ROTHAMSTED RESEARCH

Accelerating Willow Breeding and Deployment (AWBD)

Final Report from a Lot 1 Phase 1 project for the BEIS Biomass Feedstocks Innovation Programme

Contents

Key findings

Objective 1. Section 2 below. Data collation from previous breeding trials for multi-site analysis. Planned deliverable (achieved): analysis of available data on performance of current varieties at different sites for environment matching and to guide selection of genotypes for the TP. Remaining uncertainties: the analysis of data was drawn from a relatively narrow range of environments, largely RRes-bred genotypes – needs to be expanded as proposed in Phase 2.

Objective 2. Section 3 below. Generate genomic data to identify candidates and define best sequencing approach for the TP. Planned deliverable (achieved): Extensive genotype data generated, genetic diversity of available germplasm assessed, marker technologies ready to deploy efficiently at scale. Remaining uncertainties: largely mitigated

Objective 3. Section 4 below. A detailed plan for deployment of GS in willow. Planned deliverable (achieved): Deployment plan based on informed selection of germplasm, collated phenotype information and well-defined technical protocols. Remaining uncertainties: Phase 1 has enabled progression of GS as proposed in Phase 2. We envisage some evolution of the approach as analytical methods develop during the timeframe.

1. Introduction

Willows (*Salix* spp.) are a very diverse group of catkin-bearing trees and shrubs that are widely distributed across temperate regions of the globe. Some species respond well to being grown in short rotation coppice (SRC) cycles, which are much shorter than conventional forestry. Coppicing reinvigorates growth and the biomass rapidly accumulates and can be used as a source of renewable carbon for bioenergy and biofuels.

Domestication of willows for short rotation coppice (SRCw) culture is a relatively recent endeavour, with breeding programmes established in the 1980s. Willow varieties are being developed for biomass production through breeding, including at Rothamsted Research (RRes), where research on agronomy, pests and diseases has also been conducted. Willow is now grown commercially in many areas of the world including the UK, Eastern and Northern Europe, North America, India and China, all of which offer market opportunities.

UK SRCw breeding began in 1996 at the Long Ashton Research Station (LARS), it focussed on bioenergy and exploiting a large germplasm collection which begun in 1923. LARS closed in 2003. The willow germplasm collection and key members of staff transferred to RRes.

Research to develop molecular markers associated with traits of interest to the breeding programme began in 1999. Marker Assisted Selection (MAS) has been applied to good effect within the programme. Genomic Selection (GS) has been demonstrated as a highly effective method for further improving breeding efficiency when traits are under multi-gene control. In GS, a training population (TP) is established in which all loci that regulate a phenotype are in linkage with at least one

marker. The data from intensive phenotyping and genotyping of the TP are used to develop a model for predicting the genomic estimated breeding value (GEBV) of progeny. GS is now routinely used in livestock and crop breeding but the capacity of GS to accelerate variety development has not yet been applied to willow across multiple species and in a UK context.

1.1. The present situation

SRCw is currently grown on approximately 5,000 ha in the UK. This is a small area when compared to the major crops which are measured in hundreds of thousands of ha. This project report delivers the planning for two major contributions to upscaling as part of tackling the climate crisis; to accelerate breeding of improved (higher yielding, positive impact) varieties and optimum deployment of existing varieties. Knowledge of existing and near market variety responses to environments will be generated and disseminated supporting growers by de-risking immediate upscaling. Simultaneously, innovative technologies and materials required to accelerate willow breeding by GS will be identified, thereby upscaling the supply of elite planting material for the UK market more rapidly than currently possible.

1.2. The future

All crops need to be underpinned by a dynamic breeding programme. Pest and disease populations undergo genetic change exerting continuous pressure on genetic resistance mechanisms in crop plants. Abiotic factors, such as weather patterns, can be seen to be changing particularly rapidly in this period of anthropogenic derived climate change.

As is the case across agriculture, farmers are largely price takers. Therefore, at the time of writing, the financial case for willow growing is dependent on the component of income that they can influence, dry wood yield per ha. Market interventions based upon the environmental value of the crop may arise in the future to further incentivise growing.

It is clear from trials conducted over the last 20 years that greater dry wood yield potential can be gained from willow. It is equally clear that not all varieties can maintain resistance to diseases such as rust and therefore consistently achieve yield potential. The majority of crop types are bred for a specific environment, for example wheat is targeted at lower rainfall parts of the UK. SRCw has the flexibility to be targeted at a broad range of environments each of which introduces specific breeding objectives. In several environments where land use options are restricted SRCw forms an attractive additional option.

An active breeding programme is essential to exploit further yield gain, protect the potential yield and exploit new environments. This supports the overarching objective of the BFI Programme which is “to address barriers to feedstock production helping to scale up the supply of sustainable domestic biomass in the coming years.”

2. Objective one. Data collation from previous breeding trials for multi-site analysis

BEIS hold Annex 1 which contains additional graphical and tabular outputs associated with these analyses.

2.1. Identification of previous trials and initial analyses

A thorough review of previous willow trials identified 55 breeding trials planted since 1997 that could be reliably used to assess genotype-by-environment interactions. Nine of these trials included data from multiple harvest events, so that response data were available for 71 distinct harvest events. The original plan was to focus on 5 response variables – fresh weight harvested, percentage dry matter (DM) in harvested material, dry weight (DW) yield per ha per year, the number of stems per stool (plant) and the cross-sectional area of each stool. Initial analyses, using either Analysis of variance (ANOVA) or Residual maximum likelihood (REML) for mixed models (where the design structure was more complicated), allowed assessment of the patterns of variability, including identification of the need for transformation of the responses prior to analysis to satisfy the analysis assumptions, primarily of homogeneity of variance. Whilst a multi-trial analysis could have been constructed using the raw data from each of these 71 distinct harvest events, the need for different transformations suggested that a more effective approach would be to extract the genotype means and standard errors (SEMs) from each trial and perform multi-trial analyses on these data, using the reciprocal of the SEMs as weights to allow for the different levels of precision provided by the different trials (essentially equivalent to the analysis of the raw data).

The extraction of the means and SEMs produced a total of 2305 records collated across the 55 trials and 71 distinct harvest events. Of the 55 trials, only 7 (8 harvest events) had data on number of shoots per stool, albeit from large scale trials, resulting in 1159 records, and only 11 trials (12 harvest events) had data on cross-sectional areas resulting in 294 records. There were 42 trials (50 harvest events) with data on fresh weight harvested producing 1999 records and from 48 trials (64 harvest events) for both percentage DM and DW yield per ha per year, in both cases comprising 1258 records. Yields of >20 odt $ha^{-1}yr^{--1}$ were observed at number of trials sites including trial 6, 2nd harvest rotation, (Endurance, 20.42 $ha^{-1}yr^{-1}$) and trial 42, 1st harvest rotation (RR05011, (21.74 $ha^{-1}yr^{-1}$) this demonstrates the yield potential of SRCw.

The trials included a total of 714 distinct genotypes, 326 of which do not have records for either percentage DM matter or DW yield per ha per year. Of the 388 genotypes with records for these two variables, 291 appear in only 1 or 2 trials, and only 13 genotypes appear in 10 or more trials. The most frequently occurring genotypes are Resolution and Tora, which were both observed at 53 distinct harvest events for these two variables.

Where response variables had to be transformed prior to the initial “per-trial” analyses, back-transformed means and approximate standard errors were obtained using the BACKTRANSFORM procedure in Genstat. This uses a first-order Taylor

series expansion to calculate the approximate back-transformed standard errors based on the transformed means and standard errors.

2.2. Collation of environmental variables

The environment was quantified using weather data from the trial sites during the period of testing. Data were available for all but 6 trials, mostly those located outside of the UK. Some soil data were available, but for many trials collection of such data had not been a priority at the time of testing and many soil variables could not be quantified retrospectively.

Weather data were available for up to 4 different characteristics – mean temperature for the month, total monthly rainfall, total monthly radiation and total monthly sunlight hours. With trials generally harvested in the winter period between December and February, when no growth takes place, these weather variables were summarised as monthly means (temperature) or totals (the other variables) for the 9 months from March to November, inclusive, and the mean temperature and total rainfall, radiation and sunlight hours for the three-month winter period from December to February, inclusive.

For each of the 71 distinct harvest events, weather summaries were identified for 80 variables (20 for each characteristic, where available) covering the 2 years (December to November) prior to the winter period in which the harvest was taken. Whilst this doesn't cover the whole of the growth period for all harvests, a large proportion of the harvest events were following only two years of growth, so that a third year of weather would not be relevant.

2.3. Multivariate analysis of weather data

With relatively little knowledge available about the key periods when weather affects the growth of willow, principal components analysis (PCA) was used to summarise the main sources of variability of weather between the 71 distinct harvest events. Five separate PCAs were performed on different sets of the 80 variables – for the 20 mean temperature variables, for the 20 total rainfall variables, for the 40 mean temperature and total rainfall variables, for this set plus the 20 total radiation variables, and the full set of 80 variables also including the 20 total sunlight hours variables. All PCAs were based on the correlation matrix, to allow for the different scales on which the different characteristics were measured. The PCAs for mean temperature and total rainfall included data for 65 of the 71 distinct harvest events, with a further 5 harvest events being omitted when total radiation was included, and a further 9 harvest events being omitted when total sunlight hours was also included. Hence, the results based on mean temperature, total rainfall or both provide a more comprehensive assessment of the variation in environment across the harvest events. Whilst 71 distinct harvest events were identified from the 55 trials, there were 17 combinations of location and harvest year that were common across two or three different trials, so that there were only 49 distinct environments from which response data were obtained.

Biplots provide a visual summary of the results of the PCAs. Each figure shows the results for the first two principal components, plotting the distinct environments as red circles with a numeric label (overwritten where two or three trials had a common

environment), and ten of the weather variables as axes, again with a numeric label plotted at the positive end of the axis at the edge of the plot. Biplots could be constructed containing more than ten weather variable axes, but these then become difficult to decipher and interpret. Hence there are two plots for the first two PCAs (mean temperature only; total rainfall only), four plots for the third PCA (mean temperature and total rainfall), 6 plots for the fourth PCA (mean temperature, total rainfall and total radiation) and 8 plots for the fifth PCA (all weather variables).

