23/11/1975

Genetic Code

		(U)		(C)		(A)		(G)
1	Phe	2	6	Ser	4+2	10 Tyr	2	16 Cys 2
2	Leu	6	7	Pro	4	11 His	2	17 Try 1
3	ILU	3	8	Thr	4	12 Gln	2	18 Arg 4+2
4	Met	1	9	Ala	4	13 Asn	2	(Ser 2)
5	Val	4				14 Asp	2	19 Gly (4)
						15 Glu	2	

No. of aa	No. of codes	
2	1	Met, TRy
8	2	Phe, Tyr, His, Gln, Asn, Asp, Glu, Cys
1	3	Ilu
5	4	Val, Pro, Thr, Ala, Gly
7	6	Ser, Arg
23		

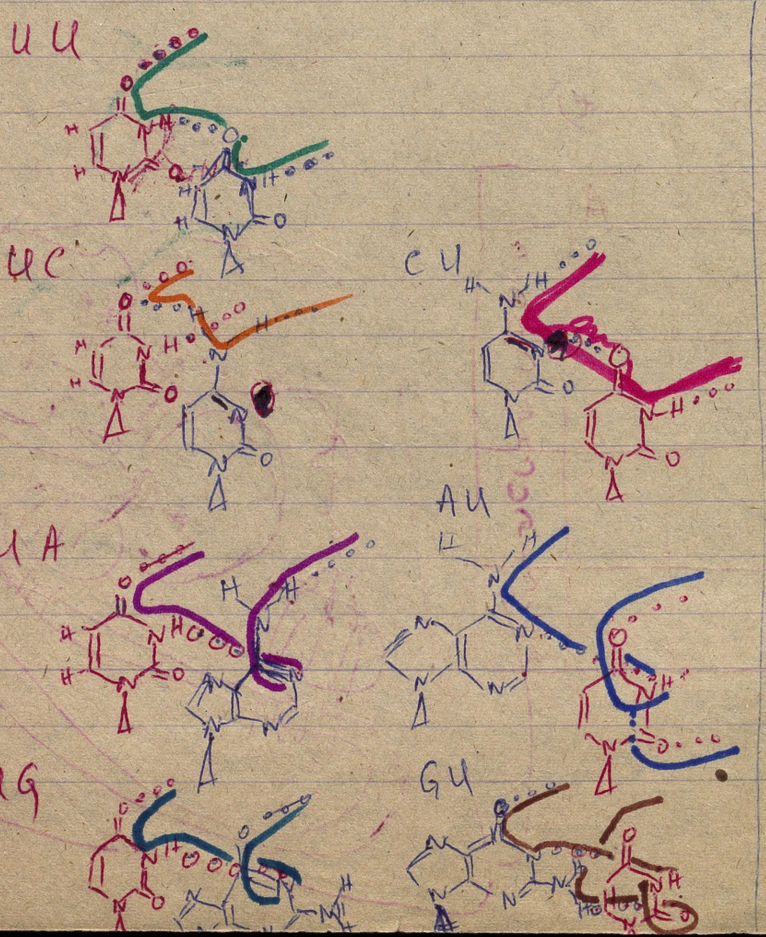
1) Primarily 4 & 2
2) ILU lost one to Met
3) 8 Doublets, 8 Quartets (hydrophobic) (hydrophilic)
1 lost to Met
Ser, Arg gained (2)

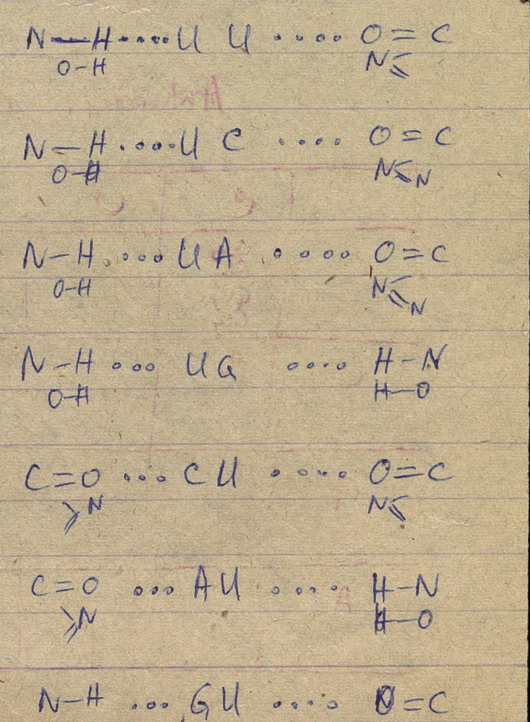
	U	C	A	G
U	Phe	Ser	Tyr	Cys
C	Leu	Pro	His	Arg
A	Ilu	Thr	Lys	Ser 2 Arg 2
G	Val	Ala	Asp Glu	Gly

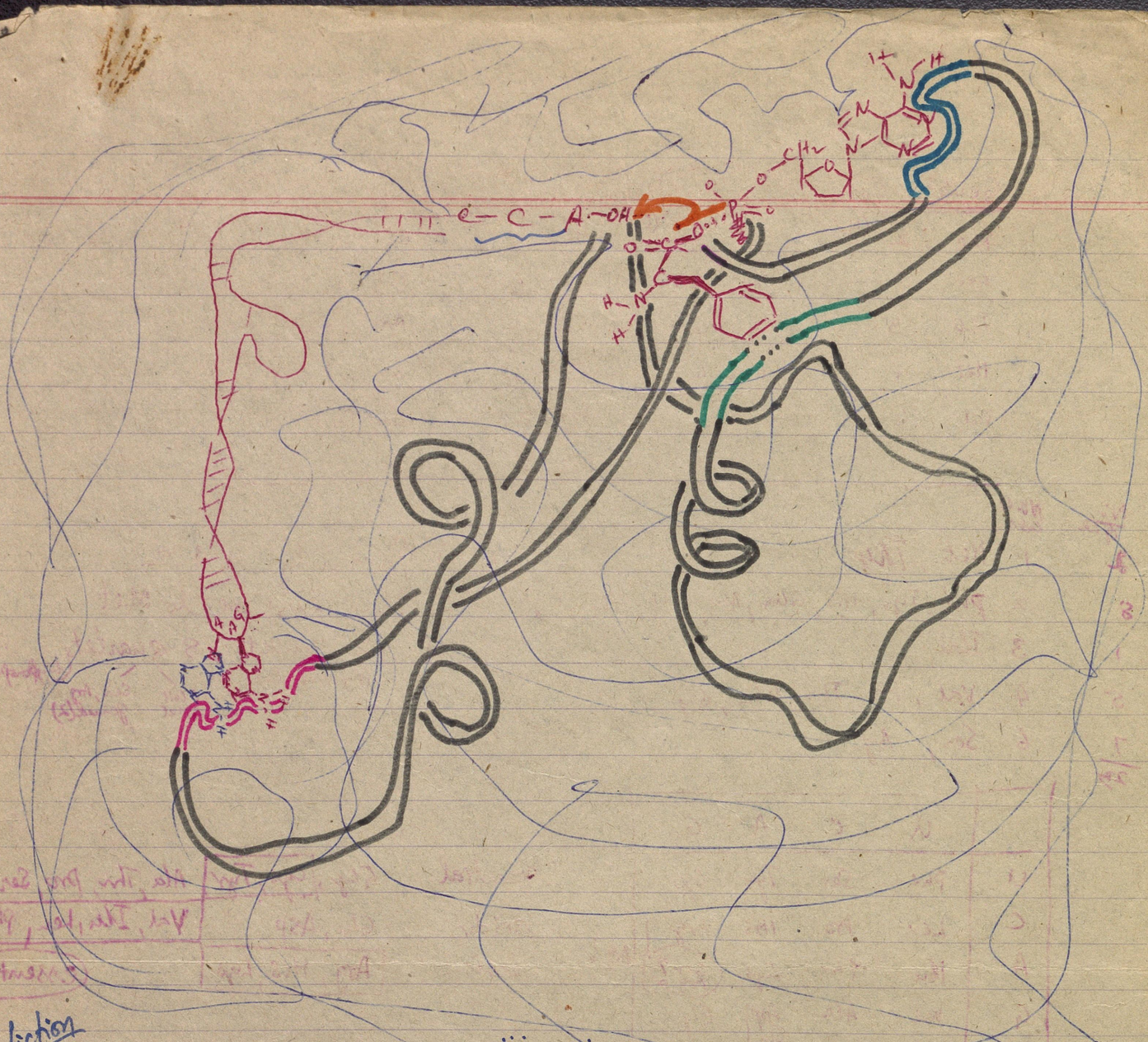
Arg?

Hydrophilic –OH neutral polar ?

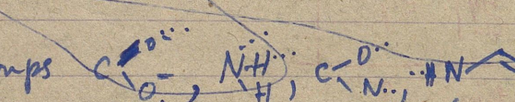
neutral: Gly, Cys, Tyr | Ala, Thr, Pro, Ser
acidic: Glu, Asp | Val, Ilu, Leu, Phe
basic: Arg, His, Lys | essential





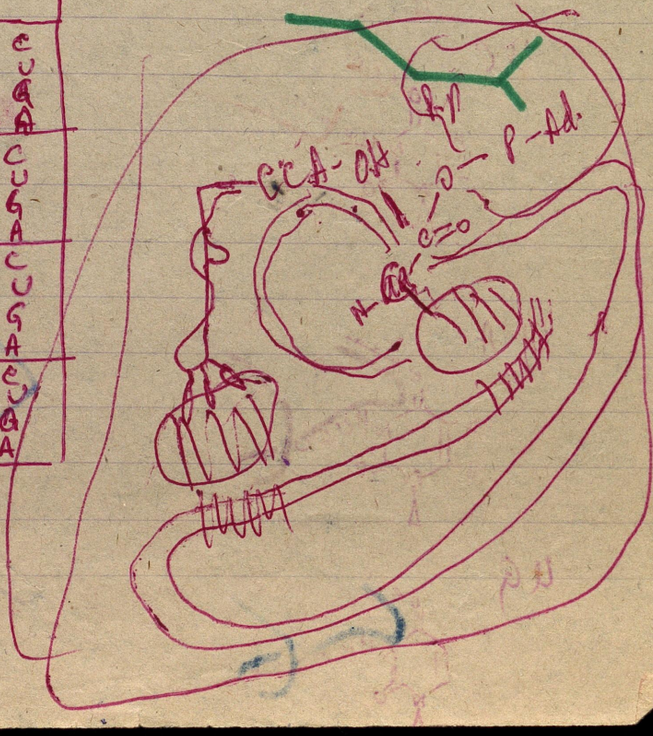


Prediction

If the side-chain groups  provide bridges in π-H pathways, modification of side-chain should prevent aa-tRNA formation ⇒ prediction.

