A genome-scale metabolic model of potato late blight suggests a photosynthesis suppression mechanism

Kelly Botero1,3, Silvia Restrepo2 and Andres Pinzón1*

From Selected articles from the IV Colombian Congress on Bioinformatics and Computational Biology & VIII International Conference on Bioinformatics SolBio 2017
Santiago de Cali, Colombia. 13-15 September 2017

Abstract
Background: Phytophthora infestans is a plant pathogen that causes an important plant disease known as late blight in potato plants (Solanum tuberosum) and several other solanaceous hosts. This disease is the main factor affecting potato crop production worldwide. In spite of the importance of the disease, the molecular mechanisms underlying the compatibility between the pathogen and its hosts are still unknown.

Results: To explain the metabolic response of late blight, specifically photosynthesis inhibition in infected plants, we reconstructed a genome-scale metabolic network of the S. tuberosum leaf, PstM1. This metabolic network simulates the effect of this disease in the leaf metabolism. PstM1 accounts for 2751 genes, 1113 metabolic functions, 1773 gene-protein-reaction associations and 1938 metabolites involved in 2072 reactions. The optimization of the model for biomass synthesis maximization in three infection time points suggested a suppression of the photosynthetic capacity related to the decrease of metabolic flux in light reactions and carbon fixation reactions. In addition, a variation pattern in the flux of carboxylation to oxygenation reactions catalyzed by RuBisCO was also identified, likely to be associated to a defense response in the compatible interaction between P. infestans and S. tuberosum.

Conclusions: In this work, we introduced simultaneously the first metabolic network of S. tuberosum and the first genome-scale metabolic model of the compatible interaction of a plant with P. infestans.

Keywords: Compatible interaction, Phytophthora infestans, Solanum tuberosum, Systems biology, Metabolic reconstruction, Flux balance analysis

Background
Plants use various strategies to resist pathogen attack and to avoid the development of diseases, such as, massive transcriptional reprogramming, production of reactive oxygen species (ROS), reinforcements of the cell wall and synthesis of antimicrobial secondary metabolites and pathogenesis-related proteins involved in resistance [1–3]. In those cases when plants resist pathogen attack, plants they are considered resistant to that specific disease, and pathogens are referred to as avirulent on that plant. This type of plant-pathogen interaction is thus known as incompatible. On the other hand, sometimes plant pathogens develop strategies to evade plant defense responses, become virulent and establish the disease. In those cases, the plant-pathogen interaction is known as compatible and the infected plant is considered non-resistant or susceptible [4, 5]. To some plant pathogens, these evasion mechanisms are, at least, partially known [1, 6], while to some others, these mechanisms remain unknown. This latter is the case of Phytophthora infestans, one of the most destructive...
pathogens of *Solanum tuberosum*. This susceptibility is triggered by effector proteins of the pathogen [1], some of which are inhibitors that target defense-related proteins, and processes such as programmed cell death in plants [7].

*P. infestans* is a hemi-biotrophic plant pathogen causing the disease known as late blight of potato, which was the same plant disease responsible for the Irish famine in the mid-nineteenth century [8]. *S. tuberosum* is the fourth leading food crop (after corn, rice, and wheat) for human consumption [9], and it is one of the most produced crops worldwide. Historically, potato late blight is the main factor affecting potato crop production [10, 11]. After its first pandemic in the middle of the nineteenth century, *P. infestans* has remained as the most destructive pathogen in plantations of this crop, leading to annual losses that would have been enough to feed several hundred million people [12]. The economic value of this loss and the cost of crop protection are estimated at $6.7 billion a year [13].

Successful biological systems analysis requires that we understand the functional interactions between the key components of biological organisms and how these interactions change in disease states. Computational modeling has been widely used in order to understand the complexity that arises from the understanding of such interactions, by means of the application of mathematics, physics, and computer science. Computational models allow us to understand how new properties emerge from the interconnection between system components and thus to study their behavior in response to stimuli and changes in the environment [14]. Even though computational modeling of potato late blight has been studied since the early 1950s, it was not until a few years ago that important molecular information to feed these models was accessible to the research community. Typically, this information has been gathered by heterogeneous approaches, which have covered several research fields such as structural and comparative genomics, protein-protein interactions and differential gene expression [15]. In particular, in the field of gene expression, several investigations have been carried out using microarray based assays on cDNA clones [16], transcriptome analysis through DeepSAGE [17] and RNA-Seq analysis [18].

These approaches have generated large amounts of data, allowing us to understand, to a certain extent, some of the main mechanisms involved in potato late blight. Nevertheless, due to the heterogeneity of this information, a full view of the disease is still fragmented. For this reason, and to get a better understanding of the molecular mechanisms underlying the compatible interaction inside the host cell, it is highly suitable and necessary to gather and integrate this information into a single computational model of the disease. This model should capture the time-dependent nature of this biological system, integrate various ranges of spatial and temporal scales, and allow us to explore the possible molecular mechanisms underlying the compatible interaction.