The biplots allow an interpretation of the relationship among the weather variables and the associations between sites and these variables. For example, positive scores on PC1 (to the right) are associated with higher mean temperatures, particularly in March, May and June (axes 2, 4 and 5), with positive scores on PC2 (to the top) associated with higher temperatures in December-February and August (axes 1 and 7) and negative scores on PC2 (to the bottom) particularly associated with higher temperatures in July and October (axes 6 and 9). Hence site/harvests with indices 46 and 40 have higher temperatures in March, May and June whilst those with indices 70, 26, 5 and 61 have low temperatures in these months.

2.4. Multi-trial analyses

Following back-transformation (where required) of genotype means and SEMs, the combined genotype mean datasets were analysed using a weighted REML approach for a mixed model. The reciprocals of the (approximate) SEMs were used as weights, to take account of the varying levels of precision of the genotype means from different trials (caused by different levels of replication, different harvested plot sizes, and different levels of between-plot environmental variability). The mixed model analysis included harvest event nested within trial as the random model and genotype as the fixed model.

From these analyses we can obtain estimates of the mean for each genotype, taking account of the differences between the harvest events and trials, together with standard errors of differences for comparing every pair of genotypes – these are highly variable as they depend both on the frequency at which each genotype appears across the combined data set, the precision (weights) associated with each mean obtained from the “per trial” analysis, and the frequency at which the pair of genotypes occur together in different trials (pairs of genotypes will generally occur together at all harvest events for a particular trial).

Figure 1 below shows the estimated means for DW yield (t/ha/year) for the 388 genotypes, sorted into ascending order, with error bars for the mean standard error of a difference (SED) for each genotype (hence large error bars occur for genotypes occurring with lower frequency and/or in less precise trials).

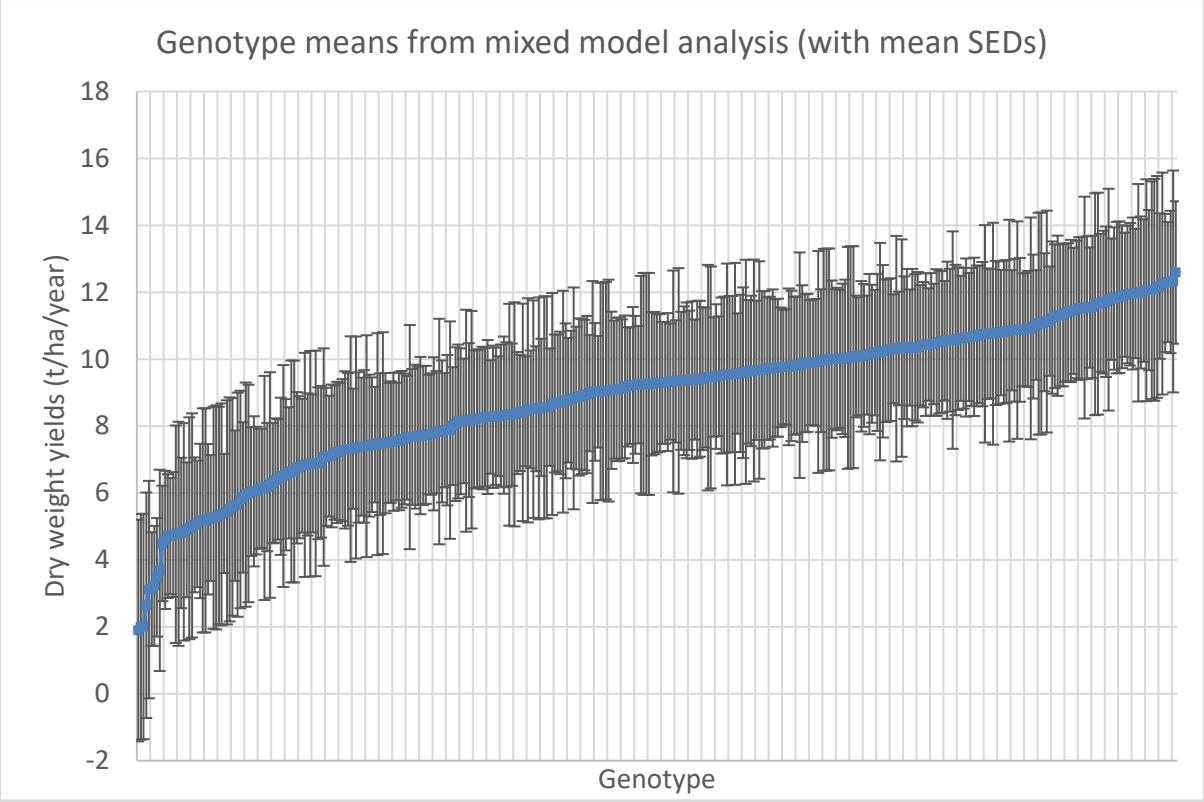


Figure 1: Genotype means from weighted REML mixed model analysis of DW yield means extracted from collated trials, with mean standard errors of differences, in rank order of genotype mean.

A second output from each of these REML analyses is a prediction of the mean response for each distinct harvest event, adjusted for the different genotypes that occur in each of the trials. These predictions can be used as a measure of the environment experienced by the genotypes in each trial, and hence as explanatory variates in a simple Genotype-by-Environment (G-by-E) analysis.

2.5. Genotype-by-Environment Analyses

Whilst the REML analyses of the collated genotype means provide information about the relative performance of the different genotypes averaged across the different trial environments, information about how individual genotypes perform across different environments can be obtained for the more frequently occurring genotypes by regressing the means from each harvest event against a measure of the environment for each harvest event. Two approaches were considered – the first used the estimated harvest event means obtained from the REML analyses to provide a measure of the trial environment and considered the genotype responses for the 54 genotypes observed at 7 or more harvest events; the second looked for more direct associations with weather variables, using the first three principal component scores from each of the five PCAs of the weather data to provide a measure of the trial environment and considered the genotype response for the 29 genotypes observed at 8 or more harvest events. Whilst it would have been preferable to use the individual weather variables rather than the principal component scores, the relatively low frequency of observations for most genotypes and the large number of potential weather variables means that this is not possible

without considerable additional information about the likely key weather characteristics and timings.

In the G-by-E analysis approach for the analyses of DW yield and % DM, respectively, genotypes with slopes close to 1.00 and intercepts close to 0.00 have an average performance across environments. Those with slopes close to 1.00 and intercepts significantly greater than 0.00 (e.g. Resolution, Endurance) perform better than average across all environments, whilst those with slopes close to 1.00 and intercepts significantly less than 0.00 (e.g. LA980289) perform worse than average across all environments. Genotypes with slopes significantly greater than 1.00 (e.g. Tora, Discovery) perform relatively better than average in better performing environments, whilst those with slopes significantly less than 1.00 (e.g. NWC885 Shrubby and Corail) perform relatively better in poorer performing environments, the intercept value providing information about the absolute performance in more extreme environments e.g. the negative intercept for Discovery indicating a poorer performance than average in poorer environments, and the positive intercept for NWC 885 Shrubby indicating a better performance than average in poorer environments – the combination of parameters for NWC885 Shrubby actually indicating a fairly consistent performance across all environments.

Similarly, summary statistics from the G-by-E analyses against the weather principal components. For these analyses, graphical presentations for each explanatory variable with the observations and fitted models adjusted for the variation in the other explanatory variables – these are, however, relatively easy to interpret in this case because the principal component score variables are, by construction, orthogonal to each other, though interpretation also requires the identification of the key weather variable contrasts associated with the principal components. So, for example, for the PCA for mean temperature, positive values of PC1 are associated with higher temperatures generally across the whole two-year period prior to harvest, with PC2 contrasting between temperatures in the year immediately prior to harvest and in the year before that (negative values associated with relatively higher temperatures in the later year, positive values associated with relatively higher temperatures in the earlier year). Genotypes with slopes significantly different from zero then indicate a stronger performance with the indicated weather conditions – for example, Tora having a positive slope for PC2 indicates a stronger performance where temperatures are relatively higher in the earlier year and relatively lower in the later year, and Stott 10 having a positive slope for PC1 indicates a stronger performance where temperatures are relatively higher throughout the two-year period.

3. Objective two. Generate genomic data to identify candidates and define best sequencing approach for the TP

3.1. Background

Under this objective we aimed to characterise all potentially useful breeding willow germplasm held by RRes with the primary objective of identifying genotypes suitable for a Training Population to underpin planned GS approaches.

A thorough understanding of the genetic diversity available in breeding germplasm is highly beneficial to crop breeding programmes. To address this, DNA-based molecular marker approaches have been widely employed to deliver a definitive measure of genetic diversity within breeding germplasm for most crops. It has long been recognised that willow is a system that can benefit from the application of these approaches due to many inherent complications of this genus. For example, willows are notoriously difficult to classify by the visual assessment of morphological characteristics using traditional taxonomic approaches. This is further complicated by the fact that there are hundreds of species and many of these hybridise readily (the majority of biomass willows bred to date are interspecific hybrids). While the diversity within the genus is a significant asset in terms of breeding, it has also hampered previous efforts to develop DNA marker assays that work across the diverse set of species of interest to breeders. Previous work in the programme, funded by the Biotechnology and Biological Sciences Research Council (BBSRC) and completed in 2008, used Amplified Fragment Length Polymorphism (AFLP) markers to largely overcome these problems as it provided a universally applicable marker system that could handle the diversity within the genus. However, the throughput that could be achieved was highly limiting, meaning that the AFLP approach could only be applied to a much-reduced subset of available genotypes.

More recently, developments in next generation sequencing (NGS) technologies have greatly increased our ability to generate vast amounts of DNA sequence information. This has revolutionised the way molecular marker studies are done, both in terms of sample throughput and the depth of information produced. As an example, Genotyping By Sequencing (GBS) is a commonly employed method that allows many (often several thousand) small regions of any genome to be sequenced in a single assay containing many individuals in a single NGS run. The selection of a subset of regions to sequence is achieved by first digesting the genome with specific restriction enzymes that cut the DNA at specific recognition sequences.

