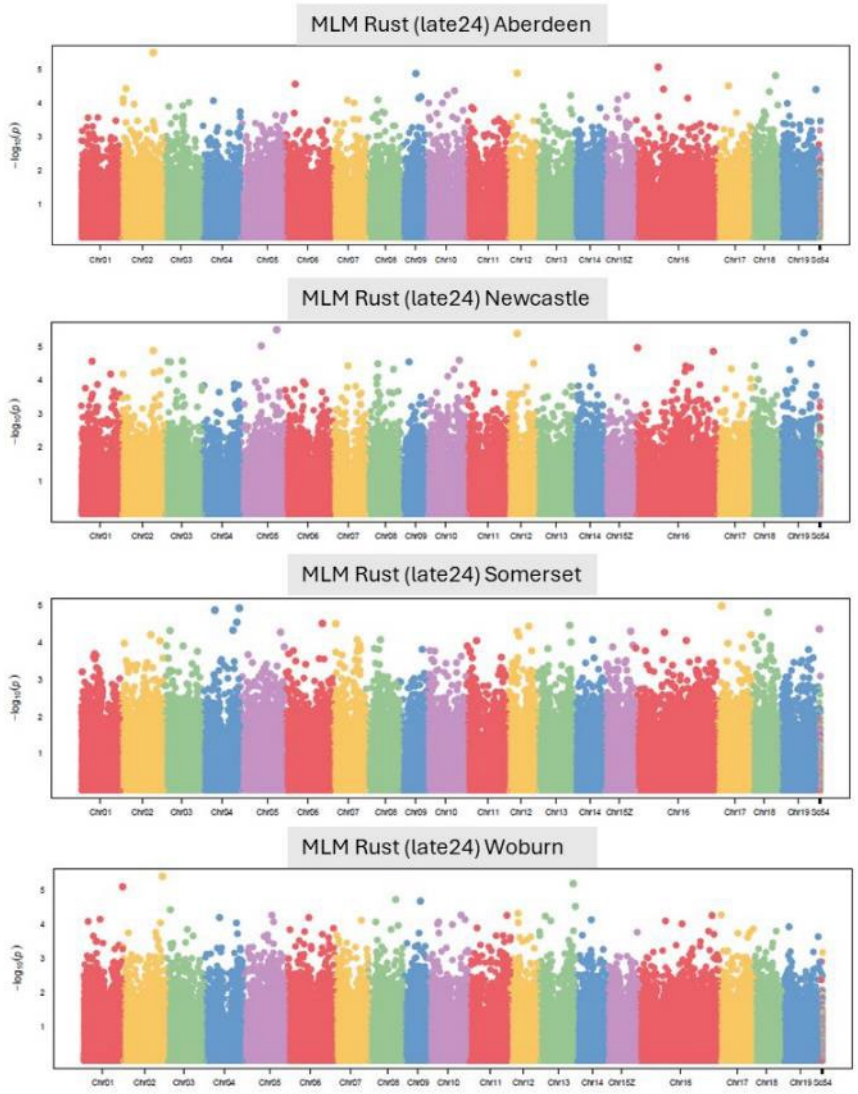


Figure 13. Manhattan plots of genome wide association mapping for late rust 2024 at Aberdeen, Newcastle, Somerset and Woburn

3.9.5 Refine markers for MAS based on known major effect loci

GWAS highlighted several putative marker trait associations for senescence, maximum stem diameter and biomass yield (non-destructive) using phenotype data collected from the different trial sites. Linkage disequilibrium decays rapidly in willow suggesting that in most instances, markers linked to the traits will be located very near to the underlying causal polymorphism. The high density of markers used here should mitigate much of the need to refine markers for future use.

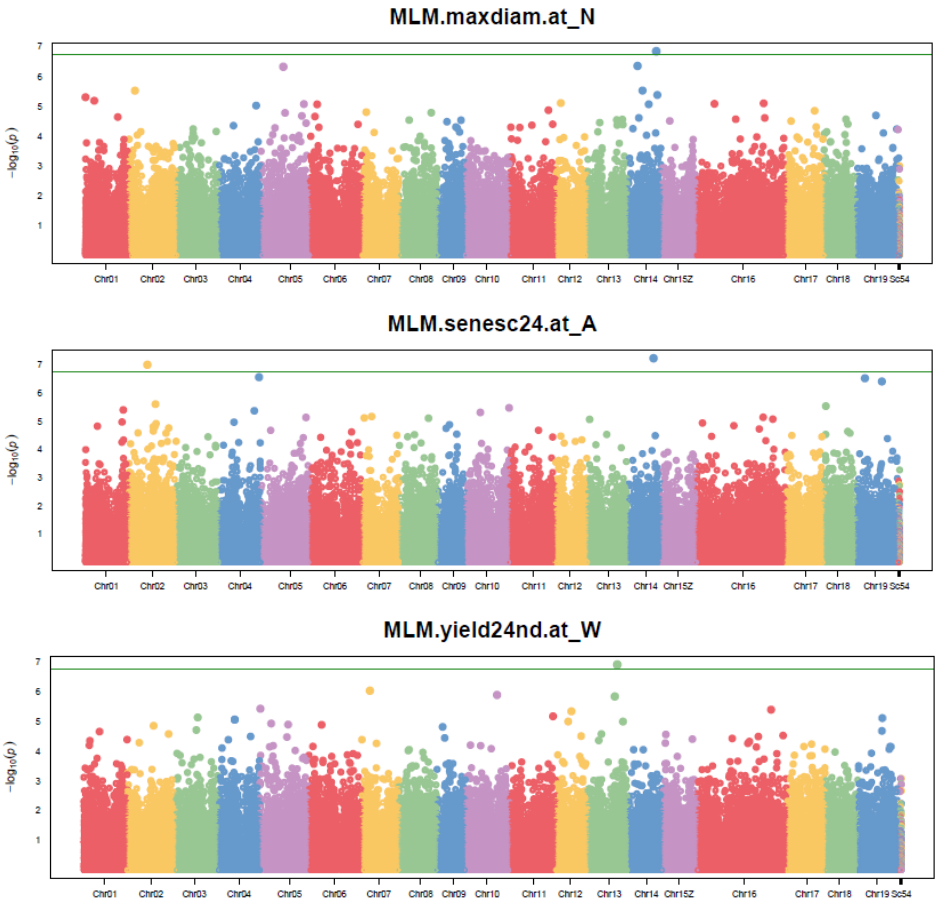


Figure 14. Example Manhattan plots from GWAS studies (2024) showing putative marker-trait associations for maximum stem diameter (Newcastle), senescence (Aberdeen) and biomass yield (Woburn)

3.9.5 Develop methodologies for identifying and confirming genotypes in the Hillsborough trial

After testing and optimisation, the panel of eight highly-informative microsatellite markers was screened against genotypes represented in the main TP collection. The markers proved to be highly discriminatory, and all genotypes could be differentiated with this set. Just four markers were able to differentiate between the vast majority of genotypes. Marker profiles were used to generate a reference database of individual tree genotype 'fingerprints'. Profiles of any unknown willow samples can then be screened against this database to determine if they are a 'match' to any known genotype or variety. An example of the results generated is provided in *Figure 15*. In this example, the profile of the Hillsborough (Northern Ireland) sample matches that from the RRes reference material and can be identified as the expected variety, 'Gudrun'.