Anticodons (rearrangement)

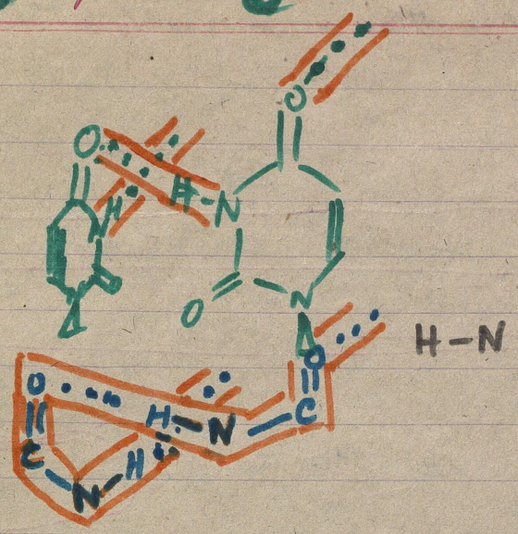
	C	U	G	A	
C	Gly Gly Gly Gly				CUGA
U	A				CUGACUGACUGA
G					
A					



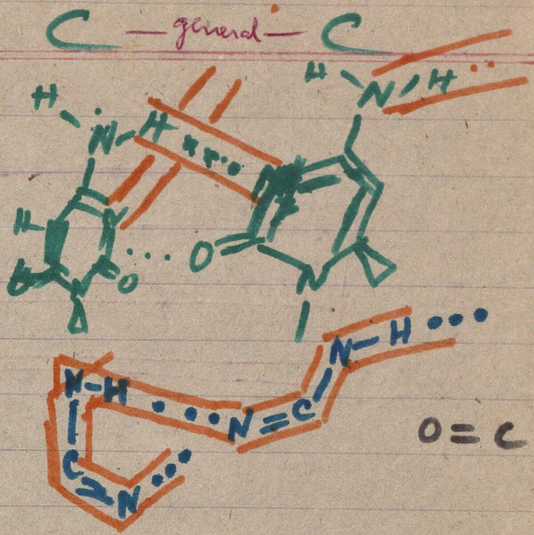
Anticodons 3
5

24/11/75

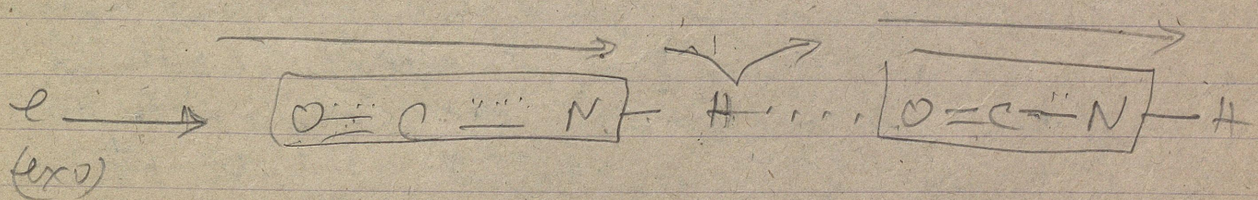
U — polar — U



C — general — C



H-N



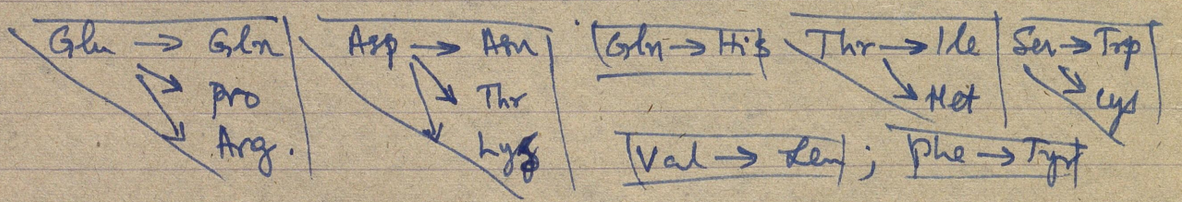
A Co-evolution theory of the genetic code

J. Fze-Fei Wong PNAS 72, 1909-1912 (1975)
Dept. Biochem., Univ. Toronto, Toronto, Canada.

Structure of genetic code was determined by the sequence of evolutionary emergence of new amino acids within the primordial biochemical system.

Code : 64 triplets
riddles of origin of code (Monod) < Stereochemical / Frozen accident

" The structure of codon system is primarily an imprint of the prebiotic pathways of aa formation, which remain recognizable in the enzymic pathways of amino-acid biosynthesis. Consequently the evolution of the genetic code can be elucidated on the basis of precursor-product relationship between amino acids in their biosynthesis. The codon domains of most pairs of precursor-product amino acids should be contiguous, i.e. separated by only the minimum separation of a single base change." — a Coevolution theory.

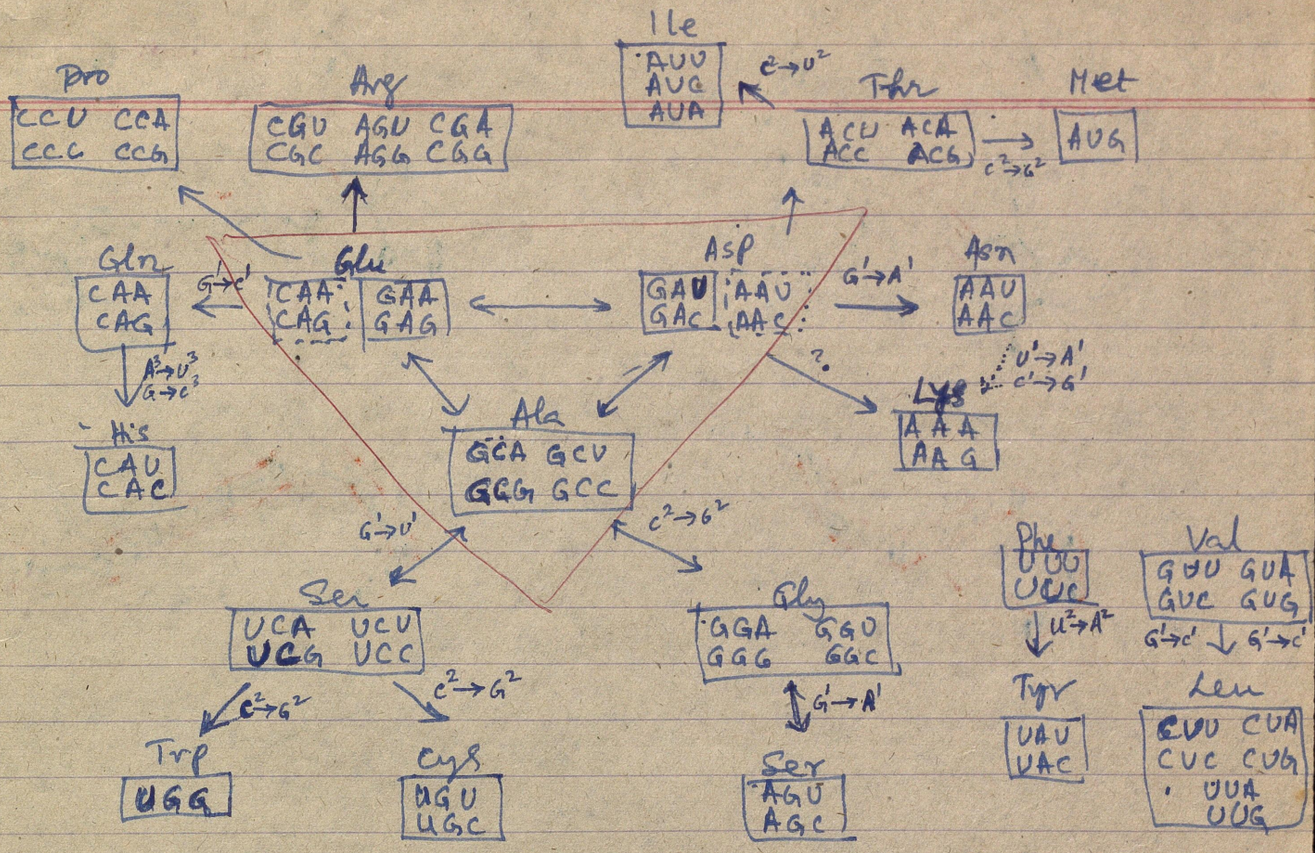


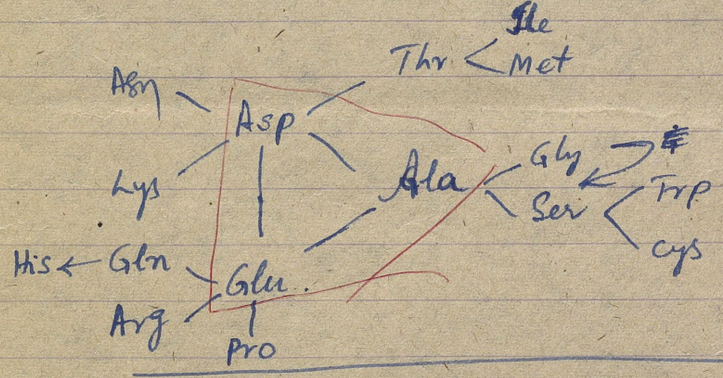
Asp → diaminopimelate
Glu → α aminoadipate
> Lys < Prokaryotes (prebiotic sign.) / eukaryotes.

Why: emergent aa Glu → Gln — CAA, CAG
Asp → Asn — AAU, AAC

few aa in primordial proteins; continuous domain
new aa: new enzymic or non-enzymic catalysts, cofactors
precursor might precede a codon to product

1) product may resemble precursor
2) conversion while still attached to adaptor
3) on adaptor conversion

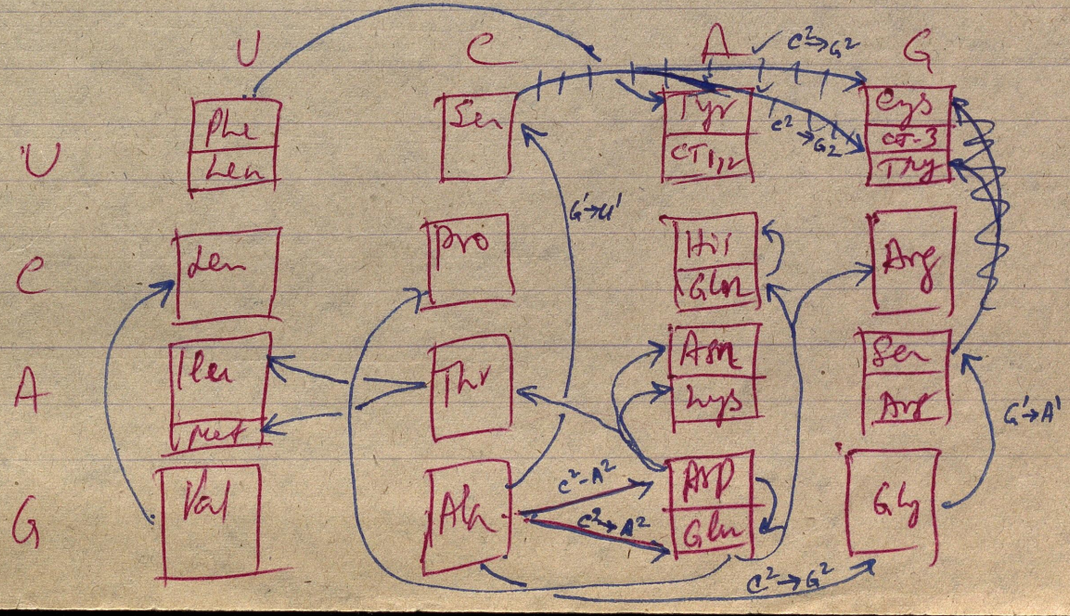




Phe → Tyr
Val → Leu

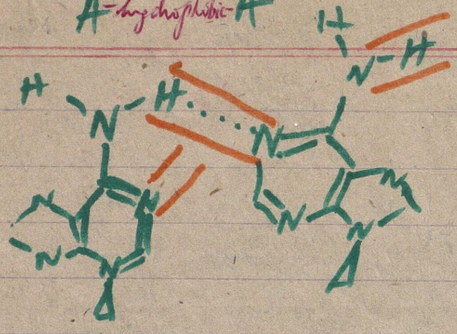
aa resemblance to t-RNA limitations :

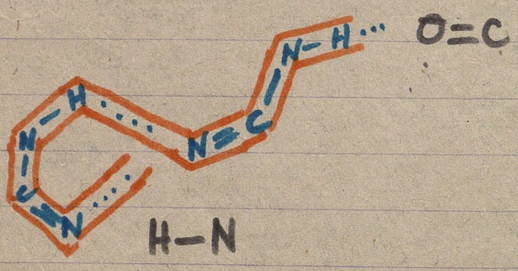
1) t-RNA primordial changed since
2) for diffnt precursor need not be dissimilar
3) 40-70% base sequences changed



3' ← 5'

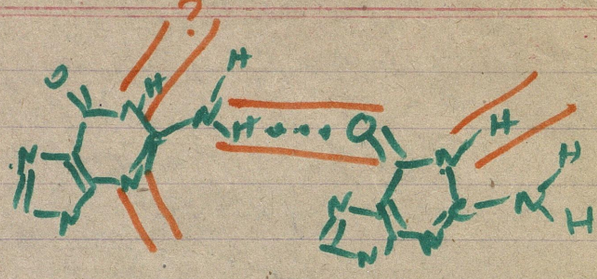
A — hydrophobic — A

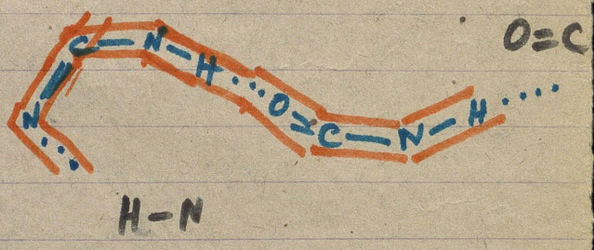




3' ← 5'

G — hydrogen bonded — G





Dec 1975

Reverse Translatase or Protein → RNA → DNA

1. Genetic code reveals an interesting feature: 2nd letter of triplet

U	= hydrophobic	A	anticodons on t-RNA
C	= neutral & hydroxy a.a	G	
A	= polar both acidic & basic	U	
G	= General, Cys, Arg, Gly, Try	C	

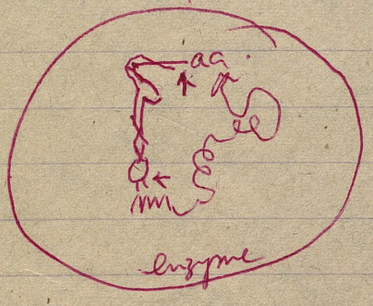
2. How does the code recognize the aa? — through t-RNA

Pro-code DNA 3'→5'
↓ code m-RNA 5'→3'
↓ anticode (t-RNA 3'→5' – aa) ← aa-c~O-3'AMP ← enzyme ↑ ATP
↓ protein N→C term aa

Recognition of both t-RNA & aa is done by amino-acyl synthetase; after aa is attached, only anticodon is recognized by m-RNA, not aa

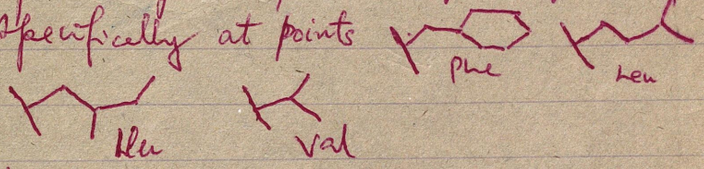
3. What is the basis of the dual recognition of t-RNA (anti-codon?) and aa; why is there a relationship between second letter of code and the nature of amino acids?