While we have adopted and developed these approaches as part of willow research at RRes, they have, to date, only been applied to specific populations or single species. In this project, we aimed to apply these sequencing approaches to screen the entire, diverse collection of useful willow breeding germplasm for the first time. This material included accessions of the UK National Willow Collection, our main breeding germplasm resource comprising of diverse species, natural hybrids and hybrids resulting from historical breeding for traditional end-uses and, more recently, biomass. The resulting data would allow us to address several issues of relevance to the development of an optimal TP. First, the data would result in an unprecedented level of understanding of the germplasm within the collection by

quantifying the genetic relationships of samples, helping to identify the major groupings of species (and hybrids) within the complex and still not fully resolved *Salix* genus. A major output would be a better understanding of the number of samples that we have belonging to each cluster. The planned analysis would also identify duplicate accessions that may be named differently but are in fact clones. Our small-scale past studies have indicated that this had happened historically as useful genotypes have been passed between growers and collections. Furthermore, willow propagates vegetatively in nature so the same genotype might have been collected multiple times. Identifying repeated genotypes would allow us to avoid redundancy in the TP that would occur if the same, but differently named, genotype was represented multiple times. The molecular analysis would also provide insight into the likely positioning of many accessions that have been difficult to place accurately within the genus – a common problem due to the difficulties of traditional taxonomic approaches with *Salix*. Similarly, the analysis will help us pick up accessions that have been misidentified previously.

3.2. Methodology

As a first step we compiled a list of available germplasm that may be of interest to future breeding efforts. A review of our NGS data generated previously for other projects identified useful data files that could be used in this study. This included a valuable set of GBS data for several hundred of our *Salix viminalis* accessions. This species is a particularly useful in breeding due to its potential for high biomass and suitable growth form. Given the large amount of useful data already available, we decided to use the same restriction enzymes to generate any new data, simplifying integration of both old and new data sets in future analysis.

Existing DNA stocks were available and used for a minority of samples requiring analysis here. Where DNA was not available, fresh leaf material was collected from plants growing in existing field trials and frozen prior to DNA extraction. Sample collection was done as early as possible after the project started to minimise issues arising from pest and pathogen damage/contamination. Also, sampling early in the season can help ensure successful DNA extractions as high levels of sometimes problematic secondary compounds are not yet present. For DNA extraction, we used DNeasy Plant 96 kits (Qiagen) according to the manufacturer's protocol. Resulting DNA quality and quantity was analysed by Nanodrop and Qubit assays.

To generate GBS data we used a two-enzyme method described by Poland *et. al.* (doi.org/10.1371/journal.pone.0032253) with enzymes *Nsi*I and *Msp*I. A semi-automated protocol where the majority of library preparation was performed on a Zephyr liquid handling robot was also developed. Resulting barcoded libraries were quantified, pooled and sequenced on an Ion Torrent Proton sequencer with sample pooled for runs accounting for differences in ploidy (where known).

After initial exploration of different bioinformatic approaches for handling and analysis of the GBS data we settled on a pipeline involving the following five stages. First, GBS sequencing adapters were trimmed using the tool *Trimmomatic* with the following parameters (minimum adapter sequence match to read = 10; minimum read length = 80). For the subsequent SNP discovery stage, we used the TASSEL v2 GBS pipeline, aligning reads to the publicly available *Salix purpurea* genome

(https://phytozome-next.jgi.doe.gov/info/Spurpurea_v5_1). After production SNP calling using TASSEL's ProductionSNPCallerPluginV2, resulting VCF files were reformatted for downstream analysis using PLINK. First, VCF data was converted into the transposed PED (tped) file format using VCFtools (http://vcftools.sourceforge.net/). SNP data in tped format were subjected to filtration in PLINK to remove loci with minimum allele frequency (MAF) < 1%, loci with > 30% missing genotypes and individuals with more than 20% missing genotypes. For diversity analysis, Identity by State (IBS) distance was calculated using PLINK's --distance-matrix flag. IBS distance matrix data was then used to generate a hierarchical cluster object using the *hclust* function in R. Finally, a dendrogram plot was generated using the R package *dendextend*.

3.3. Results

In total, 1297 accessions were included in the data generation phase of the project. For 190 of these, new extractions were not required as we already had sufficient DNA stocks in storage. Of the 1107 samples collected for new extractions, 1013 (91%) yielded DNA at levels required for downstream analysis. This success rate was deemed very good overall and was consistent with what we might have expected from this phase of the work. In our experience, some willow genotypes are consistently problematic in DNA extraction. We also expected a small number of samples might fail due to the quality of the material that could be sampled. Due to the limited duration of the project repeat extractions were not performed for failed samples. However, these extractions will be repeated in Phase 2 when fresh leaf material is available.

The method used for GBS library production performed well and libraries were produced for all suitable DNA extractions. Libraries for over 1,243 samples were sequenced and resulting data analysed and included in diversity analyses. Libraries were run on our NGS sequencer and subjected to quantification and pooling phases. Data generation and subsequent inclusion in diversity analysis was completed by the end of January 2022.

The bioinformatic analysis pipeline developed and applied to the data analysis was successful in that it allowed us to assess genetic diversity in a large number of samples that spanned the broad range of species and hybrids of potential interest. Methods performed well given the broad diversity presented.

The data generated has provided unprecedented insight into the germplasm available to the willow breeding programme. Figure 2 shows a dendrogram representing results of the diversity analysis and is included below to illustrate the output. Clear clustering of samples according to section and species is evident and is consistent with smaller datasets generated previously using low throughput marker systems. This provides support that the methodologies developed and used here, both for laboratory work and data analysis are performing well.

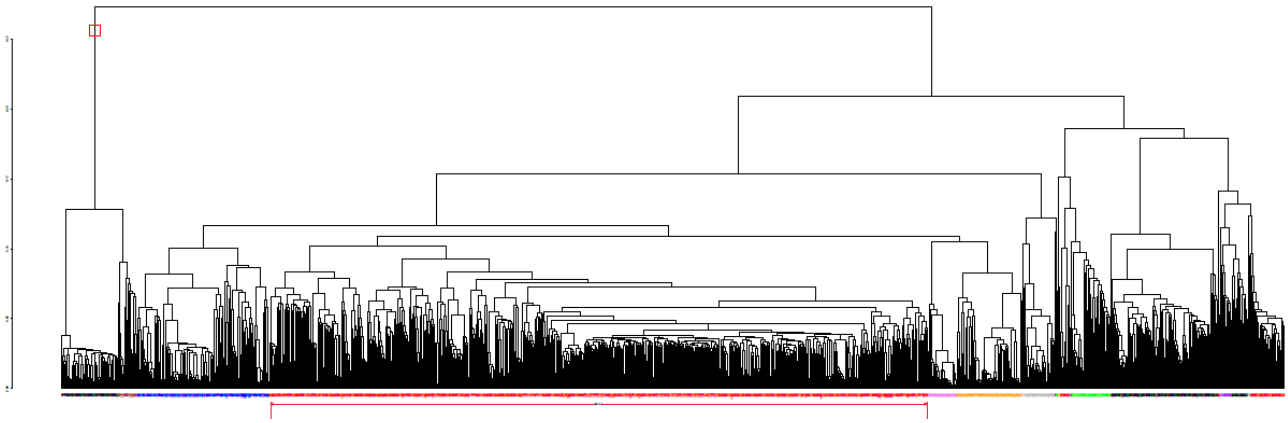


Figure 2: Dendrogram representing results of diversity analysis with TP candidates labelled in red.

The composition of the TP will integrate other breeding information, and data from WP2 but based on the molecular analysis the TP will comprise of genotypes belonging to the cluster of individuals labelled *in red* (Fig. 2). This cluster includes the progenitor species (and hybrids) of many of our elite biomass varieties, e.g. *Salix viminalis*, *S. schwerinii*, *S. dasyclados*, *S. caprea a*s well as species of interest such as; *S. aegyptiaca, S. caprea, S. cinerea, S. scouleriana, S. udensis.* Species within this cluster with no yield potential and no disease resistance such as; *S. aurita, S. atrocinerea, S. myricoides, S. myrsinifolia,* and *S. phylicifolia w*ill be excluded. The TP will be augmented with wider diversity representing key sources of rust resistance as well as elite SRCw varieties so G-by-E interactions can be further dissected.

3.4. Conclusions from molecular analysis

The molecular part of the project has achieved what we proposed in that DNA-based technologies have been used successfully to screen the majority of germplasm samples of interest in willow breeding. The molecular approaches adopted and developed have proven successful at generating data of sufficient quantity and quality for reliable analysis and the bioinformatic pipelines developed have allowed us to achieve the project objectives. Although the marker work here was only planned as an initial screen, the sequencing work has informed future sequencing strategies where a greater number of markers will likely be required. The work done here has also given us a much better understanding of the likely number of markers and data quality we can expect at different sequencing depths, something that is not straightforward to predict for restriction based GBS approaches across the diverse samples assayed here.

This phase of the project has not only delivered an improved fundamental understanding of the *Salix* genus by quantifying the relationships between different willow species but has also delivered the valuable practical and unbiased information to support the selection of the final TP. Specifically, this includes information on genotype redundancy, the true numbers of sample in different diversity groupings, likely misidentified accessions and has provided likely assignments of cryptic samples to species and hybrid groups. The dataset will also prove useful for some future analyses beyond the scope of the current project, e.g. analysis of pedigrees for some elite biomass genotypes.

Figure 3. presents diversity groups of contemporary, near market and outdated SRCw willow varieties. The analysis has the potential to ensure an optimal diversity mix is deployed in commercial SRCw plantations. The diversity analysis in combination with yield data from WP2, demonstrates a method for growers and advisors to select a 6-way varietal mixture ensuring their plantation has both genuine diversity and has optimal yield potential. Picking a six-way mix with one variety from 1a or 1b, 2a or 2b, 2c or 2d, 3, 4 & 5 could prove a simple way to demystify varietal selection.