To this end, in recent years, the scientific community has made multiple efforts to integrate functional genomic characterization and biochemical knowledge into models known as genomic-scale metabolic reconstructions (GEMR). These models are based on a detailed knowledge of the system’s individual components (functional annotation), to allow the reconstruction of the system through a bottom-up approach. The aim is to understand the properties that arise from the interactions between the metabolic system’s components in response to environmental stimuli, and their behavior through time [19–22].

During the past years, this approach has been widely implemented in plants, with the development of GEMRs for several species, including *Arabidopsis thaliana* [23–26], *Hordeum vulgare* [27], *Zea mays* [28–30], *Sorghum bicolor* [28], *Saccharum officinarum* [28], *Brassica napus* [31, 32], *Oryza sativa* [33], and *Solanum lycopersicum* [34]. In general, the development of these models can be synthesized into four fundamental steps: 1) Automatic reconstruction based on a genome annotation and biochemical databases. 2) Manual refinement of the reconstruction through a literature review with the aid of biochemical and metabolic databases. In this step, each gene and each reaction are verified, so that they are correctly located and connected. 3) Mathematical and computational formalization of the biochemical information, where the system’s specific conditions and limits are defined. This model is validated through multiple iterations and is used to prospectively simulate the system’s phenotypic behavior. 4) Verification, evaluation, and validation of the reconstruction. In other words, if we want, for example, to simulate the production of biomass precursors in an organism, during this step the reliability of the model to correctly portrait this is then evaluated. Often, this evaluation leads to the identification of incorrect metabolic functions in the reconstruction (network gaps), which are once again evaluated by steps 2 and 3. Therefore, the whole process is iterative and the model is generally susceptible to being continually refined [35].

Once the mathematical model is generated in step 3, the phenotypic predictions of the organism’s metabolic activity can be established through modelling methods based on restrictions, which include an in silico simulation of the model under inferred metabolic objectives and a set of restrictions that represent genetic or
environmental conditions [36, 37]. A widely used method is flux balance analysis (FBA), which performs the linear optimization of a metabolic network in steady-state, in order to predict an optimal set of flux values that are coherent with the maximization or minimization of the chosen objective (which depends on the purpose of the study) and under restrictions imposed on the reaction fluxes [38].

Given that genomic information used to generate the metabolic reconstructions does not consider the real expression of each gene or the subcellular localization of gene products, flux restrictions based on different “omics” data (such as transcriptomics, proteomics, and metabolomics) must be integrated into the GEMRs, in order to recreate specific metabolic phenotypes [30, 39, 40]. This approach has a powerful scope for gaining knowledge of the molecular and biochemical mechanisms of plants under specific environmental or genetic conditions. In addition, this approach allow us to contextualize high-throughput data, as well as to guide hypothesis-driven discovery or to identify novel network properties [21].

In this study, we report the first genome-scale reconstruction of S. tuberosum and generate a genome-scale metabolic model of the compatible interaction between S. tuberosum and P. infestans, through the incorporation of expression data of the pathosystem into the model. Previous work on the compatible interaction had already identified a decrease in photosynthetic activity of infected plants, however, the underpinnings behind it remain unknown [16]. Hence, this model was aimed to follow up on those findings and to try to understand the molecular bases of this typical photosynthesis turn off during plant disease. Furthermore, the model will be useful for understanding other molecular mechanisms involved during the pathosystem activity and for proposing novel directed hypotheses, guiding research to conduct metabolic engineering in the plant, and identifying emerging properties of the compatible interaction, which otherwise could not be observed through the study of individual molecules and processes.

Results and discussion

The reconstruction of the genome-scale metabolic model followed six major steps: (1) automatic reconstruction of draft network via homology searches for the identification of metabolic activities and biochemical reactions; (2) manual and semiautomatic refinement of the reconstruction (3) establishment of gene-protein-reaction (GPR) association; (4) generation of a genome-scale metabolic model in steady state; (5) incorporation of gene expression data of the compatible interaction into the metabolic model; (6) flux balance analysis using a pre-defined objective function. In addition, model predictions were contrasted against experimental observations.

Metabolic reconstruction

The draft network included 1288 reactions and 1482 metabolites, 217 of which were dead-ends that decreased the metabolic network’s connectivity. A manual search of the dead-ends metabolites allowed us to identify 448 biochemical reactions, with biological and/or genomic evidence for potato, which refined 40 metabolic pathways totally or partially, including glycolysis/gluconeogenesis, pyruvate metabolism, glyoxylate and dicarboxylate metabolism (photorespiration module), oxidative phosphorylation, carbon fixation in photosynthetic organisms, amongst others (Additional file 1).

