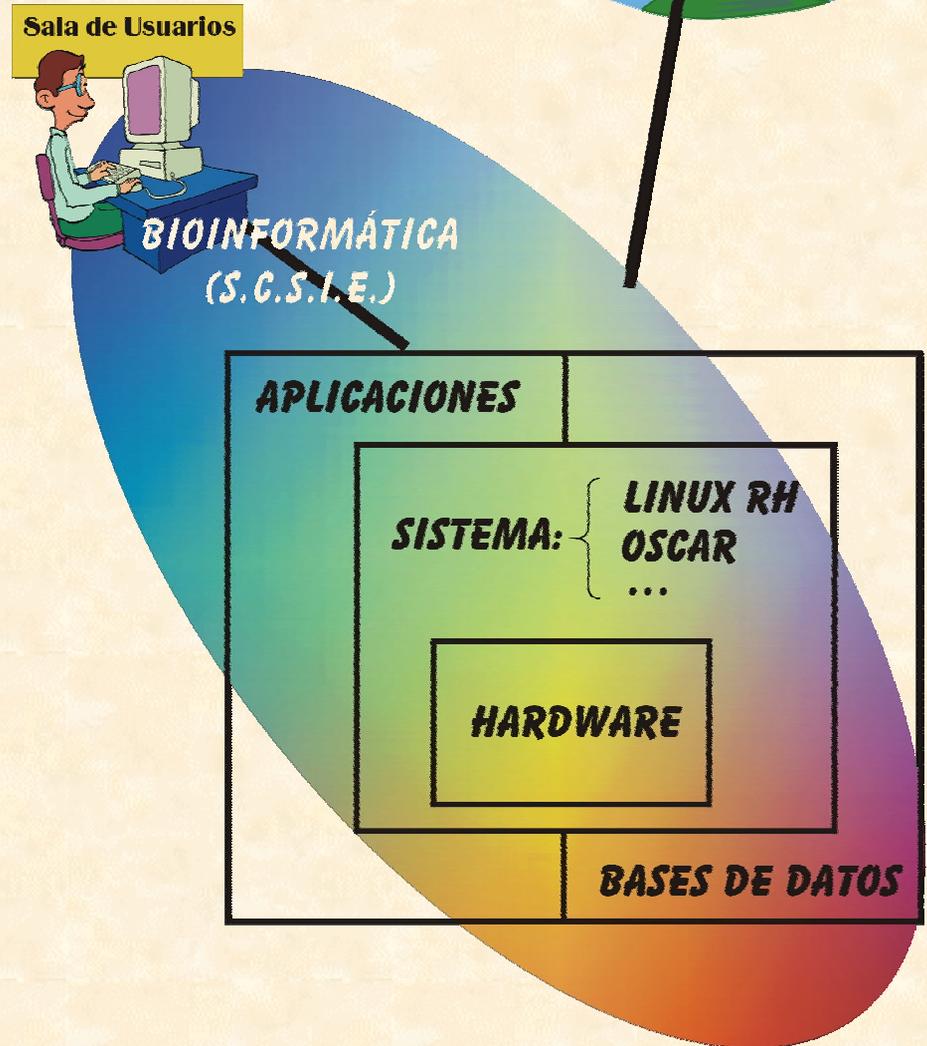


BIOINFORMÁTICA

S.C.S.I.E.

UNIVERSITAT
DE
VALÈNCIA



HARDWARE (I)

SALA DE USUARIOS

- 7 PC's
- 2 Mac's
- Impresoras
- Escáner
- Etc.



HARDWARE (II)

ALPHA SERVER

- 2 Processadores 400MHz
- Disco: 100GB
- RAM: 512MB
- Compaq Tru64 UNIX V5.1B



HARDWARE (III)

CLUSTER

- 16 nodos
 - 2*PIII 933MHz
 - 768Mb
 - 30GB
- Master
 - 2*PIII 933MHz
 - 1Gb
 - 270GB
- Switch
 - 3Com
 - 100 MHz
- SAI's
 - 5800 VA



Sistema - cluster (I)

REDHAT 7.2

OSCAR 2.2.1

- Permite montar, programar y usar **clusters** de manera bastante sencilla.
- Permite lanzar aplicaciones en **paralelo**.

Sistema - cluster (II)

- **Nuestras necesidades**: que diferentes procesos puedan ejecutarse simultáneamente sin ver reducida su asignación de recursos.
- **PBSPro** (Portable Batch System):
 - **Sistema de colas** para trabajos batch.
 - Permite que el usuario **interactúe** con sus aplicaciones por línea de comandos.
 - La **filosofía** de distribución de las tareas:
 - FIFO
 - Un trabajo por CPU
 - Para el usuario todo este sistema es **transparente**.

