

Forest Genomics Research and Development in Canada: Priorities for Developing an Economic Framework

Ilga Porth (porth@mail.ubc.ca)¹, Mark Boyland (Mark.Boyland@NRCan-RNCan.gc.ca)²,
Suborna Ahmed (suborna@alumni.ubc.ca)², Yousry A. El-Kassaby (y.el-kassaby@ubc.ca)^{1,2} and
Gary Bull (gary.bull@ubc.ca)^{2*}

Address: ¹Department of Forest and Conservation Sciences, University of British Columbia,
2424 Main Mall, V6T 1Z4 Vancouver, BC, Canada; ²Department of Forest Resources
Management, University of British Columbia, 2424 Main Mall, V6T 1Z4 Vancouver, BC,
Canada

*Correspondence: Gary Bull
Faculty of Forestry, University of British Columbia
Vancouver, British Columbia, V6T 1Z4
CANADA
Tel. 604 822 1553

ABSTRACT

Forest genomics is a relatively recent research field, and is often poorly understood both by the public and forest managers. Genomics in forestry, an expansion of forest biotechnology, seeks to develop generalized technologies for use in industrial plantations and/or natural forests as well as within process optimization, product development and international trade facilitation. With such tools it is possible to address formerly intractable issues, such as understanding the underpinnings of complex traits for conservation management purposes, improved use of forest trees as carbon sinks, feedstock for biofuels and “green chemistry” through deeper understanding

and effective utilization of forests' natural variation. Diverse end-users could benefit from genomics tools, for example, real-time detection and mapping of known and novel pathogens along with risk assessments to protect forest nurseries and natural forests from invasive pathogens and reduce economic losses associated with diseases. Since 2001, there has been approximately \$123 million investment in Canadian forest genomics research and we thought it would be helpful to summarize the various projects in Canada and the USA and identify the research priorities and potential economic implications, by (a) developing a robust typology of forest sector genomics research relevant to Canadian application, (b) categorizing each initiative for its application potential (commercial; non-commercial), and (c) demonstrating with silvicultural gain, insect resistance, and wood composition themes the application of modeling and economic analysis.

Keywords: Forest genomics, research projects inventory, thematic areas, commercial applications, social and ecological dimensions, economic analysis, pest resistance, silvicultural gain, wood composition, market access, tree improvement, tools

Introduction

Forests have important economic, ecological and social values. Forest trees contribute substantially to global carbon uptake, are an integral part of complex ecosystems, and provide benefits for human wellbeing in the health and recreational sector. Canadian forests are important for the country economy [the forest sector contributed \$19.8 billion to the Canadian gross domestic product (GDP) in 2013] but also for global carbon cycle as Canadian forests constitute roughly 10% of the earth's forests (396,000,000 hectares) (Aukema et al. 2009; Rank and Associates 2013). Therefore, healthy Canadian forests are of paramount importance locally and globally. Through the past decade, the Canadian forest sector has faced strong structural and deep cyclical challenges: low cost competition; digital media; reduced demand through global economic downturn. In addition, major insect devastations in both Eastern and Western Canada by spruce budworm, emerald ash borer, and the mountain pine beetle, respectively, affecting 28,000,000 hectares of forest land between the years 2009 - 2010 alone provided enormous environmental challenges. These challenges have combined to result in a 30% reduction to the sector's contribution to GDP since 2007 (Rank and Associates 2013) and over 130,000 job losses (Kumagai et al. 2010).

Canadian forests are mostly publically owned and therefore are subject to public forest policy with its regulations, legislation and directives. Another characteristic of Canadian forests is slow growth resulting in long rotation cycles of 80 years or longer, yet the rotation cycles are strongly species dependent. Hence, an improvement in selection gains for productivity, wood quality and other value-added characteristics, climate adaptability, insect and pathogen resistance along with shortening tree breeding cycles by genomics-informed early selection of superior genotypes with such advantageous trait characteristics would significantly increase forest

productivity (Costanza and McCord 2009). Therefore, it is worthwhile to undertake efforts to drastically shorten the breeding cycles by implementing DNA fingerprinting methods to achieve the anticipated genetic gain in tree breeding without the need to establish controlled crosses and/or progeny trials (El-Kassaby and Lstiburek 2009; Resende et al. 2012a).

The past decade has seen approximately \$123 million investment in forest genomics research (summarized in the current study, see Supplementary Material), and investment in forest genomics in Canada started in the year 1999 (Kumagai et al. 2010). However, this relatively recent research field, forest genomics, is often poorly understood by the public, including forest managers and is often confused with genetic engineering (GE). This misconception is attributed to poor communication and the dire lack of providing clear distinction between these two areas of research. Genomics technologies in forestry represent an expansion of forest biotechnology. Genomics is defined as the branch of biotechnology that deals with the DNA sequence of the entire organism's genome (chromosome set) and includes two major disciplines; namely, functional (determines the biological function of all genes and their products) and structural (determines the three-dimensional structures of proteins). Genomics' purpose in tree breeding technologies (Bhalerao et al. 2003; Costanza and McCord 2009) is foremost to describe and make efficient use of the abundant natural genetic variation present within undomesticated forest tree populations (Groover 2007) and contribute to an understanding of the genetic architecture underlying (ecological or industrial) traits of interest on a genome-wide scale (Sederoff et al. 2009; Street et al. 2006; Holliday et al. 2008; Porth et al. 2012; Prunier et al. 2013; Verta et al. 2013; McKown et al. 2014; Kirst et al. 2004; Porth et al. 2013a; Beaulieu et al. 2014). While GE is the deliberate alteration of the genetic material of living organisms involving the introduction of alien genes from other taxa resulting in the production of recombinant DNA which in turn is

used to create products (proteins) the original organism is incapable of producing without this alteration (Porth and El-Kassaby 2014). Thus, a distinction between the “improvement” achieved through GE and that of classical tree improvement (*i.e.*, traditional tree breeding) is of great importance as the former deals with introduced alien genetic material while the latter capitalizes on species’ natural variation for capturing gain. In addition, the implementation of genomics research, *i.e.* its economic potential for forestry, can be relatively wide-ranging, encompassing tree breeding and selection to non-breeding opportunities.