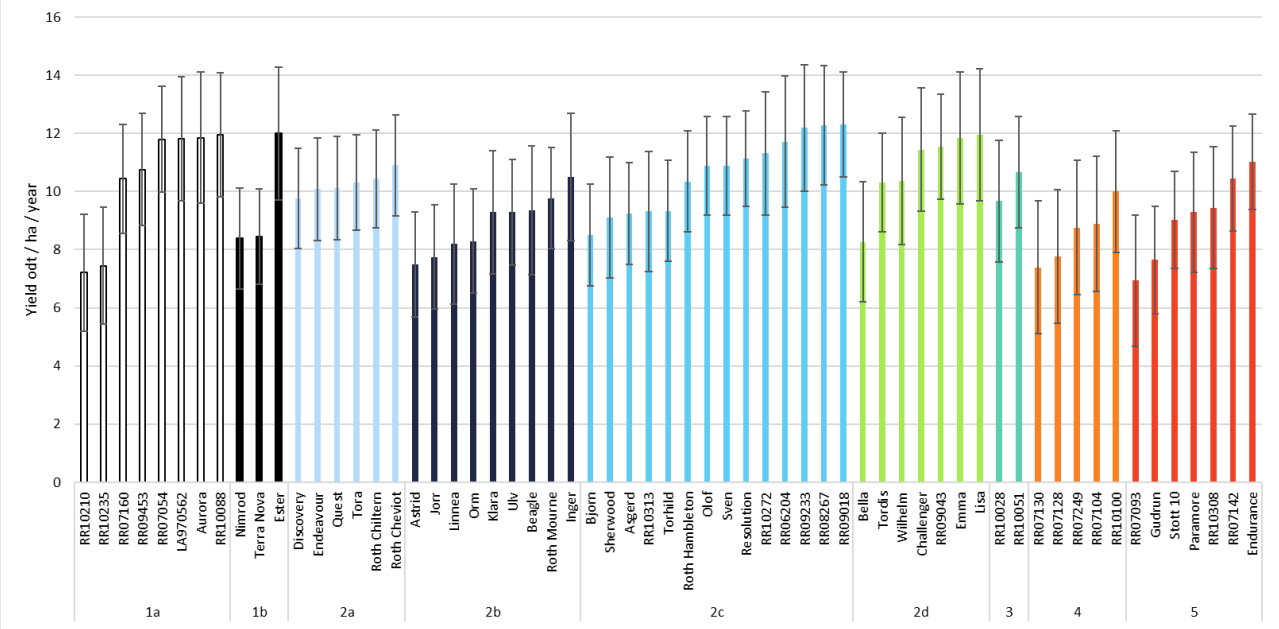


Figure 3: Yields of SRCw varieties and near market genotypes when split into diversity groups.

The molecular analysis has made a valuable contribution to informing the final selection of genotypes for both future planned GS approaches and immediate deployment of current varieties.

4. Objective three. A detailed plan for deployment of GS in willow

4.1. Where to grow the Training Population (TP)

The decision on where to plant out the TP had two components; the growing environment, and the ability to manage the site and conduct phenotypic measurements. It was not possible to consider the Lot 2 Demonstrator sites as locations for our TP as there were three competing bids each proposing to use different sites. The TP is not considered to be of great value as a demonstration to potential growers. Greater value was seen in RRes supplying the best varieties from the breeding programme to the Lot 2 Demonstrator. This would be of immediate relevance to growers and highlight the value of the output of the breeding programme.

Quality phenotyping of the TP is crucial to the success of the GS models and needs to be done at specific time points in the year. This makes it impractical to conduct at multiple sites through one team based at RRes. It was not felt that our contacts in the private sector had the skills base to conduct the intensity of measurement required. Therefore, the public sector presented the best option to host trials of the TP, with just the one exception.

4.1.1. The sites selected

The sites were selected to maximise the probability of phenotyping the TP under the following environmental conditions; cool temperatures (CT), long summer daylight hours (LSD), high temperatures (HT), high humidity (HH), drought (D) and flooding inundation (FI).

To that end we selected the following as subcontractors; Scotland's Rural University College (SRUC) at Aberdeen (CT & LSD), RRes Woburn Farm (HT & D), Somerset Willow Growers Ltd, Somerset Levels (FI), Agri-food and Biosciences Institute Northern Ireland (AFBINI) Loughgall (HH) and Newcastle University Farms (Control site). All except SRUC have direct willow cultivation experience and track record.

The principal pests of willow (grazing mammals and insects) are sporadic and unpredictable. It was not possible to choose a site specifically to increase the probability of pest damage. The interaction between willow genotype and pest will be quantified at any of the above sites where an outbreak occurs.

4.2. How to arrange the planting

The TP will be vegetatively reproduced via stem cuttings a limitation will be the number of stems available to make cuttings. Given five sites to plant and the need for 4 statistical replicates at each the plot size needs to be small. Small plots in our germplasm collection have been assessed and are capable of producing >120 cuttings (the number needed for 5 TPs).

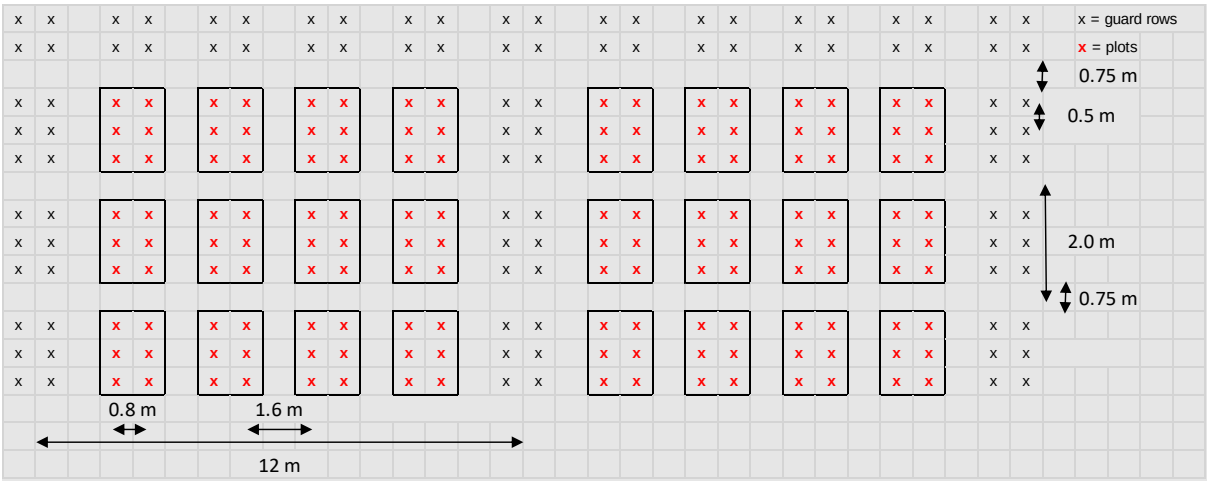


Figure 4: The proposed spatial arrangement of the genotypes for the Training Population. Each x *represents a single willow cutting. Individual plots are groups of 2 x 3 willows.*

Figure 4 shows the commonly used double row system of alternating 0.8 and 1.6 m spacing which allows machinery to drive straddling two narrow rows. Within the row cuttings are planted 0.5 m apart giving a total of 16,667 cuttings per ha. In this case the within row spacing will be increased to 0.75 m between plots making a clear demarcation between plots for those conducting phenotypic measurements. Every 4 double rows there is a tramline planted with a standard variety.

The genomics data identified approximately 600 distinct genotypes of value in the TP. Because each replicate block will be large, spatial controls will be included in each, and a sub-block structure incorporated to aid precise comparison of genotypes. The total number of plots will be approximately 2,800, surrounded by a double row of a standard variety to prevent edge effects on the outer plots.

4.3. Phenotyping

The yield of dry wood is the main criterion on which willow varieties are judged. Behind this primary trait are several components of yield such as the height, diameter and number of stems per stool. Diameter can influence the quality of the wood via a lower bark percentage. All three components influence the ability of current harvesting technologies to cut the crop. The combination of stem diameter and stem number measured at 1 m above the soil surface can be used to make a non-destructive estimate of yield. Moisture content is routinely measured to calculate a dry wood yield for comparison of varieties and because buyers usually pay on a weight of dry wood basis.

Bud burst signals the beginning and senescence the end of the growing season. The timing of bud burst needs to be late enough in spring to avoid the risk of frost damage and early enough not to waste incident sunlight. Senescence should occur before severe frost which can damage stems that have not fully senesced.

The fungal disease with the greatest potential to limit dry wood yield is a leaf rust caused by *Melampsora* spp. Less common, and with less data on yield effects, is a stem infecting rust. Mildew may also occasionally infect willow leaves.

The most reported pests of willow when grown as a crop are the invertebrates; aphids, beetles and sawflies and vertebrates such as deer, hares and rabbits. Occurrence of these pests is sporadic and largely unpredictable.

The phenotyping protocols to be deployed are listed in Annex 2.

5. Project Management

5.1. The Team

The core team will be the team that prepared the plan for implementing GS in willow breeding. Shield (IS) will act as Principal Investigator, Macalpine (WM) is the willow breeder, Hanley (SH) leads the genomic research and Mead (AM) the statistical approaches. Cerezo-Medina (SCM) will add micropropagation expertise. A new field technician, a lab technician and a specialist bio-informatician with GS experience will be recruited. These may be internal or external appointments depending upon the response to the recruitment process.

Professor Angela Karp (AK) will chair the Advisory Board. Dr Gancho Slavov has kindly indicated his willingness to continue as a member of the Advisory Board and increase his contribution in advising on statistical genomics.

Four subcontractors will be added to the team, AFBINI, Newcastle University, Somerset Willow Growers Ltd and SRUC to provide additional sites for the TPs.

5.2. Quality assurance

RRes has an excellent track record of delivering government -funded projects on willow, including leveraging additional funding. The project team are all highly trained in the skills required and the quality of their own performance is assessed by annual appraisal. This Phase 1 project leaves an organisational legacy that is valuable to carry forward to a Phase 2 project. The resources that are to be deployed to achieve the phenotyping, genotyping and statistical analysis in Phase 2 are substantial and must be optimally managed for greatest effect.

A quality assurance check will be performed annually during Phase 2 by the RRes head of Quality Risk and Assurance. A running log of concerns and/or problems observed by members of the Project Team, together with an account of lessons learned from the project that could be applied to future projects will be maintained as part of the quality assurance procedures.

5.3. Value for money

The investment made by BEIS into this programme will have lasting benefits that long outlive the project duration. To provide a relevant example from the past, funding by the DTI in 1999 enabled RRes to establish a uniquely large willow population (K8) for genetic mapping and identification of quantitative trait loci. This population provided the foundation needed for the breeding of new varieties and the identification of markers for yield and rust resistance that are now routinely used in MAS. MAS selected willows, rank 1, 3, 4, 16 & 21 in Fig 1, this demonstrates the benefits of deploying advanced breeding technologies. K8 places our UK research on the global science stage, with outputs including genetic maps aligned to poplar and a sequenced genome. The population still assists in discovery science leading

to the identification of high value products of medical significance. Willow breeding is not supported by global corporates. The market also cannot support the development of such vital genetic resources, at a time when investment is needed for expansion. This is exactly the time and manner in which government intervention is needed to invest in the foundations needed to secure the genetic tool kit that will generate varieties tailored for the bioenergy industry.

