



### Reporte Final de Avance y Logros de Proyecto\*

Project Title	Plant Genomics and Capacity Building in Latin America
Fecha de entrega al PNUD	
Semestre	Primer Semestre 2009
Realizado por	Gisella Orjeda

Award ID	00049769 (INT08K03)
Project ID	00060063
UNDAF Outcome(s):	2.1 Fortalecimiento y mejoramiento del desarrollo de mercados sostenibles de producción, capital y trabajo, con énfasis en las áreas de exclusión.
Expected CP Outcome(s):	Diseño y aplicación de esquemas para el desarrollo de micro finanzas en apoyo a pequeñas y micro empresas, considerando el aprovechamiento de las potencialidades locales.
Expected Output(s):	Transferir la tecnología desarrollada por el consorcio internacional de investigadores de América Latina y para reforzar los vínculos entre los científicos senior y jóvenes de América Latina que participan en el proyecto de la papa.
Implementing Partner:	Universidad Peruana Cayetano Heredia
Responsible Parties:	Universidad Peruana Cayetano Heredia
Award Start Date	04/02/2008
Award End Date	31/03/2009
Total Award Amount	US\$ 38,000

#### Breve descripción del Proyecto

The main objective of the proposal is to enhance the capacity in Genomics and Bioinformatics Research in Latin America by networking and replicating training exercises. This is going to be done in the framework of the International Potato genome-sequencing consortium into which all the participating countries are already involved. (See: [www.potatogenome.net](http://www.potatogenome.net)). The whole project aims to increase the flow of information between the three countries.

What we intend with the present proposal is to transfer the technology developed by the international consortium researchers to Latin America and to re-enforce the links among the senior and young Latin-American scientists involved in the potato project.

In the first place, the main young researchers assisting in the potato genome project will be trained for 2 months in both, bioinformatics and wet-lab procedures. The 6 trainees will receive training in groups of two from different countries to stimulate partnership and develop personal contact.

Once the main research assistants are trained, we need to begin planning a workshop in Lima for the partners of the project. The objective is to transfer the latest technology involved in the Potato Genome sequencing project to the maximum number of people. For this reason before the workshop begins, it is necessary for the Peruvian coordinators to do a study travel of 10 days to the University of Wageningen in Holland to be updated on the latest developments

and to planify all the topics of the workshop.

The workshop will be carried out at the Universidad Peruana Cayetano Heredia however it would be ideal to have the two main bioinformaticians of the International consortium to be trainers in Lima.

It is expected that the training activities and the contacts that this fund will produce, will improve the flow of information among participating countries and will allow for a Latin American success sequencing the genome of one of the most important crops of our region.

## I) PRINCIPALES LOGROS DEL PROYECTO

### descripción de los principales logros del proyecto durante el semestre

A Workshops was carried out at the Genomics Research Unit of Laboratorios de Investigación y Desarrollo (LID), Faculty of Sciences, Universidad Peruana Cayetano Heredia. The tile of the workshop was: Potato Genome Sequencing Consortium, Bioinformatics Workshop . It was held between the 2-6 February 2009. All the members of the Latin American Potato Consortium and participants of this project attended. A detailed report, presentations, photos and a complete list of participants is attached to the present report.

2 Chilean scientists attended a meeting and congress in Ireland and stayed at TEAGASC, Ireland with one of the International Potato Genome sequencing consortium members for training after the meeting. A detailed report is attached.

2 Brazilian scientists attended a meeting and congress in 2 European countries and received training at Wageningen University, another of the International members of the consortium. A detailed report is attached.

## II) PERFORMANCE DEL PROYECTO

Indicar el grado de avance de las actividades programadas en el Prodoc durante el semestre

<b>Resultado 1</b>	6 MSc level scientists trained in bioinformatics and wet-lab genomics		
<b>Línea de base</b>	0 MSc level scientists trained in bioinformatics and wet-lab genomics		
<b>Meta</b>	Enhance the capacity in Genomics and Bioinformatics Research in Latin America by networking and replicating training exercises		
<b>Indicador</b>	Number of scientist trained		
<b>Actividad 1.1</b>	<b>Meta</b>	<b>Indicador</b>	<b>% de logro</b>
Training on bioinformatics of 1 junior scientists from each country (Brazil, Chile and Perú) for 2 months, University of Wageningen, The Netherlands	3 scientists trained	Number of scientists trained	100%
<b>Comentarios</b>			
<b>Actividad 1.2</b>	<b>Meta</b>	<b>Indicador</b>	<b>% de logro</b>
Training on wet lab genomics of 1 junior scientists from each country (Brazil, Chile and Perú) for 2 months <University of Wageningen, The Netherlands>	3 scientists trained	Number of scientists trained	100%
<b>Comentarios</b>			
<b>% total de logro del Resultado 1</b> Promedio % de logro actividades en el semestre	<b>2009, semestre 1/2</b>	<b>Comentarios</b>	
	100%	It has been very difficult to identify young scientists with the necessary background. In Chile a call for 2 postdoctoral positions had been announced for a month in the main newspapers and no one has applied for the positions. Finally they have identified one person	