Yet, forest genomics can also develop generalized technologies that may be applied to industrial plantations as well as to natural forests. With such tools in hand it may be possible to address new or formerly intractable issues, such as the understanding of the genetic underpinning of complex traits for applications in conservation and management of natural forests, the improved use of forest trees as carbon sinks, and the use of forest trees as biofuels feedstock and in contributions to “green chemistry” (renewable biomaterials) (Sederoff et al. 2009; Groover 2007; Tsang et al. 2007). Thus, genomic basic research can provide both tool and resource development, functional genomics tools in “omics” technologies that relate to high throughput sequencing methods, bioinformatics, databases, and others, *e.g.*, with respect to the development of improved agricultural crop varieties and increased crop productivity using marker-assisted breeding (see for example “Basic Research to Enable Agricultural Development, BREAD” within the National Plant Genome Initiative, USA). Non-tree breeding related benefits from genomics research include diagnostic tools to detect and map known and novel pathogens in real time, provide risk assessments and support phytosanitation efforts. Such genomics-developed tools can generate significant and immediate economic benefits ranging from a reduction in losses due to forest diseases and protecting forest tree nurseries and natural forests from an

invasive pathogen to the support of Canadian forestry trade. Other opportunities for forest genomics with respect to tree breeding and selection are by comparison considered long-term.

Due to the lack of a clear procedure to identify and justify research priorities, “inconsistency of current funding mechanisms” for individual research projects that would best support the forest sector strategy (Rank and Associates 2013) and a lack of economics analyses supporting and driving research priorities, here, we attempted to (a) develop a robust typology of forest sector genomics research initiatives relevant to Canadian application, (b) categorize each initiative for its application potential (commercial and non-commercial, *i.e.* ecological/social, including the estimated length of time distant from real application), and (c) demonstrate with the three themes, silvicultural gain, insect resistance, and wood composition, the application of modeling and economic analysis. Thus, the overall goal of the present study is to apply an economics framework to genomics research to indicate which areas of genomics hold the most promise for contributions to the forest sector.

Determining Thematic Areas in Forest Genomics

In order to support our assessment of the subject areas of funded forest genomics projects undertaken thus far, we reviewed previous approaches taken to categorize the Canadian projects, “tools for application” - markers, diagnostics, and enzymes (Terry Hatton, unpublished), or “Healthy Forests” and “Productive Forests” (Rank and Associates 2013). Our inventory of forest genomics projects comprised 22 Canadian (<http://www.genomecanada.ca/en/centres/>) and 14 US projects (US Department of Energy 2013). Although we are aware that there exist several ways to dissect project subject areas, ultimately we felt that it was prudent to disaggregate forest genomics into different themes (or priority

areas). After a comprehensive review of all project focuses, we identified seven themes, as summarized and defined in Table 1, that we feel best describe the project focuses to which all the 36 projects we had retrieved could be assigned and that are all relevant to the Canadian forest sector (both Canadian and US based forest genomics projects: Table 2; Table S1): silvicultural gain, adaptation, insect resistance, pathogen resistance, wood composition, wood ultrastructure, and market access. We note here, however, that these existing themes within forest genomics projects could not be identified for the ethical, environmental, economic, legal and social impacts (GE₃LS) project sections. GE₃LS represent companion projects to most of the large-scale genomics projects addressing elements of ethical, environmental, economic, legal and social impacts of these genomics projects. We also observed here that, so far, no new emerging themes could be identified, but we do suspect that technology transfer will become one of the key themes in the very near future (see also below).

Table 2 represents the summary of the thematic areas used in other reports evaluating forest genomics. We aimed to provide this comparison to show how well our seven themes align with previous attempts to disaggregate genomics in a meaningful way and the extent to which our approach can be regarded as the most comprehensive while at the same time avoiding redundancy (Table 2). Another way of classifying these forest genomics projects would include the four categories: (1) reduced losses (due to insects and pests; stress-resistance), (2) increased growth, (3) wood quality, and (4) efficiency gains. Such a classification has its advantage when general risk assessment is desired across broader focus areas within forest genomics projects as almost every project focus aligns with almost every theme (Table 2 and Table S1). We included “technology transfer” as a theme as one project has already implemented a technology transfer committee. From this we assume that technology transfer is an emerging theme, which is aligned

with the community engagement priority areas outlined by the Genome BC strategy report (Table 2).

Inventory of Forest Genomics Research Based on Thematic Areas

The thematic areas pursued by these projects are one way to organize the efforts in forest genomics. These seven thematic areas (Table 2) constitute the basis of our forest genomics research inventory. Table 3 provides the summary of all the project themes that characterize the 36 identified projects relevant to the Canadian forest sector and indicates the potential application of the conducted research. The identified themes move along the supply chain from the condition of the forest to forest products to forest products trade and markets. We have chosen not to identify individual projects since our objective was not to target an individual project but to present a comprehensive summary of all the genomics projects undertaken in North American since 2002 with relevance to the Canadian forest sector.

Tables 2 and 3 also indicate the four basic tools for application of genomics as identified in our thematic review: markers, diagnostics, viral proteins (here: insecticides), and enzymes. The potential application of the tools is also summarized and it provides a further breakdown of the themes (Table 3). We also indicated the number of projects undertaken within each theme for both Canada and the USA; for Canada we also used available information to estimate what the budget allocation per theme has been thus far (Table 3). The budget for all Canadian projects adds up to a total of approximately \$123 million (Table S2). In comparison to the Canadian forest genomics initiatives, it is also interesting to note that all US initiatives focused on markers and diagnostics as tools for application thus far along with an exclusive focus on “Productive Forests” as the Genome category (Rank and Associates 2013); Table S1.