Additionally, and based on previous reports, we manually reviewed complete or partial pathways that induced a defense response in the plant, such as those related to accumulation of salicylic acid (ubiquinone and other terpenoid-quinone biosynthesis) and jasmonic acid (alpha-linolenic acid metabolism) [41], as well as other pathways related to PAMP (pathogen associated molecular patterns) signaling cascades (Additional file 1). By implementing automatic-specific-organism gap filling and semiautomatic-specific-organism gap find and gap fill, we were able to include 508 reactions in the reconstruction. The manual refinement and semiautomatic processes of the reconstruction are summarized in Table 1.

General model properties

Hereby, we present a metabolic model of S. tuberosum in SBML format, hereafter denoted PstM1 (Additional file 2). This model accounts for 2765 genes, 1113 metabolic functions, 1773 GPR associations and 1938 metabolites involved in 87 central and peripheral metabolic pathways and 2072 reactions, of which 1254 could carry a non-zero flux given different objective functions and the specified biomass components. The required

| Table 1 Summary of the manual and semi-automatic refinement connectivity reconstruction |
|-------------------------------------|----------------|----------------|----------------|----------------|
| Refinement                          | New reactions | New metabolites | New pathways   | Complete and partially complete pathways |
| Manual                              | 448           | 215            | 6              | 42             |
| Automatic and Semiautomatic         | 508           | 324            | x              | x              |
| Total                               | 956           | 539            | 6              | 42             |
enzymatic activities (according to Enzyme Commission - EC) to catalyze the reactions in the model are shown in Fig. 1. Among the 2072 reactions, 2059 represent biochemical conversions and 13 represent exchange reactions with the environment, to describe the uptake/secretion of inorganic compounds (NO₃, NO₂⁻, Nitrile, CO₂, O₂, SO₄²⁻, H₂S, PO₄³⁻, O₂, SO₄²⁻, NH₃, Fe²⁺) and light (photon). FBA solutions showed that the model was able to simulate leaf biomass synthesis, which was represented by the conversions of biomass precursors: protein (amino acids), sugars, nucleotides, cell wall (cellulose, lignin precursors) and fatty acids (hexadecanoic acid).

Comparisons with other C3 plant models
The model properties of PstM1 were compared with the versions of the genome-scale metabolic model of tomato (Solanum lycopersicum) [34], a closely-related species to potato, and other two C3 plant species, Arabidopsis thaliana [24, 25, 42] and rice (Oryza sativa) [33]. The PstM1 model contained 185 more reactions representing biochemical conversions than the initial model version of tomato. Furthermore, contained 666, 505 and 336 more total reactions than A. thaliana [24], A. thaliana [25] and rice [33] respectively. This difference is likely because we carried out a comprehensive refinement process and used different databases and resources to improve the connectivity of our metabolic reconstruction network. However, our model had 518 less reactions than the latest version of the A. thaliana model [42], which was updated from a comprehensively annotated database. In addition, our model had 71 fewer reactions than the total reactions set in the tomato model, where the additional reactions represent metabolite transporter processes. The PstM1 is not a compartmentalized model, since the main objective of this work was to evaluate the metabolic mechanisms associated to a possible decrease in photosynthetic capacity. Therefore, we focused on biochemical conversion reactions, such as light reactions and the Calvin cycle, which are catalyzed in only one more compartment than the intracellular transport process.

Our potato model contained fewer blocked reactions than the other C3 plant models compared here. However, the percentage of the reactions used for biomass synthesis in all these models was similar. The percentage for the potato model was 60.5%, compared to 56.5% for the tomato model, 67.3% for the A. thaliana model [25], and 57.3% for the rice model. During the time course of infection with P. infestans that were evaluated in this work, a total of 995 (48%) reactions in the potato model simulations could carry flux from nutrients to the specified biomass components, and in the simulations of the tomato model, between 16.8 and 17.3% of the reactions carried flux to represent three specific metabolic scenarios. It is probable that the reason for a higher number of reactions used in the potato model simulations is that the objective function was the maximization of leaf biomass synthesis, with the purpose of evaluating the possible metabolic scenarios in three moments (0, 1 and 3 days post inoculation) of the compatible interaction between potato and P. infestans. While in the tomato model, the objective function was the total flux minimization for three metabolic scenarios: heterotrophy, phototrophy and photorespiration.

Metabolic phenotype analysis of compatible interaction between S. tuberosum and P. infestans
To determine the metabolic profile for each infection time point, we performed a FBA with flux boundaries established from the gene expression data (see Methods). Our study evidenced a decrease in the objective function...
(biomass synthesis of leaf) during the compatible interaction: 28% 1 day post inoculation (dpi) and 33% at 3 dpi, from the initial time 0 dpi (Fig. 2). These model predictions can be understood if we take into account that for the first interaction days, the pathogen grows well intercellularly and then intracellularly [16, 43] mainly in the leaves [44]. We hypothesize that this pathogenic invasion causes loss of metabolic reactions capacity of the primary metabolism to synthesize macromolecules in the leaf [45–47].

