Strategies for Clonal Selection of Poplar Hybrids in Chile Based on Superior Wood Properties

Francisco ZAMUDIO; Marco YAÑEZ; Ricardo BAETTIG
Centro Tecnológico del Alamo (Poplar Technology Center)
Universidad de Talca. 2 Norte 685, Talca. P.O.Box 747

Abstract

Because the normal wood obtained today from most poplar plantings in Chile has a high degree of heterogeneity, we cannot achieve manufacturing a large and permanent volume of products derived from poplar wood with an adequate homogeneity in quality. One of the main reasons for this problem is the high probability that what is currently planted includes multiple unknown hybrids originated by natural hybridization among genotypes introduced in the past. Poplars were introduced in Chile in 1810. Until 1999, there were approximately 120 varieties in the country. Since 1999, our center has introduced more than 2,000 hybrids from different pedigrees and from the USA and Europe. Since 2002, we are establishing a nationwide clonal testing network. Some key objectives guiding the selection program are finding genotypes able to produce wood suitable for high quality end products and biomass with the right chemical composition for generating biofuels. Data collected from our clonal genetic tests are more complex to analyze because measures are doubly repeated. First, spatially because measures from two ramets from the same clone are actually repeated measures of the same genotype. Second, over time because measures of the same ramet taken at different moments are longitudinal data. Particularly, our data are taken in spatial sequence, such as a wood related trait measured within the stem of a tree at different heights or cambial age, at a particular height. Therefore, our statistical analyses are confronted with data correlated at multiple levels. In this paper, we review our current selection strategy. We discuss the linear mixed model methodology (LMM) that is suitable for the statistical and genetic analyses of repeated measures collected from clonal forest trials. This methodology permits the presence of heterogeneity of variance in the linear models related to our clonal tests and allows us to address directly the covariance structure of the data. We also explain how modeling the covariance structure significantly improves our ability to estimate the heritability and the variance of specific environmental effects, assess the presence of genotype-by-micro-environment and genotype-by-time interactions, and predict breeding values.

Keywords: Poplars, clonal-selection, wood, statistical-genetics, linear-mixed-models.

Brief Summary about Poplars in Chile

According to botanists from the University of Chile, the first poplar cultivar introduced in Chile was a hybrid of *Populus nigra var italica*, which was brought to Santiago, Chile, from Mendoza, Argentina, by Franciscan monks in 1810. Information about successive introductions of more varieties during the 19th century is lost in history. Since its arrival to the country, poplars have been part of the Chilean landscape in the cities as well as the countryside. They were planted as ornamental trees in main roads accessing many towns or as part of public recreational areas or parks. In rural areas, poplars have been used as shade and wind shelters by farmers to protect their animals and crops.

Today, we can observe poplar hybrids in very distant locations, such as the upper mountains in the Tarapaca Region, far north of Chile, and nearby Punta Arenas and the Strait of Magellan, in the extreme South of the country.

According to the National Poplar Commission (NPC), approximately 120 poplar hybrids were introduced in Chile between 1810 and 1999. During the 20th century, most of the relevant introduction of varieties was conducted by the same company that owns the largest surface of poplar plantings used for industrial purposes. This area is approximately 4,000 has and is located near the town of Parral, in the region of the country called Maule. The germplasm introduced by the company has been used mainly to provide the wood for its own industrial processing and only few hybrids have been sold to external landowners for other regions of the country. During the 40s and 50s, some public institutions and private companies imported few clones from different sources in Europe and North America; however these experiences were not well documented. It is generally accepted by the NPC that the six is the number of clones planted nowadays for commercial purposes in Chile out of the total number of clones introduced until 1999, and the total surface of poplar planting in the country is around 6,000 has.

Reasons to Think about Poplars in Chile

There are several reasons to think about poplar cultivation as an option for developing an intensive clonal forestry in Chile. Let's review some of them. First, the forest policy adopted by Chile during the last 30 years or so has mainly favored a growth model linked to large-scale investments from the private sector. Large surfaces of radiata pine and Eucalyptus plantings have been established and managed with the objective of providing wood to the industrial complexes that produce products – mainly commodities – that are the base of the Chilean forest economy (e.g. pulp and saw wood). Most of these complexes are part of the same industrial holding that also own and manage the forests. The lack of silvicultural options and the excessive concentration in the ownership, management, and the further wood processing of radiata pine and eucalypt plantings – in only two main industrial holdings – is increasingly forcing the closing up of a significant number of small size wood processing companies. The country must quickly generate forest resources with fast growing tree species alternative to radiata pine and eucalyptus, in such a way that many land owners can use their land for forest cultivation in a steady

form, and provide wood to a much more diversified wood processing industry in the future. By increasing the number of species used in forest plantings, the Chilean forest activities and silviculture will maintain a sustained development over time and the country will obtain higher economic, social, and environmental benefits.

Secondly, an economic crisis is still affecting some specific and traditional crops, such as wheat, cereals, etc. This crisis is shown by a lower growth rate than the rest of the national economy. Because Chile does not have an economy of scale and a significantly high production of grains, agriculturists and farmers cannot efficiently compete with the same type of products that are imported from other countries in South America. If alternative or complementary types of cultivars are not quickly adopted in lands used by the traditional type of crops, an important sector of the national economy will face a serious depression. This will continue involving unemployment, inadequate migration to large cities, diminution of the standard of life, and a deterioration of the economic value of specific type of agriculture soils.

The University of Talca has been involved in research with poplar hybrids since 1998. Later, in 2003, the University created the Poplar Technology Center (PTC) to lead the clonal selection program in Chile. We at the PTC, consider that the Populus genus has the greatest potential for intensive cultivation in Chile, after *Pinus radiata* D. Don and *Eucalyptus spp*. It is expected that the establishing and management of poplar plantings (a) help to resolve the problems related to the traditional agriculture and (b) give the country a diversity of forest products with a high added value. Thus, the risks of a forest economy excessively dependent on only two species, primary products, or forest commodities, should diminish.

The intensive poplar culture has several advantages in Chile. First, many agriculturists and farmers could benefit because the intensive culture of poplars requires a smaller surface of land than radiata pine plantings to be economically justified. Second, the forest management of poplars stands is compatible with several types of agricultural crops, which could allow a mixture of agricultural and forest management. Third, poplar cultivation requires a low investment to participate in the local market. Finally, poplar management involves a return to the investment that has the option to growth through the increment of the exportation of products derived from poplar wood.

The country offers good climatic and soil conditions for planting and managing more species and varieties of *Populus spp*. According to a survey conducted by the Chilean Forest Institute (INFOR: Instituto Forestal), between the VI and X administrative regions (34° and 42° Latitude South) there are several millions of hectares where poplar cultivation is feasible. However, this implies there is a wide variation of environmental conditions (soil and climate) between and within the many different sites where poplar hybrids can be established. Thus, a careful and intensive testing and selection program of new poplar hybrids has to be conducted.

