

# An Expert System for Scoring DNA Database Profiles

Dr. Mark W. Perlin  
Cybergenetics  
Pittsburgh, PA



---

---

---

---

---

---

---

---

## Outline

- STR Forensic Databases
- The Data Scoring Problem
- An Expert System Solution
- Complex DNA Analysis



---

---

---

---

---

---

---

---

## Generate STR Profiles

STR lengths

↓ PCR amplification

PCR products

↓ size separation

electrophoretic bands

↓ data acquisition

data pixels

### Data artifacts

- PCR stutter
- preferential amp
- contamination
- peak spread
- lane crosstalk
- size variation
- baseline shift
- dye bleedthrough
- size distortion



---

---

---

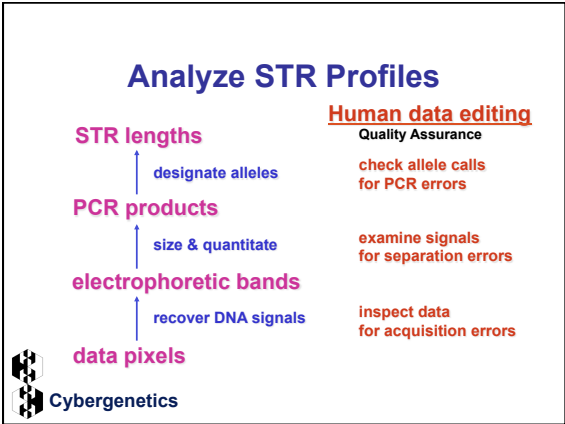
---

---

---

---

---




---

---

---

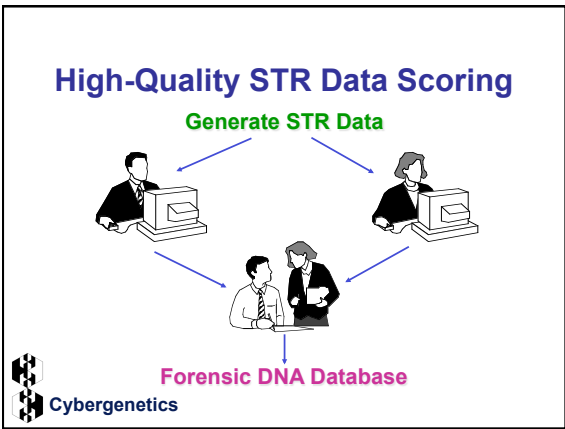
---

---

---

---

---




---

---

---

---

---

---


---

---

## Data Analysis: Labor Cost

≥ \$1.00 per experiment

LaborCost					
	A	B	C	D	E
20					
21	PEOPLE COST	\$49,152,000			
22	PER GENOTYPE	\$1.02			
23					
24	Breakdown	per person	throughput	per day	per year
25	salary	\$25,000	runs	192	48,000
26	benefits	\$2,250	genotypes	192,000	48,000,000
27	space	\$2,000			
28	computer	\$2,000			
29	software	\$10,000	Scoring	500	125,000
30	management	\$4,750	calls/person		
31	overhead	\$12,500			
32	COST	\$64,000	PEOPLE	768	
33					
34	Assumptions		Assumptions		
35	benefit rate	0.25	genotypes/run	1,000	
36	sq feet/person	100	days/year	250	
37	calls/q/day	3.25	people/call	2	
38	management rate	0.25			
39	overhead rate	0.50			
40					