5.4. Project delivery and monitoring plan

The Phase 2 project proposal presented to BEIS will be used as the primary document (Business Case) that explains the purpose of the project and the anticipated benefits. This will be accompanied by the Gantt Chart (Annex 3) indicating how the project is broken down into tasks and the interrelationship between these tasks. It indicates where the project team (above) contribute towards the project delivery, together with key milestones and deliverables. At the project onset, the PI will use this Gantt Chart to discuss and develop, with the Project Team, an interactive detailed RAG (Red, Amber, Green) Progress Monitoring Plan using publicly available software. The Gantt Chart and Progress Monitoring Plan, together with all detailed records relating to the project, will be kept throughout the project timeframe, and for a period of ten years afterwards, in a protected shared drive, with access requiring permissions from members of the Project Team. The PI will check that milestones are on track, and deliverables met on time, by reviewing these at monthly catch-up meetings with the project team leaders. The Project PI will report on progress to the Project Advisory Board. The overall metric that the Project Team will use to assess success will be yield, with genetic gains resulting in yield increases from the current 15 to a target of 20- odt $ha^{-1}yr^{-1}$ as the programme target. Yield assessments will be monitored annually, taken routinely as part of the trait assessments via non-destructive methods and as harvested biomass (destructive methods) after rotation cycles are complete. We will also monitor take- up by nurseries and areas planted with our varieties after Phase 2 of the programme has ended. Published yields of existing varieties, and any varieties that might be released from Sweden or the US programmes, will be used to benchmark the success of our varieties. However, two things should be borne in mind; our work is based upon small plots of single genotypes, field scale crops, planted as mixtures of genotypes may perform slightly differently. Our aim is to target environments that are often challenging for crop production so like for like comparisons may not be possible to make accurately.

Social Value Key Performance Indicators (SV KPI); Delivering jobs for UK citizens, supporting regional and rural economies and bolstering the UK’s reputation as a pioneer in green technologies will be assessed quarterly. A standing agenda item in meetings will discuss the delivery, progress, and improvement of SV KPIs. If needed, remedial actions to rectify low scores will be actioned by the project team. Social value outcomes will be presented in the Phase 2 final report.

5.4.1. Reporting plans

The Gantt Chart (Annex 3) identifies specific deliverables associated with each task and which will be associated with invoices to BEIS. In addition to those deliverables a written report will be presented to BEIS each quarter outlining activities during the period. The reports will also be made available to the Advisory Board. In addition to the milestones within Annex 3, RRes will:

- Meet with their Monitoring Officer quarterly to discuss project progress and highlight successes and exceptions, issues and risks. Immediate communication will occur if specific risks or issues are identified.
- Submit a project progress report every quarter covering:
    - progress against the project delivery plan and project milestones, upcoming work over the next quarter.
    - financial information (including budget spend and budget forecast).
    - updated risk and issue registers (including where risk ratings have changed, or new risks/issues have been identified).
    - any key lessons learnt during delivery.
    - progress against relevant programme KPIs.
- Speak to sub-contractors monthly and visit sub-contractors TPs at least annually.
- Engage with the successful Lot 2 demonstrator and any BEIS dissemination events.

5.5 Oversight and governance procedures

The framework and governance procedures for the project that we will adopt will be based on PRINCE2 (PRojects IN Controlled Environments) methodology. The project will appoint a Project Advisory Board, chaired by Prof. Karp, and comprise of two external experts. The PI will attend all Advisory Board meetings which will be scheduled for the key time points in the project timeframe and conducted online. In addition to reporting to the Project Advisory Board, reports on progress will be presented to the RRes Board, and/or subgroup of this Board as part of their quarterly meetings. RRes consider this to be a project of prime importance and will seek assurance that it will be delivered to the highest standards.

5.6. Financial controls

As the sole applicant RRes also acknowledges that it has overarching responsibilities for financial controls, monitoring and reporting, co-ordination, and ensuring the timely delivery of project outcomes. All financial costings have been approved though a rigorous system at RRes that is subject to annual auditing. RRes will ensure that cost accounting practices are applied in a manner consistent to how all projects are managed. A separate account code will be used for the project with sub-codes to identify specific cost areas. The specific project number assigned will be used to track all income, expenditure and budget on a transaction by- transaction basis under the Business World (Unit 4) system utilised at RRes. The authorisation of expenditure for each project will be restricted to the key staff with authorisation of sums >£10k required by the Science Director responsible for their strategic area in the institute departmental structure. The Principal Investigator will have access to

budget reports though Unit 4 at any time but will also attend budget meetings with an assigned member of the Finance department before presenting reports to the Advisory Board. RRes has a procurement manager, who will ensure that all purchases follow correct procedures. These are reviewed on a regular basis and the Finance Department exercises control over these suppliers.

5.7. Risk assessment and contingency planning

Dr Shield will manage exposure to risk using a risk and issue tracking system. A table of possible risks to successful project delivery are outlined in Annex 4. On project initiation, the Project Team will evaluate the risk management plan and agree actions to appropriately reduce the impacts of risks identified. Each risk identified will incorporate an actual or planned response and be designated an appropriate owner to monitor the threat. Issues that have affected the delivery of the project will be identified, assessed and managed at the appropriate level by the project team, the Advisory Board or escalated further.

5.7.1 Covid-19

RRes has a central Covid Response Team (CRT) chaired by the Director CEO, who have developed a detailed business continuity plan specifically for Covid-19 and have successfully managed various stages of the pandemic. The framework of controls remains available and can be implemented rapidly, should the need arise. A recent audit carried out during an unplanned visit of the Health and Safety Executive was highly supportive and found no changes needed to the institute's management of Covid-19.

6. Commercialisation

6.1. UK and international market

Breeding of willows for biomass production is a relatively recent endeavour, starting in the 1980s. Willow is now grown commercially in many areas of the world including the UK, Eastern and Northern Europe, North America, India and China, all of which offer market opportunities.

The scale at which SRCw will be scaled up within the UK, taking pressures on land-use and concerns over conflicts with food production into account, is currently being considered by the Governments Biomass Strategy.

The Climate Change Committee (CCC)'s 6th Carbon Budget report highlighted the significant potential for perennial energy crops (SRC and miscanthus) and short rotation forestry (SRF) to contribute towards carbon budget targets by increasing soil and biomass carbon stocks while also delivering other ecosystem benefits. The CCC suggested that up to 708,000 ha of land should be dedicated to perennial energy crop production by 2050. The November 2021 Biomass Policy Statement indicated that the Biomass Strategy, to be published in 2022, will establish the amount of land that could be used in the UK for perennial energy crops. As detailed in 6.5 establishment of 3,000 ha / year of RRes varieties is sufficient to sustain a continued breeding effort. The potential market size is currently difficult to determine as several barriers to uptake are still to be overcome.

6.2. Promotion of the innovation

Maximising biomass production through optimal deployment of current varieties across diverse UK environments can now be guided following our initial Phase 1 analysis, but performance data from Phase 2 field trials will be of significantly greater value as emerging, new elite varieties are included. We anticipate that ~5 new varieties which achieve yields of 15 – 20 odt ha^{-1} yr^{-1} will be available for release in 2027. Advanced, genomics -guided breeding to specifically target the required range of UK growing environments will be delivered beyond Phase 2, as a legacy benefit from the early 2030s onwards. GS technology is not itself commercially exploited - the value of this innovation lies in our capability to deploy GS in willow and in the quality, range and performance of the environment- tailored varieties arising from implementation. The project will also support security of supply and capability for agile responses to emerging breeding requirements. This “high- tech” breeding approach, together with robust, independent and freely available knowledge on performance in different environments, will be exploited in marketing.

The review of previous trials (WP2) revealed useful insights for matching varieties to environments. As this is of immediate use to those supplying cuttings and guiding planting decisions, we will disseminate through the leading actors in the field; Willow Energy, Crops4Energy and Energy Crops Consultancy, the Lot 1: Perennial Energy Crops Decision Support System (PEC-DSS): Envirocrops (if funded) and the Lot 2 Demonstrator as well as our own channels. In addition, RRes is a partner in the UKRI funded Greenhouse Gas Removal Demonstrator project on Perennial Energy Crops where data from our review will be valuable.

6.3. Interactions with suitable partners and future plans

Our plan will deliver elite high yielding resilient SRCw varieties suitable for diverse UK environments, together with knowledge for growers that will enable them to choose optimal varieties for their environments. Access to state- of- the -art technology and elite genetic lines, together with innovative multiplication methods form a strong basis for developing our grower networks, ensuring the UK is in the strongest competitive position to scale up should demand for willow increase once market conditions change. We are uniquely positioned to achieve this by building on the excellence of our science and extensive, well-established linkages throughout the value chain. Interactions with growers and advisors will be achieved through multiple routes. We will continue to work with our established licensed cutting producers, seek new multipliers, identify interested new actors through the BFI Lot 2 Demonstrator and engagement with biomass end--users.

6.4. Maximising impact and distribution

We anticipate protecting innovations using a similar model to the current RRes Willow Breeding Programme, which has released five varieties. The model protects output with Plant Breeder’s Rights (PBR) granted across Europe via the Community Plant Variety Office (CPVO). RRes willow varieties are in trials in Canada and the US where plant breeder’s IP is more directly aligned to Patenting. We will continue discussions domestically and internationally with those wishing to exploit our programme’s outputs.

A non-exclusive licence agreement has been developed for RRes bred varieties. RRes charges its licenced nurseries a royalty payment of £0.01 on every unit of multiplication (usually a stem cutting). Currently Willow Energy based in Cumbria supply RRes varieties for the UK market and Energy Crop Consultancy have indicated and interest to in cutting production. In Northern Ireland our varieties are sourced from nurseries in the Republic of Ireland.

6.5. Financial growth plan for 5 years

As stated above, we plan to release improved varieties within five years of the project end, raising revenue from the royalties on non-exclusive licenses. The perennial nature of the crop restricts royalty income to the initial planting year, unlike annual crops generating revenue for breeders each year. Our current royalty charge of £0.01 per cutting equates to £168 per ha at the usual planting density of 16,667 cuttings per ha. As a comparison, farm saved wheat seed attracts a royalty charge of £10.34 per ha (royalty rates on certified seed are confidential). As such, our royalty payment from one willow planting is approximately equivalent to 16 years royalty income from annual wheat seed sales. Plant breeding is a high-cost activity and sustaining a programme via royalty income will require market expansion.

