Bio# (Bio-Sharp)

- New project, planning stage
- Grammars of DNA and Proteins
- Experimental protocols
- Access to other languages
- Java/Mono/.NET/Perl/Python/Prolog/Ruby...

DNA as a Language

- Language analogy (genetic codes)
- Suggests how we think about DNA
- Most textbooks teach
- Popular sci. books
- But no serious linguistics
- David Searls in 1980s
- Formal DNA Grammars
- Basic Gene Grammars (our work)

DNA Grammars

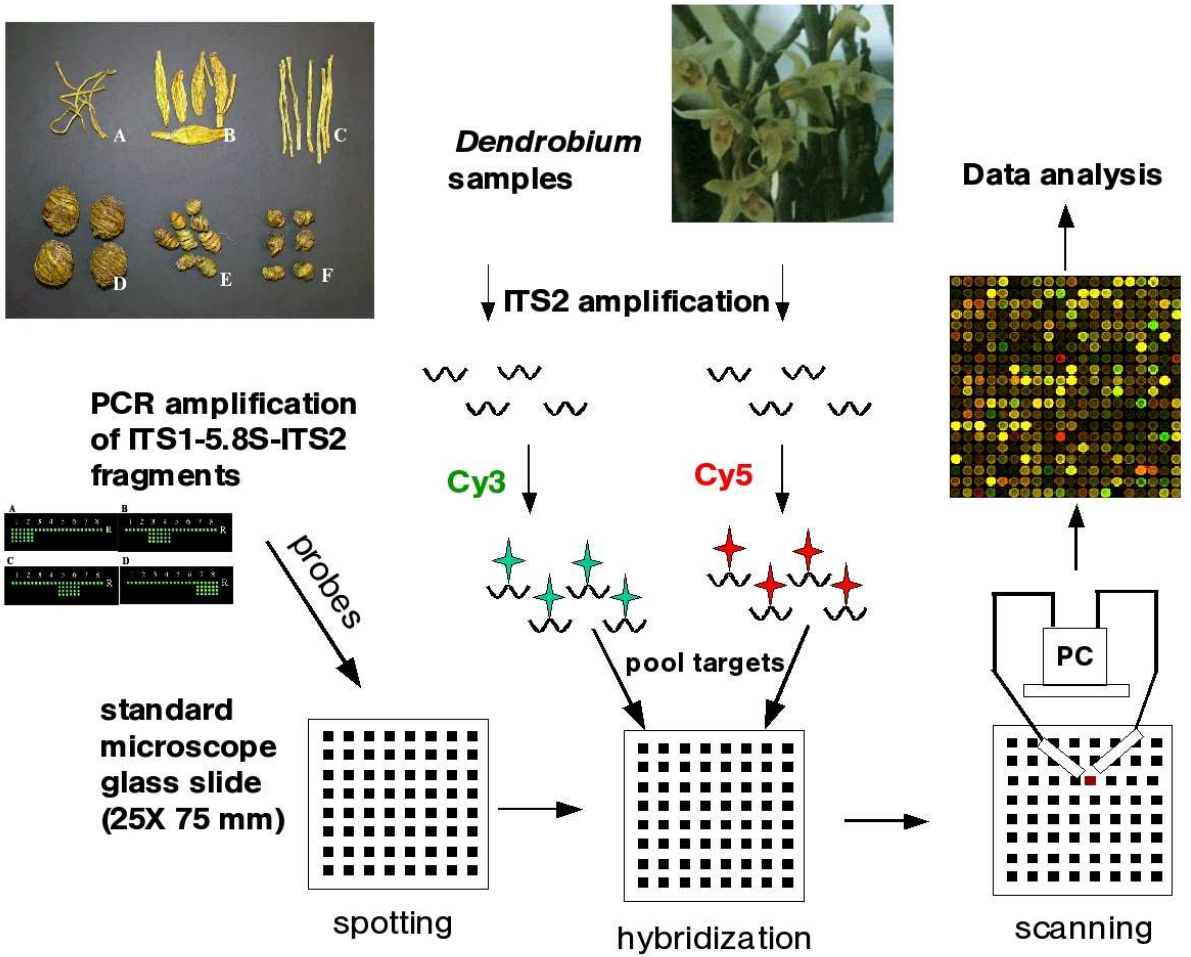
- Scientific knowledge of DNA
- Conceptual categories of DNA sequences
- Grammars for representation and reasoning
- Computational DNA grammars
- Protein grammars (harder)

Simple Grammars


```
n_rDNA_ITS ---> r18s, its1, r5_8s, its2, r28s.

its1_spA ---> patternA1, patternA2.
its1_spA ===> patternB1, patternB2, patternB3.
its2_spB ===> patternC1, patternC2, patternC3.
its2_spB ---> patternD1, patternD2.
patternA1 : hybridisable(probeA1,patternA1) ---> probeA1.
patternA2 : hybridisable(probeA1,patternA2) ---> probeA2.
....
probeA1 ---> "tgattacagacccagcccaatactttctaca".
....
```