Photosynthesis and carbon fixation metabolism

Photophosphorylation, the synthesis of ATP in photosynthesis, occurs in chloroplasts [48]. Noncyclic photophosphorylation requires photosystem I and photosystem II, and involves water oxidation, oxygen evolution and reduction of an acceptor. In contrast, cyclic photophosphorylation is driven only by photosystem I and works with light of wavelengths over 680 nm, and does not require water oxidation and oxygen evolution [49]. *S. tuberosum* PstM1 includes the definition of the following balanced photosynthetic light reactions taken from literature sources [33, 50, 51]:

\[
7 \text{Photon} + 3 \text{ADP} + 3 \text{Orthophosphate} \rightarrow 7 \text{ATP}
\]

Our results showed that the metabolic flux for noncyclic photophosphorylation slightly diminished throughout the infection time points (Fig. 2) and that the flux for the cyclic photophosphorylation reaction was consistently zero. In this steady-state model we assumed that at 0 dpi the metabolic flux in the noncyclic photophosphorylation reaction synthetized enough ATP and NAPH for CO₂ fixation; therefore, the metabolic flux in the cyclic photophosphorylation was not required [52]. The cyclic photophosphorylation reaction supplies only ATP [53] and its metabolic flux can be zero, because this reaction requires a balance in its input and output of electrons. Electron transport is zero when its components are completely reduced [48]. Our results suggest that during the interaction of the plant with the pathogen, there can be a reduction in the metabolic capacity of noncyclic photophosphorylation, and cyclic photophosphorylation could not supply the ATP deficit. These light reactions in a healthy plant convert light energy into chemical energy that is used for many cellular reactions that contribute to biomass synthesis [54, 55]. The low capacity in light reactions observed here could affect the flux of precursor reactions of biomass and, consequently, affect the capacity to synthesize biomass in the model. However, the flux of noncyclic photophosphorylation decreased less than biomass synthesis, which could be explained by the reduction of energy from photophosphorylation that is consumed by maintenance reactions different from the precursor reactions of biomass. Some of these maintenance reactions are associated with growth, for example, to maintain the electrochemical gradients across the plasma membrane, whereas others are independent of the specific growth rate of the cells [55].

The decrease in photosynthetic ability during compatible interaction between *P. infestans* and *S. tuberosum* has been reported in other studies. For example, in a previous study, the decline in the efficiency of photosystem II [56] and downregulation of genes encoding proteins involved in photosynthesis was shown [16]. During the optimization of a metabolic network in steady state, we are able to predict the reaction flux, which is the overall rate of metabolite conversion [57]. We evaluated the fluxes distribution of the fixation carbon pathway to evaluate the photosynthetic capacity of
the oxidative burst, which consists of the production of the most rapid defense responses against pathogen attack is including programmed cell death (PCD) [63]. One of the actions; 3) the regeneration of RuBP [72]. Hereafter, these stages are described for the three times of infection and the subsequent changes in the metabolic behavior are specified.

The highest flux of carboxylation of RuBP was obtained in the model optimization of 0 dpi (Fig. 3a). At this time, we observed that the RuBisCO enzyme had the highest carboxylase activity. In contrast, at 1 dpi, this enzymatic activity decreased, but increased the oxygenase activity of the RuBisCO, reducing the energy efficiency of photosynthesis in the plant [58, 59]. The subsequent metabolism of glycolate produced by the oxygenation of the RuBP is known as photorespiration, and is associated with high light intensity, uptake of O2 and increased H2O2 production [60, 61]. In the solution of this metabolic model at 1 dpi the uptake of photons is not enhanced, but H2O2 production is increased; the H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63]. One of the most rapid defense responses against pathogen attack is the oxidative burst, which consists of the production of ROS, primarily superoxide and H2O2, at the site of attempted invasion [64]. The quick generation of H2O2 is one of the major and the most stable reactive oxygen species (ROS) that regulates basic acclimation, developmental and defense processes in plants [62], including programmed cell death (PCD) [63].
Fig. 3 Metabolic flux distribution patterns of the fixation carbon pathway and some reactions associated with photorespiration and carbohydrate synthesis. 

(a) Metabolic flux distribution patterns for 0 dpi; 

(b) metabolic flux distribution patterns for 1 dpi; and 

(c) metabolic flux distribution patterns for 3 dpi.
reaction increased incipiently. Our results indicated that the loss of carbon fixation capacity of the RuBisCO and the subsequent decrease in GAP formation from PGA, can be related to the decreased efficiency of starch biosynthesis.

Previous studies have shown that starch biosynthesis regulation is linked with the expression of ADP-Glc pyrophosphorylase (AGPase) in leaves [75, 76]; likewise, AGPase activity is generally activated by PGA [77]. Therefore, AGPase activation or inactivation by PGA production allows starch synthesis to be adjusted in response to changes in photosynthesis [78]. Our model suggests that the metabolic flux of the AGPase activity decreased slightly through infection time points (view Additional file 3). This can be associated with the general decline in the efficiency of starch synthesis in response to the decrease of biosynthesis of GAP from PGA. Although the AGPase activity and starch biosynthesis both tend to decrease, they show particular tendencies that are not comparable between them, especially from 1 to 3 dpi (view Additional file 3). Hence, additional regulatory mechanisms could be required to achieve changes in the rate of starch biosynthesis, as previously reported [75, 79–82].