Alternative income generating schemes have been explored. We estimate that a conventional breeding programme without the benefit of modern molecular genetics costs ~£200k per annum to operate at RRes. The rate of genetic gain is much improved with the addition of molecular genetics underpinning the breeding programme. We have demonstrated that MAS can reduce the selection cycle by 3 years (i.e. from 12 to 9 years). Successful implementation of GS has the potential to cut another 3 years from the cycle. To achieve that rate of progress will require the method for rapid multiplication of planting material that is also to be developed in Phase 2 to build up the stocks required. Overall utilising molecular genetics raises costs to approximately £0.5M p.a. However, without molecular genetics the breeding programme would lose speed and competitiveness with others in the marketplace, and there is a risk of plantations failing due to suboptimal varieties succumbing to biotic or abiotic stress. An area of 3,000 ha of RRes bred SRCw varieties would need to be planted each year to maintain our proposed activities.

Signals for an upturn in market conditions mean that this could be a viable model. The upcoming Biomass Strategy could provide the policy certainty growers and end-users require. The high cost of gas and other common energy sources drives the search for alternatives, stimulating demand for biomass and driving prices up. In the UK the joint statement by Drax Group and the National Farmer’s Union regarding biomass and bioenergy from farms promises some stimulation of markets. The development of Bioenergy with Carbon Capture and Storage (BECCs) is foreseen to increase demand for willow biomass but BECCs installations are not in operation yet in the UK. The activity of the Demonstrator (Lot 2) and innovations funded under Lot 1 Phase 1of the BEIS BFI programme show promise to improve market conditions. It will take time to ensure that sufficient stocks of elite robust and resilient varieties are available to the industry. Given the current uncertainties over the scale of the market, our commercial plan is realistic and conservative, but also scalable. It will ensure the UK is positioned to meet the rising demand for biomass in the near future.

Basic Trait Assessment- Aphids

OBJECTIVES

To score Willow genotypes for their resistance/susceptibility to the Giant Willow Aphid, *Tuberolachnus salicis*, and the Black Willow Aphid, *Pterocomma salicis*.

OPERATORS OF THE PROCEDURE

Suitably trained personnel involved with the SRC Willow Breeding Project.

PROCEDURE

Stencils will be used to score the aphids according to the area of their colony.

Score	Aphids
0	0 cm^2
1	< 1 cm^2
2	>1 <8 cm^2
3	>8 <27 cm^2
4	>27 <64 cm^2
5	>64 <125 cm^2
6	>125 <216 cm^2

Each oval stencil will be made at the maximum of each parameter. The score will be given when the aphid colony fits inside the stencil category.

As well as giving each plant an aphid score, it should be noted if the Black Willow Aphid, *Pterocomma salicis* is present. If a score is given and no note is given, it will be presumed that the more common Giant Willow Aphid, *Tuberolachnus salicis* is present.

The assessment will be carried out twice a growing season, towards the beginning of the season (around July) when they start to become a problem, and at the end of the season, before leaf senescence (early September).

RECORD KEEPING

All data will be stored according to Data Storage for the Willow Breeding Project.

PROCEDURE REVIEW

This procedure will be reviewed as necessary.

Basic Trait Assessment- Beetles (Chrysomelids)

OBJECTIVES

To score Willow genotypes for their resistance/susceptibility to beetles.

OPERATORS OF THE PROCEDURE

Suitably trained personnel involved with the SRC Willow Breeding Project.

PROCEDURE

Damage done to the leaves by willow beetles will be scored based on a visual assessment.

Score	
0	No beetle feeding damage
1	1%
2	5%
3	10%
4	25%
5	50%
6	Almost completely covered

The assessment will be carried out twice a growing season, towards the beginning of the season (around July) when they start to become a problem, and at the end of the season, before leaf senescence (early September).

RECORD KEEPING

All data will be stored according to Data Storage for the Willow Breeding Project.

PROCEDURE REVIEW

This procedure will be reviewed as necessary.

Basic Trait Assessment- Bud Flush Timing

OBJECTIVES

To provide details of when genotypes buds flush. Early bud flush can allow plants to be injured by frost.

OPERATORS OF THE PROCEDURE

Suitably trained personnel involved with the SRC Willow Breeding Project.

PROCEDURE

Bud flush will be recorded by tri-weekly inspections. Bud flush will be recorded when the first leaf unfolds from the bud. It will be expressed as the number of days to bud flush starting from the first of January. The buds 10 cm below the terminal bud should be observed.

RECORD KEEPING

All data will be stored according to Data Storage for the Willow Breeding Project.

PROCEDURE REVIEW

This procedure will be reviewed as necessary.

Basic Trait Assessment- Branching

OBJECTIVES

To score how 'branched' Willow genotypes are. Branches are likely to be lost during mechanical harvest or if not lost are likely to result in variable chip size and are therefore undesirable as a trait for biomass clones.

OPERATORS OF THE PROCEDURE

Suitably trained personnel involved with the SRC Willow Breeding Project.

PROCEDURE

Genotypes will be given a score according to the following scale.

Score	Branchiness
0	No branching
1	Small auxiliary branches
2	Some branching, but mainly higher up the rod.
3	No true main rod, substantial branching.

RECORD KEEPING

All data will be stored according to Data Storage for the Willow Breeding Project.

PROCEDURE REVIEW

This procedure will be reviewed as necessary.

Basic Trait Assessment- Establishment

OBJECTIVES

To provide details on whole trial and individual genotype establishment.

OPERATORS OF THE PROCEDURE

Suitably trained personnel involved with the SRC Willow Breeding Project.

PROCEDURE

After their spring planting in the trial's first year, when plants are fully emerged the number of cuttings that have established will be counted. This procedure should be carried out in early July, before other trial assessments are carried out.

RECORD KEEPING

All data will be stored according to Data Storage for the Willow Breeding Project.

PROCEDURE REVIEW

This procedure will be reviewed as necessary.

.

Basic Trait Assessment- Growth Habit

OBJECTIVES

To note how suitable the genotype's characteristics are to biomass production and harvesting.

OPERATORS OF THE PROCEDURE

Suitably trained personnel involved with the SRC Willow Breeding Project.

PROCEDURE

Non quantitative, comments such as erect or sprawling will be used to describe the growth habit. It will also be noted if a clone has any trait that may be of interest other than that of biomass production, e.g. to horticulturists, twisted/distorted or coloured stems.

RECORD KEEPING

All data will be stored according to Data Storage for the Willow Breeding Project.

PROCEDURE REVIEW

This procedure will be reviewed as necessary.

Basic Trait Assessment

OBJECTIVES

To score Willow genotypes for their resistance/susceptibility leaf galls

OPERATORS OF THE PROCEDURE

Suitably trained personnel involved with the SRC Willow Breeding Project.

PROCEDURE

Damage done to the leaves by willow leaf galls will be scored based on a visual assessment.

Score	Gall
0	No gall present
1	Discrete galls along the margins of attacked leaves
2	Plants with fused galls
3	Plants with fused galls and twisted leaves

Galls are not anticipated to be as large a problem as aphids, beetles, and rust. Because of this, they will be scored when noted whilst scoring for 'the big three', aphids, beetles and rust. There will be no specific outing to the field to score solely for galls.

RECORD KEEPING

All data will be stored according to Data Storage for the Willow Breeding Project.

PROCEDURE REVIEW

This procedure will be reviewed as necessary.

Basic Trait Assessment- Mildew

OBJECTIVES

To note which willow genotypes are susceptible to mildew.

OPERATORS OF THE PROCEDURE

Suitably trained personnel involved with the SRC Willow Breeding Project.

PROCEDURE

Mildew on the leaves will be noted if present.

Mildew is not anticipated to be as large a problem as aphids, beetles, and rust. Because of this, they will be scored when noted whilst scoring for 'the big three', aphids, beetles and rust. There will be no specific outing to the field to score solely for mildew.

RECORD KEEPING

All data will be stored according to Data Storage for the Willow Breeding Project.

PROCEDURE REVIEW

This procedure will be reviewed as necessary.

Basic Trait Assessment- Shoot height

OBJECTIVES

To measure the height of willow shoots, this combined with the diameter measurements and shoot numbers will allow a non destructive yield estimate to be made.

OPERATORS OF THE PROCEDURE

Suitably trained personnel involved with the SRC Willow Breeding Project and the BBSSRC Bio-energy Centre Perennial Crops Programme.

PROCEDURE

All shoots will be measured from the four sample stools in an Obs 1 (see plot diagram below). This should be carried out in the observation trials third year (two year rods on three year roots). Present practice is to use a Senshin telescopic measuring pole readable at 1m from ground level.

An average shoot height will be calculated and used with the digital calliper data (shoot diameter and number) to calculate yield estimates.

RECORD KEEPING

All data will be stored according to Data Storage for the Willow Breeding Project.

PROCEDURE REVIEW

This procedure will be reviewed as necessary.

Senshin telescopic measuring pole supplied by;
L & M Survey Supplies
40 Westgarth Place
College Milton North
East Kilbride
G74 5NT

Basic Trait Assessment- Rust Melampsora **B**

OBJECTIVES

To score Willow genotypes for their resistance/susceptibility to rust, *Melampsora*.

OPERATORS OF THE PROCEDURE

Suitably trained personnel involved with the SRC Willow Breeding Project.

PROCEDURE

This system is designed to give a general indication of rust infections on various willow genotypes. The undersides of the leaves will be visually assessed for the presence of rust pustules, or uredinia.

Numerical scale	Description	Category
0	No rust can be detected	None
1	Leaves occasionally (usually less than 5% of the total leaves in a plant) bear a few conspicuous uredinia (often with necrosis) or bear more, but barely recognisable small uredinia	Slight
2	Leaves often (usually between 5-25% leaves) bear a few conspicuous uredinia (sometimes with necrosis) or bear more, but barely recognisable rust uredinia,	
3	Leaves frequently (usually between 25-50% leaves) bear rust pustules, up to an average of 1-2 % leaf area covered by uredinia or telia on infected leaves.	Moderate
4	Majority leaves bear rust pustules, up to an average of 5% leaf area covered by uredinia or telia, and/or plants show slight defoliation due to rust (telia can be found on fallen leaves)	
5	Most leaves bear numerous pustules, 5-25% leaf area covered by uredinia or telia, and/or plants show obvious defoliation due to rust (telia easily recognisable on fallen leaves)	Severe
6	Most leaves with high concentration of rust pustules, 25% or more of leaf area covered by uredinia or telia, and/or plants show severe defoliation due to rust (abundant telia on fallen leaves)	

The stem infections can be noted by adding a symbol (* or ^) to the numerical scales.

The assessment will be carried out twice a growing season, towards the beginning of the season (around July) when they start to become a problem, and at the end of the season, before leaf senescence (early September).

RECORD KEEPING

All data will be stored according to Data Storage for the Willow Breeding Project.