Do the H-bonding in nucleotide bases connect π-H pathways and connect to amino acid in the enzyme?



Do the nucleotide pairs ~~can~~ give bridges to fill the gap in π-H pathways?

→ Hydrophobic: provide protection for the pathways specifically at points

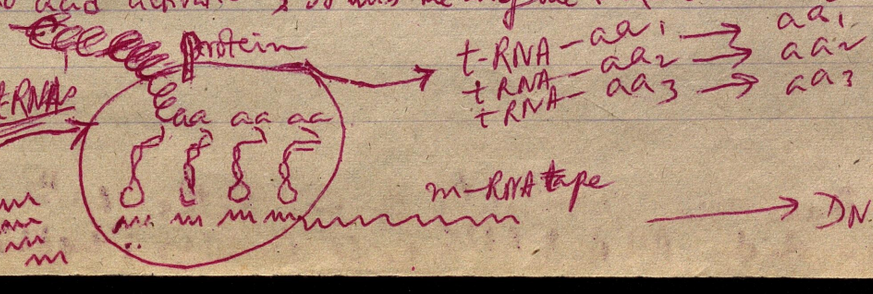


→ Neutral or OH

→ Polar: participate in bridging π-H paths

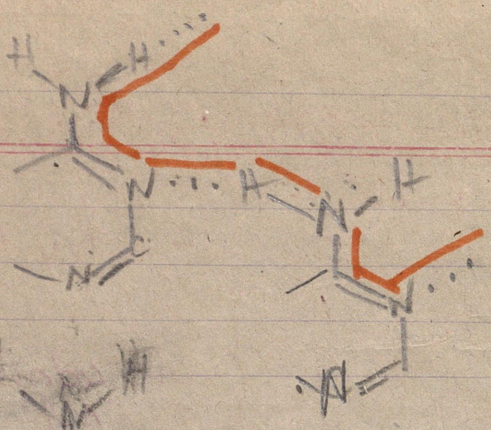
Predictions: ⊛ Termination codons: may not give the bridges
⊛ Modified aa do not react esp. c polar side chains

4. If this were to be true, the basis for mutual recognition is the interlinking π-H paths. Can this be used to reverse the system by interaction of the protein synthesizing machinery and amino acid activation? Is this the way the code evolved?

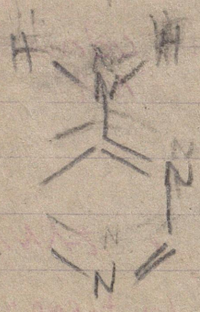


Is this the reason that triplets worked for binding t-RNAs?

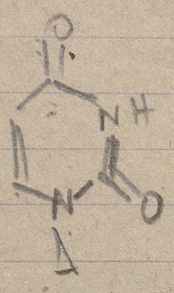
A



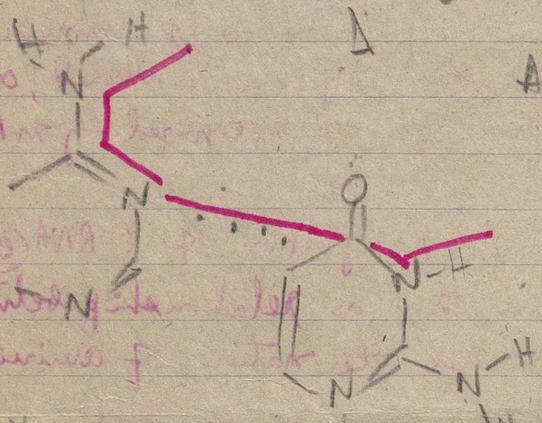
AA



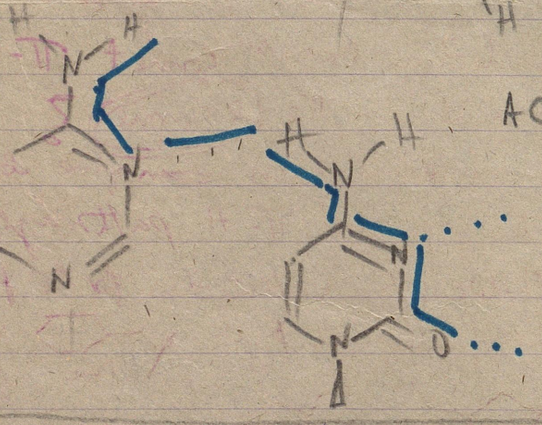
AU



?
Termination



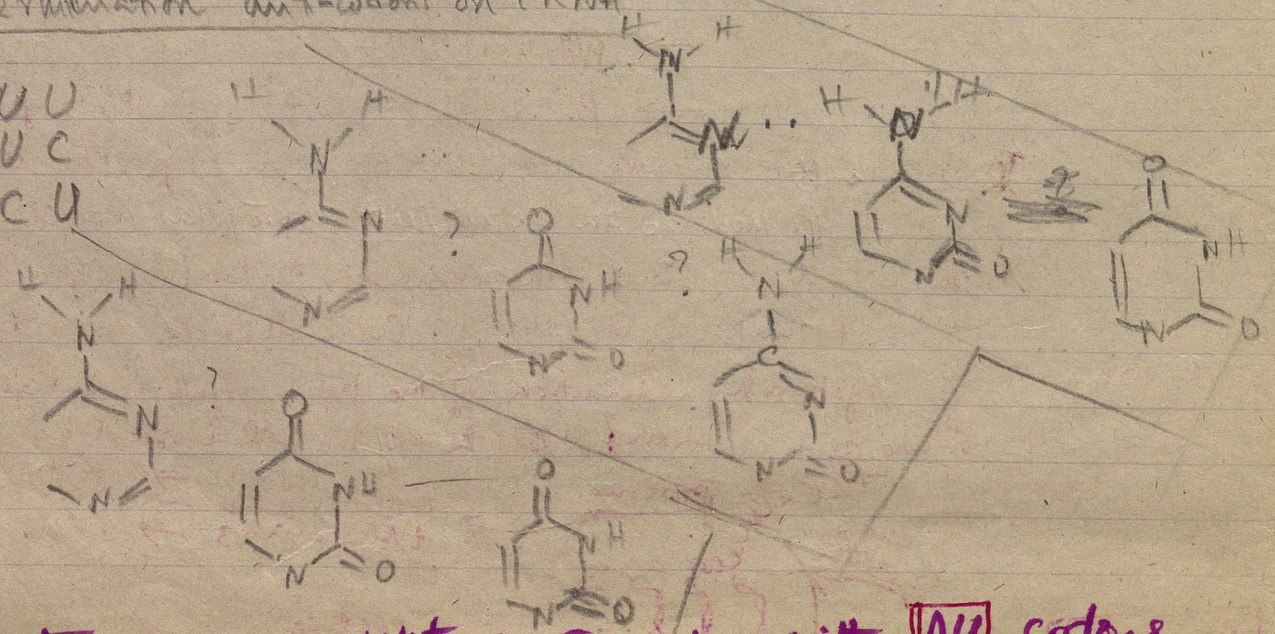
AG



AC

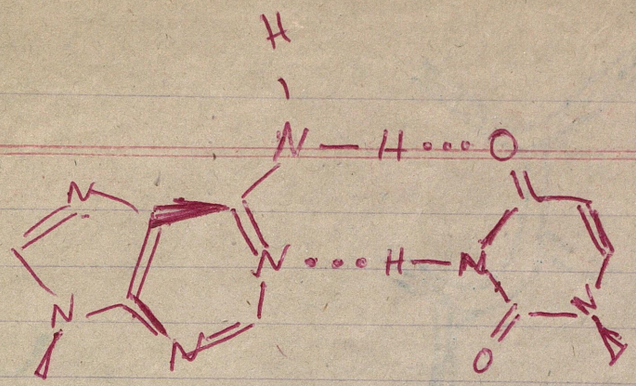
Termination anti-codons on tRNA

AUU
AUC
ACU

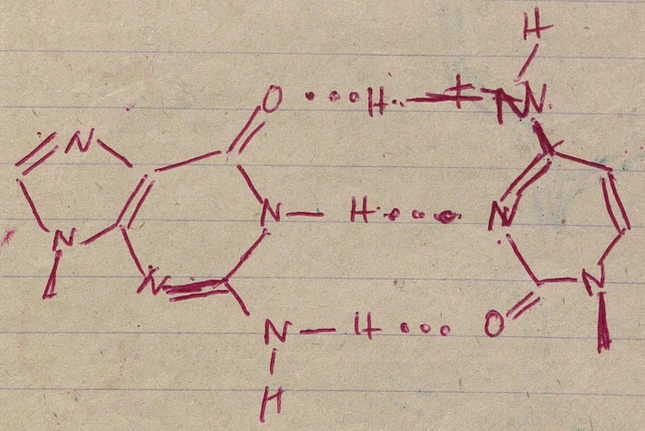


Is there no possibility of ⓔ paths with AU codons
How did AUA & AUG code for Tyr ; wobble on the first ?

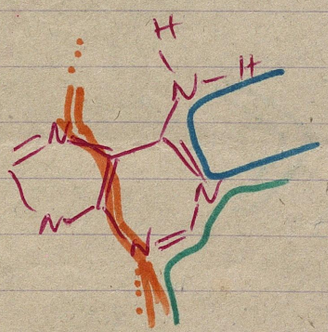
A–U



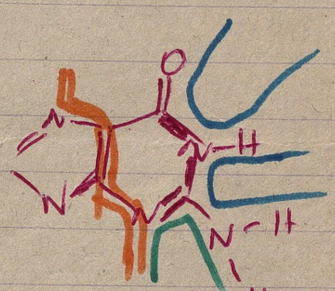
G–C



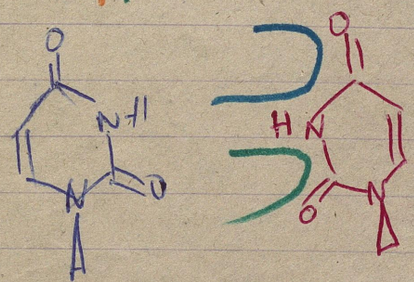
A

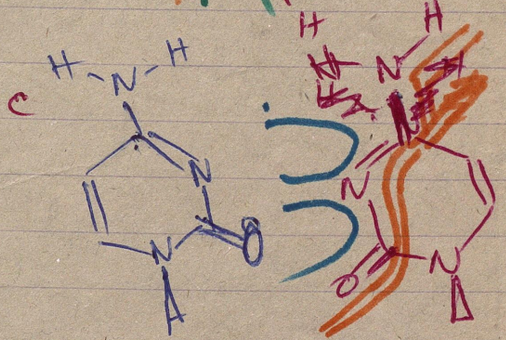


G



C





Anti-codons on t-RNA

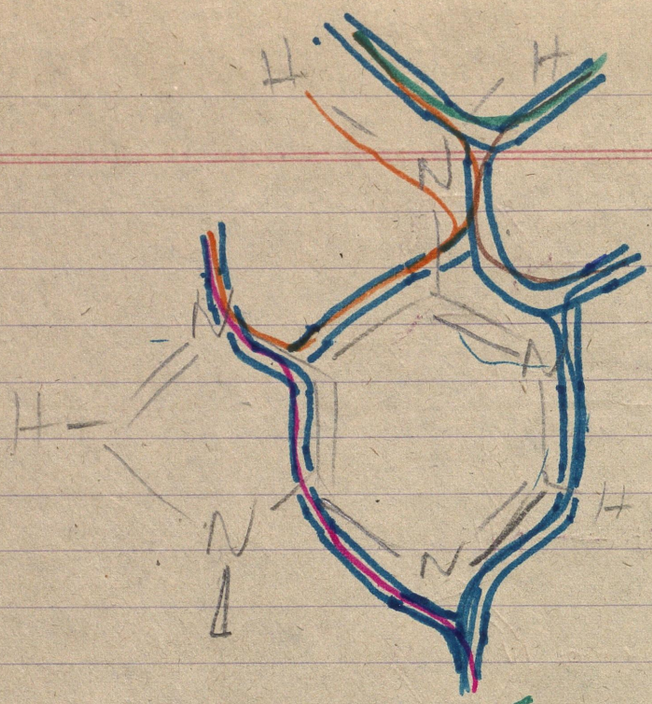
	A	G	U	C	
A	Phe Phe Leu Leu Ile	Ser Ser Ser Ser	Tyr Tyr CT_1 CT_2	Cys Cys (T3) Trp	A G U C
G	Leu Leu Leu Leu	Pro Pro Pro Pro	His His Gln Gln	Arg Arg Arg Arg	A G U C
U	Ile Ile Ile Met	Thr Thr Thr Thr	Asn Asn Lys Lys	Ser Ser Arg Arg	A G U C
C	Val Val Val Val	Ala Ala Ala Ala	Asp Asp Glu Glu	Gly Gly Gly Gly	A G U C
	Hydrophobic	neutral OH	polar	-SH Arg Gly Trp	