Table 3 further summarizes the purpose of application into three categories: upstream, downstream and non-commercial (meaning a social and ecological focus in the forest genomics research) applications. We used the symbols ‘x’ and ‘X’ to indicate the strength of the application in any of those three categories. The next column is our estimate of the years to the time when tools gained from the research could be applied. We chose to use three time periods: (a) 1-5, (b) 6-10, and (c) 10+ years. The 1-5 years should indicate near-term opportunities of application, *i.e.* these tools are close to commercial promise, in other words, they are largely ‘private goods.’ Those thematic areas that are labeled 10+ years are those that represent mostly fundamental research and the justification for such initiatives is based more on generating long-term ‘public good’ benefits; this is particularly evident for the subject area related to silvicultural gain. Naturally, those subject areas labeled as 6-10 years are those closer to commercial application and such shorter timelines of application were identified for certain initiatives related to insect and pathogen resistance and were exclusively associated with risk assessment models.

Finally, looking at the combination of application purpose and the estimated time to actual implementation in a genomics-informed operational forestry (“time to application”) does indicate that very few commercial upstream initiatives will likely be applied in under 10 years and since 2001, this is where the majority of the research funding has been allocated (Table 3). Therefore, it might be more appropriate to consider them to be non-commercial in nature, since the research is likely at the basic research stage and the ability to transform findings into commercial application is simply unknown. For the commercial downstream thematic areas we know that the funds allocated are relatively small and most initiatives undertaken so far are also basic research. Nonetheless, since they are further down the supply chain, it does seem reasonable to assume that these downstream applications, combined with advanced initiatives

related to chemistry and bio-pathways, could yield promising results. The non-commercial application purpose has been to address longer term socio-ecological challenges such as climate change adaptation. Nonetheless, these applications do have an economic impact, even if it is in the very long term, but to our knowledge, no economic assessment results are publicly available.

Figure 1 provides a summary of the estimated timeline of application, and the categories are again estimated to be short-term (1-5 years), medium-term (6-10 years) and long-term (10+ years) for the various themes pursued in genomics research. Here, we also indicated the percentage of the budget allocated for each specific theme within the graphic. Clearly, the budgetary emphasis of research funding application has been on the longer term in the areas of silvicultural gain and insect resistance; see also Table 3 and Table S2 for more details. In general, it appears that the main focus of the research efforts was related to the condition of forests defined as “Healthy Forests” and “Productive Forests” in Genome Canada terminology (Rank and Associates 2013) and as such, the timeline for application in this category was, in general, long-term (Figure 1). This can sometimes be justified given that the forests in Canada are, at large, publicly owned and governments are tackling long-term issues such as timber supply, biodiversity and community stability. The medium-term thematic areas are identified to be non-commercial in nature, and with an emphasis on ecological and/or social issues. However, it is noteworthy that initiatives which are short-term, *i.e.* their main focus is largely commercial in nature and linked to forest products as opposed to forest condition, have received relatively small budgets so far.

In concluding, our results obtained from the disaggregation of forest genomics and the subsequent summary of individual thematic areas (Table 3) are in complete agreement with the Forest Sector Challenges, Genomics Solutions report published by Genome Canada in December

2013 which generally states a “heavy focus on improved breeding opportunities” that represent generational investments, while important, they can provide less short-term benefits (Rank and Associates 2013). With this in mind, a re-thinking related to the strengthening of research emphasis on short-term benefits for the forest sector has been recently advised (Rank and Associates 2013).

Economic Analysis Review of Forest Genomics

To date, only few forest genomics initiatives have had an economic focus as part of their research, consequently, the overall economic claims related to forest genomics thematic areas were largely unsubstantiated. Here, we sought to explore themes where research in economics did occur and found the most useful economic analysis in three themes: (a) insect resistance, (b) silvicultural gain, and (c) wood composition. These themes cover commercial upstream and commercial downstream analysis classes, respectively (Table 3). We also discuss the different challenges that genomics is going to face in terms of its implementation and the economics of genomics, particularly relevant to the commercial upstream applications (wood supply, tree breeding).

Relationship of Genomics to Wood Supply

The economic studies on insect resistance and, associated with it, the timber supply were led by Dr. Olaf Schwab, currently a senior economist with the Canadian Forest Service. Figure 2 indicates for the spruce weevil the potential impact of forest related genomics research on wood supply (*i.e.* avoided losses). Given the long rotation ages of spruce trees, substantial benefits from planting genetically improved weevil-resistant stock will only become relevant 250 years from now (30% losses reduction). Moreover, the (Schwab et al. 2011) study clearly indicated

that the magnitude of the impact of planting genetically improved trees is driven largely by assumptions on variables, such as discount rates, than by the loss of wood supply by the weevil. Given the prevalence of neo-classical economics assumptions in forest decision making, discount rates are the main factor that determines timber supply in the face of using weevil resistant stock. For example, a 25% resistance increase and a 5% discount rate would translate into almost zero dollars of benefit for avoided merchantable volume losses. This was identified as the most unfavorable scenario. In contrast, a substantial 75% resistance increase and a 1% (*i.e.*, lowest assumed) discount rate could potentially translate into a \$1.1 billion benefit due to the avoided merchantable volume losses. Even planting large areas with weevil resistant trees only a small portion of merchantable volume losses can be avoided, and, as mentioned, there is the long lag time to consider. Thus, there are no immediate economic benefits to the forest sector unless we include either the Allowable Cut Effect, or make the entries into landscapes where the avoided losses were particularly important to maintaining timber supply through a pinch point in available stands for harvest. Another uncertainty in the weevil-spruce model is the cumulative merchantable volume losses (non-lethal damages) at the time of harvest. While, defect-free lumber could still be produced or damaged lumber could be used otherwise as alternative feedstock in bioenergy, the uncertainty makes investment planning difficult. Hence, there is a need for a comprehensive cost-benefit analysis as large-scale investments into weevil-resistant stock might bear a substantial risk (Schwab et al. 2011).