<b>Resultado 2</b>	Three Peruvian coordinators ready to organize a training workshop in Latin America and provide support to the various Latin-American participating scientists		
<b>Línea de base</b>	Only one coordinator has experience in High throughput genomics		
<b>Meta</b>	latest technology involved in the Potato Genome sequencing project transferred to the maximum number of people		
<b>Indicador</b>	Coordinator visit to Wageningen to coordinate the workshop and Workshop expenses and Workshops report		
<b>Actividad 2.1</b>	<b>Meta</b>	<b>Indicador</b>	<b>% de logro</b>
Study visit of 3 Senior Peruvian scientists for 2 weeks <University of	Scientist prepared to organise a Workshop in potato genomics and bioinformatics	Number of visits to the University of Wageningen by Peruvian scientists	100%

Wageningen, The Netherlands>			
<b>Comentarios</b>	2 Peruvian scientists have participated of the Potato genome sequencing consortium meeting in April 2008, One of them also participated in the steering committee meeting of the international Consortium in representation of the Latin American colleagues. Further to this 3 day meeting they have stayed for one more week to discuss the work to be done, and coordinate a workshop that will be carried out in September this year in Lima. One Peruvian scientist is the coordinator of the Latin American Potato genome sequencing consortium and she will go again in October to Wageningen and Cologne where the Solanacea conference SOL2008 will be carried out. Also she will participate in the steering committee meeting of the Potato genome		
<b>%total de logro del Resultado 2</b> Promedio % de logro actividades en el semestre	<b>2008, semestre 1/2</b>	<b>Comentarios</b>	
	100%	Activity finished before this period.	

<b>Resultado 3</b>	Eight people, trained in Lima in bioinformatics for structural genomics		
<b>Línea de base</b>	No people trained in structural genomics		
<b>Meta</b>	Improve the flow of information among participating countries		
<b>Indicador</b>	Workshop expenses, Workshop held		
<b>Actividad 2.1</b>	<b>Meta</b>	<b>Indicador</b>	<b>% de logro</b>
Ten day workshop in Lima. PI from each country and various representatives of Brazil, Chile and Peru for a hands on workshop on the techniques of the project <Universidad Peruana Cayetano Heredia>	Bioinformatics pipeline set up Workshop carried out.	Workshop expenses, Workshop held	100%
<b>Comentarios</b>	The server and computers have been installed and the bioinformatics pipeline has been set up. The Server is being accessed normally and used every day by our Chilean, Argentinean and Brazilian partners. The workshop was carried out successfully. A detailed report in Spanish is attached.		
<b>%total de logro del Resultado 2</b> Promedio % de logro actividades en el semestre	<b>2008, semestre 1/2</b>	<b>Comentarios</b>	
	100	Very successful Workshop. You may see the results looking in Genebak for potato genomic sequences produced by us and the partners using our server and the knowledge obtained during the workshop.	

Performance del Proyecto en el semestre		
<b>Total % de logro de resultados</b>	<b>2008, semestre 1/2</b>	<b>Comentarios</b>
	100	We thank the help of Mr. Alberto Suppa and Ms. Sharon Lu for the succes of the project

### III) PRINCIPALES PROBLEMAS Y OBSTACULOS EN LA IMPLEMENTACION

Describir los principales obstáculos experimentados durante la implementación. Incluir las medidas que han sido adoptadas para enfrentar las dificultades.

Descripción	Medidas adoptadas
The main difficulty has been to identify the right people to be trained as they need to have a particular background and qualifications to be able to profit of the two months in Wageningen	A new call for positions was carried out. And the call will included not only nationals resident in Chile but also Chileans living abroad
The budget envisaged for living expenses in Europe was too small to be real	Two scientists traveled instead of three.