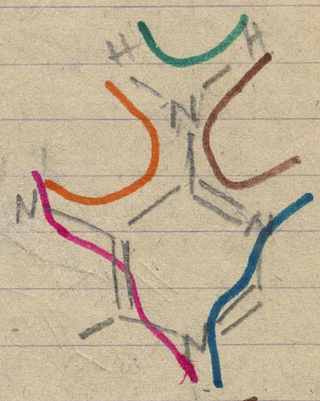
1 Phe 2 Leu 3 Ile 4 Met 5 Val 6 Ser 7 Pro 8 Thr 9 Ala 10 Tyr 11 His 12 Gln 13 Asn 14 Lys 15 Asp 16 Glu 17 Cys 18 Trp 19 Arg 20 Gly

	U	C	A	G
U	Phe	Ser	Tyr	Cys
C	Leu	Pro	His	Arg
A	Ile	Thr	Lys	.
G	Val	Ala	Asp	Gly

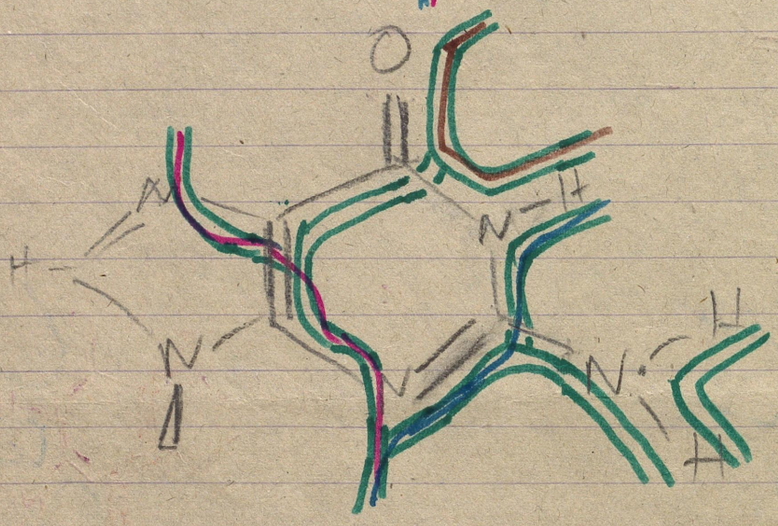
hydrophobic neutral polar

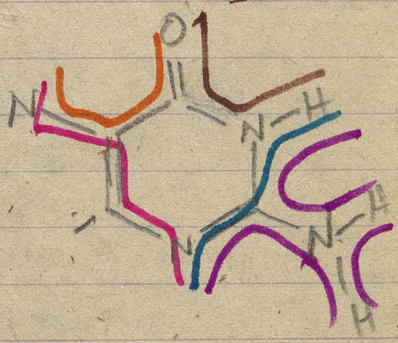
A



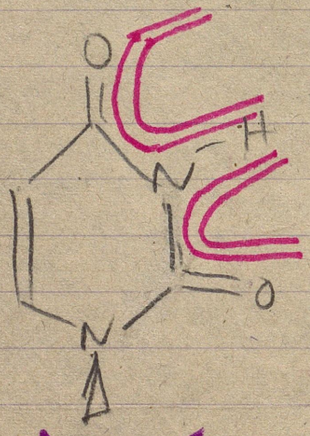


G

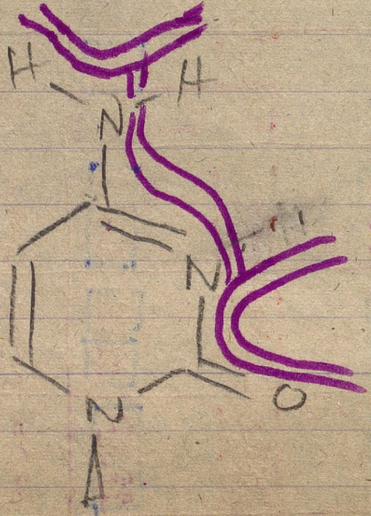




U



C



Some features:

1. 20 enzymes for 20 aa; no multiplicity of enzymes
2. same aa-enzyme for the tRNAs for each aa
   Ser (t-RNA): 3 tRNAs, all charged equally
3. each tRNA of same aa compete with each other.
4. t-RNAs of same aa in diffrt organism same anticodon and differs in ~~ges~~ nucleotides
   77% identical for Ser tRNA
5. No stereochemical similarity: charged groups, hydrophobic aa — how to distinguish 3 nucleotides
6. adaptor hypothesis of crick (1960-62)
   Dounce (1958)
7. Redundancy / degeneracy / multiplicity
   purpose! safety or specific
8. t-RNA used in heterologous system of protein synthesis
   anti codon recognized (?)
   mutation in anticodon — loss of aa charging
9. aa recognized code through tRNA

   aa — tRNA — code.

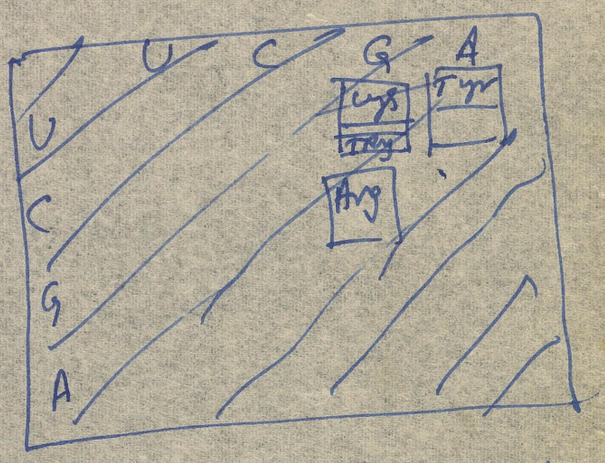
10. t RNA effects rate & specificity
    mistakes eliminated.
    Ile-activating enzyme charges Val but slow
    add tRNA(Ile) — Val released (not charged)
11. anticodon: competed with oligonucleotides 3-5 units for charging aa.
    Nirenberg, assay for codes
12. Poly Lys: (1961) synthesises poly A. (poly lys cofactor specific)

Biology & future of Man

p. 40: 20 aa — 1 nucleotide will not suffice
2 .. = 16
3 ✓

more triplets for each aa; redundant, degenerate
1961 UUU → Phe.

	U	C	A	G
U	Phe / Leu	Ser	Tyr	Cys / TRY
C	Leu	Pro	His / Gln	Arg
A	Ile / MET	Thr	Asn / Lys	Ser / Arg
G	Val	Ala	Asp / Glu	Gly



p. 46. extraordinary specificity in the attachment of aa to their correct t-RNA
each enzyme must specifically recognize an aa and its appropriate t RNA

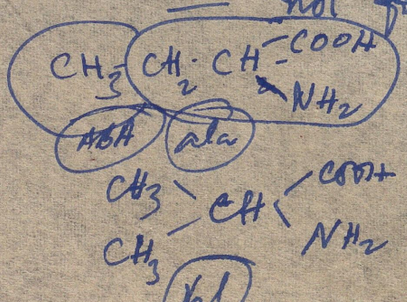
p. 48. attachment of aa to cognate t-RNA is entrusted to specific enzymes.
1 aa / 20 aa & 1 tRNA / 60 or 80. tRNA

p. 187 Origin of genetic code:

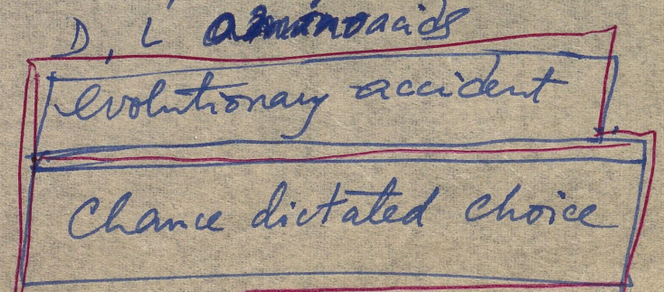
- no obvious stereochemical relationship between codon – cognate aa
- code assignments random choices
- Universality of code – stability to evolutionary changes
  not necessarily follows on stereochemical fit.
- premature to state code originated by chance alone

α-Aminobutyric acid: formed by spark discharge & metabolically — not found in proteins.

no ABA
no D-aa



D, L aminoacids



Recent Studies concerning the coding mechanism
T. H. Jukes & L. Gatlin Prog. in Nucl. Acid Research & Mol. Biol.
(J N Davidson & W. E. Cohn) II, 303 – 350 (1971)

P. 337. Tyr Suppressor tRNA : anticodon CUA } tRNA becomes charged ē Tyr
Tyr tRNA : GUA } but pairs ē UAG – chain terminator

(Ref. 4) Carbon J., Squires, C. & Hill, C.W. Cold Spring Harbor Symp Quant Biol. 34, 505 (1969)

gly-tRNA → GGA
Altered gly-tRNA → AGA (arginine) charged with glycine but low rate
– no loss of specificity for Gly in charging reaction

(Ref. 80) C.J. Burton & B.S. Hartley, BJ 108, 281 (1968).

same aa act enzyme for Met ± fMet

Met-tRNA ē Structures different
fMet-tRNA ē Charged ē Met by the same enzyme
equal length (77 Nucleotides); only 41 base are identical in homologous sites; anticodons identical.

(Ref 81) G. Gundarajas, J.R. Katze, D. Söll, W. Konigsberg & I. Lengyel, PNAS 61, 893 (1968)

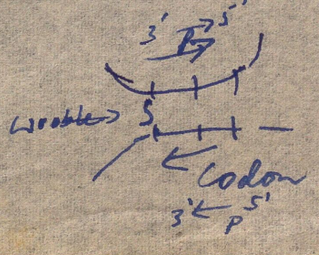
same aa act enzyme for all tRNA forms

Ser tRNA – 3 nos.; equally effective ē. tRNA synthetase
– anticodons – GGA, UGA & GCU
– kinetics similar in charging and heat inactivation
Compete with each other – UGA-tRNA competed with others for the same enzyme
– "exclude anticodon from recognition site" !!!

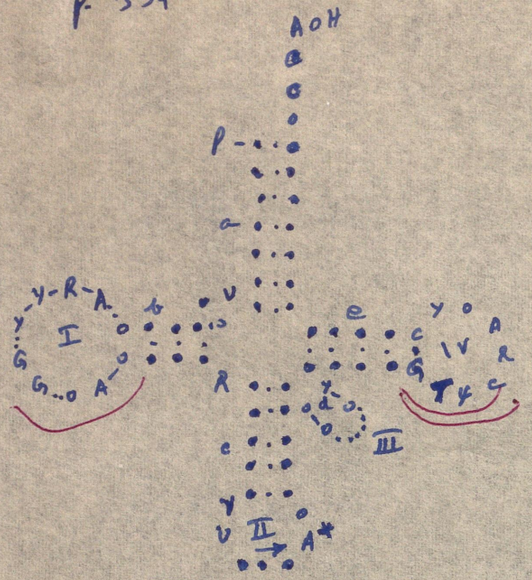
Molecular evolution of tRNA molecules:

1. Duplicate tRNA for same aa, same anticodon
   – may differ 2-3 nucleotides in same organism
2. tRNA in different organisms, ~~ē varying~~ same anticodon
3. tRNA for same aa – in pro- & eu-karyote
4. tRNA diffnt aa.

evolution has brought all tRNAs in groups 3 & 4



p. 339



I, II, III, IV loops
a b c d e helical base-paired (•·• H bonds) regions or "arms"
•—• bases in helical regions.
o unpaired bases
R; Y Pu; Py bases.
T; ψ ribothymidine; pseudouridine
* modified base
→ 3'→5'; ~ wobble.
III & d are sometimes absent.