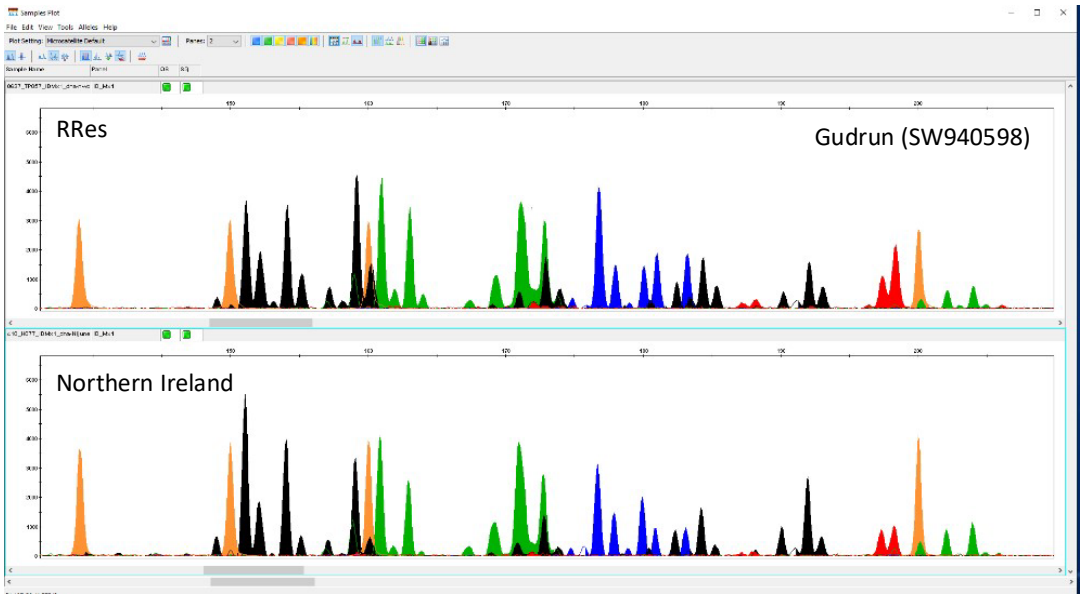


Figure 15. Example multiplexed microsatellite profiles of reference (Rothamsted) and test (Hillsborough, Northern Ireland) samples. Four different microsatellite loci are shown with different colour (black, green, blue and red), orange peaks are from internal size standard

All samples from the Hillsborough trial were screened in this way and results were compared against the reference database to identify any potential discrepancies. This was to ensure that genotypes and phenotypes were correctly paired prior to any downstream genomics prediction or association analysis. Although the vast majority of genotypes were confirmed as those expected, some discrepancies were identified and corrected. This approach also proved useful in answering questions about some potential planting errors at different sites.

Beyond the specific usage here, the creation of the improved database of willow genotype ‘fingerprints’ provides an important resource that may be used in future to identify willow individuals when required. For example, this could be important in protecting plant breeders’ rights or for confirming the identity of particular clones that may be showing symptoms of emerging diseases.

3.10 Micropropagation results

3.10.1 Establishment of material to in vitro conditions.

Establishment of a stock of mother plants.

Cuttings were successfully established in the glasshouse to produce a stock of mother plants following the methodology described in *Appendix C*. This was accomplished in two batches, the first comprising 42 genotypes in August 2022, and the second comprising 40 genotypes in April 2024.

Establishing aseptic cultures.

When establishing aseptic cultures, the first batch of 31 genotypes (11 elite and 20 test) were introduced in September-November 2022 and the second batch of 15 genotypes (1 elite and 14 test) were introduced in July 2024 following the methodology described in *Appendix C*. Not every genotype grown in the glasshouse was introduced *in vitro* due to poor growth from some mother plants or lack of workable material. The extent of available material from mother plants in the glasshouse determined the number of explants introduced. Overall, 2859 explants were established *in vitro*, averaging 63 explants per genotype.

Micropropagation assessment *in vitro*.

After 5 weeks of growth, explant contamination events (bacterial or fungal) were observed in 35 out of 46 genotypes (*Figure 16*). The average contamination rate was 11.6%. Contaminated explants were discarded, and clean, aseptic material was successfully established *in vitro* from 46 genotypes. Later, after establishment, three genotypes were lost because of endogenous contamination. This was caused by endogenous microbes that can remain undetected in tissue culture because the concentration of nutrients and physical conditions (pH and/or temperature) are not optimal for bacterial growth. When the culture conditions change during normal plant growth, they can become more favourable to microbial growth, and previously undetectable bacteria can multiply (Wotjania et al., 2005), affecting plant growth. As previously observed at RRes by Palomo-Rios et al. (2015), axillary buds can be used as an alternative to nodal segments to reduce contamination problems in establishing willow tissue cultures.

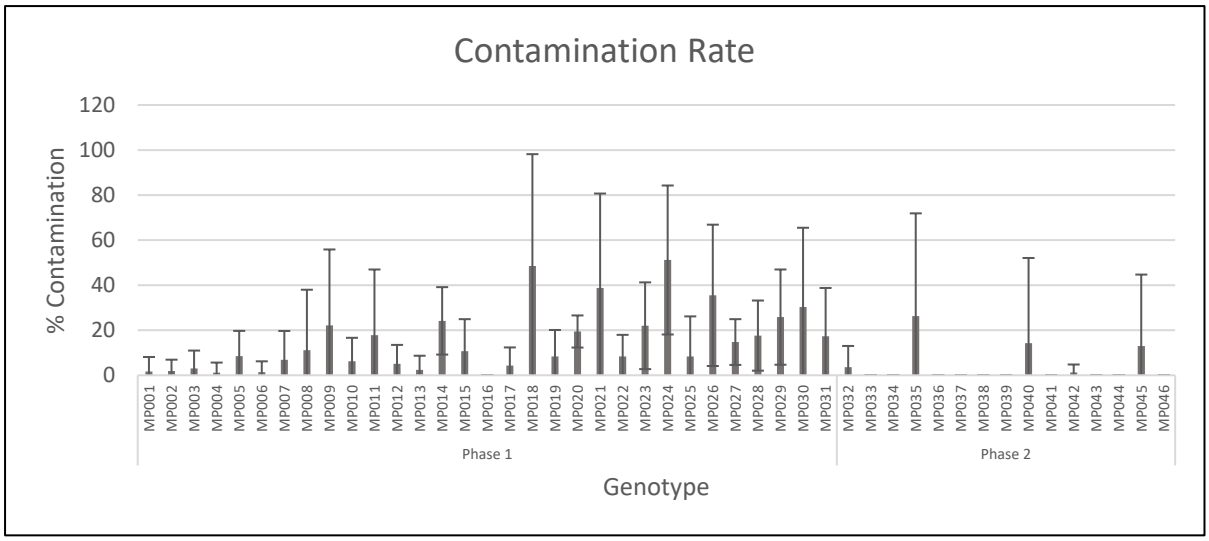


Figure 16. Percentage of explants per genotype displaying contamination after 5 weeks of growth in stock micropropagation media

Strong shooting and rooting responses were observed from most genotypes after introduction to *in vitro* conditions (*Figure 17*). Overall shooting and rooting responses were 91.6% and 93.7% respectively. For shoot length, the overall average length was 22.5 mm.