#### IV) LECCIONES APRENDIDAS\*\*

Describir brevemente las lecciones aprendidas durante el periodo considerado.

	Descripción
1	If high level training is involved it is desirable to have the people to be trained already identified when applying for funding.
2	Calculate expenses in a more realistic way
3	

\*\* El presente modelo de Reporte Anual de Proyecto tiene que ser llenado en todas sus partes que representan requisitos mínimos. El reporte puede ser complementado con información adicional a discreción del proyecto.

\* Las lecciones aprendidas se refieren a aquellos aprendizajes, positivos y negativos, basados en la experiencia, relativos a formas de resolver problemas y/o maneras de llevar adelante actividades que puedan servir a otros actores en procesos similares. Las lecciones aprendidas del IFP podrán servir como insumo para las evaluaciones de proyecto, así como para compartir con los demás proyectos que se implementan en conjunto con el Gobierno.

## Reporte de Evento Internacional

### Potato Genome Sequencing Consortium Bioinformatics Workshop

Genomics Research Unit  
Laboratorios de Investigación y Desarrollo (LID), Faculty of Sciences  
Universidad Peruana Cayetano Heredia  
2-6 February 2009

El presente reporte informa sobre la realización de un evento de entrenamiento para científicos peruanos de la UNSCH, INIA y UPCH al cual también asistieron 2 científicos chilenos, 2 brasileños y 2 argentinos financiados por sus propias agencias de financiamiento.

El evento tuvo como instructores a los Drs. Jan de Boer de la Universidad de Wageningen, Holanda y al Sr. Michael Geoffroy de la Universidad del Estado Michigan en Estados Unidos.

El evento comenzó el 2 de febrero y terminó el 6 de febrero sin embargo. La Unidad de genómica recibió a dos científicos de la Universidad Nacional de Huamanga en Ayacucho y a otros científicos Peruanos, del Centro internacional de la Papa, del instituto nacional de Innovación Agraria y participantes del proyecto en UPCH para darles entrenamiento previo al evento con el objetivo que puedan aprovechar al máximo el taller. Asimismo recibimos a un joven científico argentino durante un mes para entrenarlo previamente al taller.

#### **Objetivos**

1. Realización de un taller de entrenamiento en técnicas bioinformáticas y de laboratorio que serán utilizadas en el programa de secuenciamiento de la papa.

A continuación se presenta el cronograma del taller tal como fue entregado a los participantes

### Potato Genome Sequencing Consortium Bioinformatics Workshop

Genomics Research Unit  
Laboratorios de Investigación y Desarrollo (LID), Faculty of Sciences

### ***Monday 2 February***

- **9:15- 9:30 am Presentación de Taller: Gisella Orjeda**
  - o Each participant will present him-herself. Indicating their institute and their participation in the PGSC.
- **9:30- 9:50 am Michael Torres**
  - o Server functionality through putty-Xming and through Fedora.
- **9:50- 10:40 Jan de Boer, Wageningen University, The Netherlands**
  - o AFLP markers; the UHD genetic map
- **11:00- 13:00**
  - o AFLP markers; the UHD genetic map
  - Demos: mapping markers on the UHD map using Atari software  
finding AFLP markers in BAC sequences using Atari software
- **13:00- 14:00 Lunch. Cafeteria UPCH**
- **14:00- 16:00 Jan de Boer**
  - o The BAC library and physical map; FISH  
Demo: mapping AFLP markers on the physical map  
viewing fingerprint contigs with Atari software
- **16:00- 16:30 Coffe Break**
- 
- **16:30- 17:30 Jan de Boer**
  - o Exercise: finding markers and anchored BACs on the physical map website  
viewing fingerprint contigs on physical map website (Gelsynth)

### ***Tuesday 3 February***

- **9:15- 10:30 Jan de Boer**
  - o Use of FPC; BAC selection and tiling path extension

- **10:30- 11:00 Coffe Break**
  
- **11:00- 13:00 Jan de Boer**
  - Exercise: verifying sequenced BACs with BAC end tool
  - contig extension with BAC end tool
  - contig extension with FPC
  
- **13:00- 14:00 Lunch. Cafeteria UPCH**
  
- **14:00- 16:00 Jan de Boer**
  - Influence of heterozygosity on physical map and sequencing
  - Finding BACs in BAC pools
  