The poplar cultivation can also have a positive environmental effect. The geography in the center and south of Chile includes many rivers that cross the country from the Andes

to the Pacific Ocean. During winter, the strength of most of the rivers is so high it flows right out of the river channel, usually at corners or meanders, provoking flooding that seriously affect human populations and their properties as well as state owned land. Planting poplars along the stream banks or terraces produced by the rivers could help to protect the bank from erosion, reduce sediment and pollution from agricultural chemicals, and improve water quality.

A very promising role of poplar cultivation in Chile is in the disposal of sewage sludge and remedying contaminated soils and groundwater. Large quantities of biosolids are generated in Chile from the process of treating municipal water and there is a limited capacity to store them. At the same time, Chilean copper industry is the largest in the world. However, there are hundreds of mining tailings in the north of Chile that require to be (phyto-) stabilized in order to avoid their collapse and the dispersion of pollutants. At the same time, lecheates coming from municipal landfills are required to be treated. Plantations with poplar using biosolids could solve these environmental problems.

Finally, we strongly believe that poplar cultivars can be adequately used in Chile in generating biofuels solids and liquids. The growth observed by several cultivars under testing in several regions of Chile by the PTC looks promissory for establishing a clonal selection selection strategy aimed to identify varieties useful to generate energy from biomass processing.

The Work of PTC since 1999

The work conducted by the PRC has been strongly supported by the Chilean National Science Foundation (CONICYT: Comisión Nacional de Ciencia y Tecnología). Between 1999 and 2008, the PTC was awarded three grants from the special CONICYT program called FONDEF (Fund for development and entrepreneurship) totaling US$ 2 million approximately. Below, we give a brief description of what was done in each project.

1999 – 2002: Introducing New Germplasm

Between years 1999 and 2002, the PRC developed a first FONDEF project (coded: D98I1086). At the same time, the PTC became a member of the Poplar Molecular Genetic Cooperative (PMGC), from the University of Washington at Seattle, WA, USA. The hybridization program conducted by the PMGC involved three native species from North America (*Populus trichocarpa, P. balsamifera*, and *P. deltoides*) and polen from one species from Asia (*Populus maximowiczii*) and another from Europe (*Populus nigra*).

During this period, more than two thousand poplar hybrids from thirteen pedigrees were introduced to Chile from the PMGC clonal catalog located in Puyallup, WA. Table 1 depicts the number of hybrid according to the pedigree. Hybrids were separated in two groups and imported in two successive years. Each set of hybrids were established under a severe quarantine regime following a rigorous control by the Chilean agriculture service. This procedure was conducted in a specially design quarantine facility located within the University main campus, in Talca. Following a one-year quarantine period,

each load of clones were transferred to our clonal bank, also located within the main campus.

Table 1. Description of hybrids imported from the PMGC in July 1999 and March 2001. The initial number was 2589, however here we include only the number of hybrids that survived the quarantine period.

Type of germplasm	Taxa	Number of hybrids
Selectos parental trees	T	49
Commercial hybrids[1]	TxD	8
Experimental hybrids	DxB	6
	TxD	250
	TDxD	993
	TDxT	40
	TDxTD	213
	TMxM	13
	TMxT	20
	TMxTM	108
	TxM	54
	TxN	139
	TxT	117
	TDxTN	40
Missing identification	--	167
Total alive hybrids[2]		2217

1: generated by the University of Washington
Taxa code:
T : Populus trichocarpa
D : P. deltoides
B : P. balsamifera
M : P. maximowiczii
N : P. nigra

Most of the imported PMGC clones are experimental. However, we also introduced eight commercial clones which are currently used in the Pacific north-west in poplar commercial planting. Finally, we also introduced 49 selected parental trees used by the cooperative researchers as part of their hybridization program.

At the beginning of 2001, we also imported, from the Pepiniere Forestiere Experimentale de Guemene (GUEMENE-PENFAO), France, a set of 20 poplar clones of five different pedigrees that form what is called the "European catalog" of commercial varieties (see Table 2). Hybrids were originated in different countries in Europe (see Table 2). Parental genotypes used in the hybridization were from two European (*P. alba y P. nigra*) and two

North American *Populus* species (*P. deltoides y P. trichocarpa*). These varieties are currently used as genetic control in our clonal testing network.

Table 2. Description of imported hybrids from the French national nursery (PENFAO), in Guémené, in February 2001.

Taxa	Number of hybrids	Commercial name
A	1	Villafranca
D	2	Alcinde Carolin
DxN	10	Blanc du Poitou Cappa Bigliona Dorskamp Flevo Gaver Ghoy I-214 I-45-51 Robusta Triplo
T	2	Columbia River Fritzy Pauley
TxD	5	Beaupre Boelare Hunnegem Raspalje Unal
Total	**20**	

Number of cuttings per hybrid= 25
Total number of cuttings received from the nursery in France= 500
Code of taxa:
A : Populus alba
D : P. deltoides
N : P. nigra
T : P. trichocarpa

Today, Chile has what is probably the largest genetic base of poplar germplasm in the South hemisphere.

During this period, we also allocated resources to study few wood properties of some of the most planted poplar hybrid in Chile. The idea here was to generate comparative patterns that help to discriminate, among the poplar hybrids introduced by the PTC, those genotypes that produce a better wood quality than the currently obtained in Chile.

2002 – 2005: Initial Screening. The Beginning of Level 1 Type of Clonal Testing

Between years 2002 and 2005, the PTC developed a second FONDEF project (coded: D01I1131). During this period, we began establishing our clonal testing network. In this endeavor, we followed the approach described by Stettler et al (1992). The type of clonal trials that we established can be considered as level 1, or "nursery tests". Unrooted cuttings of 1,800 hybrids were planted during winter following a randomized complete block design, with one ramet per hybrid in each of two blocks. Each block was established in two consecutive years and the reason is double fold. First, we receive few original cuttings of each hybrid from the PMGC clonal bank, therefore during few years we have limited amount of cuttings per clone available for testing in multiple sites. Second, by planting in two different years we try to measure the seasonal effect of climatic variation with year.

Total height (m) and diameter of the stem (cm) at 40 cm above ground level were measured during dormancy following the first growing season in dominant branches, and the average recorded for analysis. We also assessed the susceptibility of each hybrid to pest and diseases present in the local environment. The full blocks were coppiced after the first growing season. Pruning was applied in the following spring to keep only the strongest dominant branches. During the second dormancy, the same traits were again measured for the second time. After the data gathering, both blocks were coppiced again and pruning was repeated in the following spring. Finally, traits were newly measured during the third dormancy on the dominant stems.