13 sites identical in all
3 always purines
5 always pyrimidines

---

[21] Sea-tRNA yeast & rat
63 identical } 77% identical
19 different }

---

phe-tRNA yeast & wheat
58 identical } 82%
13 different }

---

16 codons on left for essential aa
∴ codons with middle U + 7 other scattered
Thr, Lys & Trp.
— nothing with codons? evolution !!!

The Genetic Code (The molecular basis of genetic expression)
C. R. Woese (Harper & Row, N.Y.) 1967.

p. 23 Unlike many of his contemporaries, Crick could not be induced to make conceptual jump to complementarity between oligonucleotides and amino acids. He objected:

"I can not conceive of ... RNA or DNA acting as direct template for amino acids, or at least as a specific template... if one considers the physico-chemical nature of the amino acid side chain we do not find complementary features on the nucleic acid. Where are the knobby hydrophobic surfaces to distinguish valine from leucine and isoleucine? Where are the charged groups, in specific positions, to go with the acidic and basic amino acids? ... I don't think that anybody looking at DNA or RNA would think of them as templates for amino acids" ... (Crick as quoted in Hoagland, M., in E. Chargaff & J. Davidson, eds. "The Nucleic Acids" III, A.P. N.Y, 1960).

He countered by enunciating a radically different mechanism for protein synthesis — devoid of hypothetical direct templates: "What the DNA structure does show — and probably RNA will do the same — is a specific pattern of hydrogen bonds, and very little else ... ... each amino acid would combine chemically, at a special enzyme, with a small molecule which, having a specific hydrogen-bonding surface, would combine specifically with the nucleic acid template... ... each amino acid is fitted with an adaptor to go on to the template."

— original statement of adaptor hypothesis

p. 44. Complexes of Ribosome-mRNA-tRNA

Ribosomes in presence of poly U, specifically and exclusively absorb the tRNA for Phe. and the complex absorbed on membrane filters (Leder, P. & Nirenberg, M.W., PNAS, 52 420, 1964).

Ribosome
|||
trinucleotides
tRNA
|
a

Size of oligonucleotide : 3 but not smaller

Tri U → tRNA-phe
Tri C → tRNA-pro
Tri A → tRNA-lys

absorption by trinucleotides is as good as polymers (4 or 5)

---

p. 47 aa — more than one tRNA

leu — 5 tRNA peaks
3 clustered, respond to poly UC
1 only to poly UG
1 " poly UA & poly U

CUU, CUC, UUG & UUA.

multiplicity

— degeneracy of code – manifested in terms of at least one tRNA for each codon; more of an exception than a rule

— multiple codon recognition by a single tRNA species are now common.

---

p. 48 multiplicity of tRNA's → multiple aa activating enzymes?

no

no multiplicity of aa act. enzymes

all failed: Yamane, T. & Sueoka N., PNAS, 51 1178 (1964)

2 cases: Neurospora: phe < tRNA / tRNA > different enzymes
Asp

P. 48 Universality of genetic code

Codon catalogue universal
translation machinery ? not ?

homo-copolymeric RNA: response similar in E. coli, yeast & Rat liver

episomal fragments — in vitro / in vivo
E. coli → S. marcescens alkaline phosphatase of E. coli
(Levinthal, C., Signer, E., & Fetterolf, K. PNAS 48, 1230 (1962))

t-RNA-aa can be used in heterologous system

tRNA's: Charged tRNA prep. from E. coli, yeast .. added to Hb synthesizing system (Reticulocytes), aa on these tRNAs are correctly placed in protein
(Weisblum, B., Benzer, S., Holley, R. W., PNAS, 48, 1449, 1962)

P. 111: Tape reading process

uses base pairing mechanism
Output tape RNA synthesized from 5' → 3' end

Ribosome: 60% RNA, 40% protein → 70S
16S 0.5 s-$\times 10^6$
23S 1.10 $\times 10^6$.
→ 40 species 26 — 50S
14 — 30S

30S — binds triplet

P. 114 TRANSFER RNA:
- evolved to keep translation errors down
- implicated in regulatory processes
- specificity of aa it carries
- tRNA can tell us most about basic nature of code

composition: 20% modified nucleotides
far more compact, rigid or regular structure than other RNAs
Sharp hyperchromic response ~25% ↑ OD_{260}
tRNA alone manifests solubility in high salt concn.
resistant to endonuclease cleavage.

anticodon	IGC	GΨA	G^{ome}AA	IGA
codon	GCX	UAPy	UUPy	UCX
aa	Ala	Tyr	Phe	Ser

aa—tRNA—code

t. RNA: site to recognize some aspect of ribosome
activating enzyme (aa)
site for tRNA-tRNA interaction (common to all tRNA)
unique to each

p. 122. translation process I : aa activation.

$$aa + E + ATP \rightleftharpoons \underbrace{aa\text{-}AMP \cdot E}_{\text{enzyme bound}} + PP \quad (1)$$

$$aa\text{-}AMP\text{-}E + tRNA \rightleftharpoons aa\text{-}tRNA + E + AMP \quad (2)$$

Is tRNA in any sense recognizing the aa that it will carry? (R.B. Loftfield & E.A. Eigner, J. Biol. Chem. 240, PC 1482, 1965)

(*) tRNA effects rate & specificity →

t-RNA has a pronounced effect not only on the rate of reaction (1) but also on its specificity

Ile-activating enzyme: distinguishes Ile from Val; Val activated by the enzyme is not placed on any tRNA

✓ Val-aa-AMP — Ile act. enzyme is released to solution, specifically in the presence of tRNA Ile as free Val (Baldwin, A.N. & Berg P. JBC 241, 839, 1966).

Mistakes in charging tRNA possible but transfer slow.

aa-AMP → aa wrong; wrong tRNA

Phe-activating enzyme: activates p-F-Phe (1/10 rate of phe), p-F-phe is placed on $tRNA_{phe}$ (1/4 rate) (Flanagan, W.L. & Neidhart, F.C., JBC, 239, 1844, 1964)

p. 128. competition for aa-charging:

Hayashi H. & Miura, K., Cold Spring Harbor Symp. Quant Biol. 31, (1966).

nucleotides - ν 5 units complete aa-charging on tRNA

nucleotides of anticodon compete ∴ charging tRNA = aa

Inhibition of charging of tRNA for		Oligonucleotide A	U	G	C
Phe	UUU	++	±		−
Lys	AAA	±	++		+
Pro	CCC	−	−	++	±
Gly	GGG	−	−	−	++

p. 179

Evolution of the genetic code

complex machinery — astounding precision

aa-act. enzyme: 95% precision for aa activation; corrected further in transfer of aa to tRNA. error frequency 10^{-3}–10^{-4} mistake/charging overall.

tRNA: small, but complex structure; 20% modified post-transcriptionally by highly specific enzymes.

functionally complex: recognize aa act. enzyme
anticodon codon
ribosome (probably)
sites on other tRNA (perhaps)
aa ? it carries ?

p. 193. RNA-polypeptide associations:

direct feed-back from protein primary structure —evolution?→ nucleic acid primary structure

---

☆ Dolin M. I. BBRC 6, 11 (1961).

very interesting
poly Lys → (Lys)n synthesizes Poly A, requires polyLys. as specific cofactor

did poly Lys breakdown to Lys — not done

---

evolution of code: historical accident

4. tRNA fragments

☆ anticodon is necessary for aminoacylation

(A.A. Bayev, I. Fodor, A.D. Mirzabekov, V.D. Axelrod & L.Y. Kazaninova, Mol. Biol. 1, 414 1967).

yeast $tRNA^{Val}$

★ anticodon is dispensable

(R. Thiebe, K. Harbers & H.G. Zachau, E.J.B. 26, 144, 1972)

yeast $tRNA^{Phe}$

3/4/92

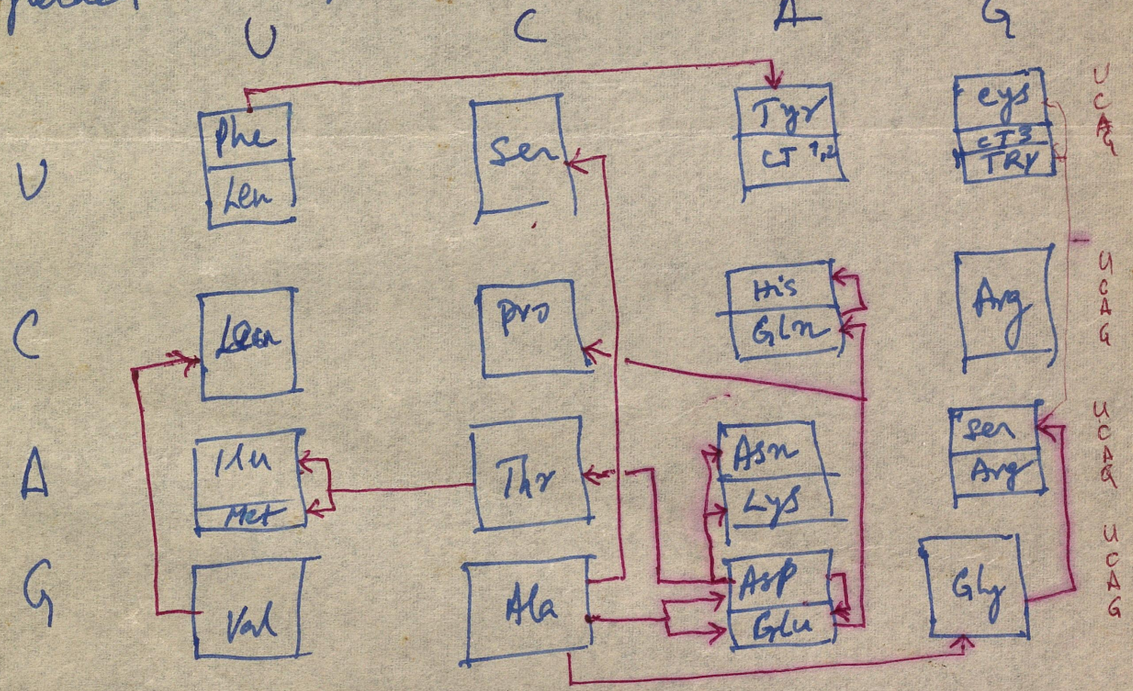
1. Distribution: 1. primarily 4 & 2
   2. Ile lost one to Met
      Cys " " " Try
   3. 3 CT.

2. no correlation of no. of codes for any type.
3. no stereochemical relationship between aa & nucleotide
4. code assignments random choices. (Stochastic)
5. Universal code – no evolutionary change
6. 64 — 3 = 61

	D	L	
	20+20	= 40	+20 (10+10 aa)
	20+19	= 39	+21 ? lost.

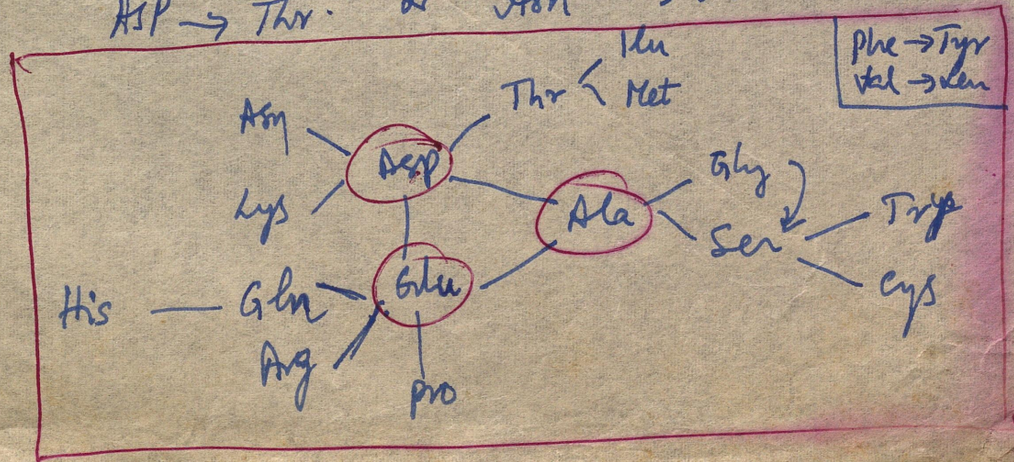
Why Gly = 4 ?