- **16:00- 16:30 Coffe Break**
  
- **16:30- 17:30 Jan de Boer**
  - Use of tomato BAC sequence data (relaxed BAC end tool)
  - Exercise: continuation with practicing FPC

### ***Wednesday 4 February***

- **9:15- 10:30 Jan de Boer**
  - Sequencing (principles of Sanger and 454 sequencing) => will be very general
  
- **10:30- 11:00 Coffe Break**
  
- **11:00- 13:00 Jan de Boer**
  - Overview of the PGSC BAC registry web site
  - Submitting sequences to the PGSC: multifasta format
  - Exercise: Mummer and Blast alignments
  - "Manual" ordering of BAC sequence fragments
  
- **13:00- 14:00 Lunch. Cafeteria UPCH**
  
- **14:00- 16:00**
  - Michael Torres: Sequence assembly process (from Sanger data – Phred, Phrap; Consed)



- 16:00- 16:30 Coffe Break
- 16:30- 17:30
- Tour Genomics Research Unit
- UPCH-Laboratorios Investigación y Desarrollo –Facultad de ciencias

### ***Thursday 5 February***

**9:15- 10:30 - Michael Geoffroy, Michigan State University, USA**

- Bambus and Extending Bacs
- 10:30- 11:00 Coffe Break
- 11:00- 13:00 Michael Geoffroy
  - Bambus and Extending Bacs
- 13:00- 14:00 Lunch. Cafeteria UPCH
- 14:00- 16:00 - Michael Geoffroy
  - Assembly exercise
- 16:00- 16:30 Coffe Break
- 16:30- 17:30 Michael Geoffroy
  - Assembly exercise

### ***Friday 6 February***

**- 9:15- 10:30 - Michael Geoffroy Michigan State University**

- Practical exercise using chromosome III sequenced BAcS
- 10:30- 11:00 Coffe Break
- 11:00- 13:00 Michael Geoffroy

- Assembly
- Scaffolding and Extension
  
- **13:00- 14:00 Lunch. Cafeteria UPCH**
  
- **14:00- 17:30 (always coffe break a las 16:00)**
  - **Participants in general:** Free to practice and to ask questions to instructors.
  - **Consortium meeting** –room 1 LID: Gisella Orjeda, Sergio Feingold, Paulo de Melo, Boris Sagredo, (others with invitation).
  
- **19:30 Cocktail at G. Orjeda’s house : Paseo de la Republica 5868, Miraflores**

Country	Funding Agency		E-mail	arrival	departu re	Hotel	Infomation flight					
	Procisur	Other					airline	arriv- time	flight number	dep- time	flight number	Status
			<a href="mailto:sofia.chapper@procisur.org.uy">sofia.chapper@procisur.org.uy</a>									
Holland		Jan de Boer	<a href="mailto:Jan.deBoer@wur.nl">Jan.deBoer@wur.nl</a>	31-Jan	07-Feb	Casa andina	KLM	16:50	743	18:50	744	OK
USA	Michael Geoffroy		<a href="mailto:geoffroy@msu.edu">geoffroy@msu.edu</a>	04-Feb	07-Feb	Casa andina	DELTA A	23:55	349	01:25	350	OK
Brazil		Paulo de Melo	<a href="mailto:pedemelo@globo.com">pedemelo@globo.com</a>	01-Feb	11-Feb	Casa andina	TAM	10:35	JJ 8066	11:45	JJ 8067	OK
Brazil	Orzenil Bonfim Jr.,		<a href="mailto:orzenil.junior@embrapa.br">orzenil.junior@embrapa.br</a>	31-Jan	07-Feb	Casa andina	TACA	21:00	TA 130	10:15	TA 131	OK
Brazil	Roberto Togawa		<a href="mailto:togawa@cenargen.embrapa.br">togawa@cenargen.embrapa.br</a>	31-Jan	07-Feb	Casa andina	TACA	21:00	TA 130	10:15	TA 131	OK
Argentina	Leandro Barreiro		<a href="mailto:leobarreiro82@hotmail.com">leobarreiro82@hotmail.com</a>	14-Jan	07-Feb	Hospedaje UPCH	LAN A	22:55	2428	12:35	4641	OK
Argentina	Sergio Feingold		<a href="mailto:sfeingold@balcarce.inta.gov.ar">sfeingold@balcarce.inta.gov.ar</a>	01-Feb	07-Feb	Casa andina	AR	22:50	1364	00:35	1369	OK
Chile	Boris Sagredo		<a href="mailto:bsagredo@inia.cl">bsagredo@inia.cl</a>	01-Feb	07-Feb	Casa andina	LAN	22:50	LA0530	07:30	LA0531	OK
Chile	Annette Fahrenkrog		<a href="mailto:afahrenkrog@inia.cl">afahrenkrog@inia.cl</a>	01-Feb	07-Feb	Casa andina	LAN	22:50	LA0530	07:30	LA0531	OK
Uruguay	Pilar Zorrilla		<a href="mailto:zorrilla@pasteur.edu.uy">zorrilla@pasteur.edu.uy</a>	01-Feb	10-Feb	Casa andina	TACA	09:40	Taca 045	21:40	Taca ???	
Uruguay	Pablo Peraza		<a href="mailto:pperaza@inia.org.uy">pperaza@inia.org.uy</a>	01-Feb	08-Feb	Casa andina	TACA	09:40	Taca 045	21:40	Taca 041	OK