PROCEDURE REVIEW

This procedure will be reviewed as necessary.

Basic Trait Assessment- Senescence Timing

OBJECTIVES

To score the senescence of willows

OPERATORS OF THE PROCEDURE

Suitably trained personnel involved with the SRC Willow Breeding Project.

PROCEDURE

Genotypes will be given a score according to the following scale.

Score	Senescence
0	No Senescence
1	25 % Senesced
2	50% Senesced
3	75% Senesced
4	No leaves

This will be recorded weekly through the autumn.

RECORD KEEPING

All data will be stored according to Data Storage for the Willow Breeding Project.

PROCEDURE REVIEW

This procedure will be reviewed as necessary.

Basic Trait Assessment- Shoots / rods per Stool

OBJECTIVES

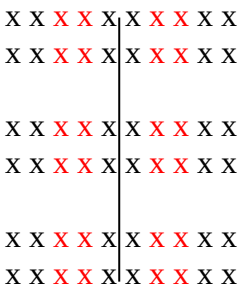
To provide details of how many shoots / rods each stool produces

OPERATORS OF THE PROCEDURE

Suitably trained personnel involved with the SRC Willow Breeding Project.

PROCEDURE

The average number of shoots per cutting will be recorded from each genotype. There are two possible timings; in the autumn of the trials first year, before first year cut back (to assess the ability to shoot without cutback) or later, at the time of the shoot height and diameter measurements (to contribute to a non destructive estimate of yield). Below shows the stools where the shoots / rods should be counted in a typical Obs1 design.



RECORD KEEPING

All data will be stored according to Data Storage for the Willow Breeding Project.

PROCEDURE REVIEW

This procedure will be reviewed as necessary.

Basic Trait Assessment- Vertebrate Grazing/Browsing

OBJECTIVES

To note which willow genotypes are grazed/browsed by vertebrates.

OPERATORS OF THE PROCEDURE

Suitably trained personnel involved with the SRC Willow Breeding Project.

PROCEDURE

Grazing/browsing will be noted if observed.

Early season assessments such as establishment should include an assessment of vertebrate grazing. Otherwise such grazing/browsing is not anticipated to be as large a problem as aphids, beetles and rust. Because of this, they will be scored when noted whilst scoring for 'the big three', aphids, beetles, and rust. There will be no specific outing to the field to score solely for vertebrate grazing/browsing

RECORD KEEPING

All data will be stored according to Data Storage for the Willow Breeding Project.

PROCEDURE REVIEW

This procedure will be reviewed as necessary.

Basic Trait Assessment- Rod Diameter

OBJECTIVES

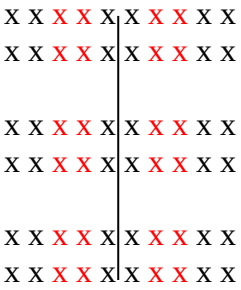
To measure the diameter of SRC willow shoots, this combined with the height measurements and shoot numbers will allow a non destructive yield estimate to be made.

OPERATORS OF THE PROCEDURE

Suitably trained personnel involved with the SRC Willow Breeding Project.

PROCEDURE

All readings will be taken 1 metre from the ground, from the four 'measured' stools from Observation 1 trials. This should be carried out in the observation trials third year (two year rods on three year roots). All shoots on a stool should be measured.



Operating the Callipers

Firstly the calliper's battery should be charged.

The callipers have one 'trigger' button on them. The 'trigger' is used to navigate round the rolling menu. Below is a map of the rolling menu.

On arrival at the trial

New UK site should be selected by pressing the 'trigger.'

At the **date** option, enter the month and year.

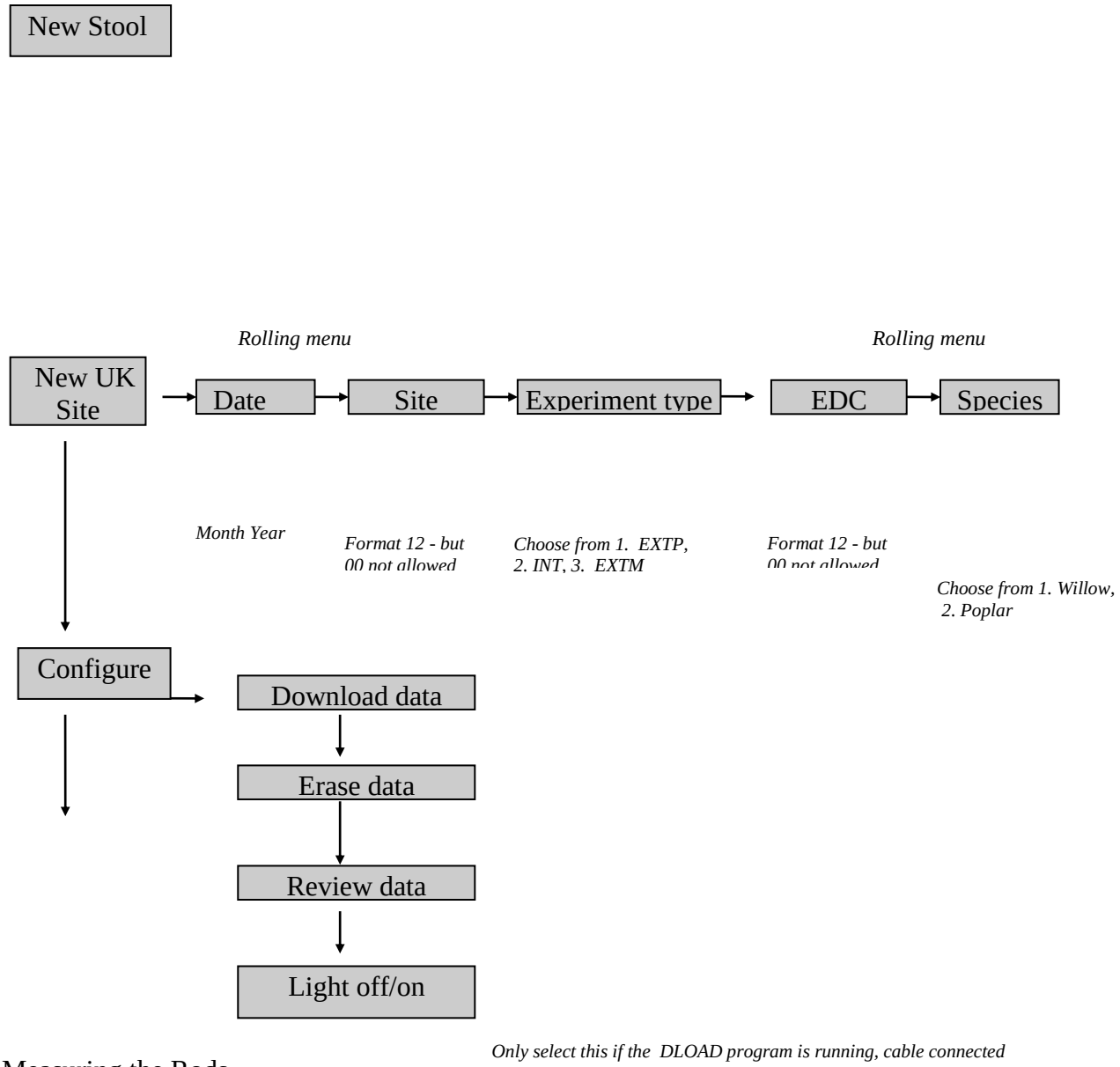
The last two digits of the unique RES trial code should be entered for the **Site number**. (e.g. 64 for the national willow collection (RES code CS/564))

EXTP should be entered for **Experiment type**.

The last two digits of the genotype number should be entered for the **EDC number**

The appropriate **Species** should then be chosen**.**

The values chosen here will now be related to all stool measurements. Stool number will automatically be set to number 1, stem to number 1, so measurements can commence immediately. Every shoot on four out of the ten stools per clone will be measured for Obs 1 trials.



Measuring the Rods

Pass the callipers smoothly over the shoot / rod to record the diameter in mm.
To commence the measurements of each stem of the stool the display will show the site number, the EDC number, the stool number, the stem number, and the measurement in mm(s).
The stem number will increment after each measurement
A beep will sound if there is a problem - eg taking a measurement too quickly

When all the stems of a stool have been measured, enter the menu section of the callipers, and choose NEW STOOL. Measure the stems of this stool, noticing the LCD display of the stool number has been incremented.

When all the stools of one clone have been measured, enter the menu section of the callipers, and choose NEW EDC. Enter the EDC number (last two digits of the genotype number) stool species.

At any point during these stages, data can be downloaded to PC, and then measurements resumed from the point left off. It is advised that downloading is done frequently to minimise potential data loss

RECORD KEEPING

Download of data from the callipers to the PC

Attach the cable from the calipers to Serial Port 1 on the computer. **With the calipers in their shut-down state**, run Command prompt. (Start →All programs→Accessories→Command prompt)

In Command prompt type; D: → cd calipers files → dload **** (*the name you want the data file to be called.

On the callipers which are plugged into the computer click **Configure** and **Download data**. Then click to ok that data has transferred successfully.

If data transfer fails for any reason (which is does very occasionally for some reason) and you are left with the message 'transferring' on the caliper screen, you may need to put a damp finger on the pins of the data transmission port to shut-down the calipers, before retrying.

