INDEX

Note: Page numbers in *italics* denote figures, tables, or photographs on the corresponding page.

A
Abel, Frederick, 52
access
federally funded research in U.S., 116
green open, 115–117
online publication, 110, 113,
115–120, 122–123
open access, 113, 115–120, 122–123
to paper in review process, 120
paywalls, 112
ACLU (American Civil Liberties Union), 50
acquired characteristics, 245
adenovirus, 201
Adler, Reid, 46
adrenaline, 40, 53–54
adult T-cell leukemia virus (ATLV), 190
A for Andromeda, 266–267
agouti gene, 254, 254–255
AIDS, 156–158, 190, 255
Albert Einstein College, 250
α-helix, 172, 273
AlphaFold, 19–22, 20
Altman, Sidney, 198
Alzheimer’s, 197
American Civil Liberties Union (ACLU), 50
American Society for Microbiology, 99
Amgen, 81
Anderson, Chris, 17
Anderson, Philip, 269
Angell, Narcia, 103
annotation, 228
antibiotic resistance gene, 41, 42, 151
APC (Article Publishing Charge), 117–118
Applied Biosystems, 47, 227
applied research/science, 24–25, 53, 54,
59–60, 62, 158
Arabian Journal of Geosciences, 100
Arabidopsis thaliana, 246, 246–247,
251, 273
Arber, Werner, 212–213
Arnold, Matthew, 6
Article Publishing Charge (APC), 117–118
artificial intelligence (AI), 10
big data, 18–19
examination of submitted papers, 145
modeling biological phenomena, 22
protein structure, 19–21, 20
use in scientific publishing, 100
arXiv system, 110–111
Asilomar meetings, 150–151, 153, 155, 156
assumptions, danger of, 73–74
Astbury, William, 174–175
ATLV (adult T-cell leukemia virus), 190
attitude, scientific, 7–8, 9, 23, 29, 78
auditing, 137–139
authorship
asterisks and footnotes, 75
CRediT, 27
honorary, 75
Jacob on process of, 66
number of authors, 34, 35, 75, 78, 87–88,
88, 139, 140, 180, 209
on patents, 56
paying to publish by authors, 116–118
responsibility in, 26–27
senior author, 75, 178–179, 187, 227
autosomes, 210–211, 217, 273
Avery, Oswald, 167, 178
B
Babbage, Charles, 136–137
Bacon, Francis, 22, 24

© 2023 by Benjamin Lewin. Published by Cold Spring Harbor Laboratory Press. All rights reserved.
INDEX

bacteriophage, 167–169, 169, 204, 212, 237, 273
Baltimore, David, 27–28, 186–187
Banbury Center of Cold Spring Harbor, 61
banding, chromosome, 215, 216–217
Banfield, Jillian, 238
Bang, Oluf, 189
Barrangou, Rodolphe, 239
base pairs, 176, 177, 268, 273
basic research, 25, 42, 53–54, 59–60, 66, 134, 158–159
Bayer, 81
Bayh–Dole Act (1980), 43
Beighton, Elwyn, 174
beliefs, triumph of data over mistaken, 11
Bell, Florence, 174
Berg, Paul, 46, 150–152, 156–157
Berget, Susan, 201
Bermuda Principles, 33, 84, 85
Big Bang Principles, 270
big data, 15
artificial intelligence (AI), 18–19
assessment of, 17
correlations, 32, 36–37
Google Flu Trends, 32
human genome sequence, 37
omics, 36
physics, 18
Big Data (Mayer-Schönberger and Cukier), 17–18
Big Pharma
corruption of medicine, 34
opinion of academics on, 53
patent strategy, 60
secrecy, reliance on, 58, 60
big science, 9, 15
authorship, 27
correlations, 17–18, 28–29
gap between original data and interpretation, 231
genome sequencing, 215–216
retractions, frequency of, 139–140
specialization, 26
Bik, Elisabeth, 145
Biochimica et Biophysica Acta (BBA), 90
Biogen, 55, 57, 60, 201, 216
biology
citation frequency of papers, 92–93
discoveries in, 265–266, 266
disdain for commercial applications, 53
intrinsic variability of living organisms, 8
number of authors on scientific papers, 34, 35
physics compared to, 8–9, 265, 266, 267–271
quality control in, 267
U.S. government funding for research, 128
BioMed Central, 113
bioRxiv server, 111
biotechnology, 53–62
academia’s relationship with, 53–62
patent and publish model, 58, 60
publication by press conference, 104
Biozentrum, 212
Bird, Adrian, 249, 252
Bishop, Mike, 189–190
blending, 166, 169, 193
Bohr, Niels, 206, 243
Botstein, David, 36, 77, 212–213, 216–217
Boyer, Herb, 41–43, 53, 55–56
Boyle, Robert, 89
Boyle’s Law, 89
Bragg, Lawrence, 172, 178
branches of science, 5
BRCA1 and BRCA2, 49–50
Brenner, Sydney, 72
Broad Institute, 156
Broker, Tom, 201
Brown, Pat, 113
Brownian motion, 18
bureaucracy, university, 131
Bush, Vannevar, 158–159
Butterworth-Springer, 90