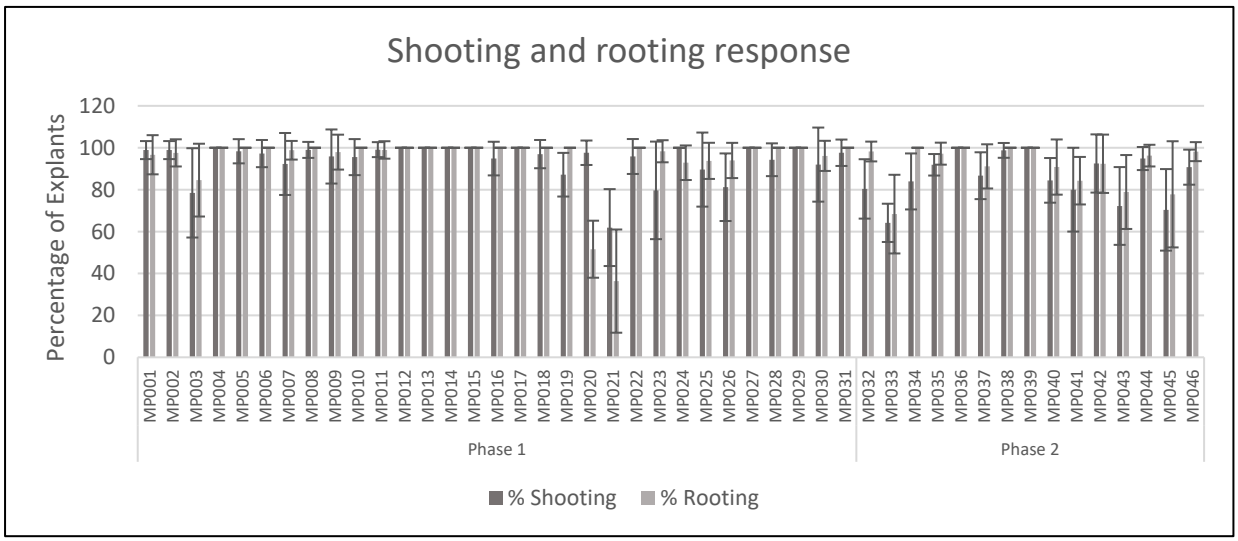


Figure 17. Percentage of explants developing new shoots and roots in vitro after 5 weeks of growth in stock micropropagation media

When establishing *in vitro* stocks, all genotypes except for MP033 displayed sufficient shoot growth for transferal. Fifteen genotypes had transfer rates of 1x or above, with genotype MP012 showing the highest rate at 5.8x (*Figure 18*). Genotypes with low transfer rates had insufficient shoot growth or high contamination levels. A genotype-dependent response was observed in the material recovery from mother plants' original explants introduced *in vitro*.

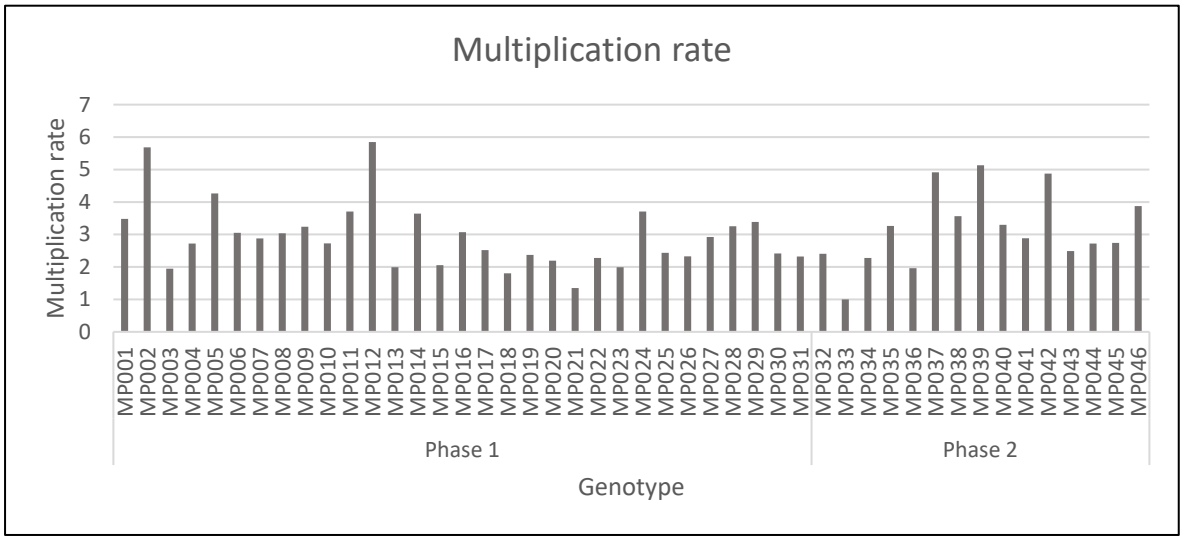


Figure 18. Multiplication rate from all genotypes during multiplication round 1

Overall, of the 46 genotypes introduced, 34 genotypes (73.9%) were micropropagated and maintained until the project end. Three genotypes (6.5% of total introduced genotypes) were lost due to endogenous contamination, and nine (16.5%) failed to grow under standard micropropagation conditions. Further studies using alternative micropropagation conditions might help to establish these non-responding genotypes and check whether the lack of response results from recalcitrance, where cells or tissues of the explants fail to respond to *in* vitro culture manipulations. Recalcitrance is a significant obstacle in the clonal propagation of many plant species, particularly in their adult phase of development (Bonja et al., 2010). It is well-known that juvenile tissues show higher morphogenetic ability than adult ones. *In vitro* recalcitrance could be avoided by using juvenile explants produced from crosses, which are the original target explants in the early stage of a breeding program. tabl

3.10.2 Optimisation of micropropagation conditions.

I*n vitro* optimisation and micropropagation for target genotypes.

Effect of basal formulation.

After 6 weeks of growth in different media, all plants developed roots, irrespective of treatment or genotype. A strong shooting response was also observed in all genotypes (*Figure 19*). Over 90% of explants developed new shoots in MS media for five out of six genotypes, whereas in WPM media, new shoots were developed by 60-80% of explants. Overall, there was a statistically significant increase in the percentage of developed shoots from explants grown in MS media compared to those grown in WPM.

The average length of these developed shoots is a good indicator of the multiplication rate. However, there was no statistical difference in the length of new shoots from explants grown in either medium. The media formulation also had no significant impact on the number of nodes on new shoots. Overall, with a higher

shooting percentage than WPM, these results indicate that micropropagation of willow genotypes in MS media results in an improved multiplication rate.

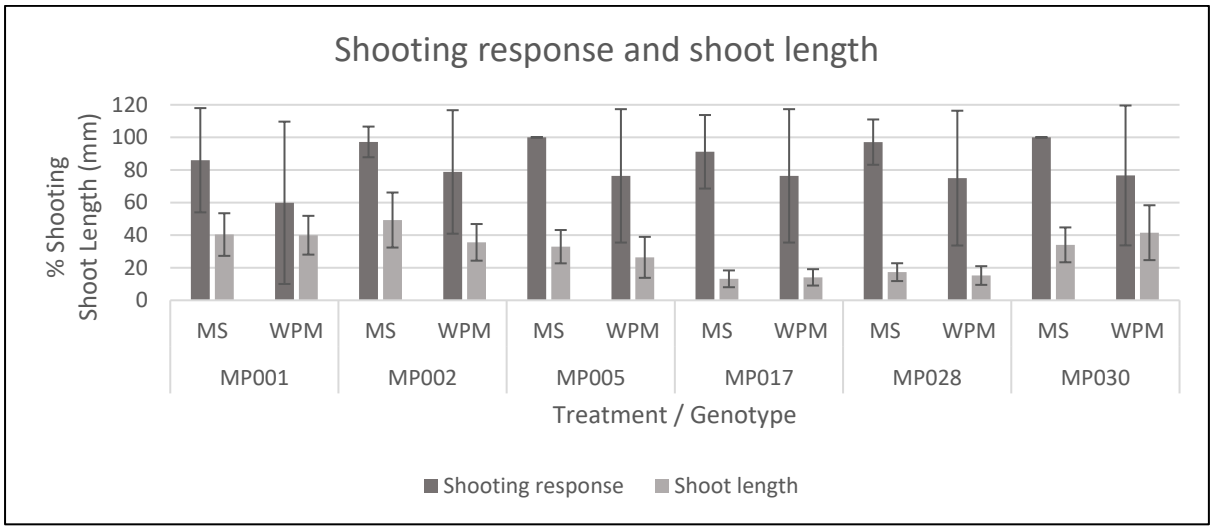


Figure 19. Shooting response (percentage of explants developing new shoots) and shoot length of new shoots (mm) after 6 weeks of growth in media with MS or WPM basal formulations

Effect of plant growth regulators.