In the Calvin cycle, the RuBP is both consumed and synthesized [83]. The synthesis of RuBP is known as re-
generation, and involves a series of reactions from GAP, that are energetically favorable and do not consume ATP or NADPH, except in the phosphorylation reaction to transform ribulose 5-phosphate into RuBP [84]. The flux distribution of these reactions was affected during compatible interaction (Fig. 4) mainly at 3 dpi, where a null flux is obtained for three reactions. The net flux of transformation of ribulose 5-phosphate in RuBP reaction decreased approximately by 55% at 1 and 3 dpi compared to 0 dpi. This can be a consequence of both, the decrease in the metabolic flux of its precursor’s reactions in the cycle as well as the decrease of the metabolic capacity of ATP production for the light reactions. In addition, we observed that at 0 and 3 dpi the metabolic fluxes of the regeneration and carboxylation of RuBP are directly proportional. At 1 dpi, the metabolic flux of the regeneration was directly proportional to the sum of the flux for carboxylation and oxygenation. Overall, these 7 results can indicate that in our model the RuBP is synthesized in the same rate that it is consumed. This observation agrees with previous studies that demonstrated in the same rate that it is consumed. This observation agrees with previous studies that demonstrated that in a photosynthetic steady state model, the rate at which RuBP is regenerated [85, 86].

Conclusions
In this work, we simultaneously introduced the first metabolic network of *S. tuberosum* and the first genome-scale metabolic model of the compatible interaction of a plant with *P. infestans*. The metabolic flux of the light reactions and carbon fixation cycle, including photorespiration and starch synthesis, suggests a suppression of the photosynthetic capacity as consequence of the compatible interaction between *P. infestans* and *S. tuberosum*. Our results suggest that the suppression of the photosynthetic capacity could be associated with a quick defense mechanism, which is not sufficient to trigger an effective defense response in the plant. The results shown here are in silico metabolic predictions, which closely match previous studies of plant physiology.

The PstM1 generated in this study can be used to simulate different metabolic scenarios of the potato plant, integrating gene expression data through constraints into fluxes of the reactions. For its part, the genome-scale metabolic model of the compatible interaction of potato plant allows the prediction of other metabolic mechanism involved during patosystem activity and is an useful tool for proposing novel directed hypotheses, guiding research to conduct metabolic engineering in the plant, and identifying emerging properties of the compatible interaction, which otherwise could not be observed through the study of individual molecules and processes. Moreover, PstM1 reconstruction alone can also be informed by new biological data in order to highlight different processes relative to potato metabolism.

Methods
Metabolic reconstruction
An initial automatic draft reconstruction was first created by means of RAVEN [87] software toolbox based on the PGSC potato genome sequence [88]. Both automatic reconstructions were then conciliated and merged through in-house scripting. This merged reconstruction was taken as a starting point for manual reconstruction and then subjected to a manual refinement based on literature and available biological data.

Metabolic reconstruction refinement
The reconstruction refinement stage was divided into six phases: (1) manual gap refinement of the metabolic network and manual refinement of the reversibility and directionality of the reactions; (2) automatic-specific-organism filling; (3) semi-automatic-specific-organism gap find and gap fill; (4) establishment of the directionality and reversibility of the reactions through Gibbs free energy of reaction value (ΔrG°) [89]; (5) inclusion of exchange reactions; and finally (6) definition of gene-protein-reaction associations (Fig. 5).

During manual refinement, the network’s connectivity was verified in a pathway by pathway basis, as well as in a metabolite by metabolite basis [35]. For this, a search for dead-end metabolites in the reconstruction was performed by means of the R package minval [90]; these metabolites correspond to those that are consumed in the set of biochemical reactions, but are not synthesized and vice versa. Based on the identified dead-end metabolites, missing reactions in the reconstruction were manually tracked within the KEGG Pathway Maps database [91, 92], which associates organism-specific genomic information to metabolic pathways maps. For all those reactions not reported for *S. tuberosum* in the KEGG Pathways database, we verified that their catalyzing enzymatic activity was reported in the Plant Metabolic Network (PMN) database specific for potato (PotatoCyc) [93]. In addition, based on biochemical literature, carbohydrate and energetic metabolic pathways were further refined. During the entire manual review process of the network’s connectivity, reactions without any type of biological and/or genomic evidence for this plant were removed, and the reversibility and directionality of 120 reactions were corrected.

