

# Challenges and Opportunities in a Mixed Strategy for Completing the Rice Genome

IRGSP Technical Meeting

June 8, 2001

Dick McCombie

CCW

# Goals of Presentation

- Assessment of situation
- Lessons from Arabidopsis project management that can help us navigate successfully
- Challenges presented by current situation
- Advantages presented by current situation
- A proposal for a successful path to the future

# Current Situation

- Majority of project initially being done in draft form
- Significant other resources being devoted to generate finished sequence
- Advantages and disadvantages to each approach, but very unlikely that this situation will change

## View for the next year

- Significant portion of the genome will be in phase II or completed
- Most but possibly not all of what is being drafted in the upcoming will be finished by the group doing the drafting
- At least 100 MB and more likely 150 MB will be drafted between the end of 2002 and sometime in 2003
- There will be an increasing number of gaps in the genome, both within clones and between clones, in all likelihood many of these will represent the most difficult areas of the genome

# Plan by country for the next year

- Japan is doing draft of most of the genome
- France and US will finish as they sequence for the next year
- China, Taiwan and Korea will draft first but finish what they draft eventually

# Lessons from Arabidopsis

- Difficult areas were put off until the end in many cases - this later caused problems
- A group claimed an area they could not finish on time, thus blocking other groups from working on this area
- As a result, the two largest groups in the AGI stopped production finishing well before the genome was completed
- When it was time to finish the part that was claimed but unfinished, the smaller groups did so, but with great difficulty and lower quality
- It was very important ultimately to count-down BACs genome wide

# Decision point - Fall 2002

- French and US groups will have finished several chromosomes
- Much of the genome will be drafted
- The amount of the genome that is uncovered will represent about 1 year of RGP draft production
- Unless a plan is in place for French and US groups to integrate their efforts much before this, they will not be able to contribute efficiently

# Problems that we will encounter

- Transfer of data files from one LIMS system to another will be difficult
- Transfer of templates from one lab to another when both use different bar coding, template preps and plasticware will be difficult if not impossible
- Approximately 3335 BACs need to be finished - current rate of 20-25 month worldwide - at this rate this will require about 11 years

# Advantages of the current situation

- The sequence is being released very rapidly
- Problem areas are being identified in the genome very rapidly since draft sequence will do this
- Time is available to start dealing with the problems in the area sooner rather than later
- Time is available to develop specialized skills and infrastructure within IRGSP for finishing the genome

# How can this help us coordinate the rice project

- Important to use the draft information to identify problem areas AND begin dealing with them
- Important that groups needed for finishing have a path to do so:
  - They finish what they draft OR
  - There is a coordinated plan for them to finish SPECIFIC REGIONS drafted by others

# Proposed changes

- A problem area working group be established to work with drafting groups to identify and assign problem areas
- Chromosome assignments for draft and finishing be disassociated
- Specific technical plans for finishing sequences drafted by other groups be developed **AND TESTED** by a second working group