Cybergenetics

---

---

---

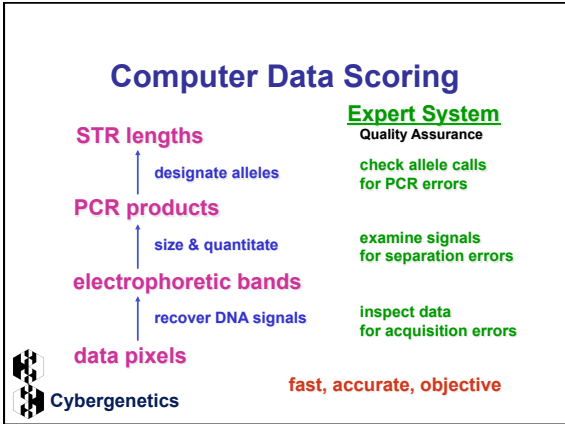
---

---

---

---

---




---

---

---

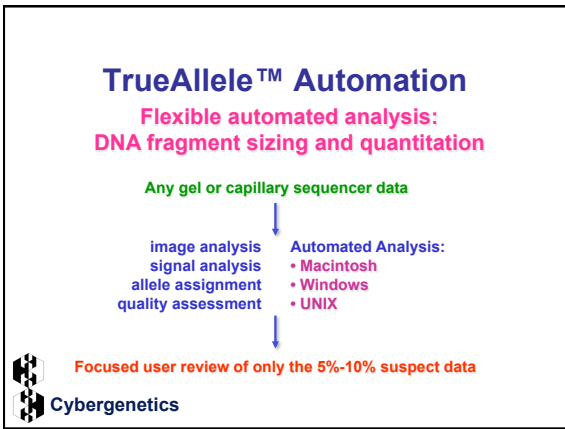
---

---

---

---

---




---

---

---

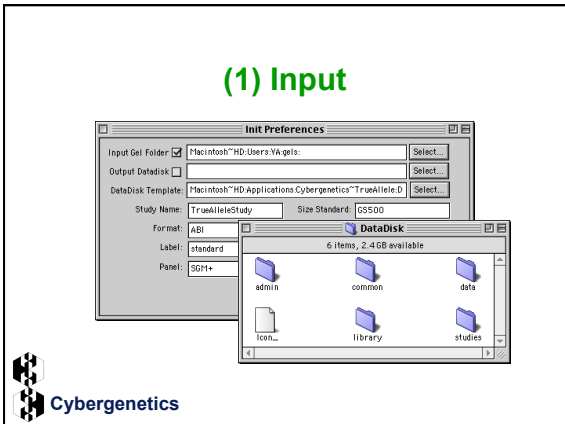
---

---

---

---

---




---

---

---

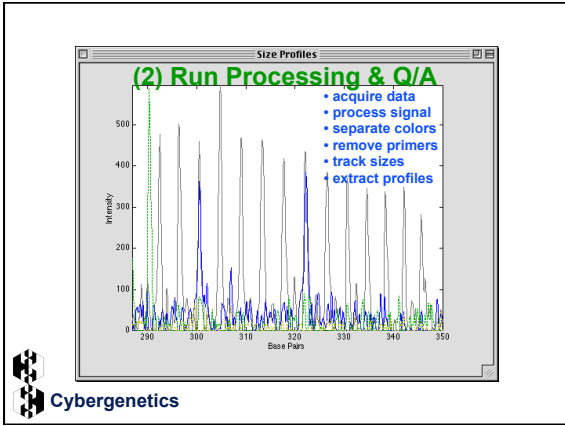
---

---

---

---

---




---

---

---

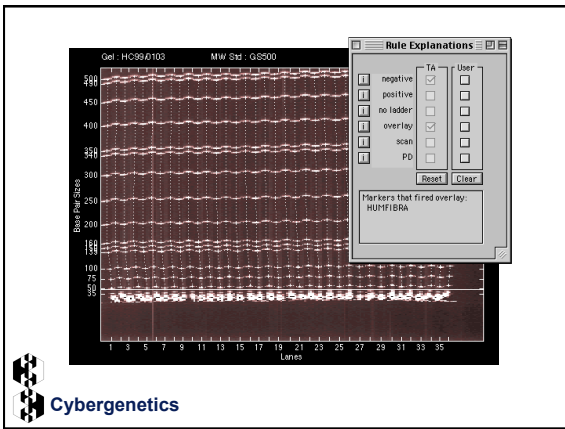
---

---

---

---

---




---

---

---

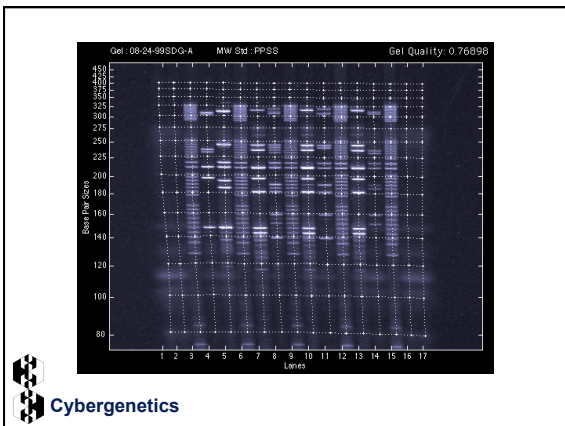
---

---

---

---

---




---

---

---

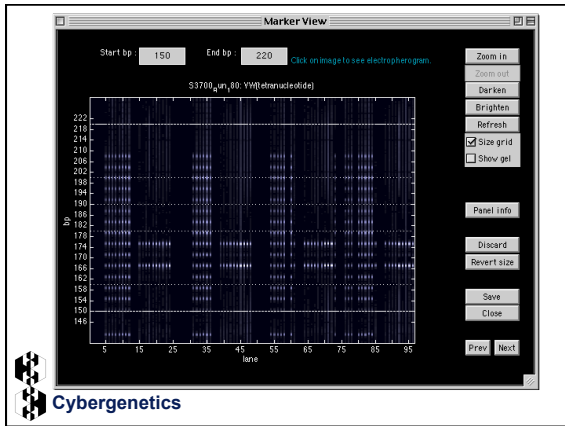
---

---

---

---

---




---

---

---

---

---

---

---

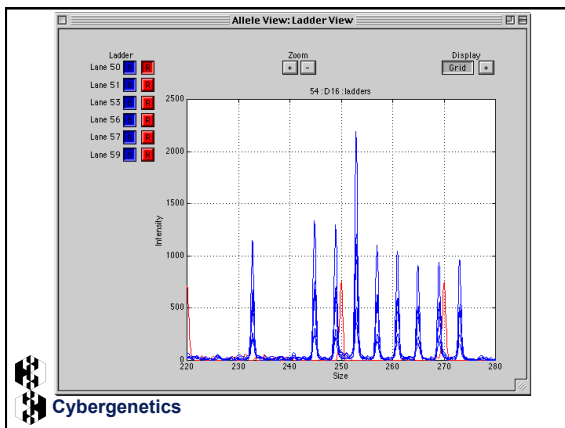
---

---

---

---

---




---

---

---

---

---

---

---

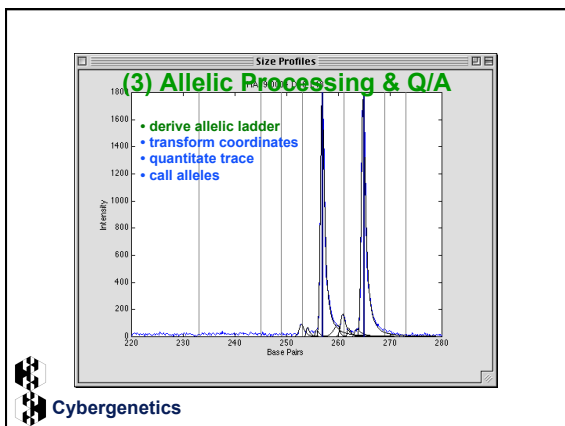
---

---

---

---

---




---

---

---

---

---

---

---

---

---

---

---

---



[overview](#)  
[advantages](#)  
[approach](#)  
[spreadsheet](#)  
[patents](#)  
[publications](#)

www.cybgen.com

[company](#) | [technology](#) | [software](#) | [support](#) | [news](#) | [contact](#)


cybgenetics

**patents**

The TrueAllele technology is protected by multiple patents. Review the claims linked to the patents below to determine whether you need TrueAllele patent protection for your genetic processes or discoveries.

Patent	Issued	Title
<a href="#">US 5,541,067</a>	Jul 1996	Method and system for genotyping