C
Caenorhabditis elegans, 255, 273
Cairns, John, 189
Caltech, 172
cancer, virus-associated, 185, 188–191
Capecchi, Mario, 130
Carnegie Institute, 169, 253
cos genes, 235–237, 273
Caskey, Tom, 219
CASP (Critical Assessment of Structure Prediction), 19
INDEX 313

Cas1 protein, 239
Cas9 protein, 237–238, 240–241
catalysis, RNA, 198, 198–200
Cavendish Laboratory, 171–172, 173
CCR5 gene, 156
cDNA, 50, 51, 57
Cech, Tom, 28, 198–200
Celera, 47, 84–85, 227–228, 231
Cell, 3, 25, 36
associated journals, 98
as brand, 99
changes after acceptance, 71
changing conclusions of paper, 107–108
competing papers, 72
cost to start, 118
CRISPR paper rejection, 237
editorial decision process, 96
Elsevier purchase of, 98
external review of papers, 71–72
first issue, 118
fraud detection, 141–143, 145–146
frequency of fraud, 141
in hierarchy of scientific journals, 91, 94–96, 97
impact factor, 94, 114
infectious protein papers, 194, 196–197
influence of, 114
Ingelfinger rule, 102
judging papers on data, 68, 74
Lander article, 156
online publication, 110
order of authors, 75
page charge, 117
paywalls, 112
press reporting of papers in, 153
rejection of papers, 92, 96
Reproducibility Project, 80
retractions, 84
review of “Atlas of Science,” 104
signed reviews, 73
speed of publication, 88–89, 103, 120, 240
splicing papers, 202, 203
test for editor applicants, 33
typeface and section placement, 91
unique materials, availability of, 83
cell, creation from DNA sequence, 267, 270
cell lines, 61, 83, 260
Cell Press, 98, 100, 114, 120–121
censorship, 77
Central Dogma, 183–186, 184, 185, 191, 197, 244
Cetus, 54–55
Chakrabarty, Ananda, 40–41, 45
chaperones, 196
Chargaff, Erwin, 175
Charpentier, Emmanuelle, 156–157, 237–239
Chase, Martha, 169
chimpanzees, 229–230
Chow, Louise, 201
chromosomes
autosomes, 210–211, 217, 273
banding, 215, 216–217
definitions, 273
genes residing upon, 164
linkage map, 211, 211–213, 215
number in human genome, 210–211
sex, 210–211, 277
Church, George, 239–240
Churchill, Winston, 72–73
citation analysis, 27, 93, 94, 97, 101, 105, 108, 114
citations
in Atlas of Science, 104
frequency of papers in biology, 92–93
hierarchy of journals by number of, 96, 97, 105
of highly cited papers, 97
impact factor compared to number of, 94, 96
importance measured by, 93–94
of method papers, 94
in online journals, 96
of preprint, 111
of retracted papers, 139
uncited papers, 100, 122
of work funded by NIH, 93
City of Hope, 55
clinical trials, failure of, 81
cloning
animals, 154–155, 259
DNA, 40–42, 41, 72, 79, 83, 149–150, 157
human, 154, 259
reproductive, 259
shotgun, 226–227, 277
therapeutic, 260
cloning vector, 40
INDEX