In order to verify that the metabolic pathways reported in this reconstruction were present in the metabolism of *S. tuberosum*, we used the KEGGREST R package, which allowed us to obtain all of the metabolic pathways reported for this organism in the KEGG pathways database.
Transformation of reconstructed network into a genome scale metabolic model

From the genome-scale metabolic reconstruction of potato, a SBML version was generated using the Python package COBRApy [100]. Consecutively, the SBML file was imported to the R package Sybil [101], where the whole set of reactions present in metabolic reconstruction was transformed into a steady state metabolic model $S_{ij} v_j = 0$. Where $S_{ij}$ is the entries of the stoichiometric matrix (Additional file 4). Rows in $S$ represent metabolites and columns represent reactions, and $v_j$ is the metabolic flux vector for each reaction. Substrates in the reaction have negative coefficients, while products have positive ones. This matrix also takes into account transport reactions across the cell membrane, which are represented as reactions interconverting intracellular and extracellular compounds. In the metabolic model, the reaction fluxes are subjected to constrains $(-v_{jmin} \leq v_j \leq v_{jmax})$ [25, 27].

Objective function was defined by the biomass synthesis reaction in the leaf $OF_{bioamass}$ (Additional file 4), previously reported in the genome-scale metabolic model of Arabidopsis thaliana [25]. This objective function was mathematically written as a combination of metabolic coefficients of the biomass composition estimated for slow-growing species [102, 103].

Incorporation of gene expression data into the genome-scale metabolic model

Gene expression data of the compatible interactions was obtained from the study of Gyetvai et al. [17]. We used the normalized libraries of the untransformed cultivar Désirée at three infection time points with $P. infestans$ (zero (0 dpi), one (1 dpi) and three (3 dpi) days post inoculation), with three biological replicates. The tag annotation was performed by BLASTN [104]. The source for the tag annotation was the available RNA sequence of “Potato 3.0” [105]. The unique genes and maximum value counts per gene were performed with the script summarization.r [106]. As the last step for building the gene expression database, all refseq gene identifiers were transformed to Entrez identifiers, by means of the R package UniProt.ws [107].

From the previously built gene expression database, we generated an expression set for each infection time point by means of the R package Biobase [108]. The gene expression values were incorporated directly into fluxes constraints of the reactions in the model using the R package ex2flux [109], which implements a method to integrate gene expression values into each GPR associated to the reactions of the model.
Metabolic flux model optimization
FBA represents a constraint-based modeling approach that allows the prediction of metabolic steady-state fluxes, by applying mass balance constraints into a stoichiometric model [38]. The reactions in the model can be represented by a linear system of equations, then, problems such as maximizing specific chemical production or growth can be solved by linear programming [21, 110]. With the purpose of obtaining a computational distribution of metabolic fluxes, FBA was employed assuming maximization of the objective function, by means of the R package Sybil (Eq. 3) maximization of the OFbioamass function to
\[
\sum_{i=1}^{n} S_{ij} v_{ij} = 0
\]
subject to
\[
-\nu_{min} \leq \nu \leq \nu_{max}
\]
To constrain the space of all the possible steady-state flux distributions in the optimization, we imposed thermodynamic constraints to reaction reversibility as well as upper and lower bounds constraints on reactions fluxes from known expression values for the particular enzyme that catalyzes the reaction.

Acknowledgements
To Daniel Osorio for your consulting in the bioinformatics packages implementation.

Funding
The biological data used in this work as input for the analysis was publicly reported in previous research. Funding for the researchers was provided by Universidad Nacional de Colombia, Universidad de los Andes and Centro de Bioinformática y Biología Computacional. The publication cost of this article was funded by Universidad Nacional de Colombia, Universidad de los Andes and Centro de Bioinformática y Biología Computacional.

Availability of data and materials
The datasets generated and/or analysed during the current study are available in the Github repository, https://github.com/kellybotero/PotatoRecon

About this supplement
This article has been published as part of BMC Genomics Volume 19 Supplement 8, 2018: Selected articles from the IV Colombian Congress on Bioinformatics and Computational Biology & VIII International Conference on Bioinformatics SoIBio 2017. The full contents of the supplement are available online at https://bmcgenomics.biomedcentral.com/articles/supplements/volume-19-supplement-8.

Authors' contributions
RS and PA conceived the original screening and research plans; PA supervised the metabolic reconstruction and supervised the analysis. BK performed most of the experiments; PA provided technical assistance to BK. BK designed the experiments and analyzed the data. BK. Wrote the article with contributions from all the authors; PA and RS supervised and complemented the writing. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
1Grupo de Bioinformática y Biología de Sistemas, Universidad Nacional del Colombia - Instituto de Genética, Calle 53 - Carrera 32, Edificio 426, Bogotá, Colombia. 2Laboratorio de Micología y Fitopatología, Universidad de los Andes, Bogotá, Colombia. 3Centro de Bioinformática y Biología Computacional, Manizales, Colombia.