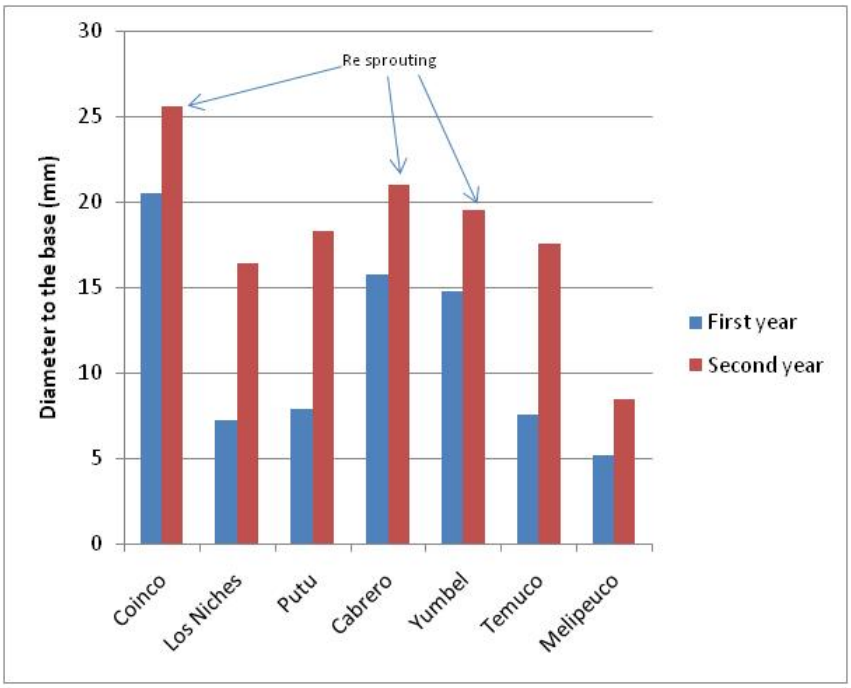
The main objective of this nursery test was to provide genetic information about initial growth (we attempted to measure repeatability of growth patterns), sprouting ability, and adaptability of the tested pedigrees to the local environment (climate, soil, pests, and diseases).

In this case, the main results were the setting up of an initial clonal testing network formed by nursery type of tests (one ramet of many hybrids in two blocks). We also generated a list of hybrids with the best performance in terms of growth rate and adaptability (tolerance / resistance to plagues and diseases) per testing area.

The network extends across 700 km of the center – south regions of the country, approximately. Two tests were planted in the Rancagua area (34°14' and 34°17' S); three tests were planted in the Talca area (35°04', 35°13', and 35°24' S); five tests were planted in the Los Angeles area (37°01', 37°13', 37°28', and 37°31' S); and two tests were planted in the Temuco area (38°39' and 38°52' S).

The use of "nursery tests" as an initial tool for clonal screening last one or two years, approximately. Most of our nursery tests were used for gathering information for three years. However, two of them were kept for five years for wood sampling purposes as well, as summarized below

Figure 1 depicts the total mean for diameter at the base of the stem and height in some of the nursery tests, and for the first growing season (in blue) and the second growing season (in red), following a coppice after the first growing season. Therefore, mean values for the second year are the means of the dominant sprouts of each hybrid. Better growth was recorded during the second growing season mainly due to the presence of a well developed root system compared to the first year. The best growth was also observed in the Rancagua area, in the trial named Coinco.



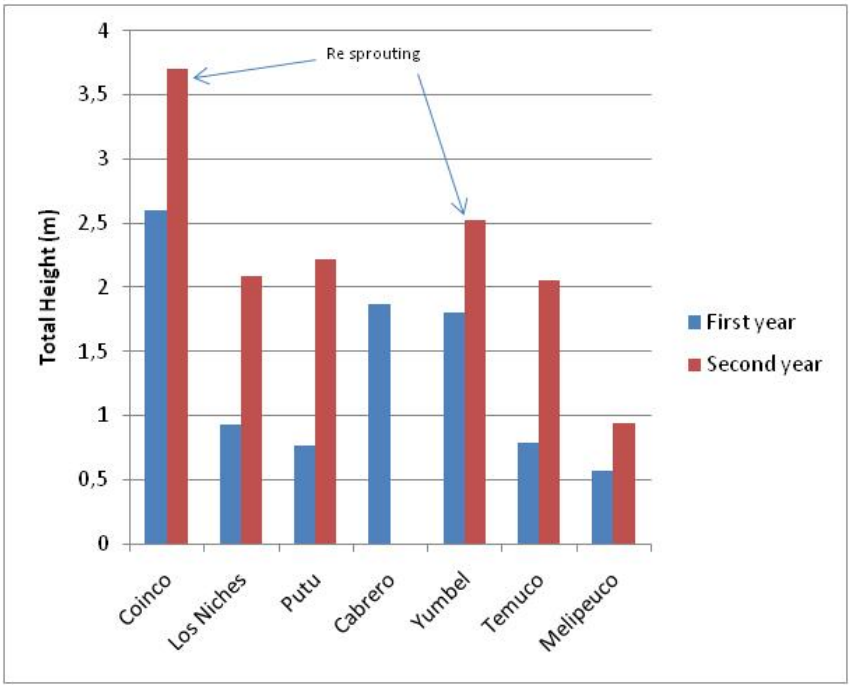


***Figure 1**. Total mean values for diameter at the base of the stem (40 cm above ground) and total height in seven nursery tests, and for the first growing season (in blue) and the second growing season (in red), following a coppice after the first growing season. Therefore, mean values for the second year are the means of the dominant sprouts of each hybrid. The Coinco trial is located in the Rancagua area; the Los Niches and Putu trials are located in the Talca area; the Cabraro and Yumbel trials are located in the Los Angeles area; and the Temuco and Melipeuco trials are locates in the Temuco area.*

2006 – 2008: Candidate Testing. The Beginning of the Level 2 Type of Clonal Testing

The ranking generated by the second project indicated the best hybrids within each nursery test and testing site, in terms of growth potential and adaptability. However, a new project was needed to initiate a second stage of clonal selection. Tests included in this stage are called *candidate tests*. They involve a large number of clones and a reduced number of ramets per clone. These tests last an average of five years, but the total testing period will depend on the type of character that is the base for selection. The evaluation of properties of mature wood will enlarge the testing period.

Between 2006 and 2008, we have been developing a third FONDEF project (coded D04I1027) that indeed allowed us to start planting our first series of candidate tests. Here, the general objective was to select hybrids with a combination of growth and juvenile wood that is suitable for industrial uses. Therefore, specific objectives were to: (1) measure the growth of the best hybrids detected from the outcomes of the previous project (FONDEF D01I1131) and (2) analyze the wood that they formed during the first three growing seasons in terms of its potential for industrial use.

The methodology included three topics. The first involved the planting of several screening clonal tests. Two are located in the Los Angeles area (37°01’ and 37°28’ S) and one in the Valdivia area (39°45’ S).