The initial optimisation experiments using genotypes MP001, MP002, MP005, MP009 and MP012 indicated different micropropagation responses to the ten treatments, suggesting that optimal conditions would differ per genotype (*Figure 20*).

Figure 20. Effect of different Plant Growth Regulator (PGR) treatments on genotypes MP001 (top) and MP012 (bottom)

Following the initial experiments with the first five genotypes, the effect of these nine BAP/NAA concentration ratios on the micropropagation response was evaluated. Under the highest concentration of BAP (treatments 4, 7,10), all five genotypes studied showed an adverse effect leading to extremely poor growth and rooting responses. Therefore, these treatment conditions were removed from the experiments with subsequent genotypes, reducing the number of treatments from 10

to 7. This resulted in saved time, a reduction in consumable costs, and a lower threshold for the number of plants required from one genotype before it can be introduced to the treatment conditions.

Overall, 30 diverse willow genotypes were studied under different PGR conditions. Results showed that the optimal combination of plant growth regulators differs significantly for different genetic backgrounds (*Figure 21*). This was observed in previous studies and would be expected in a genus as diverse as *Salix*. The final results revealed some common trends. For most genotypes, a medium cytokinin concentration (treatments 3, 6 and 9) had a negative effect on shoot length and rooting. Generally, PGR treatments 1, 2 and 8 (absence or low level of cytokinin) ranked highest for improved shoot length, with at least one of these treatments consistently ranked within the top 2 for every genotype. Optimal shoot length was achieved under the control treatment in thirteen genotypes, and under treatment 8 in nine genotypes. Treatment 2 ranked as the second-best condition for eleven genotypes. Identifying the genetic basis behind these differing responses will require further investigation in future.

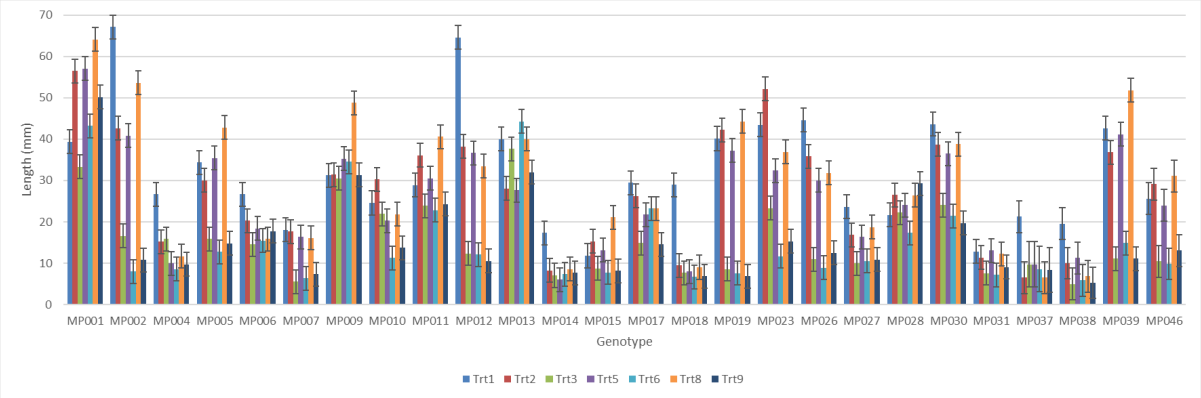


Figure 21. Average length of the main shoot after 6 weeks of growth in PGR treatment conditions

Effect of vessel type.

After 6 weeks of growth in different vessels, new shoots growing from explants were transferred to fresh stock media. Average multiplication rates were determined for plants growing in each vessel type after three rounds of micropropagation (*Figure 22*). The results indicate that plants grown in taller oval vessels yield more plants per vessel to the next round of micropropagation than those grown in standard ECOboxes. The cost and space effectiveness of both containers was estimated. Both vessel types hold the same media volume and can hold up to ten explants but require different shelf space. One square metre of shelf space can hold up to 75 tall vessels or up to 85 ECOboxes. According to the data obtained; to optimise shelf space and multiplication rate, genotypes should be grown in tall oval vessels if the multiplication rate is 0.5x higher than the rate obtained in ECOboxes.

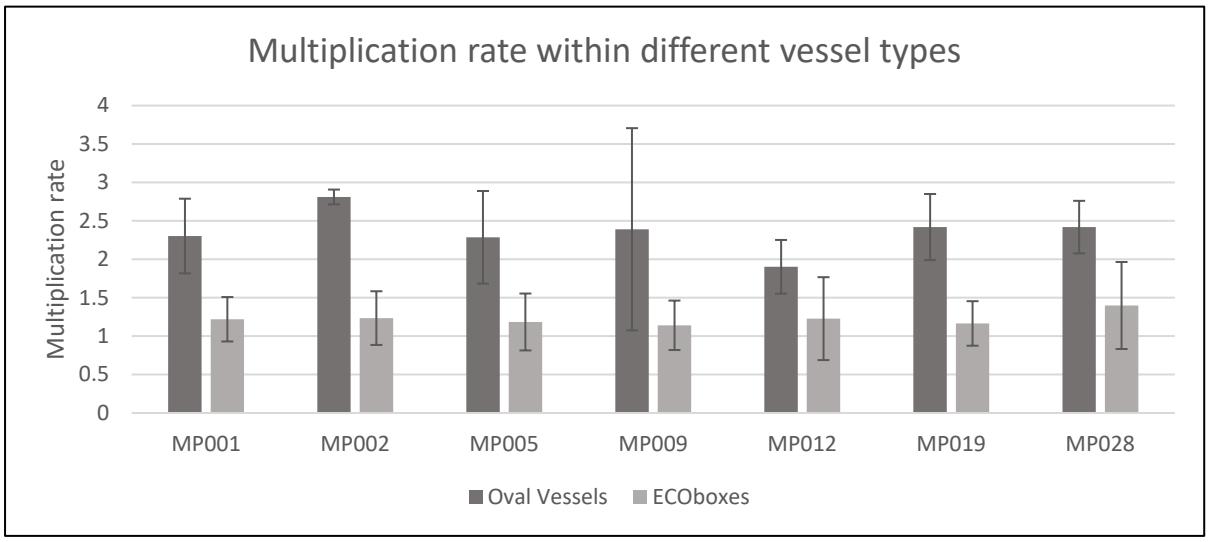


Figure 22. Comparison in multiplication rate between explants grown in tall oval vessels and explants grown in ECOboxes, after six weeks of growth

Overall, data from the optimisation experiments have informed the process for the *in vitro* multiplication of new genotypes. Using MS media with PGR treatments 1, 2, and 8 in oval OS140BOX vessels would give new material an improved multiplication rate in at least one treatment condition. If a genetic basis behind micropropagation response was discovered, the correct PGR treatment could be selected from the onset, and optimal multiplication rate could be achieved.