References
1. Jones JDG, Dangl JL. The plant immune system. Nature. 2006;444(7117):323–9. Available from: http://www.nature.com/doifinder/10.1038/nature05286. Cited 13 Nov 2016.
2. Glazebrook J. Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. Annu Rev Phytopathol. 2005;43:25–27.
3. Bellincampi D, Cervone F, Lionetti V. Plant cell wall dynamics and wall-related susceptibility in plant–pathogen interactions. Front Plant Sci. 2014;28(5):228.
4. Kamoun S. A catalogue of the effector Secretome of plant pathogenic oomycetes. Annu Rev Phytopathol. 2006;44(1):41–60 Available from: http://www.annualreviews.org/doi/abs/10.1146/annurev.phyto.44.070505.143436.
5. Tyler BM. Molecular basis of recognition between plant pathogens and their hosts. Annu Rev Phytopathol. 2002;40:137–67 Available from: http://www.annualreviews.org/doi/pdf/10.1146/annurev.phyto.40.120601.125310. Cited 3 Oct 2015.
6. Rojas CM, Senthil-Kumar M, Tzin V, Mysore KS. Regulation of primary plant metabolism during plant-pathogen interactions and its contribution to plant defense. Front Plant Sci. 2014;5(February):1–12 Available from: http://journal.frontiersin.org/article/10.3389/fpls.2014.00017/abstract.
7. Ali A, Alexandersson E, Sandin M, Reijós S, Lennman M, Hedley P, et al. Quantitative proteomics and transcriptomics of potato in response to Phytophthora infestans in compatible and incompatible interactions. BMC Genomics. 2014;15(1):1–18.
8. Ribeiro D, Ribeiro O, Erwin D. Phytophthora diseases worldwide. St. Paul: American Phytopathological Society; 1996. p. 561.
9. Bollinder MA, Kätterer T, Ploepau C, Börjeson G, Parent LE. Net primary productivity and below-ground crop residue inputs for root crops: potato (Solanum tuberosum L.) and sugar beet (Beta vulgaris L). Can J Soil Sci. 2015;95(2):87–93 Available from: http://pubs.acs.org/doi/10.1021/css400141f.
10. Li J, Zhu L, Gu G, Zhan X-B, Lin C-C, Zheng Z-Y. Curdlan β-1,3-glucosylsaccharides induce the defense responses against Phytophthora infestans infection of potato (Solanum tuberosum L. cv. McCain G1) leaf cells.
51. Allen JF. Photosynthesis of ATP-electrons, proton pumps, rotors, and poise. Cell. 2002;110(3):273–6.

52. Allen JF. Cyclic, pseudocyclic and noncyclic phosphorylation: new links in the chain. Trends Plant Sci. 2003;8(8):151–5.

53. Arnon DI. The chloroplast as a functional unit in photosynthesis. In: Ruhland W, Ashby E, Bonner J, Geiger-Huber M, James WD, Lang A, et al., editors. Handbuch der Pflanzenphysiologie. Berlin, Heidelberg: Springer Berlin Heidelberg; 1960. p. 773–829. Available from: http://link.springer.com/10.1007/978-3-642-94978-4_28. Cited 13 Nov 2016.

54. Berg J, Tymoczko J, Stryer L. The Light Reactions of Photosynthesis. In: Berg J, Tymoczko J, Stryer L. Biochemistry 5th edition. New York: WH Freeman; 2002. p. 1050.

55. Stephanopoulos GN, Aristidou AA, Nielsen J. Metabolic engineering. Principles and methodologies. Metabolic engineering. Principles and methodologies. San Diego: Academic Press; 1998. p. 22–66.

56. Eckardt NA. Photoperoxidation Revisited. Plant Cell Online. 2005;17(8):2139 Available from: http://www.ncbi.nlm.nih.gov/pubmed/1598190. Cited 3 Dec 2016.

57. Dangl JL, Jones JDG. Plant pathogens and integrated defense responses to environmental stresses. Acta Biochim Pol. 2007;54(1):39 Available from: http://www.ncbi.nlm.nih.gov/pubmed/17163992. Cited 13 Nov 2016.

58. Dangel JL, Jones JDG. Plant pathogens and integrated defense responses to environmental stresses. Acta Biochim Pol. 2007;54(1):39 Available from: http://www.ncbi.nlm.nih.gov/pubmed/17163992. Cited 13 Nov 2016.

59. Eckardt NA. Photoperoxidation Revisited. Plant Cell Online. 2005;17(8):2139–41 Available from: http://www.plantcell.org/cgi/doi/10.1105/tpc.105.035878. Cited 13 Nov 2016.

60. Noctor G, Arisi A-CM, Jouanin L, Foyer CH. Photorespiratory glycine pyrophosphorylase is activated by posttranslational redox-modification in response to light and to sugars in leaves of Arabidopsis and other plant species. Plant Physiol. 2003;132(2):3838–49 Available from: http://www.ncbi.nlm.nih.gov/pubmed/12972664. Cited 2 Dec 2016.