Sistema - cluster (III)

Manejo del PBSPro:

- **Lanza tu trabajo:** granjasb> codeml
97.granjasb.scsi.uv.es

- **Los resultados en:** codeml.output

- **Ver el estado de los trabajos:**

```
granjasb> qstat
```

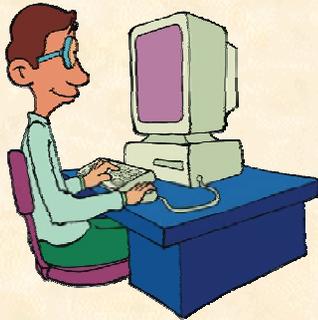
<u>Job id</u>	<u>Name</u>	<u>User</u>	<u>TimeUse</u>	<u>S</u>	<u>Queue</u>
97.granjasb	codeml.script	pepito	00:00:29	R	cola

- **Eliminar un trabajo en ejecución:**

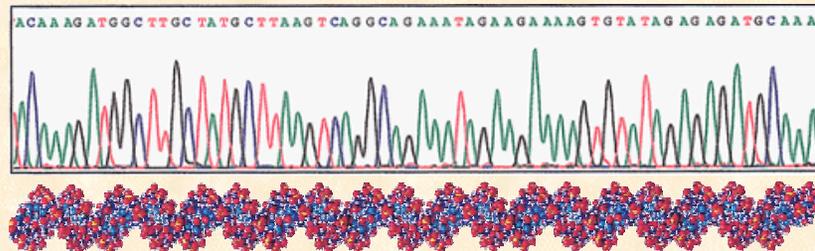
```
granjasb> qdel 97
```

Intercomunicación

- **DATOS:** FTP, SAMBA
- **CONEXIÓN:** TELNET, SSH



**SECUENCIACIÓN
S.C.S.I.E.**



Aplicaciones (I)

GCG

- The Wisconsin Package
- Paquete de 130 programas.
- Manipulación y análisis de secuencias de ADN y de proteínas:

Aplicaciones (II)

GCG

- **Comparación:** Gap, compare, ...
- **Búsqueda en bases de datos:** BLAST, Fasta, ...
- **Edición de secuencias:** SeqLab, ...
- **Evolución:** Distances, ...
- **Ensamblaje de fragmentos:** GelMerge, GelAssemble, ...
- **Búsqueda de genes y reconocimiento de patrones:** Consensos, ...
- **Importación/Exportación:** FromEMBL, ToStaden, ...
- **Mapeado:** Map, PeptideSort, ...
- **Selección de “Primers”:**Prime
- **Análisis de proteínas:** Motifs, ProfileScan, ...
- **Estructura secundaria de RNA:** Mfold, StemLoop, ...
- **Traducción:** Translate, ExtractPeptide, ...
- **Utilidades:** Reverse, Corrupt, ...

Aplicaciones (III)

EMBOSS

- "The European Molecular Biology Open Software Suite "
- Paquete de 100 programas.
- Básicamente pensados para la biología molecular:

Aplicaciones (IV)

EMBOSS

- **Alineamiento** de secuencia
- **Búsqueda** rápida en BBDD con patrones de secuencia.
- **Identificación** de motivos de proteínas, incluyendo análisis de dominios.
- Análisis de **patrones** de secuencias de nucleótidos.
- Análisis de utilización de **codones** para genomas pequeños.
- Herramientas de presentación para **publicaciones**.
- Etc.

Aplicaciones (V)

Aplicaciones comerciales

Aplicaciones Freeware

- **Clustalx:** Multiple Sequence Alignment Program.
- **DNArates:** Maximum likelihood estimates of the rate of nucleotide substitution at each position.
- **FastDNAmI:** Improved version of PHYLIP DNAML.
- **Molphy:** Program package for MOLEcular PHYlogenetics.
- **Paml:** Phylogentic Analysis by Maximum Likelihood.
- **Phylip:** Phylogeny Inference Package.
- **Phylowin:** A graphic tool for molecular phylogenetic inferences.
- **Puzzle:** Maximum likelihood analysis for nucleotide, amino acid, and two-state data.
- **Staden :** Staden is a suite of programs for sequence analysis.
- **Etc.**

Base de datos (I)

Tipos:

Secuencias de ADN: EMBL.

Secuencias de Proteínas: SwissProt, Pir, Tremol.

Motivos de secuencia: Prosite.

Familias de proteínas: Pfam.

Clasificaciones funcionales de proteínas: Enzyme.

Genomas: Ensembl, GeneCards.

Mapas

Estructuras: PDB, SCOP.

Bibliografía

Base de datos (II)

Almacenamiento local:

- Crecimiento → Disco
- Actualizaciones automatizadas

Consultas:

- **SRS:** Sequence Retrieval System: buscar palabras clave en las BBDD de secuencias de DNA y proteínas (EMBL, SwissProt, ...).
- **Entrez:** Idem con GenBank.
- **BLAST,**
- **FASTA:** Búsqueda por similitud de las secuencias.
- Etc.

contacto

<http://scsie.uv.es/bioinf/>