Cost Effectiveness of Genomics

The planting of genetically improved trees can have significant financial benefits, provided that the costs associated with generating the planting material are not too high (Petrinovic et al. 2009). Thus, to justify the application of genomics, the costs attached would

266 have to be assigned over a very large sale and further analysis of the benefits of the time savings
267 is warranted. Due to the long rotation cycles in Canadian forestry, conducting economics
268 analyses on “breeding-objective traits” is difficult (Ivković et al. 2010). A preliminary appraisal
269 of the costs and benefits with regards to genomics was recently conducted. These economic
270 studies connected to silvicultural gain were led by Dr. Nancy Gélinas, a professor of forest
271 economics at Laval University. She modified the methods in Dr. Harry Wu’s study on Australian
272 pine (Ivković et al. 2010) for Canadian boreal forest conditions, specifically with a focus on
273 volume and wood quality in Eastern white spruce. Accelerated breeding is able to cut the
274 duration of propagation and breeding phases by half (Resende et al. 2012b). Seedling production
275 costs are higher when genotyping and somatic embryogenesis (SE) are involved in seedling
276 production but will be largely offset by more volume, a shortened timeline and production of
277 higher value products. The integration of SE can optimize the benefits from the time gain for
278 rotation such that nine years can be saved to obtain improved seedlings compared to traditional
279 tree breeding (Gélinas unpublished). The optimal economic rotation age could also be reduced
280 by up to nine years for genetically improved white spruce (Petrinovic et al. 2009). Traditional
281 selection delays the rotation by 20 years. Genomic prediction, an approach originally developed
282 for dairy cattle breeding ((Meuwissen et al. 2001); (VanRaden 2008)), can substantially reduce
283 the entire tree breeding cycle, as the lengthy progeny testing phase is replaced by the prediction
284 of the total genetic value of the non-progeny tested population based on genotypic (DNA
285 marker) information (Grattapaglia and Resende 2011). Here, the effects of all available genome-
286 wide genetic information are simultaneously tested on the trait of interest. This approach is
287 particularly valuable for low heritable traits and late expressing traits, as it allows for accurately
288 evaluating traits at an early age. Additionally, the application of genomic selections offers a new

dimension to selection through substantial increase in selection intensity and thus a corresponding increase in the response to selection which is expected to be even greater than that obtained through the application of genomic selection in traditional breeding programs (as poor performers can be eliminated already at the zygotic embryonic stage). In summary, the introduction of genomic selection with emphases on large selection intensity coupled with SE is expected to produce gains far exceeding those obtained through traditional breeding (El-Kassaby, unpublished). Costs for whole-genome sequencing and genotyping have dropped almost exponentially over the past 10 years (Poland and Rife 2012), yet phenotyping for important wood characteristics is still tedious and costly (Porth et al. 2013b). While in traditional tree breeding every single tested tree requires phenotyping, with genomic selection, in the end, only a fraction of trees needs to be phenotyped for the proper development of accurate genomic-based prediction models as progeny testing can be omitted. Currently, there is no published bio-economic modeling for forest trees that compares the benefits of genomic selection to traditional breeding.

Genomics to Products

The final examples of economic analysis relate to the potential of breeding for lignocellulosic traits as well as improvement of value-added processing of wood fiber. The economic studies for particular aspects of bioethanol production (commercial downstream) were carried out by Dr. Catalin Ristea (Ristea 2014) and Dr. Jamie Stephen (Stephen 2013).

Using poplar as a feedstock, Ristea (2014) compared different land types for biomass production (non-irrigated, irrigated; low, medium, high productivity) and demonstrated that for the idle treatment scenarios the decrease in the net average area used was more substantial (Table

S3), highlighting the impact gained by improvements in both biomass yield and conversion yield.

Stephen (2013) summarized the production costs for bioethanol that relate to lignocellulosic feedstock and enzymes and further included opportunities for genomics research to reduce such costs that relate to feedstock and enzymes (Table S4). For example, the reduction in enzyme loading and of unproductive binding of the cellulase enzyme to lignin is considered key in R&D for the optimization of cellulose hydrolysis. This can be achieved by development of novel cellulase enzymes and custom designed enzyme cocktails for specific feedstocks as well as an optimized feedstock with reduced lignin content which should facilitate feedstock pretreatment and subsequently improve saccharification through reduced cell wall recalcitrance (Van Acker et al. 2013; Rico et al. 2014; Wilkerson et al. 2014). In conclusion, these studies (Ristea 2014; Stephen 2013) suggest that an integration of genomics has high potential in bioprocessing improvement of wood fiber.

Conclusions

The present review on forest genomics attempted to improve rigor and structure to the discussion over the use and application of genomic tools to the forest sector. Given that over \$123 million were spent since 2001, it is appropriate to reflect on what the focus has been in terms of areas of research and its potential application. In order to lay the foundation for a discussion on what the future areas of focus should be this review is intended to provide a summary for decision makers.

So far, relatively little of the genomics budget has been allocated to economic analysis. Given the preponderance of projects that could be considered basic research, it is appropriate, to

reflect on the degree to which future projects will have a different focus on commercial application.

It is fair to say that thus far, the budgetary focus has been on the public forest resources, not the private goods associated with private forest lands or forest products. However, at some point, it needs to be assessed how the science of genomics can be associated more with potential commercial benefits. Balancing resource allocation will become a great challenge as we move forward with forest genomics research. It is important to highlight the new funding initiatives created by Genome Canada and Genome BC that are rooted in involving the “end-user” as (an) integral partner(s) in these funding opportunities. These initiatives have the potential to assess the usefulness of and drive the immediate integration of genomics research to application. These include the User Partnership Program (UPP), Genomic Applications Partnership Program (GAPP) and the Strategic Opportunities Fund for Industry (SOFI) [www.genomebc.ca/research-programs/opportunities/current-funding-competitions/].