3.10.3 Demonstrating rapid scale-up of elite material.

Following the traditional micropropagation approach on agar-based micropropagation media, work to demonstrate rapid scale-up was completed with a subset of micropropagated lines taken to cutting production (MP001, MP002, MP005, MP007, MP009, MP010, MP011, MP012, MP013). The willow material was introduced *in vitro* in September 2022 and subjected to two cycles of micropropagation before it was transferred to soil conditions by the end of January 2023. A total of 394 plants from ten different genotypes were successfully acclimatised to *ex vitro* conditions, with a survival rate above 93%. This material was later transferred to the nursery for cutting production.

Temporary immersion systems for rapid scale-up of elite material.

Preliminary experiments showed that whole plants divided into two explants were optimal for TIS culture, using fifteen plants (30 explants) per vessel. In further TIS experiments with different concentrations of cytokinin (BAP), healthy and well-elongated shoots were produced using a low cytokinin concentration. A detrimental effect on shoot proliferation was observed when using a high cytokinin concentration; small, stunted shoots showing leaf malformation were produced, and the rooting response was repressed. Therefore, a lower cytokinin concentration was chosen as the optimal for micropropagation using RITA® TIS. In rapid scale-up tests, multiplication factors obtained ranged from 2.2x to 9x, with an average number of shoots per vessel ranging from 62 to 135 (*Figure 23*). The percentage of shooting

observed differed between genotypes, varying between a minimum of 53.2% in MP009 to a maximum of 100% observed in MP001 (*Figure 23*). The average shoot length ranged from 6.2mm to 23.5mm in MP010 and MP011, respectively.

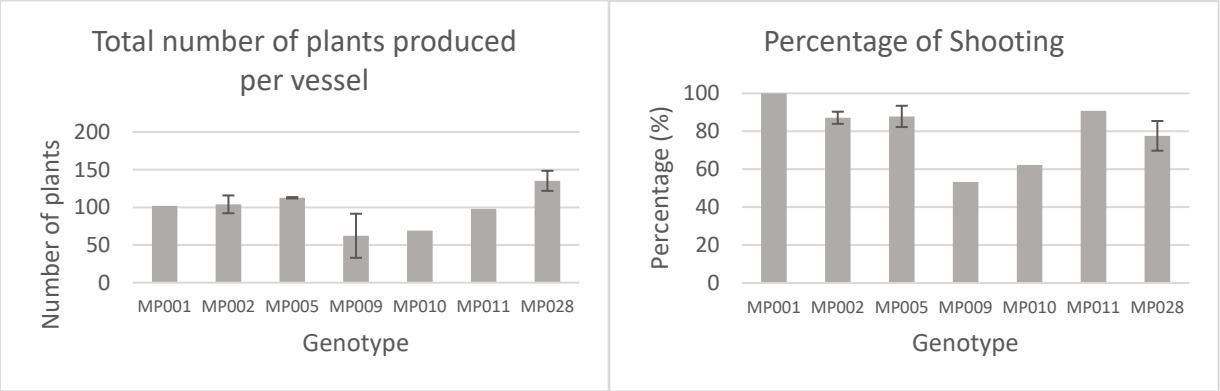


Figure 23. Number of shoots and shooting percentage observed from each genotype after five weeks' culture in RITA® TIS

As previously observed in micropropagation experiments, optimal micropropagation conditions differ significantly for different genetic backgrounds. Productions carried out with material from MP002 and MP028 yielded 419 (four vessels) and 815 (six vessels) shoots, respectively, in a single multiplication cycle of five weeks. Shoots recovered were successfully rooted in stock micropropagation media and moved to soil. The full-scale demonstration of rapid scale-up of elite material, using the TIS to propagate diverse willow genotypes, has been demonstrated.

Cutting production for mass deployment to yield trials.
Plants produced to demonstrate rapid scale-up of elite material were acclimatised in the glasshouse and were then transferred to the nursery in April 2023, where they were monitored and allowed to grow as a source of cuttings for future assessments of performance in a Biomass Connect field trial. From 24 plants per genotype initially transferred to the nursery, the number of cuttings produced ranged from 130 to 250, with a total of 1094 cuttings produced from six biomass genotypes (*Figure 24)*.

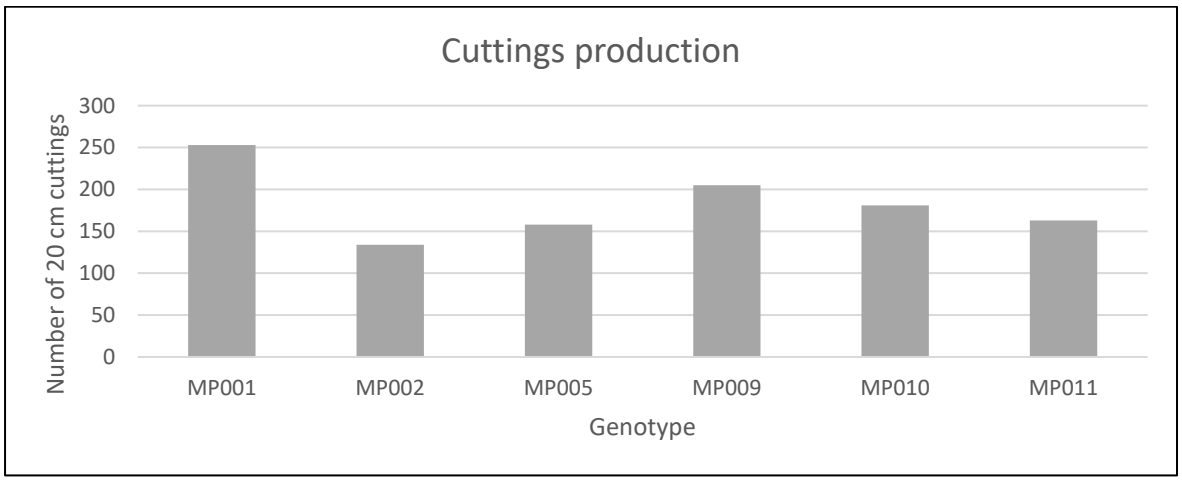


Figure 24. Number of cuttings produced from 24 plants per genotype transferred to the nursery

A total of 84 cuttings produced from micropropagated biomass genotypes were planted in a field trial in Aberdeen in June 2024 (*Figure 25*). Height data was collected in September 2024, which indicated normal growth of all genotypes except for MP002 (*Figure 26*). The reason for this will require further investigation in future.

Figure 25. Aberdeen micropropagated cuttings field trial (establishment year)

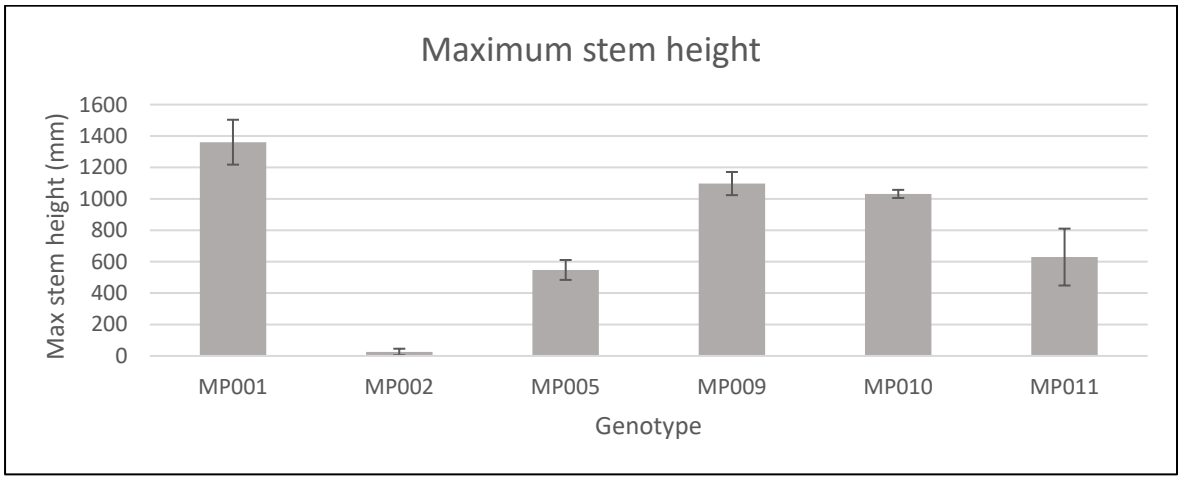


Figure 26. Maximum stem height of plant transferred to Aberdeen field trial

Model for rapid up-scaling using micropropagation.
Based on micropropagation data gathered across all stages of the process, using both traditional micropropagation and TIS, models were designed to estimate RRes capacity to produce willow cuttings for two different purposes: field trials and mass production.