Microarray Experiments

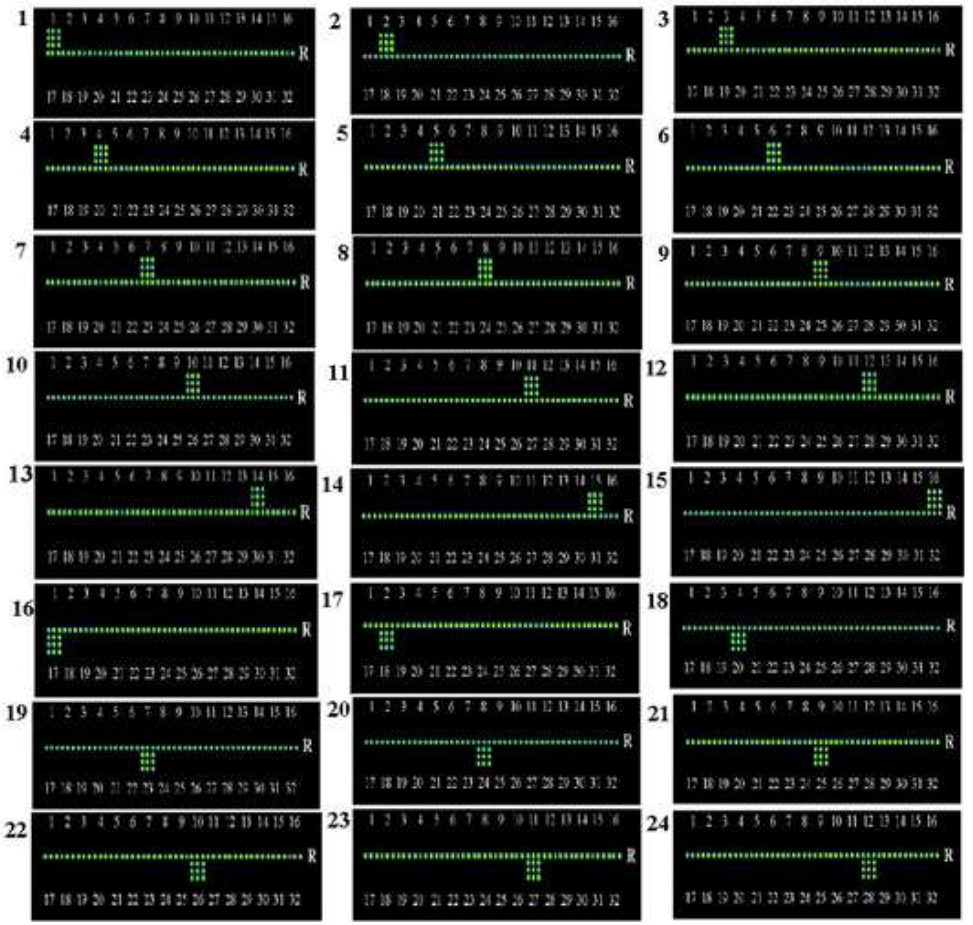


Probe Specificity Test

A Simple Grammar and Experiment Result

```
its2_Dbrymerianum ---> pat(1).
its2_Dacinaciforme ---> pat(2).
its2_Dtrigonpus ---> pat(3).
its2_Dloddigesii ---> pat(4).
its2_Dlongicornu ---> pat(5).
its2_Dsalaccense ---> pat(6).
its2_Dthysiflorum ---> pat(7).
....
its2_Dsulcatum ---> pat(24).


pat(X):
 hybridisable(pat(X),probe(Y))
---> probe(Y).
```



Gr is Abstract Spec

- Grammars (K/M/T/H...)
- Protocols
  - Material
  - Apparatus / servers
  - Method / workflow
  - Expt design constraints
- Expt design
- Operationalisation
- Execution

Biosequence Experiments

- Physical or computational
- Microarray toolkit design
- Bioinformatics database search
- Computational sequence analysis/synthesis
- Need sequence knowledge & KR
- Coordination of experimental protocols
- Simpler if we use grammars?
- Higher level spec

Language Issues

- Never easy
- Syntax – Grammars, XML?
- Incorporate experimental protocols
- Web/Grid Srv/.NET/Mono/Java/Parrot
- Ontology issues

Contributors

- DNA Grammars
  - Siu-wai Leung (Edinburgh University, UK)
  - Chris Mellish (Aberdeen University, UK)
  - Dave Robertson (Edinburgh University, UK)
  - *Bioinformatics*, 17(3):226-236, 2001
- DNA Microarrays
  - Pang-chui Shaw (Chinese University, HK)
  - Yanbo Zhang (Chinese University, HK)
  - *Planta Medica*, 69(12):1172-1174, 2003
- Grammars for Microarrays
  - *ISMB04* Poster F-63
- Bio# Implementation
  - Collaborators welcome!