An average of one hundred hybrids was selected from the nursery tests planted in the different combinations of soil and climate during the previous FONDEF project. Selection criteria were based on growth rate and adaptability (including resistance-tolerance to plagues and diseases). Ten cuttings per hybrid were planted on ach testing site. The experimental design consisted in a randomized block design, with four cuttings per hybrid and block. Four tests were established in equal number of testing sites during year 2006. The number of common clones in two or more trials was variable. The second topic involves the analysis of wood formed during the span of the project, which is juvenile type of wood. Mechanical, physical, and chemical properties are to be measured in samples collected from different testing site and hybrid. In 2006, preliminary wood determinations were conducted with samples collected from few nursery tests planted in 2002. At the end of 2008, one cutting will be cut per hybrid and block from each candidate test planted on 2006, and wood analyses will be conducted. The statistic analyses will be oriented to measure: (1) the variation of wood properties between and

within hybrids and (2) the degree of repeatability of wood properties over time and trial location. The third topic consists in evaluating the hybrid susceptibility or tolerance to the attack of pathogens (insects and rust).

Expected results of the project are: 1) information about some relevant wood properties of the tested hybrids; 2) a ranking of the best hybrids per tested environmental condition, based on their juvenile wood and growth rate; 3) a network of clonal tests aimed to select poplar varieties useful for industrial uses; and 4) information about the degree of tolerance to the attack of pathogens among the selected hybrids.

The main impacts of the project are expected to be: (1) a significant increment in poplar wood of high quality, suitable for the industry of wood transformation; (2) a significant reduction in the risk investment in managing poplar stands planted with our selected hybrids; (3) an improvement in the economic value to land with no present use, or underused; and (4) an increment in our understanding of wood formation in selected hybrid poplars (with known pedigree).

Preliminary results obtained from our clonal testing network indicate that some poplar cultivars can growth faster and better adapted than radiate pine and eucalyptus trees to areas where the soil presents some limiting factors to these two types of trees. For example, one of the places with the greatest potential for poplar cultivation seems to be the sandy and humid soils in the Los Angeles and Yumbel area (37° S). Figure 2 depicts the total mean values for diameter at the base of the stem and total height in three candidate trials; two located in the Los Angeles area (Yumbel and Los Angeles) and one located in the Valdivia area.

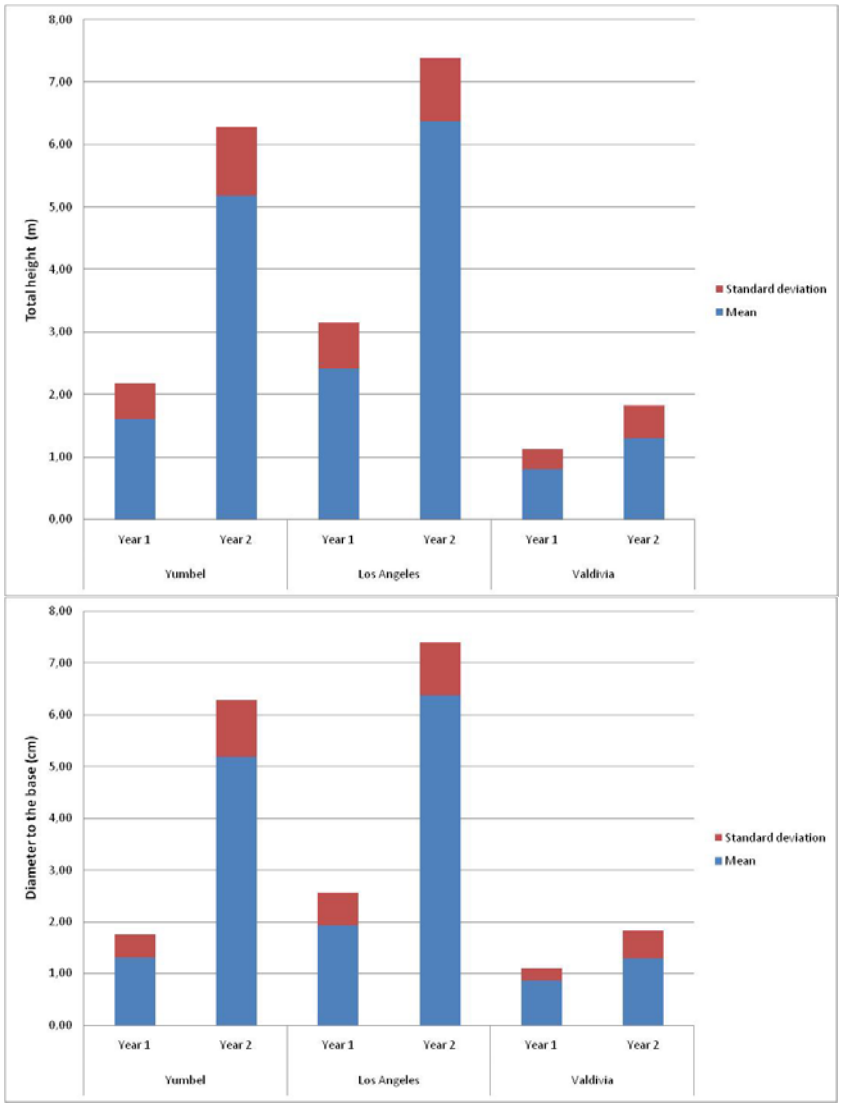


***Figure 2**. Total mean values for diameter at the base of the stem (40 cm above ground) and total height in three candidate trials, and for the first two growing season. The mean for year 2 is related to the cumulative growth. The Yumbel and Los Angeles trials are located in the Los Angeles area and the Valdivia trial is located in the Valdivia area.*

We will normally measure wood properties in candidate trials. However, we have assessed wood density and few mechanical properties in wood samples obtained from some of our first nursery tests, which were thinned after the third year of growth. The results obtained from these tests will only be considered only for comparison purposes once we obtain more informative results about wood properties from our candidate trials. In a nursery test, the spacing between trees is usually much reduced because the test is expected to last only two or three years. However, our results will be useful to know the wood properties of a genetic base submitted to minimum silviculture practices. Average wood densities recorded by all taxa tested in our trial named "*Coinco*" (near Rancagua)

are shown in Figure 3. Histogram of wood density recorded in four nursery trials are shown in more detail in Figure 4.