61. Foyer CH, Noctor G. Oxygen processing in photosynthesis: regulation and metabolic engineering. Principles and methodologies. Metabolic engineering. Principles and methodologies. San Diego: Academic Press; 1998. p. 22–66.

62. Foyer CH, Noctor G. Oxygen processing in photosynthesis: regulation and metabolic engineering. Principles and methodologies. Metabolic engineering. Principles and methodologies. San Diego: Academic Press; 1998. p. 22–66.

63. Klessig DF, Durner J, Noad R, Navarre DA, Wendehenne D, Kumar D, et al. Enhanced turnover of reactive oxygen species by expression of up-regulated ADP-glucose pyrophosphorylase in Arabidopsis thaliana. Plant Cell. 2006;18(11):11–11

64. Kowallik V, Kopka J, Asamizu E, Willmitzer L. Generation of oxygen reactive species by expression of an ADP-glucose pyrophosphorylase in Arabidopsis thaliana. Plant Physiol. 1999;120(1):115–126. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10570216. Cited 13 Nov 2016.

65. Lack A, Evans DE. In: Owen E, editor. Plant biology. 2nd ed. New York: Taylor & Francis group; 2005. p. 351.

66. Lack A, Evans DE. In: Owen E, editor. Plant biology. 2nd ed. New York: Taylor & Francis group; 2005. p. 351.

67. Lark J, Botero et al. BMC Genomics 2018, 19(Suppl 8):863 Page 43 of 99
91. Kanehisa M, Goto S, Hattori M, Aoki-Kinoshita KF, Itoh M, Kawashima S, et al. From genomics to chemical genomics: new developments in KEGG. Nucleic Acids Res. 2006;34(Database issue):D354–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16581885. Cited 15 Nov 2016.

92. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res. 2012;40(D1):D109–14. Available from: http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkr988.

93. Plant Metabolic Network (PMN) [Internet]. Available from: http://pmn.plantcyc.org/POTATO/organism-summary on www.plantcyc.org. Cited 8 May 2016.

94. Tenenbaum D. KEGGREST: Client-side REST access to KEGG. R package version 1.14.0. 2016.

95. Botero K, Osorio D, Gonzalez J, Pinzon A. g2f: Find and Fill Gaps in Metabolic Networks [Internet]. The Comprehensive R Archive Network. 2016 [cited 2016 Sep 1]. Available from: https://cran.r-project.org/web/packages/g2f/index.html

96. Botero K. Sot_g2f [Internet]. Bogota, Colombia; 2016. Available from: https://github.com/kellybotero/PotatoRecon/blob/master/Code/Sot_g2f.R

97. Flamholz A, Noor E, Bar-Even A, Milo R. eQuilibrator--the biochemical thermodynamics calculator. Nucleic Acids Res. 2012;40(Database issue):D770–5. Available from: http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkr988. Cited 13 Nov 2016.

98. Caspi R, Billington R, Ferrer L, Foerster H, Fulcher CA, Keseler IM, et al. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res. 2014;44(D1):D471–80.

99. Botero K, Osorio D. Reversibility [Internet]. Bogota, Colombia; 2016. Available from: https://github.com/kellybotero/PotatoRecon/blob/master/Code/Reversibility.R

100. Niemann GJ, JBM P, Eijkel GB, Poorter H, Boon JJ. Differential chemical allocation and plant adaptation: A Py-MS Study of 24 species differing in relative growth rate. Plant Soil. 1995;175(2):275–89. Available from: http://link.springer.com/10.1007/BF00011364. Cited 13 Nov 13.

101. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215(3):403–10. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2231712. Cited 10 Jul 2014.

102. Gentleman R, Carey V, Morgan M, Falcon S. Bioconductor: Base functions for Bioconductor. 2016.

103. Ebrahim A, Lerman JA, Paisson BO, Hyduke DR. COBRApy : COnstraints-based reconstruction and analysis for Python. 2013

104. Poorter H, Bergkotte M. Chemical composition of 24 wild species differing in relative growth rate. Plant Cell Environ. 1992;15(2):221–9. Available from: http://doi.wiley.com/10.1111/j.1365-3040.1992.tb01476.x. Cited 13 Nov 2016.

105. Gelius-Dietrich G. sybil - Efficient Constrained Based Modelling in R. BMC Syst Biol. 2013;7(1):42. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3843580&tool=pmcentrez&rendertype=abstract.

106. Carlson M. R Interface to UniProt web services. 2014.

107. Flamholz A, Noor E, Bar-Even A, Milo R. eQuilibrator--the biochemical thermodynamics calculator. Nucleic Acids Res. 2012;40(Database issue):D770–5. Available from: http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkr988. Cited 13 Nov 2016.

108. Tomar N, De RK. Comparing methods for metabolic network analysis and an application to metabolic engineering. Gene. 2013;521:1–14. Elsevier B.V. Available from: https://doi.org/10.1016/j.gene.2013.03.017.