clustered regularly interspaced short palindromic repeats. See CRISPR
Cohen, Stan, 41–43, 55, 150
cold fusion, 82, 103, 105
Cold Spring Harbor Laboratory, 111, 150, 188, 188–189, 201, 203, 222–223
Cold Spring Harbor Symposium, 102, 186, 188, 216, 222
Coleridge, Samuel Taylor, 22
Columbia University, 175
commercial applications, 53, 58–59, 62
common cold, 153
communality, 4
competition
CRISPR, 238–241
dNA structure discovery, 171–181
eyearly biotech companies, 60–61
gene editing, 238–241
as intrinsic part of science, 179–180
Nobel Prize, 106
priority disputes, 52
rapid publication, 186–187
review of scientific paper, 70–72
scientific journals, 95
selection of data, 136
computers, increase in power of, 9–10
certainty level, 82–83
competition of interest, 4, 77
consciousness, 205, 271–272
constructivism, 24
containment facilities, 151–152
contracts, 26
controlled experiment, 8
controls
data validity, 15
in question-driven science, 17
cooking, 137
Correlations
among genome sequence patterns, 37
big data/big science, 17–18, 28–29, 32, 36–37
cause and effect, 249
suspicion of, 17
Council of Advisors on Science and Technology, 156
Council of the American Society for Biochemistry and Molecular Biology, 132
COVID, 37, 77, 103, 152–154, 225, 255
CpG islands, 230, 252, 273
CRediT (Contributor Roles Taxonomy), 27
Crick, Francis
birth anniversary, 3
Central Dogma, 183, 186
consciousness, 271
criticism of amateurs, 73–74
on discovery of splicing, 204
DNA structure, 171, 173–181, 175, 209
on dogma, 183
Nobel Prize, 106, 178
panspermia, directed, 266
on physics and chemistry, 265
What Mad Pursuit, 173–174
Crick Institute, 181
CRISPR, 10, 233–242, 273
artificial, 237
base-pairing, 176
competition, 238–241
discovery of, 101, 234
ethical issues, 155–156
Nobel Prize, 156
papers published, 236–241, 240
papers rejected, 235, 237–238
patents, 47, 233, 236, 241
power of technique, 155
specificity of the system, 236–237
criticism in science, 70, 72–74, 76, 119
crystallography, 19–21, 171–172, 174, 176–177
culture, scientific, 3, 5–7, 23
Current Contents, 93, 122

D
Danisco, 236–237
Darwin, Charles, 270
data
duplication of, 137–139
fabrication, 137–138
fudging, 137, 165–166
ownership of, 173
selection, 135–137, 140, 143, 166
theory trumped by, 28
triumph of, 5, 11, 52, 256, 265
viewed with skepticism, 145
database
data mining, 33, 36
Lander’s defense of, 241–242
MEDLINE, 93
PDB, 86, 276
as primary source, 33, 84–86
PubMed, 87, 92, 111, 115, 138, 276
Retraction Watch, 138–139
data-driven science, 17–18, 29, 31, 36, 93, 177
data mining, 17, 31–37
Davis, Ron, 149–150
Dean, Caroline, 248, 256
deduction, 15, 22, 185
DeepMind, 19–20
Delbrück, Max, 135, 144, 204
demethylases, 135, 144, 204
Denisovans, 231
Department of Energy (DOE), 218
Dewar, James, 52
difference, 25, 36
disinterestedness, 4
DNA
A-form and B-form, 173
Avery experiments, 167, 178
bacteriophage, 167–169, 169
base-pairing, 176, 177, 268
Central Dogma, 183–186, 184, 185, 191
cloning, 40–42, 41
definition, 273
double helix, 173, 175–177, 176, 178–181, 210, 251, 266
as genetic material, 167–170, 194, 210–211
Hershey–Chase experiments, 169
information in, 210
Meselson–Stahl experiment, 179
model, 175, 176, 179
Neanderthal, 37, 230, 321
plasmids, 42, 150–151, 276
replication, 268
structure, 171–181, 209, 268
synthetic, 57
thread of life, 10
transcription into RNA, 184–187
unifying force in modern biology, 11
DNA fingerprinting, 213–215, 214, 221, 274
DNA polymerase, 223
DNA sequence
cell creation from, 267, 270
patents, 40, 45–51, 48, 51
release to public databases, 33, 48, 84–85, 85
DNA sequencers, 224, 227
DNA sequencing
automated center, 226
cost and speed, 224–225, 225
defined, 274
early history, 223–224, 224
human genome sequencing, 215–219, 221–231
large-scale, 209–210
next-generation, 224, 276
Nobel Prize, 216
read-length, 224–225
shotgun, 226–227
spin-offs, unanticipated, 225
doctrines, 193–206
DOE (Department of Energy), 218
dogma, 183–192
Central Dogma, 183–186, 184, 185, 191, 197, 244
dogmatic phase of molecular biology, 205
epigenetics contradiction of, 247–249
as impediment to progress, 256
nucleic acid as the genetic material, 193
nucleotides in DNA, 251
prion behavior, 195
reactions catalyzed by enzymes, 199
rejection of papers conflicting with, 91–92
threat when theory becomes, 204
timeline of twentieth century, 196
time required to overthrow, 191
Dolly the sheep, 154–155, 259, 260
Donohue, Jerry, 181
INDEX