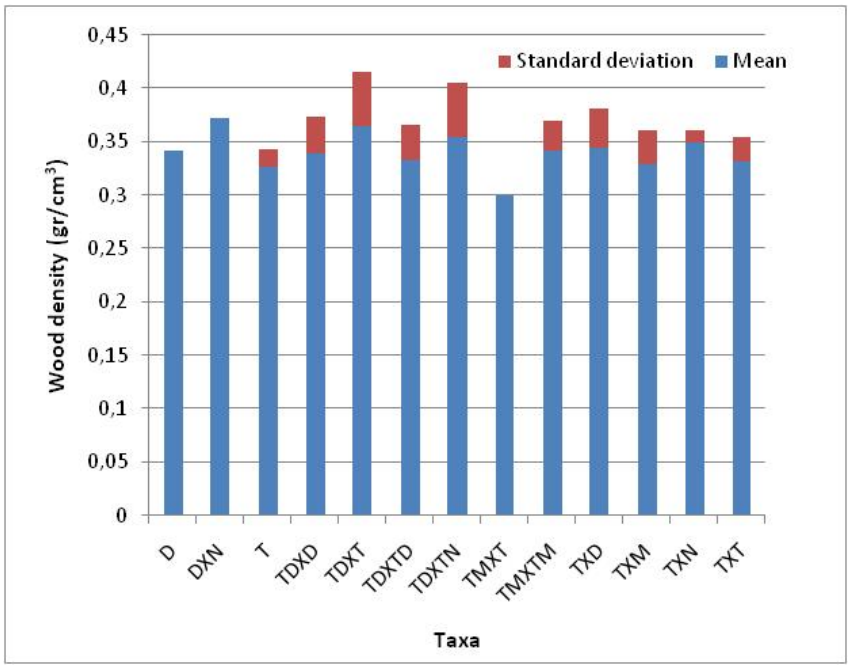


Figure 3*. Average wood density for taxa included in nursery trial coded "Coinco", and located in the Rancagua area. Spacing in the trial was 0.6 x 1 m.*

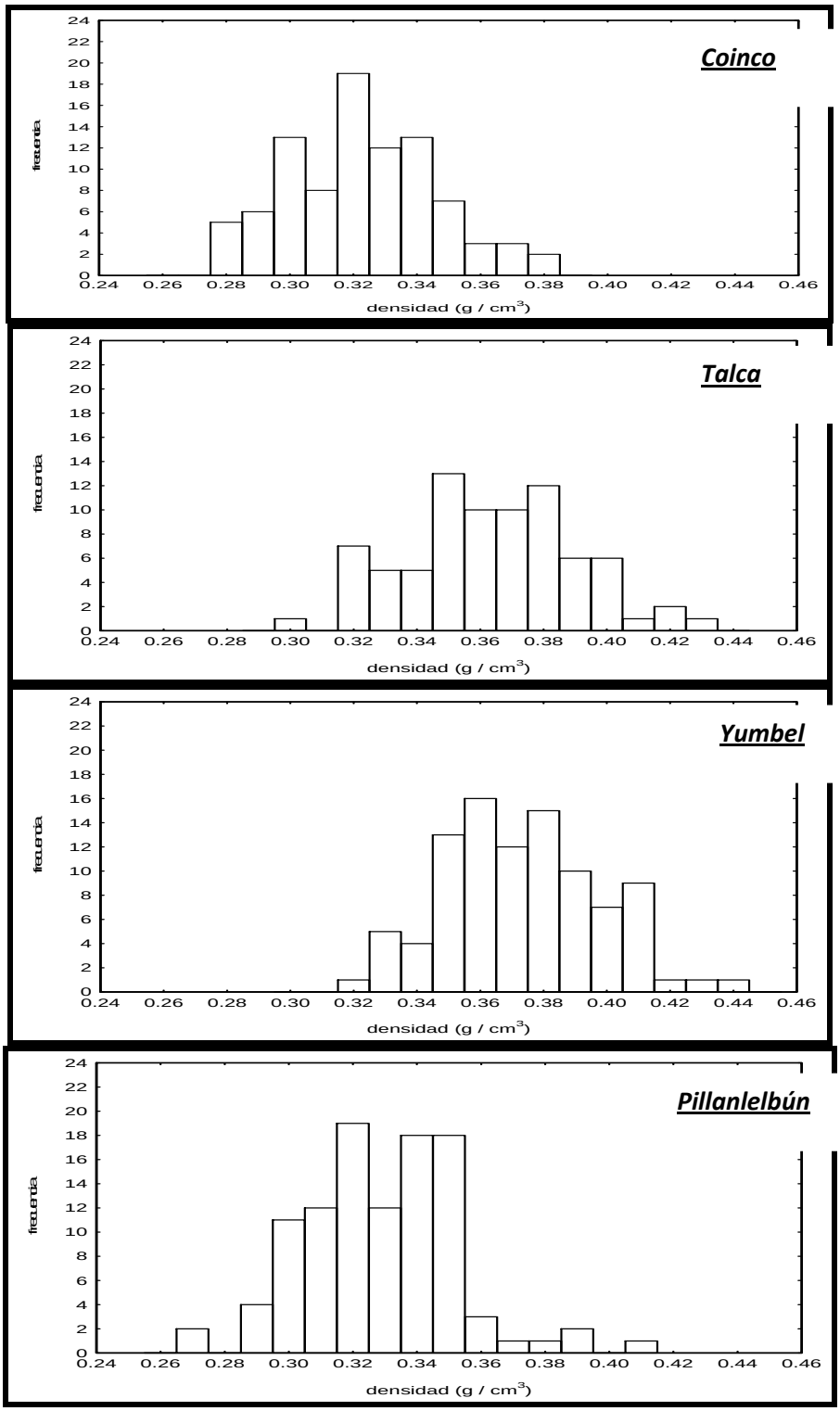


Figure 4*. Histograms of wood density measured in wood samples obtained in four nursery tests after thinning at age 4. Spacing in each trial was 0.6 x 1 m. These results are used only for comparison purposes as an indication of wood density under constrained spacing.*

Strategic plan and research lines in PTC

In the PTC, we consider that our mission is aim to generate the necessary technologies and adequate ideotypes to develop and utilize the future poplar plantings from Chile, in an ecologically and economically sounded form, with a positive impact in the private sector, the State, and the community. We will achieve this with the contribution of a multidisciplinary and unique working group committed to excellence and constant innovation.

Consequently, our strategic plan for the period 2009 – 2018 is oriented to develop the following research lines:

1). Line of research N° 1: Selection of superior poplar varieties capable to produce very high wood quality products.

2). Line of research N° 2: Selection of superior poplar varieties suitable for biomass production, suitable for: a) bioenergy generation and b) carbon sequestration.

3). Line of research N° 3: Selection of superior poplar varieties suitable for phytoremediation and phytoestabilization of: a) mining tiles and (b) biosolids generated from the process of treating municipal water.

4). Line of research N° 4: Selection of superior poplar varieties specially adapted for river restoration.

The first line of research is a continuation of our selection efforts started on 2002. The proposed strategy for going forward in our efforts to select clones with superior quality wood is described below.

Research Line N° 1. Strategy for Clonal Selection of Poplar Hybrids in Chile Based on Superior Wood Properties

Main Hypotheses

The main hypotheses that sustain this research line are:

1) The genetic control is expected to be moderate for growth related traits (this means a broad sense heritability near 0.4) and moderate to high for wood related traits (broad sense heritability between 0.4 and 0.7).

2) The predicted genetic gain based on clonal selection is expected to be in the order of 10 % for growth traits and 30 % for wood related traits.

3) The genetic correlation between wood properties and either growth or adaptive traits (such as pest and disease resistance) is low or no significant ($r_g < 0.3$) at any age of assessment.