For mass production, the multiplication data produced at all stages (stages 0 to 4 and cutting production) for six elite biomass genotypes was used with the aim to produce as many cuttings as possible (*Appendix C, Figure 7*). From 120 initial explants, after four cycles of micropropagation of six weeks, the total number of plants produced was estimated to be 9,231, ranging from 782 (MP011) to 2,147 (MP002) per genotype. The estimated cutting production from this amount of material after one more year was 67,171 cuttings in total (*Figure 27*).

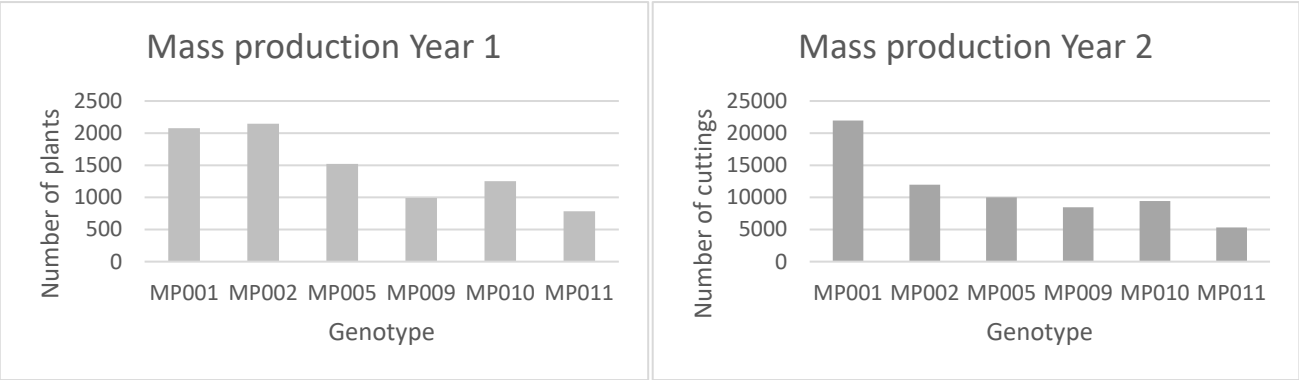


Figure 27. Predicted number of plants (year 1) and cuttings (year 2) generated from the mass production model

An alternative production model was designed that focused on generating material from a broader range of genotypes for use in field-based yield trials (as would be required in a GS-led breeding programme). Here, data from micropropagation optimisation experiments (using 30 genotypes) was used. Production was designed to take place in one year and five months (*Appendix C, Figure 8*). Assuming a minimum of 640 cuttings would be needed for field trials in four locations using four replicates (n=40), two micropropagation cycles would be required to produce sufficient material. The initial number of explants needed per genotype was calculated based on the expected multiplication factor to produce between 120 and 140 plants per genotype after two micropropagation cycles. After two cycles, a total of 3,821 plants from 30 genotypes would be produced. One hundred explants from each genotype were considered for the next stages of rooting and acclimatisation, leading to a final cutting production of 722 cuttings per genotype after material is established and grown at the nursery (a total of 21,652 cuttings).

4. Quantitative key performance metric

The quantitative key performance metric selected was the attainment of high yields thereby maximising profitability for the grower and minimising the land area required to make a meaningful contribution to UK net zero targets amongst competing land use demands. Within the timeframe of the project, this was only achievable by selecting the right genotype for the right environment. Breeding objectives towards this goal will take many years, albeit fewer than previously once GEBVs are deployed in the programme.

Almost all varieties (genotypes) currently marketed in Europe were included in the TP and planted out in the varying environments, described above. This provided data regarding their performance in environments that had not been recorded previously. Albeit with a caveat that the timeframe of the project only allowed for a non-destructive estimate of yield after one growing season in the coppiced form. The absolute yield and rank order of genotypes may change within the first harvest rotation of 3 years. In addition, there is an increase in yield observed in the second and subsequent harvest rotations of SRCw crops.

The yields estimated from the genotypes at Aberdeen were most promising. Overall, 12 genotypes were estimated to have produced a yield in excess of 15 t ha^{-1} DM. Previous breeding trials conducted in the south of the UK rarely achieved 15 t ha^{-1} DM and Newcastle (chosen for similarity to an average site) and Somerset (south of UK) were unable to produce such high yields. The field chosen at Aberdeen was typical of the area suggesting a wider potential for strong crops from northeast Scotland. The BFI Demonstrator Biomass Connect planted two sites in Scotland with SRCw, yield data comparable to the Aberdeen site will be available soon.

The work presented here demonstrates many opportunities for breeding for greater yield potential protected from losses due to pests and diseases. AWBD Phase 2 allowed the application of Genomic Selection with rapid multiplication of breeding material to be applied to willow breeding, this has moved Technology Readiness Levels (TRL) from TRL4 to TRL 5. Phase 2 funding has delivered GEBVs. These have yet to be deployed in breeding due to the length of the willow growth cycle. Therefore, we consider Phase 2 have progressed the GS technology in willow to TRL5 (Pilot Scale, about to be applied and thoroughly tested on the target breeding).

5. Contribution to UK sustainable biomass supply

The UK is land limited relative to population size and demand for resources from the land such as food, fibre, fuel and biodiversity. AWBD's objective is to generate knowledge that optimises the use of land for biomass production and deploys the latest technology in genomics and plant breeding to maximise the fuel supply from one of the key perennial energy crops, willow. This is to be achieved by maximising yield potential whilst protecting that potential from losses due to pests and diseases.

Prior to this project, yields of 20 t ha^{-1} yr^{-1} dry matter had been achieved twice in trials. The national average yield was below 10 t ha^{-1} yr^{-1} dry matter. Crop modelling had predicted that such high yields should be possible more regularly by matching the genotype to its preferred environment.

In Phase 1 of the BFI Programme a Meta-analysis of previous SRCw trial data had been conducted. The results were used to produce a 'Growers guide to short rotation coppice willow (SRCw) varieties for biomass' Leaflet (Appendix E, Figure E3), which included a table for varietal selection. A PDF version was made available on the RRes, Biomass Connect and FarmPEP websites. To date, it is possible to identify 1,131 downloads and 450 hardcopies being viewed. Hardcopy was taken and distributed at events such as the Low Carbon Agriculture Show, NZIP & Biomass Connect Showcases and Groundswell (Appendix E. Table E2). Attendance at these events was essential to maintain our profile as active and applying the latest technologies to crop improvement, and to gauge the market conditions.

Previous crop trials had been conducted in a narrow range of environments. This work has expanded that range whilst remaining aware of the practicalities of growing a SRCw crop. Extreme slopes, rocky ground etc. where a mechanised crop production would not be possible and niche environments representing no more than a few thousand hectares were excluded.