Country	Institute	E-mail	Name
Peru	UPCH	<a href="mailto:gorjeda@upch.edu.pe">gorjeda@upch.edu.pe</a>	Gisella Orjeda
Peru	UPCH	<a href="mailto:lguzmanfe167@yahoo.es">lguzmanfe167@yahoo.es</a>	Frank Guzman
Peru	UPCH	<a href="mailto:mike.cb.pe@gmail.com">mike.cb.pe@gmail.com</a>	Michael Torres
Peru	UPCH	<a href="mailto:rob_luc@hotmail.com">rob_luc@hotmail.com</a>	Roberto Lozano
Peru	UNMSM	<a href="mailto:olga_ponce_unmsm@yahoo.es">olga_ponce_unmsm@yahoo.es</a>	Olga Ponce
Peru	UNSCH	<a href="mailto:labtomasevich@yahoo.es">labtomasevich@yahoo.es</a>	Tomas Miranda
Peru	UNSCH	<a href="mailto:german012@latinmail.com">german012@latinmail.com</a>	German de la Cruz
Peru	CIP	<a href="mailto:e.mihovilovich@cgiar.org">e.mihovilovich@cgiar.org</a>	Elisa Mihovilovich CIP
Peru	INIA	<a href="mailto:fserna@inia.gob.pe">fserna@inia.gob.pe</a>	Fernando Serna
Peru	INIA	<a href="mailto:kevivanco@gmail.com">kevivanco@gmail.com</a>	Kelly Vivanco INIA
Peru	INIA	<a href="mailto:jbenavides@inia.gob.pe">jbenavides@inia.gob.pe</a>	Jorge Benavides INIA

# Travelling report to Wageningen/Holland and Stockholm/Sweden

## Orzenil Bonfim and Roberto Togawa

### 21-29 June 2009

---

## Prospective visit to Wageningen Genomics Facility

---

**22-26 of June 2009**

We had the opportunity to meet the Bioinformatic group at Wageningen Genomics Facility. Our contact was Dr, Roeland van Ham, Cluster Leader for Applied Bioinformatics. He introduced their facilities and the projects that they are involved. Our particular interest was about the system that was developed to perform BAC genome annotations called Cyrille2 and their strategy for the potato genome analysis.

Dr. Mark Fiers the main author of the Cyrille2 pipeline system [1] was also present on the first day and we had the chance to discuss about further collaborative development on Cyrille2 with more integrated modules.

The hands on was with Erwin Datema one of the developers and co-author of Cyrille2. He showed the system design architecture and the ways to use the different modules. Also he explains how to configure and run the different programs on the pipeline. The practical work was to create a small pipeline using a BAC downloaded from NCBI (Potato Chromosome III) and annotate it. At the end we could visualize the annotations made by the system. Another aspect was the need for the use of a grid computing and this is important for the parallelization to run the annotation process. They are using the Sun Grid Engine [2].

The visit was very productive and we are currently installing the Cyrille2 system in our machines at Embrapa/Brazil. The first step will be the installation of the infrastructure (Sun Grid Engine) and then the actual installation of Cyrille2. Mark Fiers already create a repository for Cyrille2 on the sourceforge website [3] to share the development and we are very excited to see the system working in our facilities.

[1] Mark WEJ Fiers, Ate van der Burgt, Erwin Datema, Joost CW de Groot and Roeland CHJ van Ham. High-throughput bioinformatics with the Cyrille2 pipeline system. *BMC Bioinformatics*, 2008, 9:96.