4). On the average, genotype by macro-environment interaction is moderate in growth and wood related traits. If we express this interaction in terms of the genetic correlation between gene effects in different environments, this correlation is approximately 0.3.

4b). Genotype-by-micro-environment interaction in growth and wood related traits could be significant in particular testing sites. Expressed as the correlation between the within-clonal effects from two different blocks, this correlation could be less than 0.1 or even negative within some specific environments.

5). The age-age genetic correlations are not significant, or genotype-by-time interaction has a moderate effect on growth and wood formation. Expressed in terms of the genetic correlation between clonal effects from different assessment time, this correlation is 0.2 and 0.3.

These hypotheses are based on the current literature, our previous research experience with radiate pine and *Pinus tecunumanii*, and, more recently, the analysis of data collected in our clonal testing network of new poplar varieties.

Main Goal

The main goal of this research line is to select clones with the adequate combination of wood properties and growth pattern capable to support a wood transformation industry linked to poplar wood.

Specific Objectives

The specific objectives are:

1) To determine patterns of wood formation with cambial age and their relationship with growth patterns.

2) To determine the interaction patterns between wood formation and changes in environmental conditions, at the micro and macro level.

3) To determine or classify poplar hybrids according to: (a) their growth and wood formation patterns and (b) their adaptability to local (specific) or general environmental conditions.

4) To select clones with the best pattern of wood characteristics (across cambial age) and the fastest and most stable growth rate to be placed in the best environmental conditions available for poplar cultivation.

5) To select clones with the best pattern of wood characteristics (across cambial age) to be placed on environmental conditions with few limiting factors.

Characters of Interest

The main characters of interest will be:

1. Wood density,
2. Mechanical properties (MoE, MoR), and
3. Microfiber angle.

Ancillary traits will be:

1. Growth rate (diameter and height),
2. Pest and disease tolerance (particularly we will look for tolerance to the attack of rust and aphids), and
3. Propagation capability.

Selection Strategy

If we consider that the initial population size under testing is N_0, the size of the selected population at selection time $t=1$ will be $N_1 = a_1 N_0$, where a_1 is the selection intensity, and $0 < a_1 < 1$. If the size of the selected population at selection time $t=2$ is $N_2 = a_2 N_1$, with $0 < a_2 < 1$, where a_2 is the new selection intensity, therefore the size of selected population at selection time $t=s$ will be $N_s = \prod_t^s a_t N_0$, with $0 < a_t < 1$.

Poplar hybrids to be included in the selected population at any selection time will be those that maximize the genetic correlations: (a) between growth and wood related traits at the age of selection; (b) between growth increments from different growth periods until the age of selection; and (c) between wood characteristics measured at different cambial ages that make up the selection age.

Plan of Action

Clonal testing procedure

The following is the schedule for our clonal testing procedure:

1). 1999 – 2002: initial introduction of germplasm from the PMGC and Europe. This first step is already over, even though, the PTC is planning to continue introducing new hybrids between years 2009-2014. The specific objective of the future new introductions will be to enlarge the genetic base that will support the future continuation of our clonal selection program.

2). 2003 – 2005: initial screening: the beginning of the level 1 type of clonal testing. During this period, we established the first set of tests normally called "nursery tests". Each test involved many hybrids, one cutting per hybrid and block, few blocks (only two in our case).

3). 2006 – 2008: Candidacy testing: the beginning of the level 2 type of clonal tests. Here, we began planting a large number of clones and small number of ramets per clone.

4). 2009 – 2014: Clonal performance testing: we will begin a more extensive level 3 type of clonal testing of better candidate clones. The number of clones will be small and the number of ramets per clone will be large. During this period, we will also continue establishing new candidate tests in environmental conditions different from to those involved in trials planted between 2006 – 2008. We will also introduce new hybrids from abroad to enlarge our genetic base as well as establish a new set of nursery tests with this new germplasm, after finishing the required quarantine procedure.

5). 2014 – : Compatibility trials: in this year, we will begin to establish the level 4 type of clonal tests to identify sets of clones that can be advantageously grown in sequenced mixtures. We will also continue establishing new candidate tests and clonal performance, based on the clonal selections conducted with the hybrids introduced between 2009 – 2014.

Genetic and statistical analyses

The statistical and genetic analyses can be outlined as follows:

1). Growth and adaptability traits will be measured in a yearly bases. Wood properties will be measured using both destructive and non destructive procedures. In both cases, we will determine a protocol of wood sampling based on the growth rate observed in each testing site. Wood assessment will follow standard and internationally valid testing protocols.

2). Genetic control will be assessed for growth, adaptability traits, and relevant wood properties (physical and mechanical properties. As the age of trees increases, we will have a better chance to get measures of the genetic variation of wood characteristics with cambial age. For example, with the help of increment borers, we will measure the genetic variance of the density or the microfiber angle at the ring level.

3). Genetic gain prediction for relevant traits and related response to selection will also be conducted. The different estimate of genetic control will be useful to predict the genetic gain that is possible to obtain by performing several clonal selection criteria. Normally, genetic gain prediction is conducted in breeding populations which are sexually propagated. In fact, the basic genetic gain prediction equation is a regression model that correlates the breeding value of a descendant to the phenotypic value of one of or both parents. Genetic gain modeling can also involve assessments of different selection options, such as performing family and within-family selections simultaneously. This

concept must be carefully extended to the case when we attempt to predict the breeding value of a genotype by observing its phenotypic values of its clone. At the same time, we will also predict the related response in wood quality traits to selection performed based on growth.

4) In parallel to the analysis of related response to selection, the genetic correlation between growth and wood properties at different cambial ages will be assessed.

5) The presence of the genotype-by-macro-environment interaction will also be measured if number of common clones in two or more testing sites is significant. This analysis will involve two types of approaches. First, we will assume that genes affecting the clonal performance in more than one site are the same; therefore, a trait measured in different testing sites is the same trait. Second, we will assume that genes affecting the clonal performance in more than one site are different; therefore, a trait measured in different testing sites can be considered as different traits (Falconer and Mackay 1996). The assessment of the genotype-by-micro-environment interaction will be mandatory and performed in every trial. This analysis will be based on the estimation of covariances between intra-clonal (or residual) effects from two different blocks within a site. If the performance of a plant from a clone in one block is highly correlated to the performance of another genetically identical ramet in a second block, therefore the genotype-by-micro-environment interaction is null or negligible (Zamudio et al 2008).

6) The presence genotype-by-time interaction will also be measured in all trials. The analysis will be based on the estimation of covariances of clonal and within-clonal effects measured and compared at different ages, for growth traits, or cambial ages, for wood traits. We will emphasize the work with growth increments. If the covariance of clonal effects for different growth increments is zero or even negative this implies the presence of a heavy genotype-by-time interaction. We will use the approach used by Zamudio and Wolfinger (2002).