Doolittle, Russ, 33
double helix structure, 173, 175–177, 176, 178–181, 210, 251, 266
The Double Helix (Watson), 67, 172–173
Doudna, Jennifer, 156, 237–240
Dreams of a Final Theory (Weinberg), 9
Dutch Famine, 254

E
e-BioMed, 110, 113, 274
economics, reproducibility problem in, 81–82
The Economist, 153
Eddington, Arthur, 200
Edman method, 223
educational system, 6–7
Einstein, Albert, 69–70, 200
Eisen, Michael, 113, 118
electronic publication, 109–124
electrophoresis, 213, 274
eLife, 96, 97, 118–119, 239
Ellermann, Vilhelm, 189
Elsevier, 90, 97–98, 100, 112, 114, 121
EMBO Journal, 117
emergence, 269–271
ENCODE database (Encyclopedia of DNA Elements), 36, 274
enzymes
dogmas of, 196
reaction catalysis, 198
restriction, 212–214, 219, 233, 235–236
ribozymes, 198, 198–200
epidemiology, 32
epigenetic marks, 249, 253, 260
epigenetics, 10, 205, 243–256, 266, 268
definitions of, 248–249, 249, 274
Dutch Famine, 254
FLC (FLOWERING LOCUS C), 246–249, 247, 251
imprinting, 251–254
incomplete knowledge of, 255–256
in maize, 253
methylation, 37, 249, 251–254, 256
in mice, 253–254, 254
number of papers on, 249
social, 244
terminology, 253
transgenerational inheritance, 255
X-inactivation, 250–251, 277

Errington, Tim, 80
ES (embryonic stem) cell, 260–262, 274
Escherichia coli, 42, 56, 150, 234, 237, 274
e-science, 109–124
ethics, 149–159
CRISPR, 10
germline editing, 155–157
human cloning, 259
stem cells, 257, 261
European Molecular Biology Organization, 117
evolution, 229, 247, 270
exon, 51, 201, 274
Experimental Procedures, 68, 91
experimental science, controls as feature of, 8, 17
expressed sequence tags (ESTs), 46

F
fabrication, 137–138
faking data, 137, 140, 142–143
falsification, 137
feline leukemia virus, 189
Feynman, Richard, 23
fingerprinting, DNA, 213–215, 214, 221, 274
Fisher, H.A., 165
FLC (FLOWERING LOCUS C), 246–249, 247, 251
Fleischmann, Martin, 82
folding, protein, 195–196
forgery, 136, 142–143
Franklin, Rosalind, 171, 173, 176–178, 180
fraud, 135–148
auditing, 137–138, 137–139
congressional investigation of, 186
continuum of behavior, 137
dealing with, 145–146
defense to accusation of, 143
detection of, 137–138, 141–147
duplication of data, 137–139
fabrication, 137–138
faking data, 137, 140, 142–143
habitual, 143
intuitive approximation, 22
presumption of innocence, 140
prominent cases, 147
reactions/response to, 144–145, 147–148
retractions, 138–147, 140

© 2023 by Benjamin Lewin. Published by Cold Spring Harbor Laboratory Press. All rights reserved.
scientific paper as, 65
as solitary affair, 139–140
susceptibility of scientific endeavor, 143
tribal nature of science, 24
fudge factors, 136
fudging data, 137, 165–166
funding, 127–134
duration of, 131–132
of fraud perpetrators, 147–148
of human genome sequencing, 218–219
by lottery, 133–134
National Institutes of Health (NIH)
grants, 43, 49, 55, 58, 93, 115, 127, 128,
129–134, 147, 219–220
people-based, 132
publication influence on, 108
by U.S. government for research in life
sciences, 128