Our overall goal was to develop an improved approach to the economic assessment of forest genomics. We contributed to the discussion by:

1. Developing a typology of genomic projects. In our view, our typology is robust and forward looking.

2. Categorizing the existing projects in terms of their potential application (commercial up-stream; commercial down-stream; non-commercial, *i.e.* social and ecological) and the length of time from real application. Clearly, one role of future GE₃LS projects would be to provide an assessment of the economic impacts for these potential applications.

3. Reviewing and commenting on the existing economic analysis conducted so far under three themes: silvicultural gain, insect resistance and wood composition.

356 In many respects, the economic analysis will be challenging since time matters and
357 predicting the commercial application of basic research in economics is notoriously difficult for
358 all researchers concerned. Nonetheless it must be done. To make headway, it will be imperative
359 to develop a systematic approach and systems analysis to encourage a dialogue between
360 geneticists, economists and other scientists.

REFERENCES

- Aukema, B., Bohlmann, J., Bonfils, A.-C., Doucet, D., Elouafi, I., Erbilgin, N., Seguin, A., and Smith, S. 2009.** Canadian Forest Health Genomics: Canadian Strengths Address Forestry Challenges. Canadian Genome Centres.
- Beaulieu, J., Doerksen, T., Clement, S., MacKay, J., and Bousquet, J. 2014.** Accuracy of genomic selection models in a large population of open-pollinated families in white spruce. *Heredity* **113**(4): 343-352.
- Bhalerao, R., Nilsson, O., and Sandberg, G. 2003.** Out of the woods: forest biotechnology enters the genomic era. *Curr Opin Biotechnol* **14**(2): 206-213.
- Costanza, A., and McCord, S. 2009.** Forest Biotechnology and its Responsible Use [online] Institute of Forest Biotechnology. http://www.responsibleuse.org/resources/Forest_Biotechnology_and_its_Responsible_Use.html
- El-Kassaby, Y.A., and Lstiburek, M. 2009.** Breeding without breeding. *Genet Res* **91**(2): 111-120.
- Grattapaglia, D., and Resende, M.D.V. 2011.** Genomic selection in forest tree breeding. *Tree Genet Genomes* **7**(2): 241-255.
- Groover, A.T. 2007.** Will genomics guide a greener forest biotech? *Trends Plant Sci* **12**(6): 234-238.
- Holliday, J.A., Ralph, S.G., White, R., Bohlmann, J., and Aitken, S.N. 2008.** Global monitoring of autumn gene expression within and among phenotypically divergent populations of Sitka spruce (*Picea sitchensis*). *New Phytol* **178**(1): 103-122.
- Ivković, M., Wu, H., and Kumar, S. 2010.** Bio-economic Modelling as a Method for Determining Economic Weights. *Silvae Genet* **59**: 77-90.
- Kirst, M., Myburg, A.A., De Leon, J.P.G., Kirst, M.E., Scott, J., and Sederoff, R. 2004.** Coordinated genetic regulation of growth and lignin revealed by quantitative trait locus analysis of cDNA microarray data in an interspecific backcross of eucalyptus. *Plant Physiol* **135**(4): 2368-2378.
- Kumagai, K., Culver, K., and Castle, D. 2010.** Public-Private Research Collaborations in Canadian Forestry Genomics: Knowledge Management and Innovation. *The Integrated Assessment Journal* **10**(1): 15.
- McKown, A., Klápště, J., Guy, R., Gerald, A., Porth, I., Hannemann, J., Friedmann, M., Muchero, W., Tuskan, G., Ehlting, J., Cronk, Q., El-Kassaby, Y., Mansfield, S., and Douglas, C. 2014.** Genome-wide association implicates numerous genes underlying ecological trait variation in natural populations of *Populus trichocarpa*. *New Phytol.* doi: 10.1111/nph.12815.
- Meuwissen, T.H.E., Hayes, B.J., and Goddard, M.E. 2001.** Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**(4): 1819-1829.
- Petrinovic, J.F., Gelin, N., and Beaulieu, J. 2009.** Benefits of using genetically improved white spruce in Quebec: The forest landowner's viewpoint. *Forest Chron* **85**(4): 571-582.
- Poland, J.A., and Rife, T.W. 2012.** Genotyping-by-Sequencing for Plant Breeding and Genetics. *Plant Genome* **5**(3): 92-102.