One major outcome of the work has been to identify potential for high yield in the northeast of Scotland. Genotypes were identified that will be investigated further for growing on low productivity, drought prone land. Whilst very high yields may be unachievable on such land, breeding SRCw varieties tolerant of such conditions may provide an attractive option for growers faced with few financially viable alternatives. Nutrient cycling within a perennial crop such as SRCw should be more efficient than annual cropping. The sandy soil at Woburn is vulnerable to erosion. In September 2024, there was an extreme rain event (Appendix B, Table B1 and Figure B1). The SRCw site showed no sign of erosion, whereas previous less intensive rain events had caused severe erosion to arable fields. Previously, there had been very little data from the north of the UK. The outcome has been a large step forward in breeding for maximum yield in different environments.

6. Impact on Greenhouse Gas emissions

6.1 Opportunity for domestic sustainable biomass

In the early years of the willow breeding programme the target was fossil fuel replacement in combustion to steam electricity generation. Full replacement of, and co-firing with, coal allowed biomass to conveniently fit into the energy mix. In early autumn 2024 the UK saw the closure of the last coal fired power station. Now the target for large scale biomass production is the development of BECCS (Bioenergy with Carbon Capture and Storage). The infrastructure is not yet well developed but the development of the biomass supply side must continue in parallel with the engineering to be ready for deployment. Ultimately this will bring negative emissions, where a key Greenhouse Gas, CO_2, is removed from the atmosphere in combination with energy provision. Plant breeding is a long-term activity, however, the high yields already indicated in northeast Scotland are a strong positive component for a future BECCS scheme linked to CO_2 storage in North Sea gas wells.

6.2 BECCS

Assessing the scale of these negative emissions is challenging until it is known what scale of BECCS engineering projects will come online. Evero intend to be the first to implement BECCS and give an example of quantification of CO_2 removal. Their 20.5 MWe power station in NW England is targeting *ca*. 218,000 tonnes CO_2 captured per year. The power station is fuelled by 113,500 tonnes of waste wood. If SRCw were to directly replace that wood, it would require ca. 12,000 ha of land at current yields of 10t ha^{-1} DM and ca. 8,000 ha if we consistently achieved 15t ha^{-1} DM in the near future by matching the genotype to the environment. Drax power station in North Yorkshire utilises almost as much wood in a week as Evero in a year. To supply that power station for a year from SRCw would require more than 400,000 ha at 15t ha^{-1} DM to produce 4% of UK electricity consumption. Drax rather modestly states a potential to ultimately capture 8 M t CO_2 per year, a useful component of the Government's 20 – 30 M t per year target for 2030.

6.3 Impact of Transportation requirements

A second challenge is how competitive energy crops such as SRCw, produced in the UK or elsewhere, will be with an established forestry-based virgin wood and/or waste wood supply chain. Willow growing at scale will be spatially dispersed, logistical and transport costs will limit the proportion of SRCw going into any one facility. The willow breeding programme can contribute to competitiveness by generating a supply of high yielding varieties along with knowledge of where best to exploit their potential. The current Key Performance Metric of achieving consistently high yields by placing the right variety in the right environment is the next step on a path to maximising competitiveness and de-risking the decision to plant.

6.4 Self-supply

The market for heat at small and medium scale is ever present. SRCw fits well into a self-supply / smaller scale system especially for those not on the gas grid, often replacing fossil mineral heating oil. Increases in the price of gas since 2022 have made biomass more price competitive within the gas grid. Biomass boilers are expensive to install and maintain but buffer against oil and gas price volatility. Manufactured "heat logs", largely comprised willow wood chip, have been popular with users of multi-fuel stoves. They produce lower particulate emissions than many log supplies and may replace coal. This continues the theme of fossil fuel replacement and makes a useful contribution to net zero targets. However, it is unlikely to generate sufficient new plantings to support the breeding programme.

6.5 Non energy

Composted willow has been used as a substitute for peat compost. It is reported that sales of such compost in Denmark are strong. There is interest in planting willow as a grazing supplement. This has been shown to reduce nitrous oxide emissions from urine patches and potentially reduce methane emissions directly from ruminant animals. Both contribute to the UK Net Zero targets via a non-energy route but are not a high value market.

6.6 Soil carbon

When SRCw is planted on former arable land there is potential for increases in soil carbon content (Gregory *et al*. 2018). However, when planted on former grassland the effect may be the opposite. Unpublished data from the BBSRC Perennial Energy Crops for Greenhouse Gas Reduction project shows CO_2 emissions from soil continuing 3 years after ploughing in grass and planting SRCw. There has been interest in SRCw in paludiculture, especially in peat soils. The site on the Somerset Levels within this project provided valuable information on genotype response to high water tables including flood inundation. Reports from Ireland where SRCw was planted on worked out peat bogs indicated potential for acid tolerant genotypes. Across all land types there are questions over the permanency of any carbon incorporated into the soil system by SRCw. A future change of land use could release much of the carbon accumulated if sensitive management regimes are not adopted.

7. Commercialisation

7.1 Strategy

RRes is a world-renowned breeder of SRCw for biomass. Varieties are licenced to multipliers who usually also provide planting and other management services. RRes earns a royalty on each sale of willow cuttings from its varieties (genotypes). In 2021, a minimum 3,000 ha of new planting with RRes varieties was needed to achieve an income that would support the advanced breeding programme. The Climate Change Committee (CCC) recommended 23,000 ha of energy crop planting each year up to a total >700,000 ha by 2050. There would be competition between crop types (especially miscanthus and SRCw) and between SRCw breeders, however, 3,000 ha from a total of 23,000 ha seemed achievable in 2021. To date, post 2021, there has been little SRCw commercial planting so the opportunity to generate royalty income has not been realised.

7.2 Market Segmentation

The willow breeding programme sits at the starting point of a supply chain. It is highly dependent on factors further along that chain to drive demand without many routes to influence those factors. Biomass clearly has a role to play in achieving Net Zero in the UK and in many other countries. It has been successfully deployed by several companies in the energy sector. However, the biomass utilised is not derived from perennial energy crops. Forest co-product and waste wood dominate the combustion market whilst annual crops supply the biological conversion technologies.

Those supply chains for woody biomass are well-established and set the price. By comparison SRCw growers are small scale and dispersed. The SRCw sector will need to reform some of the aggregating type organisations that existed 15 years ago but fell by the wayside as markets failed to develop. Terravesta perform this function for the miscanthus crop.

7.3 Market Failure

Appendix E. Table E1, shows the typical Gross Margin for each hectare of SRCw planted in England. After 5 years at a price of £40 t^{-1} DM delivered the crop has not repaid the establishment costs. The delivered price of SRCw wood chip would have to be £62 t^{-1} DM to return a positive gross margin in year 5. If yields were consistently raised to 15 t ha^{-1} DM the price of £40 t^{-1} DM would produce a positive gross margin in year 5. In Scotland a positive gross margin would not be generated until either yield increased to 15 t ha^{-1} DM at £40 t^{-1} DM or the price delivered to £54 t^{-1} DM. All of the above scenarios ignore fixed costs and the cost of financing a negative gross margin for 5 years.

Those growing for self-supply may be comparing these costs against expensive and volatile heating oil and gas prices. However, such plantings are unlikely to generate sufficient revenue to support the breeding programme but may contribute usefully to the target income. The breeding programme is highly dependent upon deployment at scale, such as with BECCS, to be self-sustained. Presently, for those looking at a commodity crop market, the alternative crop types are far more attractive practically and financially.