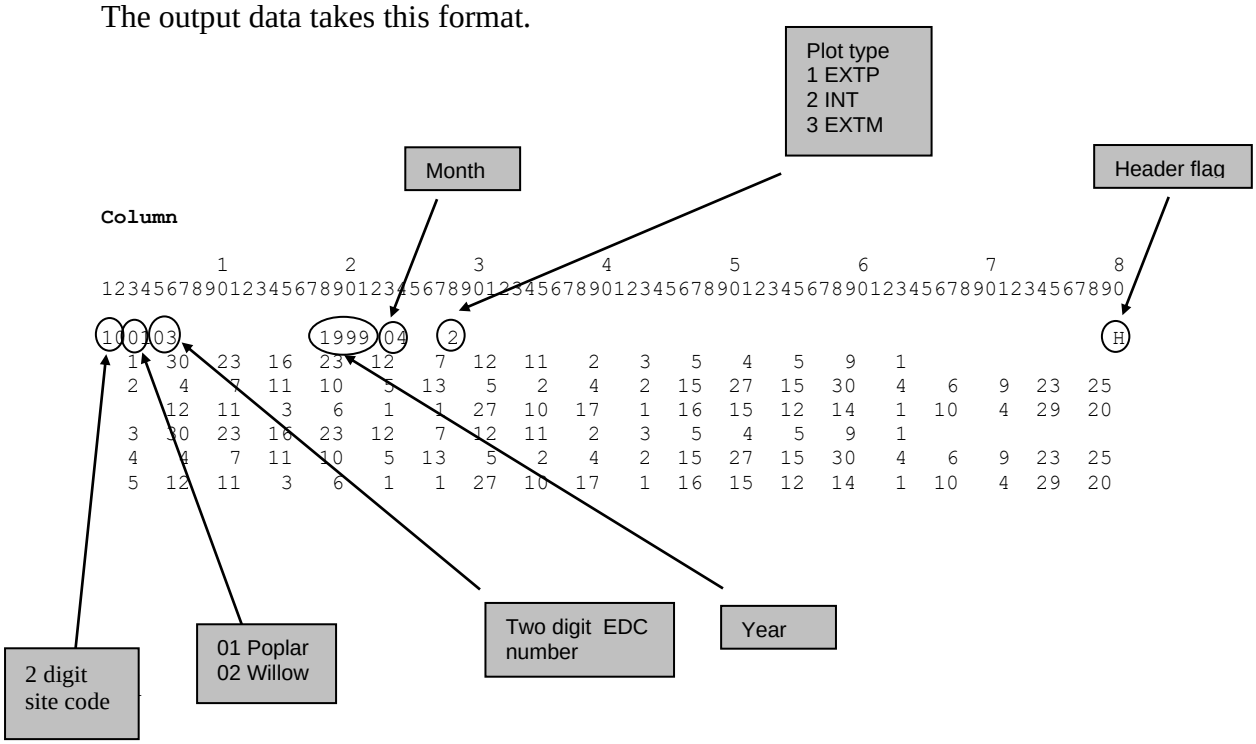
The data will now be in D: calipers files.

To get the data in the desired format in exel the file, **** needs to be opened in notepad (open: all files). This notepad file then needs to be saved as a text Doc. file. This text Doc. file can be opened in excel using the auto wizard which comes up, to provide the desired column/row divides.

After successful data transfer, data can be erased from the callipers. **Configure**, **Erase data.**

Output data format

The output data takes this format.



This data will then be stored according to Data Storage for the Willow Breeding Project.

PROCEDURE REVIEW

This procedure will be reviewed as necessary.

Accelerating Willow Breeding and Deployment (AWBD) Phase 2 GANTT

Task	Description	2022									2023													2024													2025			Staff	Deliverable
		Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar		
1	**Project management**																																						**IS**	
1.1	Monthly meetings																																						IS, SH, WM, SCM	Minutes
1.2	Advisory Board meetings																																						AK & IS	Minutes
1.3	Final report																																						IS, SH, WM, SCM	Report
2	**Planting out the Training Population**																																						**WM**	
2.1	Site visit, GPS map, recommendations for land prep.																																						WM & IS	Field plans
2.2	Land prep, cover crop, marking out of site. Soil sampling.																																						Subcontractors	Images
2.3	Cutting production																																						Field Tech	Images
2.4	Destruction of cover crop, planting.																																						Subcontractors	Images
2.5	Site maintenance. Weed control.																																						Subcontractors	Images
3	**Phenotyping the Training Population for GS**																																						**WM**	
3.1	Introduction to protocols, training																																						WM, Field Tech, Subcontractors	
3.2	Establishment counts & initial vigour score																																						Subcontractors & Field Tech	Data sets
3.3	Pest and disease scoring																																						Subcontractors & Field Tech	Data sets
3.4	First year cutback, incl. yield measurement.																																						Subcontractors & Field Tech	Data sets
3.5	Supporting data, meteorology, soil water monitoring																																						Subcontractors & Field Tech	Data sets
4	**Genomics and molecular breeding**																																						**SH**	
4.1	Build required genome resources for target species																																						Lab Tech & Bioinformatician	New references for TP species
4.2	High-density TP genotyping required for GS models																																						Lab Tech & Bioinformatician	List of marker variants in TP for GS
4.3	GS data analysis																																						Bioinformatician	GS models developed with available trait data, first GEBVs
4.4	Convert disease resistance info to markers for MAS																																						Lab Tech	New markers resistance gene content (ensure future resilience)
4.5	Refine markers for MAS based on known major effect loci																																						Lab Tech	Markers for known major performance loci (quick-wins)
5	**Biotechnology to enable rapid GS-based breeding**																																						**SCM**	
5.1	Establishment of material *in vitro* conditions																																						WM & SCM	Key genotype stocks established *in vitro*
5.2	*In vitro* optimisation and micropropagation for target genotypes																																						SCM & Lab Tech	Refined methods representing TP species and hybrids
5.4	Demonstrate rapid scale-up of elite material (from MP to cuttings)																																						SCM & Lab Tech	5 genotypes from MP established on nursery.
5.5	Cutting production for mass deployment to yield trials																																						SCM & Lab Tech	Cutting material available (1000+ per 5 genotypes)
6	**Interactions, dissemination.**																																						**IS**	
6.1	Lot 1 projects, Lot 2 Demonstrator and general dissemination																																						IS & WM	Annual report

Accelerating Willow Breeding and Deployment Risk Register

Risk *(Identify and describe all key project risks, including: financial, technology, supply chain, regulatory, etc.)*	Type of risk	Risk owner	Probability *1 - Low* *2 - Med* *3 - High*	Impact *1 - Low* *2 - Med* *3 - High*	Overall risk rating *(Probability × Impact)* *0 - 1 Low* *2 - 4 Med* *5 - 9 High*	Mitigation actions *(Describe the actions taken or planned responses to reduce the impact and/or probability of the risk)*	Residual risk rating, after mitigation applied *(Probability × Impact)* *0 - 1 Low* *2 - 4 Med* *5 - 9 High*
Change of the start date	Implementation	PI & RRes Finance	3	1	3	Delivery plan for the project allows for flexibility should the start date be later than expected	1
Lab tech recruitment	Implementation	PI, Co-Is & RRes HR	1	2	2	Previous recruitment campaigns for similar positions have resulted in high numbers of suitable applicants. Worst case scenario; possibility to reallocate internally.	1
Field tech recruitment	Implementation	PI, Co-Is & RRes HR	2	2	4	Previous recruitment campaigns for similar positions have resulted in lower numbers of suitable applicants than desirable. Worst case scenario; possibility to reallocate internally.	2
Delays in implementation leading to project not meeting its targets in the time period	Implementation	PI & PM	1	3	3	The Steering Committee will convene monthly to review progress with the PM. Advisory Board will have oversight.	1
Availability of key personnel and resources, and adaptation to loss of key personnel	Implementation	PI, PM & Advisory Board	1	3	3	The project teams of staff are skilled and qualified researchers. Within RRes there are other staff who could be deployed on the project if required. Identification of deputies to key individuals will be assigned at kick off meeting so tasks can be re-allocated between staff for short-term cover if needed.	2
Failure of sub-contractors to achieve their objectives for the project.	Financial, Legal	PI, PM & Co-Is	1	3	3	Sub-contractors have been carefully selected from our understanding and experience of organisations with a reputation for delivering on such projects as this.	2
Weather, pests and diseases	Implementation	PI, Co-Is & sub contractors	1	3	3	Although we aim to quantify the effects of weather, pests and diseases on the Training Population, extreme events or combinations of events may render the planting useless, setting the project back by 1 year.	3
Failure to deliver all deliverables may result in BEIS seeking to reclaim the costs of the project	Financial	PI & PM	1	3	3	BEIS are funding this project at 100% FEC. RRes is the sole applicant & acknowledges it has overarching responsibilities for financial controls, ensuring delivery of project. This project has been planned with achievable milestones / deliverables with an experienced project team. The duration of the project has been taken into account to mitigate any risk of not delivering the project.	1
Not understanding or following BEIS T&C's	Legal	PI & PM	1	3	3	Before contract signature T&C's will be discussed with Legal Team and appropriate modification made. All team members to be made aware of specific clauses that may impact the completion of the project by the PI .	1
Unable to deliver work plan	Technical	PI, PM & Co-Is	1	3	3	All project members have proven experience and the technical skills required to conduct the entire project effectively. If additional skills are required, get input from colleagues within RRes due to our collaborative culture.	1
Potential for any conflict of interests	Implementation	PI, PM & Advisory Board	1	3	3	Good project management during scoping stage ensuring all issues and requirements are captured early and discussed with Project Advisory Board initial meeting. Effective communication throughout the project team and in communications with other partners, projects and the BEIS Monitoring Officer. Timetables and task progress, and key decisions regularly reviewed byPI and brought to the Project Advisory Board and Monitoring Officer if they arise.	1
Access to stakeholders	Information	Co-I, PI & Advisory Board	1	3	3	The team have good networks across the supply chain from grower to contractor to end user. These important networks will be nurtured from the initiation of Phase 2	1
Failure to obtain all data sources or obtain permissions to analysis all data. Failure to process correctly or adequate backups made.	Information	PI, PM & Co-Is	1	3	3	Results will be reported at regular project meetings and sent to BEIS when requested. All staff are adequately trained and experienced for the analysis being undertaken in this review. If other sources of data are to be included in the study, appropriate permissions will be gained	1
The risks associated with COVID 19	Health and Safety	All	1	3	3	The risks from Covid 19 have diminished as widespread vaccination has been achieved. Should the situation worsen the protocols developed during early 2020 can be rapidly re-implemented.	1
The risks associated with COVID 19 when selecting Training Population sites	Health and Safety	CO-I	1	3	3	Robust Covid-19 Risk assessments with appropriate controls will be put in place for each visit, and signed off by our internal safety committee and the relevant project partners organizations.	1
Risk of not selecting optimal training population (TP) genotypes	Technical	PI & Co-Is	1	3	3	The teams intimate knowledge of willow germplasm will mitigate associated risks and Phase 1 has allowed planning time for deployment in Phase 2	1
Issues connected to IPR or copyright.	IPR and Innovation	PI	1	3	3	All data is available for analysis. Appropriate discussion within team and BEIS will be informed if potential issues arise.	1
Innovations identified during phase 1 that cannot be completed in 3 years	Implementation	PI	1	3	3	We have designed project to produce outputs within the project timescale. We have already confronted the risk posed by a late start and the lack of flexibility in the planting time for the Training Population. Additional benefits; despite value being identified in Phase 1, this means than some benefits will not be realized in their entirety within Phase 2.	4