- Porth, I., and El-Kassaby, Y.A. 2014.** Current status of the development of genetically modified (GM) forest trees world-wide: a comparison with the development of other GM plants in agriculture. *CAB Reviews* **9**(8): 12.
- Porth, I., Klapste, J., Skyba, O., Friedmann, M.C., Hannemann, J., Ehling, J., El-Kassaby, Y.A., Mansfield, S.D., and Douglas, C.J. 2013a.** Network analysis reveals the relationship among wood properties, gene expression levels and genotypes of natural *Populus trichocarpa* accessions. *New Phytol* **200**(3): 727-742.
- Porth, I., Klapste, J., Skyba, O., Lai, B.S.K., Gerald, A., Muchero, W., Tuskan, G.A., Douglas, C.J., El-Kassaby, Y.A., and Mansfield, S.D. 2013b.** *Populus trichocarpa* cell wall chemistry and ultrastructure trait variation, genetic control and genetic correlations. *New Phytol* **197**(3): 777-790.
- Porth, I., White, R., Jaquish, B., Alfaro, R., Ritland, C., and Ritland, K. 2012.** Genetical Genomics Identifies the Genetic Architecture for Growth and Weevil Resistance in Spruce. *Plos One* **7**(9): e44397.
- Prunier, J., Pelgas, B., Gagnon, F., Desponts, M., Isabel, N., Beaulieu, J., and Bousquet, J. 2013.** The genomic architecture and association genetics of adaptive characters using a candidate SNP approach in boreal black spruce. *BMC Genomics* **14**(1): 368.
- Rank, D., and Associates. 2013.** Forest Sector: Challenges, Genomic Solutions. Genome Canada.
- Resende, M.D.V., Resende, M.F.R., Jr., Sansaloni, C.P., Petroli, C.D., Missiaggia, A.A., Aguiar, A.M., Abad, J.M., Takahashi, E.K., Rosado, A.M., Faria, D.A., Pappas, G.J., Jr., Kilian, A., and Grattapaglia, D. 2012a.** Genomic selection for growth and wood quality in *Eucalyptus*: capturing the missing heritability and accelerating breeding for complex traits in forest trees. *New Phytol* **194**(1):116-128.
- Resende, M.F.R., Jr., Munoz, P., Acosta, J.J., Peter, G.F., Davis, J.M., Grattapaglia, D., Resende, M.D.V., and Kirst, M. 2012b.** Accelerating the domestication of trees using genomic selection: accuracy of prediction models across ages and environments. *New Phytol* **193**(3): 617-624.
- Rico, A., Rencoret, J., del Rio, J.C., Martinez, A.T., and Gutierrez, A. 2014.** Pretreatment with laccase and a phenolic mediator degrades lignin and enhances saccharification of *Eucalyptus* feedstock. *Biotechnol Biofuels* **7**(1): 6.
- Ristea, C. 2014.** Modeling the Net Greenhouse Gas Balance of Projects that Display Gasoline with Wood Ethanol from Short Rotation Tree Plantations. Unpublished draft of PhD dissertation. Forestry, UBC, Vancouver.
- Schwab, O., Maness, T., Bull, G., Welham, C., Seely, B., and Blanco, J. 2011.** Modeling the timber supply impact of introducing weevil-resistant spruce in British Columbia with cellular automata. *Forest Policy Econ* **13**: 61-68.
- Sederoff, R., Myburg, A., and Kirst, M. 2009.** Genomics, domestication, and evolution of forest trees. *Cold Spring Harbor symposia on quantitative biology* **74**: 303-317.
- Stephen, J. 2013.** The Variability of Lignocellulosic Ethanol Production as a Business Endeavour in Canada. Unpublished PhD Dissertation. Forestry, UBC, Vancouver.
- Street, N.R., Skogstrom, O., Sjodin, A., Tucker, J., Rodriguez-Acosta, M., Nilsson, P., Jansson, S., and Taylor, G. 2006.** The genetics and genomics of the drought response in *Populus*. *Plant J* **48**(3): 321-341.
- Tsang, A., Douglas, C., Mohn, B., Bocking, S., Einsiedel, E., Facchini, P., Larouche, A., Mansfield, S., Page, J., Regan, S., Singh, J., Smith, M., and Weselake, R. 2007.**

- Genome Canada-Strategic Research Theme: Securing Canada's Future Bio-Based Economy through Genomics. The Biofuels and Bioproducts Genomics Working Group. Biofuels and Bioproducts Genomics.
- US Department of Energy, O.o.S. 2013.** Plant Feedstock Genomics for Bioenergy Joint Awards 2006-2013 [online]. *Edited by* National Institute of Food and Agriculture United States Department of Agriculture, <http://GenomicScience.energy.gov/research/DOEUSDA/>.
- Van Acker, R., Vanholme, R., Storme, V., Mortimer, J.C., Dupree, P., and Boerjan, W. 2013.** Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in *Arabidopsis thaliana*. *Biotechnol Biofuels* **6**(1): 46.
- VanRaden, P.M. 2008.** Efficient Methods to Compute Genomic Predictions. *J Dairy Sci* **91**(11): 4414-4423.
- Verta, J.-P., Landry, C.R., and MacKay, J.J. 2013.** Are long-lived trees poised for evolutionary change? Single locus effects in the evolution of gene expression networks in spruce. *Molecular Ecology* **22**(9): 2369-2379.
- Wilkerson, C.G., Mansfield, S.D., Lu, F., Withers, S., Park, J.Y., Karlen, S.D., Gonzales-Vigil, E., Padmakshan, D., Unda, F., Rencoret, J., and Ralph, J. 2014.** Monolignol Ferulate Transferase Introduces Chemically Labile Linkages into the Lignin Backbone. *Science* **344**(6179): 90-93.

TABLES

Table 1: Thematic themes identified among 36 forest genomics projects

Theme	Description
Market access	Research allowing (immediate) diagnosis of tree health (pathogen diagnostics).
Climate change adaptation	Research on assessment of the adaptive potential including stress-resistance in forest tree populations; knowledge-based directives (assisted migration).
Insect resistance	Research on bark beetles, weevils, budworm; development of insecticides.
Pathogen resistance	Research on plant pathogens, especially fungi and insect associated pathogens (e.g. MPB associated blue stain fungus).
Wood composition	Research focused on the percentage cellulose versus lignin, but also on altered lignin structure where applicable (easier breakdown of lignocellulosic feedstock for bioethanol production).
Wood ultrastructure	Research on fiber properties important for pulp and paper as well as timber industry; wood density is also related to wood ultrastructure and an important determinant for yield (e.g. biofuel lignocellulosic feedstock).
Silvicultural gain (breeding and selection gain)	Research on tools such as MAS (marker assisted selection); research aims at identifying primarily genetic markers underlying traits of interest; development of diagnostic tools for breeding targets to facilitate market access. It also refers to general efforts to shorten breeding cycles and increasing selection intensity in trees (genomic selection, e.g.); conditions to support improved gain (soil microbial conditions; hybrid breeding); access to improved management tools.