[2] <http://www.sun.com/software/sge/index.xml>

[3] <http://sourceforge.net/projects/cyrille/>

## 17th Annual International Conference on Intelligent Systems for Molecular Biology

---

**27-28 of June 2009**

17th Annual International Conference on Intelligent Systems for Molecular Biology had place at Stockholm, Sweden, on June 27, and 28. The conference was hosted by the International Society for Computational Biology (ISCB) and the European Conference on Computational Biology (ECCB) and provided a forum regarding the latest developments in bioinformatics and opportunities to participate in special interest group meetings. Embrapa researchers participation took place at the following events:

- Data and Analysis Management Satellite Meeting, June 27. Oral presentations on bioinformatics data sharing, workflow development and reuse, tool development for data mining and datawarehouse, standards and infrastructure for managing data, metadata and grid/cloud computing.
- Metagenomics, Metadata and MetaAnalysis Special Interest Group, June 27. Oral presentations on standards and infrastructure for managing data and metadata towards enhanced discovery of genomic and metagenomic data.
- Short Read Sequencing Special Interest Group, June 28. Oral presentations on sequence assembly from short reads (sequence reads from Next Generation Sequencing Technologies), SNP discovery, variation discovery, RNA sequencing and metagenomics.

Regarding Potato Genome Project activities on sequence assembly and bioinformatics analysis the conference contributed particularly to enhance our knowledge

of most powerful and up to date methodologies to deal with sequencing projects. Due to availability of some special tools discussed in the events which researchers participated, new softwares and approaches could be disseminate within the Consortium.

Reporte al Fondo Pérez Guerrero Fondo Pérez Guerrero proyecto "Genomics and Capacity Building in Latin America.

**Actividad:** Asistencia de 2 científicos chilenos al Taller de Coordinación y entrenamiento del Consorcio Internacional de Secuenciación del Genoma de la Papa.

**Fecha y lugar de realización:** 10 al 13 de junio 2009, Carlow, Irlanda

Los Dres. Boris Sagredo y Nilo Mejía, investigadores del Instituto de Investigaciones Agropecuarias (INIA, Chile), especialistas en genómica vegetal asistieron representando a Chile a un Taller de coordinación y entrenamiento del Consorcio Internacional de Secuenciación del Genoma de la papa conocido por su siglas en inglés PGSC (Potato Genome Sequencing Consortium). Este taller y entrenamiento se realizaron en el Centro de Biotechnology Programme del Crops Research del TEAGASC, Carlow, Irlanda entre los días 10-12 de Junio 2009. Asistieron más de 27 investigadores, representando a los 16 países miembros del PGSC. Adicionalmente, los investigadores se quedaron a un día de capacitación en el TEAGASC, actividad coordinada por el Dr. Dan Milbourne.

En este informe reportamos (I) los alcances del taller de coordinación que fue orientado a discutir e informar sobre los avances y estado de secuenciación del genoma de la papa, y reportamos (II) los detalles del entrenamiento para el genotipado de alto rendimiento para análisis genéticos y su integración con recursos genómicos. Se espera aplicar el genotipado de alto rendimiento para lograr una incorporación efectiva del INIA en la publicación del genoma de la papa que estará lista para fines de este año.

#### I. Taller de coordinación.

Se analizaron los principales avances en la obtención de la secuencia de cada uno de los 12 cromosomas que posee esta especie. Para este proyecto de secuenciación se están utilizando dos individuos diploides, DM y RH. El primero es un doble monohaploide homocigoto obtenido por cultivo de gametos de un individuo *Solanum phureja*, una especie diploide. El segundo un diploide heterocigoto, que tiene un 75% de *S. tuberosum* y un 25% de *S. phureja*.

##### 1. Estrategia de secuenciación BAC x BAC (ejemplo del cromosoma 5)

Los principales logros fueron alcanzados por el grupo holandés del WRU. Los resultados más relevantes se describen a continuación. A partir de un mapa meiótico ultra saturado (Ultra High Density Map) se logró anclar 1.628 clones BAC mediante marcadores de tipo AFLP posicionados en el cromosoma 5, en términos del camino más corto (Minimal tiling path) esto implica 1 a 3 clones BACs anclados por marcador de AFLP. A pesar de esta gran saturación se detectan dos copias de cromosomas representados por Contigs de FPC (Finger Printing Contigs) debido a la heterocigocidad



