

2006. 12. 12. NationalGrid Tae Hoe Koo

## Contents

- 1. The Project Introduction
- 2. Biochip Analysis Process
- 3. The Application of Distribution Analysis

# 1-1. Final Target

#### Biochip Distribution Analysis System

(Medical Diagnosis Analysis Base, Commercialization)

# And Technology & Contimization of BT, IT Technology

Normal
Ultra-small
Bio scanner

Unified
Analysis
of
Bio information

of
Distribution
Environment

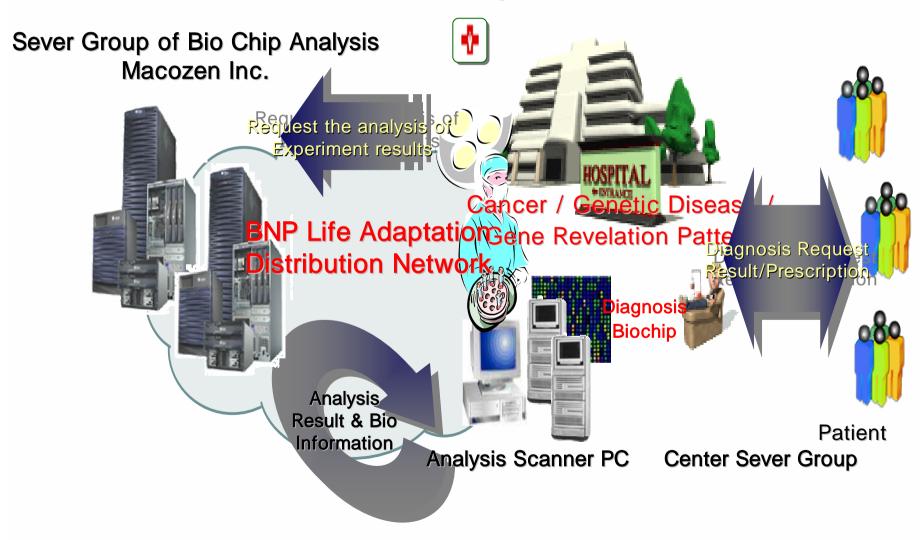
Analysis
Application of
Bionics
Algorism

