Camb. Phil. Soc. 1969?

The Two Languages of Biology

apologies — should be molecular languages.

Two "languages" are nucleic acid and protein.

Why "languages" — perhaps scripts a better word.

Setting like scene

Our written language:

① linear sequence (~~but~~ cut up into lines and pages for convenience)

② standard set of symbols — like letters. The detailed order of the letters conveys the information.

Same way with nucleic acid and protein.

both polymers — ie linear — uniform backbone.

symbols are the side-chain side-groups.

- 4 for NA
- 20 for protein.

Specificity depends on the detailed sequence of the side groups.

Basic roles

NA = the genetic material or its working copy

(Subsidiary use to form certain very special structures)

~~proteins~~ not enzymes.

proteins are the working biological ~~molecules~~ catalysts.

– enzymes (Kendrew + Perutz)

also used extensively for structures (eg. hair keratin) and other specific purposes (eg anti-bodies).

flow of sequence information — detailed

5

explain arrows

comment on 1 or 2 strands for NA.

Size of protein say 200

gene part of a NA molecule.
from 1000 to several million.

CENTRAL DOGMA — acquired characteristics.

Basic chemistry

NA — DNA formula ——— S

connect on RNA

can be double-helical ——— S

base-pairs ——— S

Thus molecules does one job — complementary recognition — very well. Poor for specific structures (Reasons: — only 4 bases.

— stacking + double helix is restrictive)

is not at all versatile.

S

proteins — backbone, to show how different (leave to Kendrew + Perutz).

Replication of the genetic material

eg. DNA copying.

known in outline — so users have parsing

— so idea very well known.

details still uncertain. which enzymes? [[Kornberg Cairns.

untwisting?

DNA often circular (no ends)

not much known about control mechanisms.

Repair mechanisms — extensive.

DNA → mRNA

enzyme known, but still being characterised.

σ factors ~~explicit~~ which give specificity

only copies part of DNA

but start & and stop signals not yet known

mRNA → protein — elaborate machinery → minimum ~ 100 proteins (one enzyme) NA: 2 sig ~30 small.

SINGLE-STRANDED (– DNA)

Genetic code ——— S

degeneracy, especially third place.

fairly universal (but not start signals)

not give history of code here, but indirect

merely indirect methods. { stripler binding; artificial messengers; mutants – phase shift or base analogs

SANGERS methods (for RNA)

But now. Mr. Peter Jeppesen, phage R17 ——— S

part of coat protein.

when is between genes? R17 again.

start signal in bacterial AUG (formyl met)

(GUG) Met in mid message. Val

~~but AUG also met~~ but

Must should be a recognition sequence (special protein to do this) factor

Dr. Joan Steitz. ——— S

the start of 3 seperate proteins.

doesn't or not it unambiguously.

Bad gaps in R17 (Dr. Jack Nicols)

c.f.

double end signal ? :

gap of about 30, unpublished.

N.A. as structure

r RNA . little known

t.RNA. eg. tyrosine. ——— S

anticodon { GUA ↓ CUA

SMITH et al

BRENNER proof of anticodon

clefoire structure. mixed molecules

can be co-crystallize Dr. Jacques FRESCO

general formula

(Pr) Michael Levitt. + base pairs ——— S

——————— S

his model

Remaining problems

DNA in higher organisms

why so much

why "repeats" ? — reserve of genetic material

— control mechanisms (cheap)
(Britten)

Amplification — rRNA in ?

— I DNA ? Eugene Bell.

---

the two families — advantages

origins?

Astronomy, Size.

place head ——— S

your DNA — all cells.

small part of you.

push right across solar system.

Genes

Nucleic acid.
(DNA & RNA)

polynucleotide chain
double

~ 4 nucleotides
(phosplate - sugar - base)

A : T (U)

G : C

Enzymes
Protein

protein

polypeptide chain

~ 20 amino acids.