Table 2: Large scale and follow-up Canadian research projects in forest genomics

Theme	Genomic Tools for Application	Genome Canada Category	Genome BC Category	Projects Focus
Silvicultural gain (breeding and selection gain)	markers; diagnostics	Healthy Forests, Productive Forests	Healthy Forests, Productive Forests	Wood Quality; Reduce Losses; Increased Growth; Efficiency Gain
Adaptation	markers; diagnostics	Healthy Forests, Productive Forests	Healthy Forests, Productive Forests	Wood Quality; Reduce Losses; Increased Growth; Efficiency Gain
Increase insect resistance	markers; diagnostics	Healthy Forests	Healthy Forests	Wood Quality; Reduce Losses; Increased Growth; Efficiency Gain
Increase pathogen resistance	markers; diagnostics	Healthy Forests, Productive Forests	Healthy Forests, Productive Forests	Wood Quality; Reduce Losses; Increased Growth; Efficiency Gain
Wood composition	markers; enzymes	Healthy Forests, Productive Forests	Downstream Application	Wood Quality; Reduce Losses; Increased Growth; Efficiency Gain
Wood ultrastructure	markers; diagnostics	Productive Forests	Downstream Application	Wood Quality; Reduce Losses; Increased Growth; Efficiency Gain
Market access	diagnostics	Healthy Forests	Healthy Forests	Reduce Losses
(Technology Transfer)	-	-	Engaging Communities	-

Table 3: A comparative summary of classification of forest genomics projects

Theme	Tools for Application	Potential Application of the Research (can be more than one row per theme)	Total number of Canadian Projects per theme	Approved budget of Canadian Projects (million)	Number of US projects per theme	Application Purpose			Estimated time to application (years), 1-5, 6-10, 10+
						Commercial up-stream	Commercial down-stream	Non-commercial (social and ecological)	
Silvicultural Gain	Markers	Identify fast growing varieties for biofuel feedstock	7	\$2.5	8	x			10+
		Develop genomic methods and tools to enhance genetic selection and breeding		\$22.7		x			10+
		Develop environmental genomic tools for improving forest management practices		\$3.9		x			10+
		Identify desired mutants based on genotype of targeted genes to generate novel varieties suitable as bioethanol				x			10+
		Develop specialized bioenergy tree cultivars				x			10+
		Develop varieties with superior biomass feedstock potential				x			10+
		Understand how endophytic bacteria alter plant growth and productivity, to ultimately manipulate plant performance for feedstock production				x			10+
		Improve hybrid breeding of tree varieties that are used to produce wood and bioenergy				x			10+
		Improve yield by capturing hybrid vigor				x			10+
		Understand phytochrome-mediated responses to competition to maximize carbon capture per unit of land area for increased biomass production				x			10+
	Diagnostics	Develop genomic tools to enhance forest productivity		\$3.8		x			10+
		Identify germplasm with unique genotypes and increased biomass yields for tree breeders				X			10+
		Apply hyper-accelerate breeding using genome-wide selection, develop new breeding strategies				X	x		1-5yrs
Adaptation	Markers	Assess microbial diversity for	7	\$5.2	4	x			10+

Theme	Tools for Application	Potential Application of the Research (can be more than one row per theme)	Total number of Canadian Projects per theme	Approved budget of Canadian Projects (million)	Number of US projects per theme	Application Purpose			Estimated time to application (years), 1-5, 6-10, 10+
						Commercial up-stream	Commercial down-stream	Non-commercial (social and ecological)	
		sustainable use of forest biomass resource							
		Investigate tree interactions with soil microbiome				x			10+
		Investigate plant performance in association with microbial communities				x			10+
	Diagnostics	Investigate physiological traits for adaptation and biomass productivity		\$2.2		x			10+
		Assess the adaptive potential in forest tree populations; provide knowledge-based directives (on assisted migration)		\$5.7		x		X	1-5yrs
		Generate genomic tools to study mycorrhizal ecology				x			10+
		Develop varieties that are abiotic stress tolerant and grow on marginal land				x			10+
		Develop robust biomass productivity under marginal conditions				x			10+
		Provide knowledge about dormancy induction				x			10+
Insect Resistance	Markers	Develop and integrate genomics for ecological risk models	7	\$7.8	0			X	6-10yrs
		Identify forest health markers to support breeding programs		\$25.3		x			10+
	Viral proteins	Generate insecticides using naturally occurring insect viruses		\$4.8		X			1-5yrs
	Diagnostics	Increase insect resistance		\$3.8		X			10+
Pathogen Resistance	Diagnostics	Identify and monitor forest pathogens using DNA-based diagnostic tests	5	\$2.1	0	X	X		1-5yrs
	Markers	Develop and integrate genomics for ecological risk models		\$7.8				X	6-10yrs
		Develop strategies to more quickly detect and monitor rust, and to more effectively prevent infection				X			10+
		Determine pathogen species, identify pest susceptible genotypes in breeding trials		\$2.5		X			10+

Theme	Tools for Application	Potential Application of the Research (can be more than one row per theme)	Total number of Canadian Projects per theme	Approved budget of Canadian Projects (million)	Number of US projects per theme	Application Purpose			Estimated time to application (years), 1-5, 6-10, 10+
						Commercial up-stream	Commercial down-stream	Non-commercial (social and ecological)	
Wood Composition	Markers	Develop allelic markers in marker assisted breeding for accelerated feedstock improvement	4	\$4.7	6	x			10+
		Generate knowledge on genomics of wood formation		\$5.4		x			10+
		Identify alleles with breeding values				x			10+
		Generate knowledge regarding the potential of the protein-protein interactions relevant to biomass production				x			10+
		Generate knowledgebase about genes for effective manipulation of lignocellulosic traits to facilitate ethanol production				x			10+
		Identify biochemical functions of acyltransferases in polysaccharide acetylation, lignol biosynthesis, and phenolic compound modification				x			10+
	Diagnostics	Generate cultivars to produce high energy yields ready for deployment				X	x		1-5yrs
	Enzymes	Identify fungal enzymes for breakdown of wood		\$2.2		x	x		1-5yrs
		Isolate and identify novel enzymes for biorefining to cellulosic ethanol	4		0	x	x		1-5yrs
Wood Ultrastructure	Markers	Develop allelic markers in marker assisted breeding for accelerated feedstock improvement		\$4.7		x			10+
	Diagnostics	Develop diagnostic markers associated to high stiffness and dimensional stability		\$3.8		x			10+
Market access	Diagnostics	Identify and monitor forest pathogens using DNA-based diagnostic (phytosanitary status) tests	1	\$2.1	0	X	X		1-5yrs

FIGURE CAPTION

Figure 1: Graphic summary of the emphasis of the funded genomic projects in Canada based on the estimated time to application of Canadian projects per defined theme.