Synergy
Effects
Of H/W, S/W
Unified Solution

## 1-2. Concept Map of Development System

Project Inroduction

#### Hospital/Diagnosis Center/Research Center

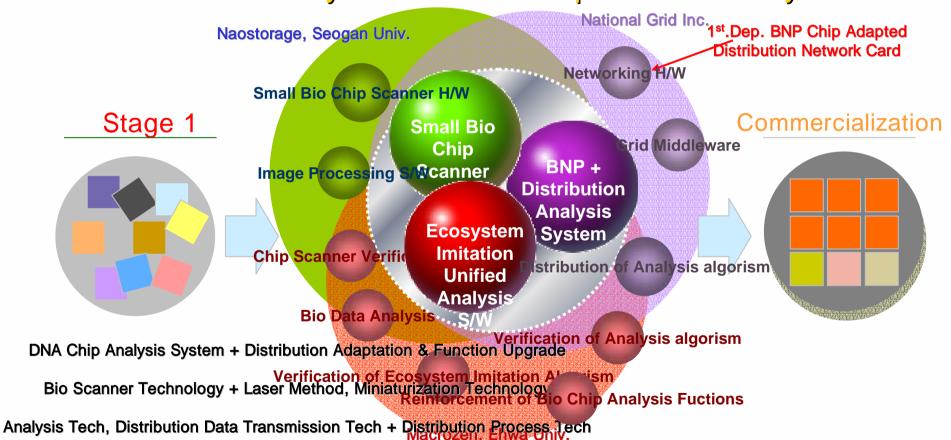


## 1-3. Stage & Related Institutions

Project Inroduction

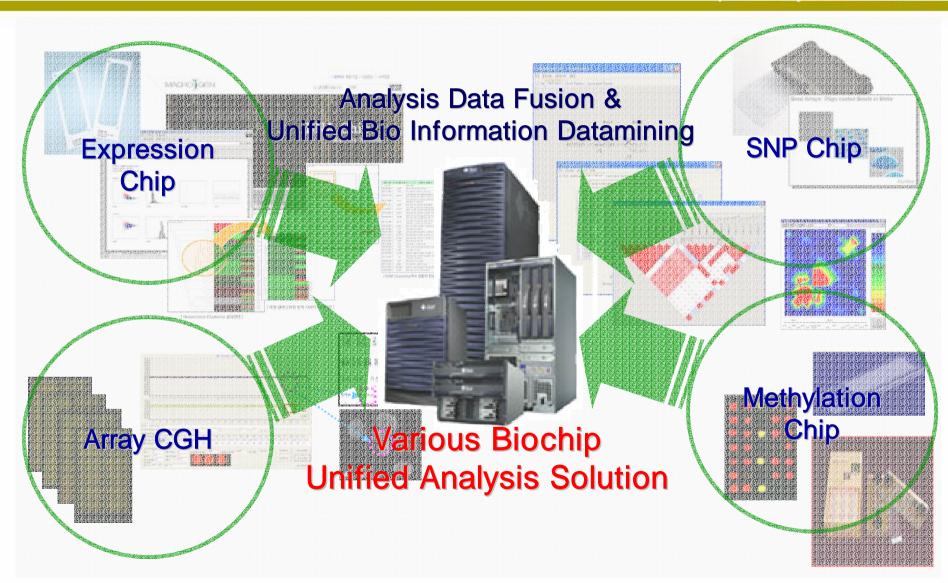
#### Stage 2

#### Life Network System of Bio Chip Unified Analysis



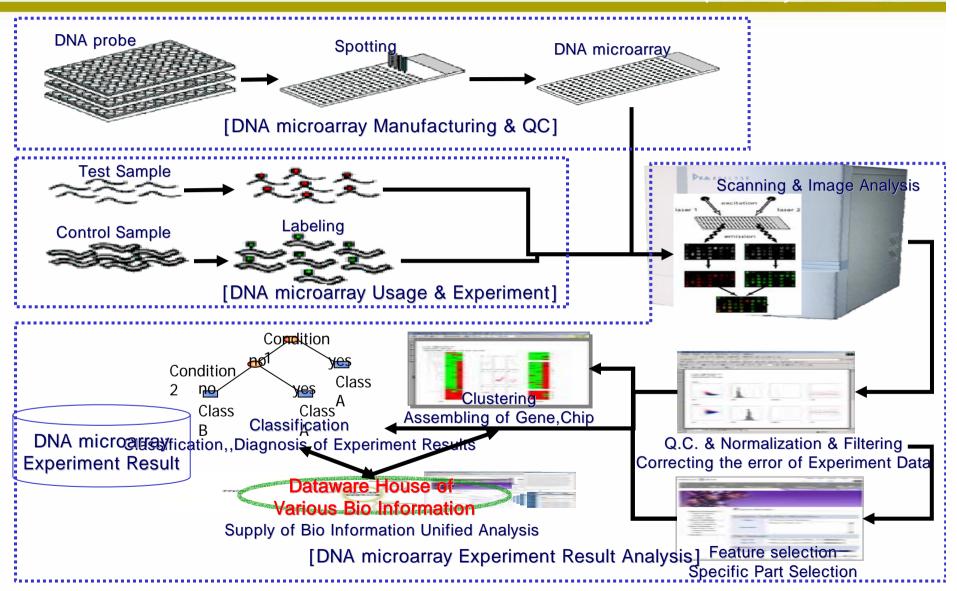
#### Biochip Analysis Process

## 2-1. Biochip

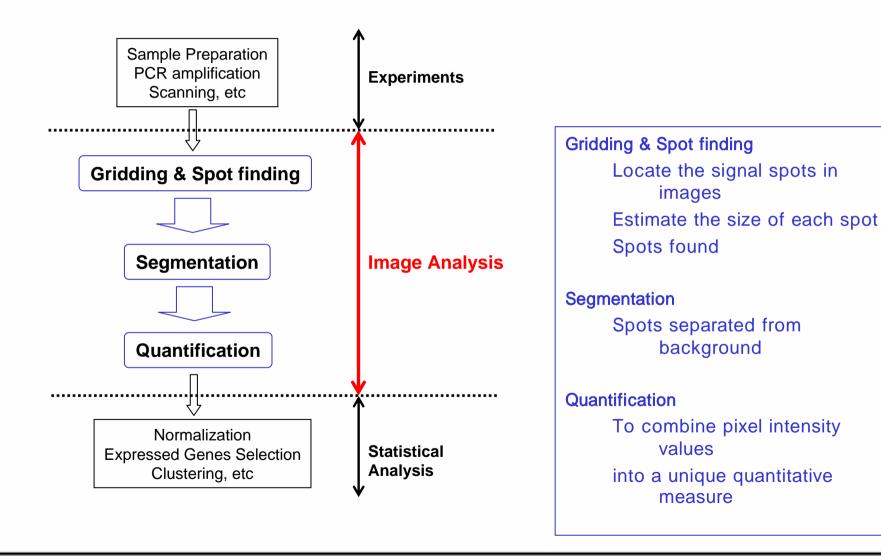


#### 2-2. Total Process

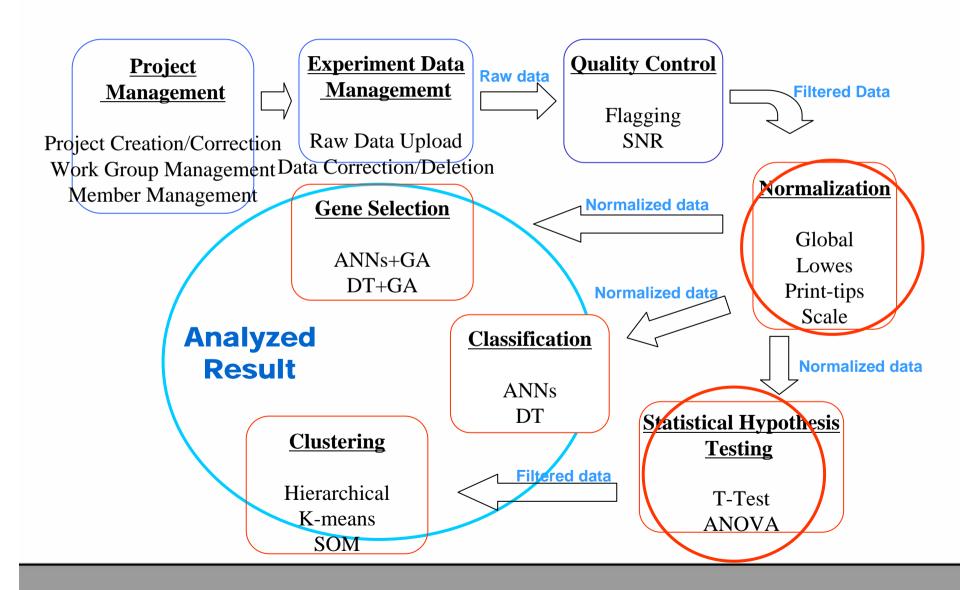
#### Biochip Analysis Process



## 2-3. Image Processing

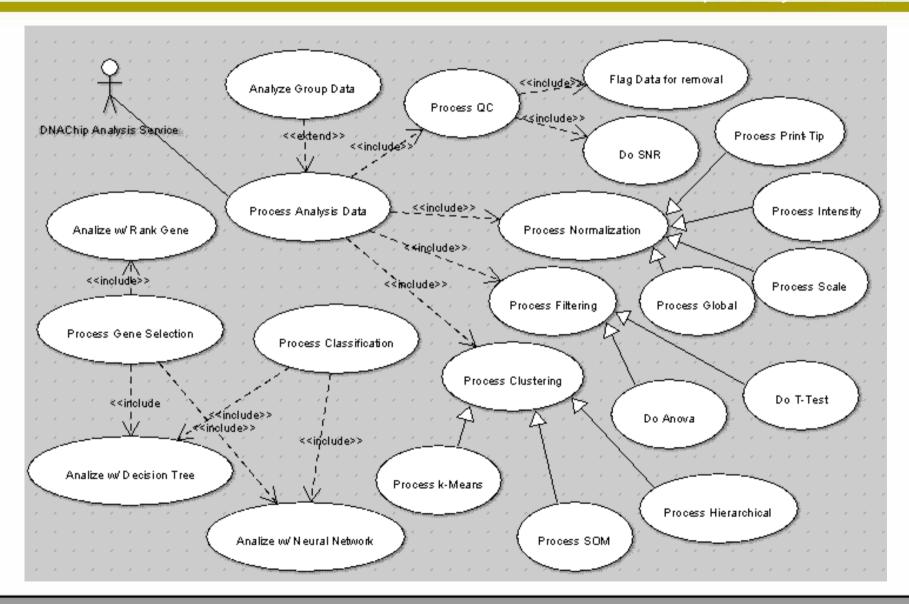


## 2-4. Statistical Analysis



# 2-5. Use case diagram

#### Biochip Analysis Process



- Lowess Normalization & Distribution Adaptation of T-Test