<a href="#">US 5,580,728</a>	Dec 1996	Method and system for genotyping
<a href="#">US 5,876,933</a>	Mar 1999	Method and system for genotyping
<a href="#">US 6,054,268</a>	Apr 2000	Method and system for genotyping

© Cybgenetics 1999-2000. All rights reserved.



Cybgenetics

---

---

---

---

---

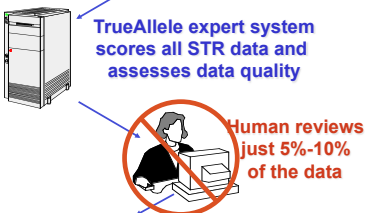
---

---

---

**Automated Scoring: FSS/UK**


Generate STR Data



TrueAllele expert system scores all STR data and assesses data quality

Human reviews just 5%-10% of the data

UK National DNA Database



Cybgenetics

---

---

---

---

---

---

---

---

**NIJ Validation Project**

Automatically rescore 30,000 samples


Florida database

- ABI/310
- ABI/3700

Virginia database

- Hitachi/FMbio2

Compare with CODIS profiles



Cybgenetics

---

---

---

---

---

---

---

---

## Complex DNA Analysis

Automatically resolve data artifacts

- Remove PCR stutter
- Adjust relative amplification
- Quantitate DNA band overlap
- Resolve DNA mixtures

Mathematical model



---

---

---

---

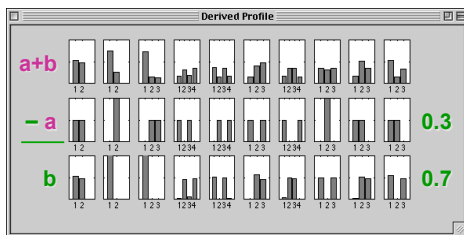
---

---

---

---

## Mixture Deconvolution



---

---

---

---

---

---

---

---

## Automated DNA Profiling

STR Data Pathway

- Prepare Sample
- Amplify DNA
- Separate Sizes
- Designate Alleles
- Match Database



---

---

---

---

---

---

---

---