Coevolution: Wong PNAS (1975)
yielded a code pair or word ~~to~~ each product



one base change in First or Second letter

Asp → Thr. or Asn → Thr. Gln → Thr.



14/4/72

Aminoacyl-tRNA Synthetases

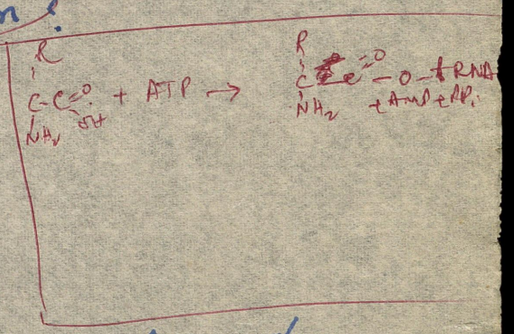
D. Söll & P.R. Schimmel in The Enzymes (PDB) Vol-X, P. 489-538 (1974)

Specificity (P. 506): Requires free NH_2 group; carboxyl not reqd.
- tyrosinamide or tyrosinol - act as competetors.
- Ile tRNA activates Val and Leu; Val. is not transferred to tRNA; hydrolyzed.

Recognition problem: (P. 522)

1. General sites: Common to all tRNAs - wide divergence
2. Specific sites: anticodon?

Ap. site for anticodon ✓



1. sequencing of isoacceptor tRNAs
2. mutant tRNAs
3. Chemical modification
4. dissected molecule
5. Heterologous aminoacylation
6. topology of aminoacyl tRNA Synthetase · tRNA complexes

1. $tRNA^{met}$ - $tRNA^{fmet}$ differ 36/77 (more close to Val!) equally well with one enzyme. Enzyme same tRNA different
2. Suppressor tRNA: partial or complete loss
   i Su^{+}_{III} tRNA ⇏ tRNA mutants with altered aa specificity.
   - nucleotide analysis revealed tRNAs Su^{+}_{III} $tRNA^{Tyr}$ is altered by a different single base near the acceptor end.
   - aminoacylated by Glu in vitro & in vivo and also by Tyr by Tyr. tRNA Synthetase (no! same enzyme)
   ii mutation in anticodon E.Coli $tRNA^{Trp}$ → new amber suppressor tRNA, aminoacylated by Gln.

3. C → U (bisulfite catalyzed)
   Val tRNA loses biological activity when anticodon sequence IAC → IAU
   (R.W. Chambers, S. Aoyagi, Y. Furukawa, D.S. Bhatnagar, H. Zawadzka Fed Proc. 23, 585 (1973))

mutation of anticodon - act. lost

! Misaminoacylation of chemically modified tRNA has never been observed.

Katz Top. Biochem 73

AA activation Reaction catalyzed by Met-t-RNA synthetase
S. Blanquet, G Fayat & J.-Pierre Waller
Lab. d'Eng. du CNRS 91190 Gif-sur-Yvette, France
JMB 94 1-15 (1975)

Atl-Mg: all 3-P.
Glassman, Klopman, Cooper Biochem. 12, 2430-2437 (1973)
Kuntz GPP & Swift TJ Fed Proc, 32, 546 (1973)

On the origin of the genetic code and the stability of the translation apparatus

G. W. Hoffmann JMB 86 349-362 (1974).

The Amino acid activation reaction catalyzed by methionyl-tRNA synthetase.

S. Blanquet, G. Fayat & J. Waller J MB, 94, 1-15 (1975)

Aminoacylation: Review: Söll D. & Schimmel, P.R. The Enzymes (PD Boyer) 1D, A.P. (1974)

Adenosine & 8-Amino adenosine: competitive inhibitors of ATP-Mg^{++} in ATP-PP exchange reaction

---

ATP-Mg^{++} Structure in solution.

Glassman T.A., Kloprman G. & Cooper, C
Biochemistry 12, 5013-5019 (1973)
Kuntz GIP & Shift TJ Fed Proc. 32, 546 (1973)
Mg reacts with all 3 P, with N_1 or N_3 of Pyrimidine ring ??

Yeast Phenylalanyl-tRNA Synthetase: Evidence for the triggering of an AMP-ATP exchange by tRNA

P. Remy & J. P. Ebel: Inst. de Biol Mol. et cellulaire
15 Rue Descartes 67000 Strasbourg, France
FEBS Letts 61, 28-31 (1976)

aminoacylation of $tRNA^{Phe}$ by yeast phe-tRNA synthetase (PRS)
- an intermediate complex different from aminoacyl adenylate-enzyme complex.
→ AMP–ATP exchange should occur

$$ATP + Phe + PRS \rightleftharpoons Phe\text{-}AMP\text{-}PRS + PP_i$$
$$Phe\text{-}AMP\text{-}PRS \rightleftharpoons Phe\text{-}x\text{-}PRS + AMP$$

Such an exchange has never been observed in the absence of tRNA, and easily observed in its presence but does not demonstrate the presence of Phe-PRS

★ tRNA may be required to induce the conformational change of protein, triggering the formation of aminoacyl-

non-acylable $tRNA_{phe}$ – terminal adenosine replaced by 2'deoxyadenosine is able to induce the AMP-ATP exchange.

Ref.: 1. Murayama A., Raffin JP, Remy P & Ebel J.P. FEBS Lett. 53, 15-22 (1975) (first report on complex)
2. Fasiolo, F. & Ebel, J.P., EJB 49, 257-263 (1974) (prep of enzyme from yeast)
3. Maas, R.S. & Novelli, G.D. ABB 99, 48-53 (1961) (assay phe-tRNA synthetase)

Tamilnadu Outbreak 1994 (31.7.94 onwards)

Pasumpon District (near Madurai)

152 Goats affected *

35 " died. *

Total Goat Population : 500 goats

* This no. may be have more ~~before 150~~ over the next month, as there was no follow-up.

21 semantic designations

64 places.

21 places must have unique designations

1st place can have any of 21 designations

2nd place any of twenty, thus it is sampling without replacement.

Thus it becomes $21 \times 20 \times 19 \ldots \times 2 \times 1$

or $!21$

The first 21 places can be filled up in $!21$ possible ways.

There is no restraint on the remaining 43 places, $\therefore$ there is sampling with replacement.

43 places can be filled with 21

be argued that natural selection may not have played a role in leading to fixation of one genetic code (it may be considered an allele among many) in a population of primordial organisms that initially had multiple codes. Genetic drift and perhaps a headstart effect (when a sub-population with a particular code might have cornered a large amount of resources, that aided in rapid population expansion) might be called in. Thus the current genetic code may have become fixed entirely due to population dynamics phenomena operating in small populations.

Informed Consent form

- **Title:-**Studies on Cathepsin C gene mutations in Indian patients with Papillon lefevre syndrome.

- **Investigators:- Dr. Yogesh S Shouche (National Center for Cell Sciences, Pune)**

  Dr.Nihal Devkar (Rural Dental College, Ahmednagar)

- **Scientific background and the possible benefits that would be obtained from the study:** -The purpose of this study is to look for mutations in the cathepsin C gene of Indian Patients with Papillon Lefevre Syndromes. Novel mutations can be located through sequencing of the cathepsin C gene and the pathogenecity of particular mutation can be accessed by looking for biochemical characteristics of the protein coded by this gene.

- **Confidentiality: -** Confidentiality of records (if required), regarding the identification of the patients will be maintained.

- **Risk of participation: -** An explanation of any involvement of risk in obtaining samples will be given. The samples will be obtained through doctors who are treating the patients.

- **Types of samples and frequency of collection: -**As far as possible, both samples will be collected only once from the patient, incase cell lines are not established, samples may be collected again. The samples collected may include:

  a. Peripheral blood and
  b. Skin biopsy

  The patients will be given an explanation about whom to contact if they have any pertinent questions about the research.

- A statement of any explanation of the purpose and duration of research will be made.

- A description of any benefits to the subject will be explained.

The purpose of this study is to understand the pathogenic ~~mutate~~ mitochondrial DNA mutate in Indian population. The scientific information and interpretate of this study will be published in scientific journal.

semantic designations as follows

$$21_1 \times 21_2 \times 21_3 \cdots\cdots \times 21_{42} \times 21_{43}$$

or 21^{41} ways.

∴ The total no. of genetic codes possible with 21 semantic designations (read amino acids & STOP) for a total of 64 places (read codons) is

$$!21 + 21^{43}$$

Of such a large number, only one genetic code has ultimately survived.

Given that mitochondria and other organelles have codes that are different from the standard code slightly, and yet function as efficiently (or perhaps more) as the standard code, it may

Here is an expansion and hopefully clarification of the model I presented in footnote #14 of the paper.

① Basically, a mothers fitness, ~~W~~ W, in terms of number of her gametes depends on the frequency of sib matings, F_1; daughters outcrossing, F_2; and sons outcrossing, F_3 times the relative value (relatedness) of each, R_1, R_2, R_3 respectively. $W = R_1F_1 + R_2F_2 + R_3F_3$. F_1, F_2, F_3 are a function of local conditions and are, potentially at least, under the mothers control.* R_1, R_2, R_3 are properties of the population a result of its mating structure and unless a mother knows whether she is sib mated or not (I assume a mother can not tell) the relatedness values are not something a mother should be expected to respond directly (ie facultatively) to.

Now given: N = the total number of individuals in a local mating populat
m = the proportion of those individuals which are offspring of the mother whose fitness we will calculate
p = the proportion of that mothers offspring which are sons
q = the proportion of all other offspring which are male
n = foundress number (in most of the following derivation $m = 1/n$ unless explicitly stated otherwise)

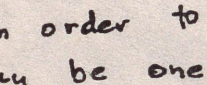
$F_1 = Nm(1-p)$ [#of daughters] $Nmp/(Nmp + N(1-m)q)$ [prop of sons in ♂ population]

$F_2 = Nm(1-p)\ N^{(1-m)q}/(Nmp + N(1-m)q)$ (daughters out)

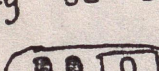
$F_3 = N(1-m)(1-q)\ Nmp/(Nmp + N(1-m)q)$ (sons out)

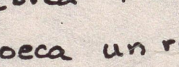
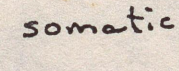
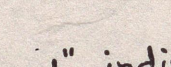
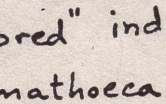
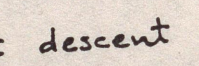
– N's cancel in prop of sons / unrelated mates in the total male population –

In order to figure R_1, R_2, & R_3 we consider that any given gravid female may be one of two types either

1)  ("outbred" individual)
with gamete in spermathoeca unrelated to either somatic gamete

or

2)  ("inbred" individual)
with gamete in spermathoeca related by direct descent to one of two somatic gametes

	for outbred mother		for inbred mother	
sib mated daughter		$R_1 = 2/3$		$R_1 = 3/3$
out crossed daughter		$R_2 = 1/3$		$R_2 = 2/3$
out crossed son		$R_3 = 1/3$		$R_3 = 1/3$

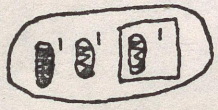
* thru control of sex ratio of her offspring

In any population a gravid mother has a probability, i, that she is inbred. Thus for an average mother in a population

(2) $R_1 = (2/3 + (1/3)\,i)$
$R_2 = (1/3 + (1/3)\,i)$
$R_3 = (1/3)$

We find i in the following manner:

If we number the gametes there are four possible types of gravid females in the population:



= A_1, all gametes identical by descent

= A_2, mother homozygous with unlike gamete in spermathoeca

= A_3, mother heterozygous with gamete in spermathoeca identical to one of her two gametes.

