

The genetic code

Jessica Marshall

Science Writer

After the double-helical structure of DNA was revealed in 1953, researchers had before them the chance to unlock an utterly compelling puzzle, one for the ages and fundamental to the processes of life: What code linked DNA sequences to each of the 20 amino acids that comprised proteins? Every biology textbook now lists which three RNA bases correspond to which amino acid, but in 1961, that was still a mystery.

This page of Marshall Nirenberg's laboratory notebook from 1965 shows the answers. Nirenberg reported the first decoded amino acid, based on the finding that poly-uracil RNA incorporated only phenylalanine, in the classic 1961 PNAS paper featured in this issue. After that, the race was on to identify them all, a task Nirenberg finished in 1966 when this chart was finally complete.

Nirenberg took a novel yet simple approach to cracking the problem. Published papers had predicted it would take 20–30 years to crack

the code using genetic methods. Nirenberg's insight was to use biochemistry instead.

"It was incredibly simple," says C. Thomas Caskey, now at Baylor College of Medicine in Houston, Texas, who began working in Nirenberg's laboratory at the National Institutes of Health not long after the code was broken. "It gave you a direct readout and it could be reproduced very easily." Such simplicity was a hallmark of Nirenberg's scientific approach, he adds. "He always wanted an assay where you could reach into the freezer, pull out the components, and start blasting away on the reaction."

The method used a cellular extract containing intact biochemical machinery for protein synthesis. Nirenberg with Heinrich Matthaei prepared 20 different vials of this extract, each with a different radiolabeled amino acid added. To each was added the RNA sequence to be tested—in the first case it was chains of uracil—and let the protein synthesis reaction go. At the end

of the experiment, Nirenberg poured each mixture through its own nitrocellulose filter, which bound the peptide chains. The filters that came back provided the amino acid smoking guns.

To determine the precise order of the bases and prove that the code was based on triplets, the team required synthesized RNA triplets. Other National Institutes of Health researchers chipped in to help with this tricky synthesis, to beat the competition and be the first to fully break the code. "He made no incorrect assignments," Caskey adds. "He was a very meticulous record keeper and believed in doing all assays in triplicate and carrying them out multiple times."

The notebook also shows that three triads failed to link to any amino acids in this assay; these proved to be signals for protein termination. Uncovering their role was one of Caskey's contributions once he joined the group, as well as showing that the code was universal across species. "It's sort of fun to think back over it. Most young people don't even know how it was done and yet they use the code every day. It was absolutely fun to be racing through these discoveries at the time."

	ALA	ARG	ASP	ASN	CYS	GLU	GLN	GLY	HIS	ILEU	LEU	LYS	MET	PHE	PRO	SER	THR	TRP	TYR	VAL
UUU	0.11	0.30	0.03	0.05	-	0.03	0.07	0.03	0.04	0.03	0.24	0.11	0.10	0.03	0.09	0.26	0.13	0.05	0.05	0.01
UUC	0.01	0.29	0.03	0.10	-	0.01	0.13	0.25	0.04	0.06	0.32	0.08	0.01	0.08	0.08	0.09	0.09	0.04	0.05	0.02
UUA	0.01	0.10	0.02	0.06	0.16	0.02	0.03	-	0.05	0.02	0.00	0.08	0.05	0.05	0.04	0.08	0.03	0.05	0.04	0.00
UUG	0.22	0.02	0.07	0.05	-	0.02	0.08	0.04	0.05	0.02	0.39	0.06	0.17	0.02	0.07	0.21	0.10	0.05	0.01	0.01
UCU	0.01	0.29	0.05	0.05	-	0.06	0.19	0.18	0.00	0.05	0.11	0.09	0.17	0.02	0.07	1.27	0.07	-	0.02	0.04
UCC	0.01	0.33	0.01	0.02	-	0.03	0.16	0.21	0.04	0.00	0.06	0.09	0.02	0.02	0.09	0.54	0.05	0.03	0.04	0.02
UCA	0.01	0.08	0.03	0.05	-	-	0.20	0.17	-	0.01	-	0.00	-	0.02	0.05	0.12	-	0.03	0.01	0.06
UCG	0.20	0.00	0.00	0.07	-	0.01	0.23	0.44	0.01	0.02	0.00	0.04	0.11	0.03	0.06	1.09	0.09	0.02	0.05	0.03
UAU	0.02	0.41	0.04	0.03	-	0.00	0.11	0.23	0.04	0.01	0.24	0.09	0.03	0.27	0.00	0.02	0.05	0.01	0.81	0.03
UAC	0.01	0.37	0.02	0.02	-	0.03	0.16	0.18	0.04	0.04	0.21	0.06	0.04	0.23	0.02	0.00	0.11	0.02	0.56	0.00
UAA	0.02	0.41	0.01	0.07	-	0.02	0.20	0.21	0.08	0.00	0.05	0.10	0.09	0.38	0.02	0.02	0.03	0.00	0.01	0.03
UAG	0.07	0.06	0.01	0.12	-	0.04	0.00	0.01	0.02	0.02	0.09	0.05	0.02	0.17	0.00	0.00	0.05	0.03	0.03	0.00

An excerpt from the pages of Marshall Nirenberg's laboratory notebook from 1965. See Fig. S1 for complete image. Image courtesy of the National Library of Medicine.