- Distribution Borker Server, Distribution Labor PC Environment
- 3.2GHz CPU, 2G Memory
- Windows 2000 OS
- Distribution Broker Server 1, Distribution Labor PC 3
- Lowess Algorithm
  - Sample Count: Reference 1, Test 3
  - Row Count: 15644
  - Option : Delta 0.0

Smothing Parameter 0.66

Iteration 10

- T-Test Algorithm
  - Number of Sample per Group: 3 (Total 6 samples)
  - Number of Sample per row: 33,012
  - The part of Distribution Adaptation: The most time-consuming Data Reading Part
- The effect of Distribution Analysis Adaptation
  - Long analysis time is required to perform the analysis of large volume Bio Data.
  - Able to reduce the analysis time by the ratio of the number of work node
  - More effective for repetition work
  - The knowledge of analysis process is required to adapt distribution Algorithm

## 3-2. Lowess Normalization

The Adaptation of Distribution Analysis

Distribution Adaptation of Lowess Normalization Total - Before Distribution Adaptation 760 Sec 1'st Data 2'nd Data 3'rd Data **DB** Data Lowess 10 Sec 250 Sec 250 Sec 250 Sec Selected Lowess Lowess Completed Lowess Lowess Lowess Lowess **Local PC CPU** CPU Usage( 22:40:30 22:40:40 22:40:50 22:41:00 22:41:10 22:41:20 22:44:30 22:49:00 22:49:30 22:52:50 22:53:00 22:53:10 22:53:20 22:53:30 22:53:40 22:45:00 Time Time Time - System Usage — Worker Usage System Usage — Worker Usage System Usage — Worker Usage System Usage - Worker Usage