= A_4, mother heterozygous with unlike gamete in spermathoeca

A_1, A_2, A_3, A_4 are the proportion of each type in the population

$A_1 + A_2 + A_3 + A_4 = 1$

$[A_1 + A_3] = i$ (both fullfil the condition of "inbred" mother)

now if offspring are sib mated

all A_1 offspring are A_1 type
all A_2 " " A_3 type
$1/4$ A_3 " are A_1 type
$1/4$ A_3 " are A_2 type
$1/2$ A_3 " are A_3 type
$1/2$ A_4 " are A_3 type
$1/2$ A_4 " are A_4 type

if offspring outcross assume they recieve unrelated sperm

all A_1 offspring outcrossed are A_2 type
all A_2 " " A_4 type
$\frac{1}{2}$ A_3 " " A_2 type
$\frac{1}{2}$ A_3 " " A_4 type
all A_4 " " A_4 type

given n as foundress number (to begin with consider it constant)

$\frac{1}{n}$ individuals are sib mated and
$n-1$ " " outcrossed

$\phi = 0$ ie no contribution

So

③ $A_1(t+1) = \frac{1}{n} A_1(t) + \phi A_2(t) + \frac{1}{4n} A_3(t) + \phi A_4(t)$

$A_2(t+1) = \frac{n-1}{n} A_1(t) + \phi A_2(t) + \frac{2n-1}{4n} A_3(t) + \phi A_4(t)$

$A_3(t+1) = \phi A_1(t) + \frac{1}{n} A_2(t) + \frac{1}{2n} A_3(t) + \frac{1}{2n} A_4(t)$

$A_4(t+1) = \phi A_1(t) + \frac{n-1}{n} A_2(t) + \frac{n-1}{2n} A_3(t) + \frac{2n-1}{2n} A_4(t)$

at equilibrium $[A_1 + A_3] = i = 1/(2n-1)$

There is a matrix M_1 which is appropriate to describe transitions between generations with 1 foundress. There is $M_2, M_3 \cdots M_n$ appropriate for 2, 3, ... n foundresses. If there is a distribution of foundress number in one generation such that some proportion, p_1, of ~~[illegible]~~ mating populations are founded by single foundresses, p_2, by 2 p_3 by 3 · etc. then the matrix appropriate to describe the transition is $p_1 M_1 + p_2 M_2 + p_3 M_3 + \cdots p_n M_n = M_T$

(here I assume equal female production ~~from~~ for all foundress numbers, the data I've collected is conflicting on this point. In some instances more total ♀s come from single foundressed figs and in some instances fewer total ♀s come from " " ".)

Further if distributions vary through time the average degree of inbreeding through time is very closely estimated by taking the composite matrix using the distribution ~~fo~~ of foundresses through time (again $p_1 M_1 + p_2 M_2 + \cdots p_n M_n$) and ~~using that matrix to~~ computing the equilibrium $[A_1 + A_3]$ ~~i is~~ $1/(2n-1)$.

A few results:

1) Average foundress number is not necessarily a good indicator of level of inbreeding.

eg. If all figs have 3 foundress always the proportion of ~~inbred~~ sib-mated individuals each generation $\frac{1}{n} = 1/3$. $\frac{1}{2n-1} = 1/5 = .2 = i$

If 1/3 of all figs have 2 foundresses, 1/3 have 3, and 1/3 have 4 then $(\frac{1}{3} \times \frac{1}{2}) + (\frac{1}{3} \times \frac{1}{3}) + (\frac{1}{3} \times \frac{1}{4})$ ~~one~~ or 13/36 are sib mated each generation
$(13/36 > 12/36)$ $\frac{1}{n} = \frac{13}{36}$ $\frac{1}{2n-1} = .2203 = i$

Finally 1/2 of all figs have 1 foundress and 1/2 have 5 foundresses

[left margin, written vertically:] → $\frac{1}{2} + \frac{1}{10} = \frac{6}{10} = \frac{1}{n} \therefore \frac{1}{2n-1} = .4286 = i$

arithmetic

In all three cases the mean foundress number is the same but the level of inbreeding in the population is different and depends on the distribution.

So the proportion of sib mating in a population which in turn controls the level of inbreeding is given by

$$p_1\left(\tfrac{1}{1}\right) + p_2\,\tfrac{1}{2} + p_3\,\tfrac{1}{3} \cdots p_n\,\tfrac{1}{n} = \frac{1}{n_T}$$

the level of inbreeding $i = 1/(2n_T - 1)$

Now to incorporate the level of inbreeding into the degree of relatedness R_1, R_2, R_3 and put those into the equation

$$R_1 = \frac{2}{3} + \frac{1}{3}\frac{1}{(2n_T-1)} = \frac{4n_T-1}{3(2n_T-1)}$$ (relatedness of sib crossed daughters)

$$R_2 = \frac{1}{3} + \frac{1}{3}\frac{1}{(2n_T-1)} = \frac{2n_T}{3(2n_T-1)}$$ (relatedness of out crossed daughters)

$$R_3 = \frac{1}{3} = \frac{2n_T-1}{3(2n_T-1)}$$ (relatedness of out crossed sons)

$$W = R_1F_1 + R_2F_2 + R_3F_3$$

$$W = \frac{4n_T-1}{3(2n_T-1)}\,Nm(1-p)\frac{mp}{mp+(1-m)q} + \frac{2n_T}{3(2n_T-1)}\,Nm(1-p)\frac{(1-m)q}{mp+(1-m)q} + \frac{2n_T-1}{3(2n_T-1)}\,(1-m)(1-q)\frac{mp}{mp+(1-m)q}$$

divide thru by $N/3(2n_T-1)$

W' = fitness (gametes identical by descent exported per individual in the mating population, the efficiency with which the population is exploited as a resource)

then differential with respect to p, set equation = 0
the result is equation #1 in footnote 14
set $q = p$ and $n_T = n$ (all populations with n foundresses)

$m = \frac{1}{n}$ solve for p

and

$$p = (2n-1)(n-1)/n(4n-1)$$

which is the Hamilton Taylor-Bulmer result for constant # of foundresses with equal contribution.

Similarly, given R_1, R_2 & R_3 and m (proportion of sites, each mother contributing equally could use $1/n$)

$$p = \frac{R_2(m-m^2) - R_1(m-m^2) - R_3(1-m)^2}{-R_1m^2 - 2R_1(m-m^2) - R_2(1-m)^2 - R_3(1-m)^2}$$

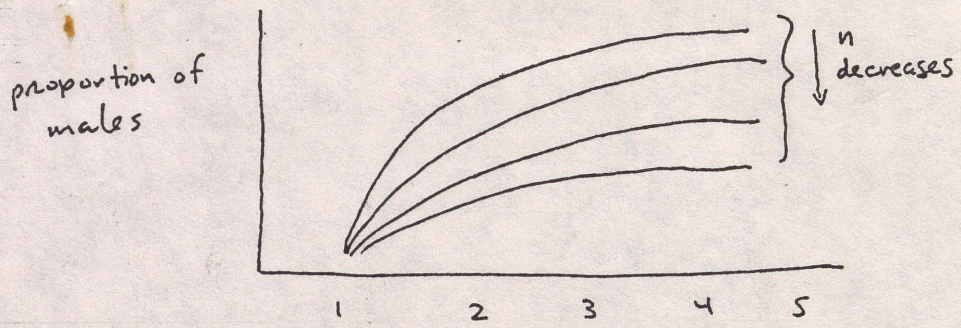
gives the ESS sex ratio appropriate for any degree of inbreeding and given

(5)

another special case is

$$(1-m)(2n-1)/(4n-1)$$

which is the equation which predicts the nested curves



$m = \frac{1}{\#\text{ of foundresses}}$

n = ~~arithmetic~~ harmonic mean number of foundresses

Two errors in the paper:

#1 in figure ~~two~~ three the three graphs on the right are mislabelled The y axis reads "Proportion of Males in brood". It should read "arc sin transformation of proportion of males in brood".

#2 In footnote 20 the analysis was performed using SAS General Linear Models program

---

Finally

I now have similar data on 11 species of fig pollinating wasps. The same patterns hold up across these species IE both number of foundresses in a fig (LMC) and harmonic mean number of foundresses in a species explain significant portions of the overall variation ($p < .0001$) of sex ratio. However 8 of the 11 species show lower, more female biased sex ratios than predicted. Also 2 species do not show much if any evidence for sex ratio adjustment. These two species only very rarely experience more than 1 foundress. I tentatively conclude that no adjustment occurs to the 2 foundress condition since it is so rarely encountered in these two species.

Sincerely yours

Mike Hamilton

Nucleocapsids of Rinderpest virus

Length of Particle	1157 ± 22nm
# of helical turns	185 ± 12
Pitch of helix	6.5 ± 0.6nm
external diameter	17nm
diameter of hollow core	5nm
Subunit angle	60°
helix angle	12°.

RIBONUCLEOPROTEIN PARTICLE
(Subviral Particle)
RNA Genome (15,000 bases) in the central core
& contains one major protein N which covers the RNA
& 2 minor proteins

Changes in the e.m. structure of Nucleocapsids from persistently infected cells

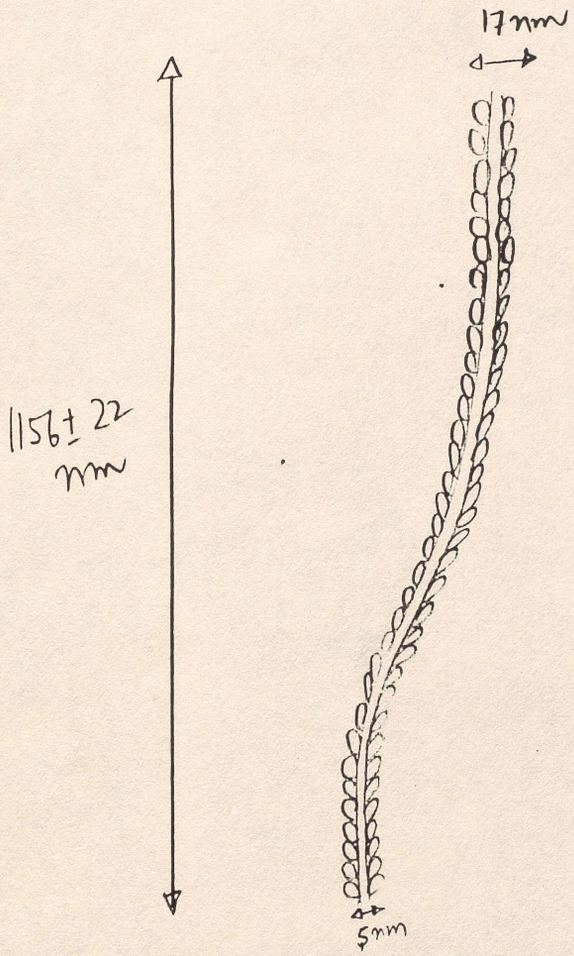
Pitch of the helix 5.53 ± 0.19 nm

Helix angle 8°.

Conclusions drawn on the str:

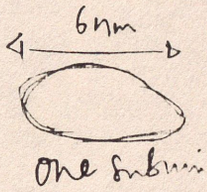
the nucleocapsids of persistent virus is more condensed.

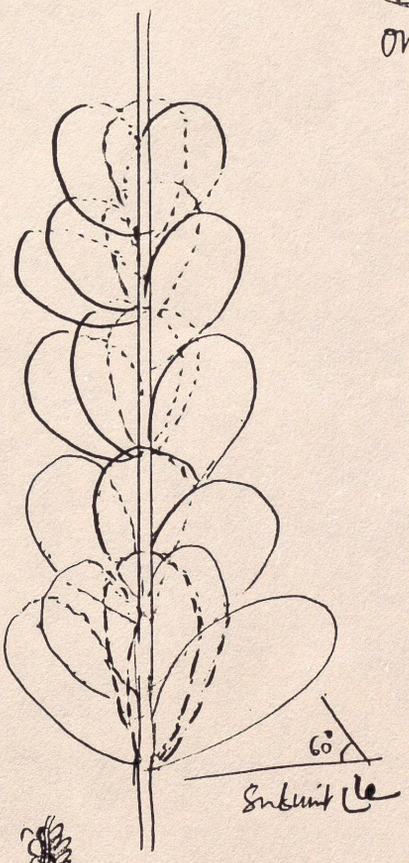
Two dimensional Cartoon of
a three dimensional Nucleocapsid Particle as seen on
an electron microscopic grid



herring-bone appearance (See next Page for e.m.)