In 2024, RRes commissioned the National Non-Food Crops Centre to compile an independent review of market conditions for a willow breeding programme. The conclusion was that it is unsustainable at this time.

7.4 Competing land use

The Sustainable Farming Incentive (SFI) is the mechanism to replace Basic Payment Scheme (BPS) as it is phased out in England. Very few SFI options are directly applicable to SRCw. However, the SFI includes many options that generate far greater income than SRCw and with greater certainty. For those wishing to plant trees there is an Agroforestry SFI option and the England Woodland Creation Offer (EWCO) makes large payments to those choosing to plant new woodlands. EWCO includes supplementary payments for ecosystem services that could be equally or better achieved using SRC. SRC is not eligible for either scheme. The devolved administrations have different but similar schemes.

7.5 Routes to market

Previously the focus for the RRes SRCw breeding programme was biomass to energy. During the period 2008 – 2025 there was interaction with 25 private sector companies, 8 UK, 13 European, 3 North American and 1 in China. In total 29 yield trials and 8 multiplication fields were planted, and 5 licences signed by these companies.

Unfortunately, interest declined in the latter years. Whilst biomass to energy and BECCS remain an important longer-term target, alternative end uses were investigated. For some time, there had been an interest in making a peat alternative compost for domestic and larger scale growers. This can be a demanding market requiring a high-quality product. At this time, it is unlikely to generate new planting. SRCw wood chip can be drawn from existing crop as it justifies a greater price than biomass to energy.

The chemistry of willow wood has generated a lot of interest. Early investigations highlighted the tannin component as of particular interest. Feeding willow to ruminant animals has been shown to reduce nitrous oxide emissions from urine patches and potentially reduce methane directly emitted by the animal. Feeding of exotic animals led the way with interest at zoos. More recently, RRes supplied planting material for cattle and sheep farms in the Yorkshire Dales. This is unlikely to generate large scale plantings as it is a supplementary feed not the principal forage fed to the animal.

The greatest potential to generate planting and therefore income to the breeding programme is by increasing the wood chip sale price. Peat replacement compost began this process, but much wider options are needed. RRes has a programme of research under the banner Green Engineering in which high value products are being targeted from willow chemistry. Some may be very high value, but potential low volume, such as pharmaceuticals, others greater volume (chemical building blocks replacing oil derived compounds) and contributing to a greater price. It is likely that these value streams will result in a residue which could be used for biomass to energy. The commercial reality for these products will take a few more years to realise.

The knowledge gained via the AWBD project will be valuable in directing breeding for such a multi-purpose crop. High yield, protection from pests and diseases and suitability for different environments will be incorporated using the data generated here. GS will play a large part in breeding for these new higher value traits and micro-propagation will be essential to commercial development.

During the course of this project, RRes has maintained its profile as a source of competitive willow varieties through attendance at many events, see Appendix E. Table E2 for details. Working closely with Biomass Connect, Envirocrops and the Centre for High Carbon Cropping projects has greatly assisted in maintaining that profile. The ‘Growers guide to short rotation coppice willow (SRCw) varieties for biomass’ (see Appendix E, Fig. E3) was distributed and discussed at many of these events.

8. Conclusion

The AWBD project has delivered against several ambitious project objectives and in doing so has laid strong foundations that will underpin the future development and successful deployment of SRC willow as a biomass crop in the UK.

The establishment of the network of AWBD trial sites, coupled with the extensive and robust phenotyping program, has delivered a hugely important resource for deployment and future improvement efforts. For the first time, we have robust data that will inform decisions on precision deployment of particular genotypes in different environments with a view to maximising yield – a major objective of the project.

The AWBD project has also generated unprecedented genomic resources for willow. The sequence resources created are, to the best of our knowledge, unrivalled worldwide and have the potential to underpin many aspects of willow improvement in future. They also have the potential to inform many aspects of fundamental plant biology research.

Combining the phenotyping and genomic resources developed within AWBD has enabled us to demonstrate the feasibility of a genomic selection-based willow breeding programme. The datasets developed here mean the UK is uniquely poised to take advantage of these innovations should there be an upturn in interest and an increase in planting. Innovations around micropropagation mean we are ready to rapidly upscale new varieties (or existing ones) as required.

This multidisciplinary project has successfully brought together experts in agronomy, phenotyping, genetics, bioinformatics and statistical genomics, molecular biology and tissue culture, in addition to training new project staff. The project has helped to maintain the UK expertise and competitive position in willow growing, breeding and research.

In the short term there is additional performance data from currently available varieties to be disseminated. Of particular value is the data from sites further north than previously available, where performance appears to have been particularly strong with high yields and low pest and disease incidence. Generally, the willows deployed in the Training Population were very tolerant of seasonal flooding and so offer great potential for combining biomass production with geo-engineering projects.

In the longer term an invaluable and world leading data set has been generated on which to base further crop improvement via plant breeding using the latest technologies of Genomic Selection. This will be a vital addition to the work on high value extractives from willow being conducted at RRes, Further investigation of the potential for drought tolerance identified in this project warrants prioritisation.

9. References

The Committee on Climate Change 6th Carbon Budget
https://www.theccc.org.uk/publication/sixth-carbon-budget/

BEIS 2021 Biomass Policy Statement
https://www.gov.uk/government/publications/biomass-policy-statement-a-strategic-view-on-the-role-of-sustainable-biomass-for-net-zero

Biomass Strategy (2023) https://www.gov.uk/government/publications/biomass-strategy

Bonga, J.M.; Klimaszewska, K.K.; von Aderkas, P. (2010) Recalcitrance in clonal propagation, in particular of conifers. *Plant Cell Tissue Organ Cult*., **100**, 241–254.

Catt JA, Weir AH, Norrish RE, Rayner JH, King DW, Hall DGM, Murphy CP (1980) Soils of Woburn Experimental Farm III. Stackyard. *Rothamsted Report for 1979*. Part **2**. pp 5-39.

French BK, Legg BJ. Rothamsted irrigation 1964-76. *J Agric Sci* 1979;**92**:15-37.

Gregory AS, Richter GM, Dungait JAJ, Shield IF, Cuniff J, Macalpine WJ, White RP, Murray PJ, Joynes A and Karp A (2018) Species and genotype effects of bioenergy crops on roots, carbon and nitrogen in temperate agricultural soil. *Bioenergy Research Volume**11** Issue2 Page 382-397* DOI10.1007/s12155-018-9903-6

Lloyd, G. and McCown, B.H. (1980) Commercially-Feasible Micropropagation of Mountain Laurel, Kalmia latifolia, by Use of Shoot-Tip Culture. *Combined Proceedings-International Plant Propagator's Society*, **30**, 421-427.

Murashige, T. and Skoog, F. (1962), A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum*, **15**: 473-497. https://doi.org/10.1111/j.1399-3054.1962.tb08052.x

Palomo-Rios, E, Macalpine, W, Shield, I, West, J, Amey, J, Hanley, S, Krygier, R, Karp, A & Jones, HD. (2015) Efficient and genotype-independent method for rapid multiplication of disease-free willow clones via *in vitro* propagation. *Canadian Journal of Forest Research*. Vol: 45 Issue: 11 Pages: 1662-1667

Penman HL. Woburn irrigation, 1960-8. V. Results for leys. *J Agric Sci* 1970;75:75-88.

UK Government (2024) Clean Power 2030 Action Plan: A new era of clean electricity
Wojtania, A., Puławska, J., and Gabryszewska, E. (2005). Identification and elimination of bacterial contaminants from Pelargonium tissue cultures. *J. Fruit Ornament. Plant Res.* **13**, 101–108.