7). Clonal selection will be based on the best linear unbiased prediction (BLUP) of the breeding value for each genotype under testing. Complementary, we will use the selection index methodology to classify our clones based on different sources of information, such as type of traits (growth, adaptability or wood characteristics), age of evaluation, and environmental conditions.

Expected Results

Because of our clonal testing and selection procedure, we are expecting to select clones with at least one of the following characteristics:

1) A stable growth through ontogeny. Each clone should contribute to make the ring-to-ring genetic correlation to be significantly high $\left(r_{g_{tt'}} \geq 0.7\right)$.

2) Stable wood properties from ring to ring. Clones should also contribute to increase the ring-to-ring genetic correlation $\left(r_{g_{tt'}} \geq 0.7\right)$.

3) A high intra-clonal correlation between growth and relevant wood properties. Clones should contribute to increase the genetic correlation between both traits to at least $r_g \geq 0.6$

4) A stable growth and homogeneous wood properties regardless of the site. Clones should not contribute to the genotype-by-macro-environment interaction but they should contribute to increase the genetic correlation of a trait measured in two different sites, which will be measured as the correlation between clonal effects measured in two sites. These clones will be considered to be universal type of clones.

5) A stable growth and wood properties at specific sites. They will contribute to the genotype-by-macro-environment interaction. However, they will show minimum signs of contributing to the genotype-by-micro-environment interaction. Their correlation between intra-clonal effects from different blocks will be positive and higher than 0.5. These clones will be considered to be site-specific type of clones.

Restrictions to the Statistical and Genetic Analyses

The statistical and genetic analyses of data from our clonal tests will become complicated because of several reasons. First, data collected from different genetic entries established in field trials are unbalanced due to a normally expected mortality rate or the presence of an uneven number of genotypes for each taxon included in each clonal trial. Second, we will assess our experimental units (ramets from different clones) several times during the life span of clonal trials. Different characters will be measured at regular or irregular intervals. In general terms, data in which the response variable is observed in sequence on the same subject are called repeated measures. Some measurements will be cumulative in nature. For instance, we will assess diameter and height of trees in yearly bases and later on determine the annual growth increments. Therefore, every new growth measurement will carry the previous record plus the increment occurred between the two measuring times. The mathematical representation of this type of trait is very simple

$$Y_2 = Y_1 + \Delta Y \qquad [1]$$

where Y_1 is the trait measured at "moment" 1, Y_2 is the trait measured at "moment" 2, and $\Delta Y = Y_2 - Y_1$ is the increment in growth between moments 1 and 2. Few questions are to be asked after we observe expression [1]. The first is about the variability of Y_2, how large is it? Part of the theoretical answer to the question is given by the following equation

$$\begin{aligned} Var\left(Y_2\right) &= Var\left(Y_1 + \Delta Y\right) \\ &= Var\left(Y_1\right) + Var\left(\Delta Y\right) + 2Cov\left(Y_1, \Delta Y\right) \end{aligned} \qquad [2]$$

Expression [2] is a reminder that as we gather information about the trait, its variability at later age will depend on the association between the cumulative growth at an earlier age and the increment recorded between both ages. Therefore, variability at later ages can decrease simply because of a negative association between growth increments. A second question is related to this association. If variability of Y_2 is lower than Y_1, is it due to a negative genetic correlation between cumulative growth? Or is it due to non-genetic effects? Part of the answer can be obtained by observing another related expression

$$\begin{aligned} Cov(Y_1, Y_2) &= Cov(Y_1, Y_1 + \Delta Y) \\ &= Var(Y_1) + Cov(Y_1, \Delta Y) \end{aligned} \qquad [3]$$

Comparing [2] and [3], we predict that the variability of cumulative growth and the association between two cumulative growths are functions of: (a) the variability of growth increments and (b) the association between themselves. Therefore, a key part of our work will be to carefully assess how much of these functions are due to genetic effects and non-genetic effects. The application of this knowledge will be crucial for developing our clonal selection program. For example, the right combination between a stable or homogeneous growth pattern and adequate wood properties can be essential to produce high quality timber or biomass suitable for biofuel generation. Besides, if we must establish commercial plantings of forest trees in heterogeneous environmental conditions – for example, due to particular fluctuations in soil properties – detecting genotypes with stable growth can be determinant for recovering unused or degraded land.

There is other type of trait that will be measured several times but it is not cumulative in nature. Repeated measures can also be taken in spatial sequence, such as a wood related trait measured within the stem of a tree at (a) different heights or (b) different ring numbers (cambial age), at a particular height (Zamudio et al 2002; Zamudio et al 2005). This is the case for several and relevant wood properties. For instance, by working with microdensity profiles, we will assess specific gravity at different ring numbers or cambial ages. Eventually, we could measure the cellulose and lignin content at different ring sections of the stem. Relevant properties at the cell level are also non-cumulative over time, but they can be assessed periodically, such as microfiber angle, fiber length, and cell wall diameter, etc.

A third reason that will complicate our analyses is due to a relevant statistic characteristic of working with clones. Data will be doubly repeated. First, data will be spatially repeated because two ramets of the same clone are two identical copies of the same genotype. Thus, planting two or more cuttings of the same genotype in one or more environments is equal to working with "*genetically*" repeated copies of the same genotypes, and measures from two ramets from the same clone are actually two repeated measures of the same genotype. Second, data will be repeated over time because measures of the same ramet taken at different moments are longitudinal data. Therefore, measuring growth – or wood properties – at different ages on the different ramets (trees) of a clone, planted in different locations, generate data that are repeated in two senses, "*over the time and over the space*" Our research will confront data correlated at multiple

levels and we cannot use the traditional approach for measuring statistical differences between clones or the interaction between clones and any sort of treatment applied to our trials.

There are several statistical methods used for analyzing repeated measures over time. Littell et al (2006) mention for example: 1) separate analyses at each time point, 2) univariate analysis of variance, 3) univariate and multivariate analyses of time contrasts variables and 4) mixed model methodology. Moser et al (1990) also mention time series analysis. There has been considerable work on linear mixed models (LMM) in the past twenty or so years (McCulloch and Searle 2001), which has been enhanced by the development of computing hardware and software. As a result, current LMM methodology not only permits the presence of heterogeneity of variance in the linear model (still assuming normality) but also allows the researcher to address directly the covariance structure. Modeling the covariance structure of the data can improve our ability to analyze repeated measures data by providing valid standard errors and efficient statistical tests.

A fourth reason to complicate our data is related to the underlying probability functions of the traits that are the base of the clonal selection and the random effects in the linear mixed model, which is linked to the different experimental designs used in our testing program. From a probability point of view, we will be facing different type of data. There will be continuous measurements where our observations may, in theory, fall anywhere on a continuum, for example, specific gravity, biomass content, length, volume, etc. Some traits will follow a normal distribution. However, others might follow a gamma or a Weibull probability distribution. There will be other traits where measurements will be expressed on a non-continuous scale. For instance, the responses or measurements used to evaluate adaptability may be alive or dead, or presence or absent of plagues or diseases. This dichotomous trait is also called a binary random variable. "*Success*" or "*failure*" is used as generic terms for the two outcomes of what is usually known in statistics as a "*Bernoulli trial*". The joint probability density function (*pdf*) of N Bernoulli trials is the Binomial distribution. If there are more than two categories, the variable is called "*polychotomous*", "*polytomous*" or "*multinomial*". These are "*nominal*" classifications. Multinomial data can follow a Multinomial *pdf*.