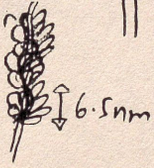
185 sub units for the unit length of core RNA





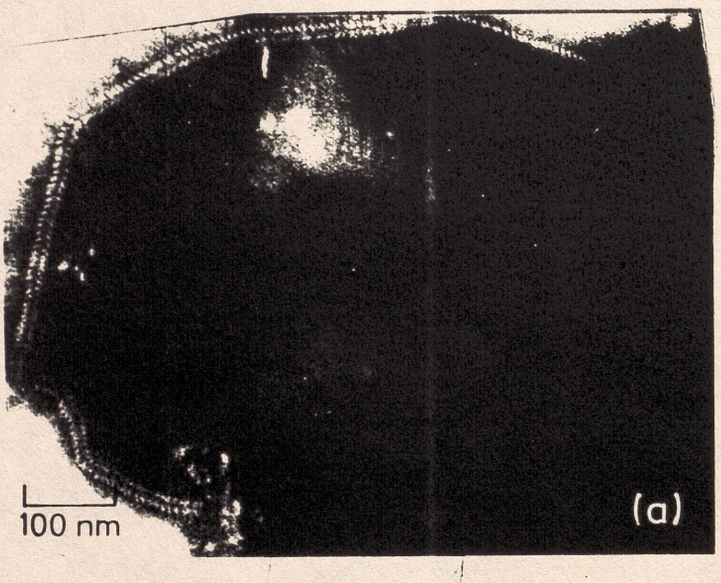
Subunit is N Protein (Mr 66,000)
it is positioned at
an angle of 60° to the
long axis of the
helix

Helix Le
12°

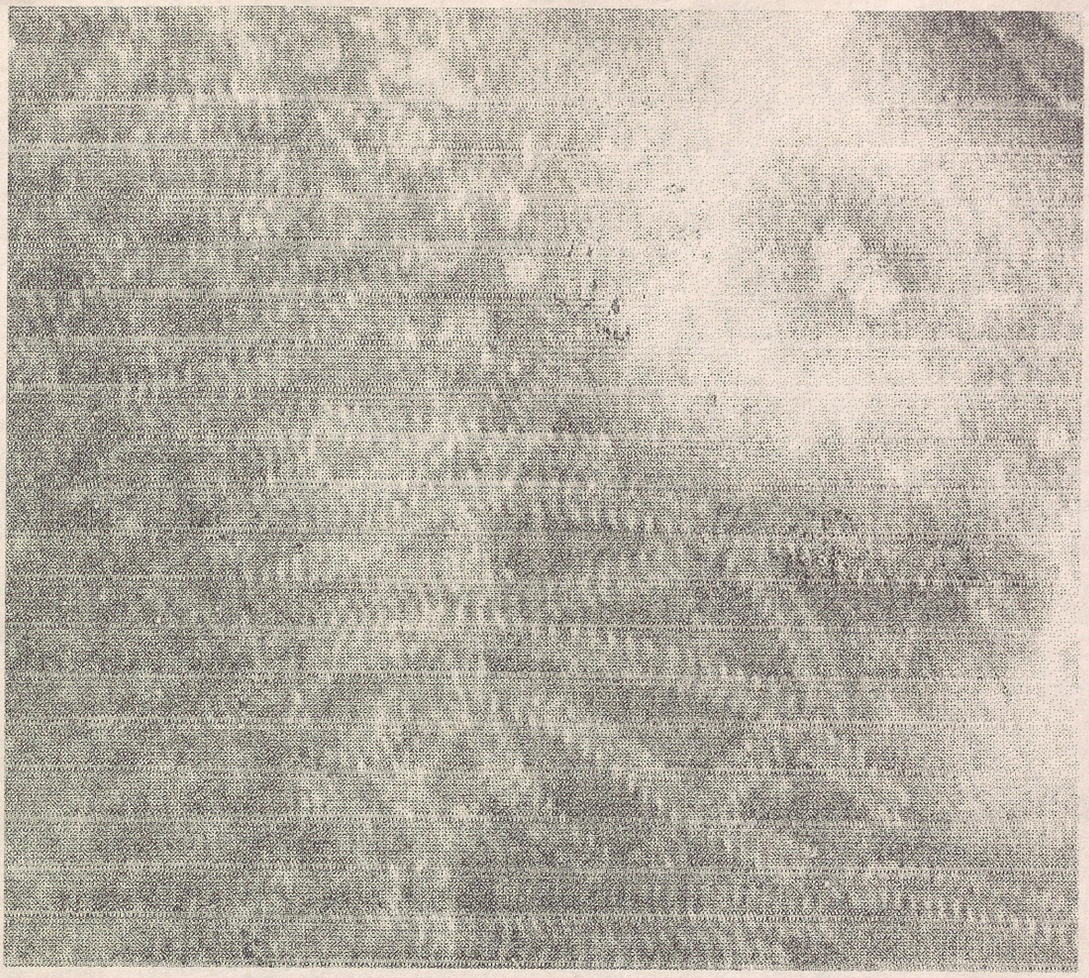


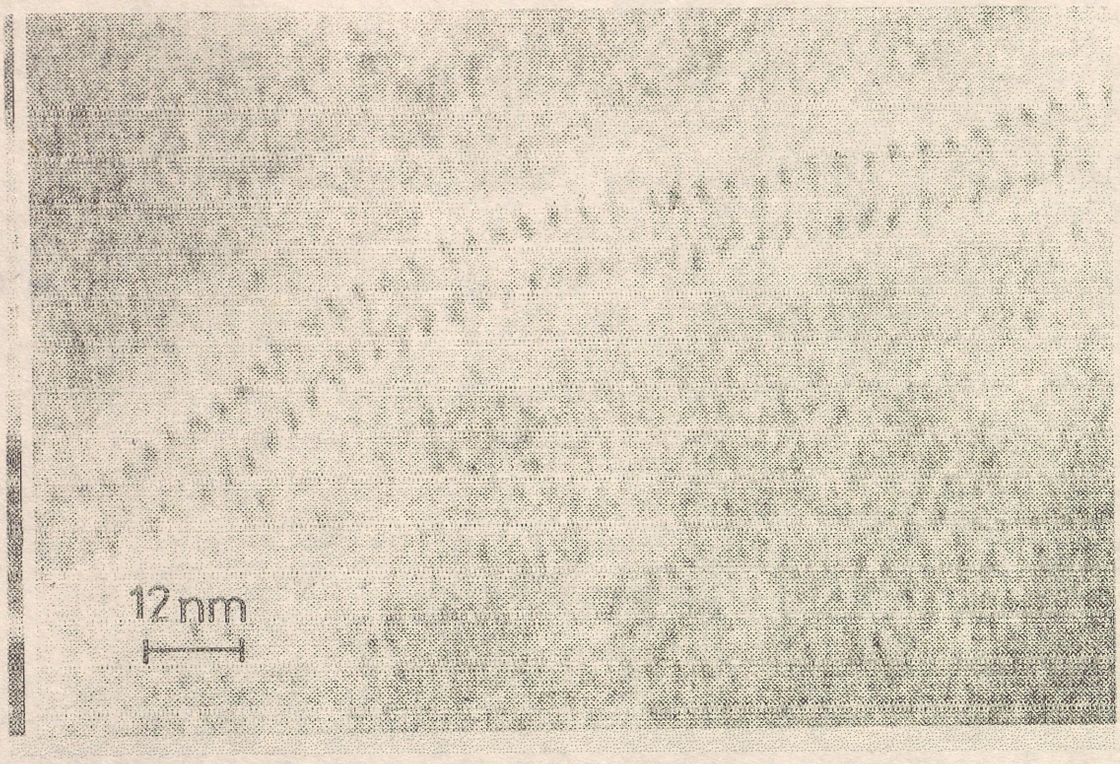
View from top of 3rd

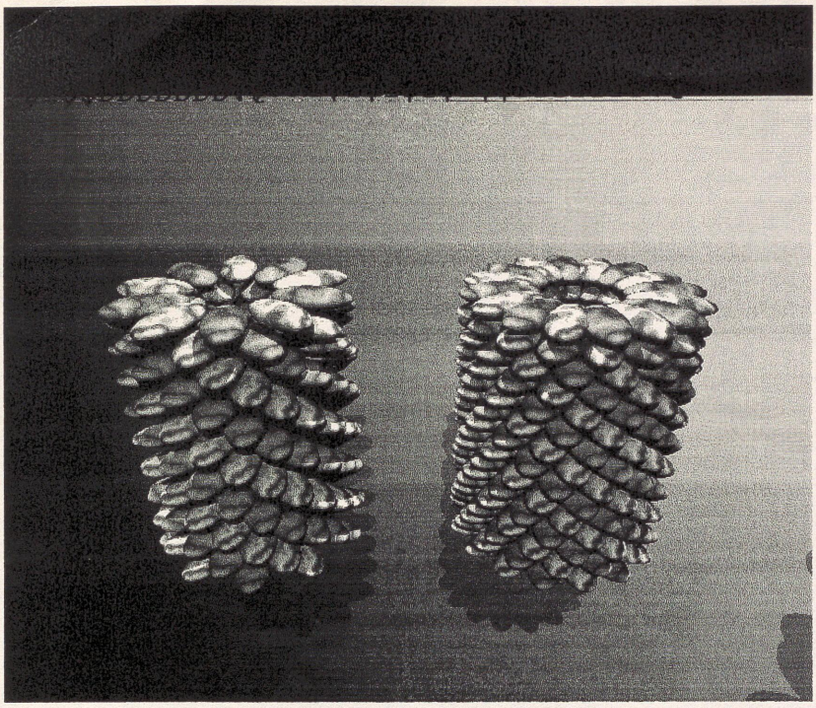
My own crude drawing of a 3d-view!


100 nm
(a)


20 nm




12 nm



A note on comparison of two groups.

We have N stations, divided into ~~two~~ groups in two different ways.

I : k groups with sizes $n_1, \ldots n_k$

II : l groups with sizes $m_1, \ldots, m_l$.

Corresponding to every way of grouping we associate a number denoting mismatches for that way, $m(\text{I})$ and $m(\text{II})$.

We usually expect to have the following:

~~(i)~~ smaller the number of groups, larger will be the mismatches.

i.e. $k < l$ will lead to $m(\text{I}) > m(\text{II})$.

Now the problem is of comparing these ~~to~~ two ways and decide which one is better.

At an outset we see that we have two ~~conflict~~ conflicting demands : we like to have a small number of groups and also we

like to have a small number of mismatches.

{If we aim at only having a small number of mismatches, disregarding the number of groups we can have N groups, each with size 1. This will have no mismatches}.

There is no way of comparing two different of ways of ~~one~~ grouping unless we introduce costs associated with mismatch and number of groups.

Indeed we have ~~two~~ different kinds of mismatches and different costs may be associated with these different kinds.

Thus the cost associated with way I of grouping is $c(I) + c'(k)$. Here $c(I)$ is the cost because of mismatches and $c'(k)$ is the cost because there are k groups. Similarly we have $c(II) + c'(l)$ as cost associated with the second way of grouping.

One can find the number of mismatches for different groups through simulation. Perhaps we can assume multivariate normality and known mean vectors and variance-covariance matrices, {as the estimates are based on a large data} while doing the simulation.

Study of coherence of a meteorological zone through simulation

1. There are 50 rain gauge stations in karnataka and there is data on rainfall at each station for 80 years.

2. We can assume that mean annual rainfall at a stn is known as also its variance and covariances between rainfall values at any two stations are also known.

3. Thus we have a vector ($\underline{X}_{50\times 1}$) valued random variable following (we assume) a 50-variate normal distribution with mean vector $\underline{\mu}_{50\times 1}$ and variance-covariance matrix $\underline{\Sigma}_{50\times 50}$ ~~by~~ both $\underline{\mu}$ & $\underline{\Sigma}$ known.

4. Given a realisation of this vector for any one year we have two methods of dividing this vector into subvectors (or zones). M method divides $\underline{X}$ into 3 zones. SG method divides $\underline{X}$ into 10 zones.

5. The zoning is fixed ie every coordinate is placed into a certain zone regardless of the actual value of the rainfall at that station.

6. It is desirable to have as large a zone as possible and also as coherent a zone as possible. These two are conflicting demands. In general smaller zones are more coherent. We wish to maximise a suitable function of coherence and zone size. This cannot be chosen easily.

7. It may nevertheless be of interest to see how the M & SG methods behave in a probabilistic sense. In particular how often do we get a case in which a zone average is quite high (considering its distribution) while some individual values are quite low (w.r.t. their distribution).

8. We may call a zoning method (or the resulting zone) more coherent if such undesirable events have low probabilities.

9. To get these probabilities may involve rather complex numerical integration. Instead a simulation approach could be simpler.

10. Perhaps each of the previous 80 years could be taken up one at a time and occurrences of single/multiple mismatches could be counted.

11. Alternatively we could generate realisations of the random variable in para 3. This can be done in 2 steps.

Step 1 - generate 50 independent standard normal (Normal with mean zero, variance 1) variates. Call their vector Y. We now transform it.

Step 2 - Consider $X = AY + \mu$

$E(X) = \underline{\mu}$ covariance matrix of X is AIA' I being cov. matrix of Y.

Thus A must be such that $AA' = \Sigma$ the known matrix.

12. Every $\underline{X}$ vector generated is in a sense a new year's rainfall story. Now we can count mismatches in the M & SG schemes.