Figure 2: Effect of genetic resistance to the spruce weevil on merchantable losses (Schwab et al. 2011)

Supplementary Material

Table S1. Inventory of thematic areas: 14 US research projects in Forest Genomics

Table S1 represents the number of USA projects that fall under the 7 themes defined in this study. Also, these projects were aligned with four categories: (1) reduced losses (due to insects and pests; stress-resistance), (2) increased growth, (3) wood quality, and (4) efficiency gains as the project focus and with “Productive Forests” as the Genome category.

Theme	Tools for Application	Category	Project Focus
Silvicultural gain (breeding and selection gain)	markers; diagnostics	Productive Forests	Efficiency Gain, Wood Quality, Increased Growth, Reduce Losses
Adaptation	markers; diagnostics	Productive Forests	Efficiency Gain, Reduce Losses
Increase insect resistance	-	-	-
Increases pathogen resistance	-	-	-
Wood composition	markers; diagnostics	Productive Forests	Efficiency Gain, Wood Quality, Increased Growth
Wood ultrastructure	-	-	-
Market access	-	-	-

Table S2. Estimated time to application and percentage of allocated budget of Canadian projects

Investment into 22 Canadian projects was merged together by defined 7 themes in Table S2. The time of application categories are estimated to be: short-term: 1-5 yrs, medium: 6-10 yrs and long-term: 10+. The amount of budget was equally divided into number of areas it covered, if a project covered more than one theme. See also Figure 1 for a graphical representation of this table.

Theme	Tools for Application	Estimated time to application (years), 1-5, 6-10, 10+	Approved budget of Canadian Projects (million)	Budget (%)
Silviculture Gain	Markers	10+	\$2.5	2.03
		10+	\$22.7	18.46
		10+	\$3.9	3.17
	Diagnostics	10+	\$3.8	3.09
Adaptation	Diagnostics	10+	\$5.2	4.23
	Markers	10+	\$2.2	1.79
		1-5yrs	\$5.7	4.63
Insect Resistance	Markers	6-10yrs	\$7.8	6.34
		10+	\$25.3	20.57
	Viral proteins	1-5yrs	\$4.8	3.90
	Diagnostics	10+	\$3.8	3.09
Pathogen Resistance	Diagnostics	1-5yrs	\$2.1	1.71
	Markers	6-10yrs	\$7.8	6.34
		10+	\$2.5	2.03
Wood Composition	Markers	10+	\$4.7	3.82
		10+	\$5.4	4.39
	Enzymes	1-5yrs	\$2.2	1.79
Wood Ultrastructure	Markers	10+	\$4.7	3.82
	Diagnostics	10+	\$3.8	3.09
Market access	Diagnostics	1-5yrs	\$2.1	1.71
SUM			\$123	100

Table S3: C-BOS model results by land unit scenario: improvements in both biomass yield and conversion yield; idle treatment allowed (Ristea 2014)

Three of the land types were non-irrigated (LU1, LU3, and LU5) and three were irrigated (LU2, LU4, and LU6). For the non-irrigated types, LU1 had high productivity, LU3 medium productivity, and LU5 low productivity. Similarly for the irrigated types, LU2 had a high productivity, LU4 medium productivity, and LU6 low productivity.

Land unit type	Max annual area	Net average area	Site prep. unit cost	Land rent unit cost	Biomass production unit cost	Biomass transport unit cost	Harvested biomass (stem+bark+branch)
	[ha/yr]	[ha/yr]	[\$/t]	[\$/t]	[\$/t]	[\$/t]	[t]
LU1	0	0	0	0	0	0	0
LU2	325	144	0	11.58	93.21	45.97	166,670
LU3	71,443	59,288	0.29	12.84	77.11	42.89	49,647,185
LU4	0	0	0	0	0	0	0
LU5	14,612	4,515	1.00	12.47	94.41	28.67	2,919,349
LU6	0	0	0	0	0	0	0

Table S4: Lignocellulosic ethanol production cost for base case 200,000 bdt/yr biomass processing facility (modified after Jamie Stephen (Stephen 2013))

FPU...filter paper units; FPU g⁻¹ protein: protein activity

bdt...bone dry ton

† total costs for enzymes: at 600 FPU g⁻¹ protein activity, 20 FPU g⁻¹ cellulose, assuming 42% feedstock cellulose content, and yield of 250 L EtOH bdt⁻¹ biomass requires 56 g protein per L EtOH; 10 FPU g⁻¹ cellulose loading results in \$4.40 kg⁻¹ protein

Facility Overview		Research Opportunities for Cost Contribution Reduction (potential Genomics contribution)
Plant Capacity	75.1 ML yr ⁻¹	-
Ethanol Yield	321 L EtOH bdt ⁻¹	Wood quality (composition and ultrastructure)
Unit Installed Cost	\$2.15 L ⁻¹ at yearly capacity	-
Total Capital Cost	\$161,500,000	-
Feedstock		
Feedstock Cost	\$95.87 bdt⁻¹	Improved productivity of woody feedstocks through selective breeding or genetic engineering
Non-Feedstock Operating Costs		
Enzymes	\$2.20 kg⁻¹ protein† ; 600 FPU g ⁻¹ protein; 20 FPU g ⁻¹ cellulose	Identification or development of novel cellulase enzymes; custom designed enzyme cocktails for specific feedstocks
Revenues		
Specific revenue (year 1)	\$0.59 L ⁻¹	Co-product credit (lignin and C5 sugar-derived products)