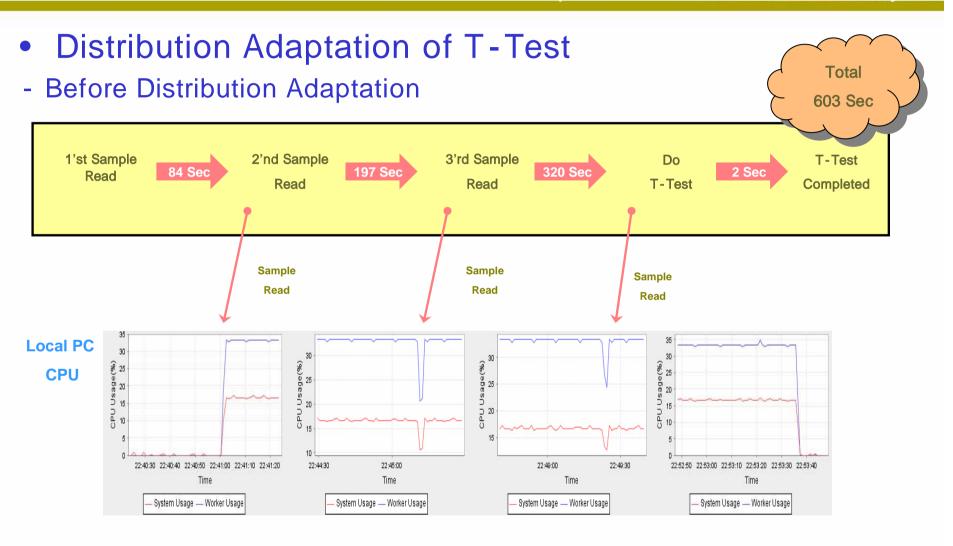
#### 3-2. Lowess Normalization

The Adaptation of Distribution Analysis

Distribution Adaptation of Lowess Normalization - After Distribution Adaptation Total 265 Sec 1'st Data Lowess **DB** Data Job Lowess 2'nd Data Lowess **Grid Job** 10 Sec 3 Sec 250 Sec 2 Sec Completed Completed Selected 3'rd Data Lowess Transfer data to Broker Lowess Transfer Data to Labor 1,500 Analysis **Network** 1,250 1,250 1,000 1.000 750 22:00:00 22:00:00 22:09:30 22:10:00 22-12-10 22-12-20 22-12-30 22-12-40 22-12-50 22-13-00 Time Time Received - Sent Received - Sent Received Sent **Distribution Server CPU** Use of 3 Nodes . Reduction 22:12:00 22:12:10 22:12:20 22:1 of the analysis time by 1/3 System Usage - Worker Usage System Usage - Worker Usage System Usagi

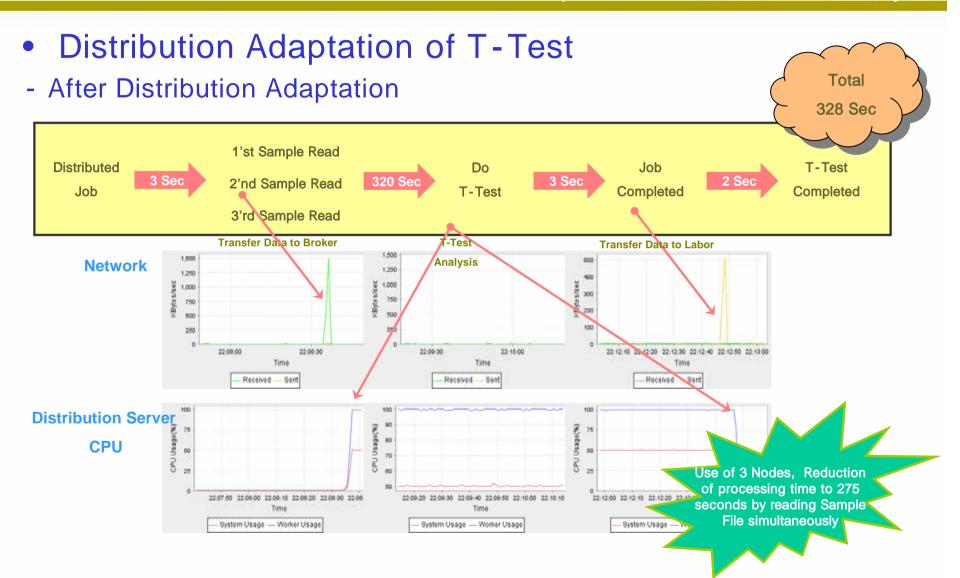
## 3-3. T-Test

#### The Adaptation of Distribution Analysis



## 3-3. T-Test

#### The Adaptation of Distribution Analysis



## 4. Contributor

- Grid R&D Center, NationalGrid Inc.
  - Jong Hyun Youn
- Bioinformatics Research Center, Macrogen Inc.
  - Kyong Oh Yoon
  - Dae Hyun Kim
  - Keun II Lee