We will also have the possibility to record measurements that can be classified as "*ordinal*", in which there is some natural order or ranking between the categories, for example: no attack, low, medium, and severe attack. Nominal and ordinal data are sometimes called categorical variables and the number of observations, *counts* or *frequencies* in each category are recorded. The number of times an event occurs has been a common form of data collected in our testing program in the clonal bank and our testing network. For example, we have recorded the number of branches or cuttings of certain dimensions produced by different hybrids in the clonal bank during a growing season. We have also measured the number – or proportion – of branches per cutting, or number of cuttings per hybrid in a clonal trial, attacked by aphids. We will continue with this type of assessment in the future. The Poisson joint *pdf* is often used to model count data.

The LMM theory provides us with a robust methodology to work with repeated measures over time, space or both. However, data and the random effects must follow a multivariate normal pdf, which means that measurements taken at different time and space must be normally distributed. The analysis of repeated measures by using LMM methodology implies that we have to estimate matrices of variance-covariance components. The methodology includes two basic steps. First, we must select several possible variance-covariance structures for random factors in our linear model. The selection procedure uses likelihood ratio tests and information criteria. Once we select variance-covariance matrices, we test the fixed effects in our model by also using a test based on the likelihood of the model. There is no more use of the F distribution for testing hypothesis.

The use of the Generalized Linear Model (GLM) theory is based on the principle that there is an *exponential family* of distributions that includes the Binomial, Poisson, Normal, Gamma, and other joint pdf (McCullagh and Nelder 1997). All these specific distributions involve parameters that are specific cases of a more general probability function. The dependent variable follows any specific pdf, however the expected mean value of the variable is transformed by using a "*link function*", and the transformed mean value is the variable expressed as a linear function of quantitative *explanatory (predictor* or *independent)* variables, or qualitative explanatory variables also called *factor*. For any trait, that follows a normal *pdf*, the link function is a relation one to one; there is not actual transformation.

The modern development of the GLM methodology started in the early 80's. At the beginning, developers of the methodology combined in the new methodology the different statistical procedures suitable for non-normally distributed data. The classic use of GLM is still linked to regression analysis and data modeling based on fixed effects. Parameters in the model are derived by using maximum likelihood estimation. Usually, alternative and simplified models are also stated and used to test the adequacy of our model of main interest. Testing this hypothesis is based on likelihood ratio tests and the use of a statistics called the "*deviance*".

Clonal testing involves some new challenges to the use of GLM methodology. There are circumstances when the dependent variable (that follows any member of the exponential family different from the normal *pdf*) must be modeled as a function of fixed (covariates or factors) and random effects. Now, we have a "*generalized linear mixed model*" (GLMM). Therefore, we have to estimate parameters related to the *pdf* of the dependent variable as well as the *pdf* of the random effects in the model.

When we establish a clonal test, we consider the clonal effects as random and assume that they follow a normal *pdf*. If the test is planted in only one location, we usually consider block effects as fixed. However, within-clonal residual effects from two different blocks will be correlated. If we work with traits that are normally distributed, everything is straightforward. We estimate variance-covariances and test hypotheses by using the LMM theory summarized above. However, in our selection program we will also measure characters which are not normally distributed. We will assess the presence or

absence of rust or insect attack in each ramet of a clone in a block, and this trait will follow a binomial distribution. We will also count the number and type of branches per clone in a block, the number of sprouting buds per ramet, or the number of aphids in a particular surface area in a sample of leaves from our clones. As mentioned before, count data follows a Poisson *pdf*.

The real challenge will be the analysis of characters which generate data which are dichotomous (binary), counts, or percentages, and these response variables follow non-normal *pdf*. However, these traits will be explained by random factors that will follow a Normal *pdf*. When we establish a clonal test, we set up an experimental design with a related linear model. We include the clonal effect and reasonably assume it follows a Normal *pdf*. Then, we decide to study traits for which we have good antecedents that the genetic control is explained by the additive effect of several quantitative loci. However, we measure a "*response variable*" which is not quantitative and does not follow a normal *pdf*. How to proceed with the statistical and genetic analysis?

Recent advances in the GLMM theory are allowing us to analyze models that involve response variables and explanatory factors that follow different *pdf*'s of the exponential family. To model the variances and covariances in the GLMM, we will have to use "*generalized estimating equations"* (also called GEE). These equations are based on the same mathematical principals used by the maximum likelihood procedure involved in the GLM and LMM methodology (Hardin and Hilbe 2003). However, the GEE includes more than one *pdf* in the equation system that must be resolved by mathematical procedures based on iteration methods.

References

Falconer DS, Mackay TFC (1996). Introduction to quantitative genetics. 4rd Ed. Longman Sci. And Tech. Harlow, UK. 438 pp.

Hardin, J.W. and Hilbe, J.M. 2003. Generalized estimating equations. Chapman & Hall / CRC. Boca Raton. 222 pp.

Littell RC, Milliken GA, Stroup WW, Wolfinger RD, Schabenberger OL (2006). SAS® System for Mixed Models, Cary, NC: SAS Institute Inc.

Moser EB, Saxton AM, Pezeshki SR (1990). Repeated measures analysis of variance: application to the tree research. Repeated measures analysis of variance: application to tree research. Can. J. For. Res. 20: 524 – 535.

McCullagh, P and Nelder, J.A. 1997. Generalized linear models. Second edition. Chapman & Hall / CRC. Boca Raton. 511 pp.

Stettler, R.F.; Bradshaw, H.D.; Zsuffa, L. 1992. The role of genetic improvement in short rotation forestry. In: Ecophsiology of short rotation forest crops. Ed: Mitchell, C.P.; Ford-Robertson, J.B.; Hinckley, T.; Sennerby-Forsse, L. Chapman & Hall. Pp 285-308.

Zamudio. F.; Wolfinger, R. 2002. Growth Increments and Stability over Time in Fast Growing Forest Tree Species. Canadian Journal of Forest Research. 32 (6): 942-953

Zamudio F, Rozenberg P, Baettyg R, Vergara A, Yañez M, Gantz C 2005. Genetic variation of wood density components in a radiata pine progeny test located in the south of Chile. Annals of Forest Science 62 (2): 105 - 114.

Zamudio, F.; Wolfinger, R.; Stanton, B.; Guerra, F. 2008. The use of linear mixed model theory for the genetic analysis of repeated measures from clonal tests of forest trees. I. A focus on spatially repeated data. Tree Genetics & Genome. 4 (2): 299-313.