del genotipo utilizado (RH). Debido a la misma heterocigocidad no homogénea, la librería BAC establecida representa apenas un 65 % del cromosoma. Esta falta de cobertura también se ve reflejada a nivel de las secuencias obtenidas y ensambladas, 47 de 60 Mb. Esta falta de cobertura se debe a la falta de integración entre el mapa genético y mapa físico en regiones homocigotas (altamente similares). Paralelamente, se abordó la secuenciación del genoma completo (Whole Genome Sequencing) del genotipo RH, alcanzando una cobertura de 95x, sin embargo, es necesario un mejor mapa físico (Whole Genome Profiling) antes de integrar el mapa físico (BACs) con las secuencias. Finalmente se discutió acerca de la pertinencia de continuar con los esfuerzos de secuenciación BAC x BAC dado el cambio en las tecnologías de secuenciación de alto rendimiento y dadas las dificultades que representa el genotipo RH.

2. Secuenciación completa del genoma (WGS) de genotipo experimental homocigoto DM.

Mediante la plataforma Solexa del BGI principalmente se alcanzado una cobertura cercana al 60x. El ensamblado de contigs ha permitido ensamblar megacontigs o scaffolds que cubre el 85 % del genoma, y gracias a los BAC End Sequences de los BACs de RH se ha logrado ensamblar Super Scaffolds. Se ha logrado una integración parcial entre el genoma RH y el DM (hasta 95 % de identidad entre regiones génicas), hay ausencia de integración en las regiones ricas en elementos repetitivos. Preliminarmente se han anotado 22 mil genes y se ha detectado un evento de duplicación genómica

**II. Coordinación y entrenamiento para la integración de la secuencia genómica con el mapa genético.**

El Centro Internacional de la Papa, CIP, ha desarrollado una progenie que tiene como uno de sus progenitores al genotipo experimental DM . Esta progenie ((DM x DI) x DI), que cuenta con unos 200 segregantes deberá ser utilizada para mapear caracteres de interés en un mapa de alta densidad, y para integrar el mapa genético con las secuencias genómicas (Scaffolds y Super Scaffolds). Dadas las restricciones de tiempo principalmente, se ha decidido abordar el mapeo genético a partir de un sub conjunto de la población. Este nivel de mapeo será instrumental para integrar el mapa genético con las secuencias genómicas, pero se requerirá abordar el mapeo de alta densidad con toda la progenie para analizar caracteres de interés. Para ambos objetivos se mapearán al menos 1.000 microsatélites (SSRs) identificados en las secuencias genómicas derivadas de la WGS de DM.

Se recibió un entrenamiento teórico para participar de este genotipado a la par de los otros miembros del consorcio que tienen amplia experiencia y lograr resultados en el mismo plazo de tiempo. La realización del genotipado del sub conjunto de la progenie mediante técnicas de alto rendimiento involucra principalmente el uso combinado de electroforesis en geles de poliacrilamida denaturante (PAGE) para detectar los microsatélites polimórficos en la progenie derivada de DM, y el genotipado mediante electroforesis capilar (secuenciador de Applied Biosystems). El rendimiento de la electroforesis capilar se incrementará diseñando una estrategia que permita resolver los SSRs mediante carga en

multiplex (uso combinado de hasta 4 fluoróforos distintos). El objetivo de este entrenamiento es que entre los países sudamericanos (Argentina, Chile y Perú) se alcance el objetivo de genotipar 20 % de los SSRs. Además, la información generada, servirá para asistir a los programas de mejoramiento genético de las diversas instituciones involucradas.

En el caso de Chile en particular, los investigadores aportaran con mas marcadores mediante la incorporación del análisis de Diversity Arrays Technology (DArT) en los experimentos de mapeo de alta resolución. Sin embargo, esta tecnología es efectiva en lograr generar información muy relevante para los objetivos estratégicos de mapeo genético.

### **Resultados y logros:**

Gracias al entrenamiento, y la captación de recursos e infraestructura nos sumaremos al esfuerzo de construir capacidades en el área de la genética genómica y bioinformática en Chile.

Por otro lado, este taller de entrenamiento y coordinación nos permitirá participar en la elaboración y discusión del manuscrito que reportará los alcances de la secuenciación del genoma de la papa. Para hacer más eficiente nuestra participación, los investigadores Chilenos se incorporaron a los sub-grupos de discusión de mapeo, bioinformática y difusión.