G
Galileo, 52
Gallo, Bob, 52, 72, 189–190
Garfield, Eugene, 93–94, 104–105
gel electrophoresis, 213, 223, 274
gene
colinearity with protein product, 200
cost of, 10
imprinted, 251–252
inactivation of, 258
interrupted, 193, 201–203
Mendel and concept of, 164, 166
number of, 221, 222, 228, 228–229
patents for, 47, 48, 49–50
physical basis, 166–170
size, 167
split, 130, 201–202, 277
synthesis, 57, 59
term introduced, 166
gene editing, 10, 47, 155, 176, 233–242. See
ealso CRISPR
gene expression
definition, 274
epigenetic control of, 248
resetting the state of, 259
Genentech, 34, 54–60
management style, 58
South San Francisco campus, 56, 59
gene targeting, 130
The Gene (Mukherjee), 250
gene therapy, 233
genetic code, 176, 184–185, 199–200, 274
genetic engineering. See also recombinant
DNA
costs over, 155
ethical considerations, 150–151
genetics
DNA structure, 171–181
history, 163–170
Mendel, 163–166
Schrödinger, 167
genome, 221–231
amount of DNA in, 222
definition, 274

genome sequences
correlations among, 37
human and chimpanzee compared,
229–230
human and Neanderthal compared, 230
as tool, 32
genome sequencing
cost and speed of, 224–226, 225
human genome sequencing, 215–219,
221–231
genomics, 36
Genomic Sciences, 47
Geological Society of London, 69
germline cells, 274
germline editing, 149, 155–157
Gilbert, Wally, 57, 59, 79, 216–219, 223
Gonda, Matthew, 26
Goodfield, June, 39
Google Flu Trends, 32
Gordon Research Conference on Nucleic
Acids, 150–151
Gosling, Raymond, 173, 174, 176, 180
grant applications
gaming the system, 130
proportion funded, 129, 130, 133
rejection, 133
review of, 93, 129–130, 132–133
writing, 128–130, 133
grants
duration of, 131
funding by lottery, 133–134
National Institutes of Health (NIH), 43,
49, 55, 58, 93, 115, 127, 128, 129–134,
147, 219–220
university’s monetary share of, 131
human cloning, 154
human embryo research, 154
Human Genome Project, 221–223
accuracy, 228
Bermuda Principles, 33, 84, 85
big science approach, 227, 229
Celera competition with, 227–228, 231
Craig Venter, 46, 226–227
data mining, 32–33, 36
Eric Lander, 156, 227
funding, 16
Jim Watson, 209, 222–223
open access, 33, 84
path to, 216–219
peer review, 132
reaction to patenting, 46, 48
reporting on, 153
Human Genome Sciences, 48
human genome sequence, 47, 84–85
human genome sequencing, 215–219,
221–231. See also Human Genome Project
annotation, 228
by Celera, 227–228, 231
cost, 224–226, 225
draft sequences, 228
human heritable germline editing (HHGE),
157
humanities
citation frequency of papers, 92
difference between science and, 6–8, 25
human T-cell leukemia virus (HTLV), 190,
275
Hunkapiller, Michael, 47
hybridomas, 43, 44, 275
hypothesis
falsifying, 16, 23, 67, 183
Karl Popper on, 193
Medawar genesis of, 65
testing, 3, 5, 15–16, 22, 66–67, 179
hypothesis-driven science, 16, 18, 28–29,
32, 36, 77, 93, 177, 241, 258, 265, 271
hypothesis-obsessed science, 84
hypothesizing after results known
(HARKing), 66
I
IEGs (Information Exchange Groups), 111
immune system, bacterial, 235, 238, 242
impact factor
Cell, Nature, and Science journals, 94, 114
correlation to retraction index, 141
defined, 94
of different publication sources, 99–100, 114
eLife, 118
number of citations compared to, 96
use to assess tenure candidates, 108

Imperial College, 131
imprinting, 251–254
incomprehension of science, 6–7
Incyte, 48
independent variables, 243
individual contributions, emphasis on, 16, 28, 34
induced pluripotent stem (iPS) cells, 261–262, 275
induction, 15, 22
industry, academia’s relationship with, 53–62
influenza, 32
Information Exchange Groups (IEGs), 111
Ingelfinger, Franz, 102
Ingelfinger rule, 102
Ingram, Vernon, 211
ingredients, unique, 83
inheritance
theory of, 243, 247
transgenerational, 255
innovation, dependence on small-scale science, 27
Institute for Scientific Information (ISI), 93, 94, 104
Atlas of Science, 104
Institut Pasteur, 190
insulin, 42, 54–55, 57–60, 79, 223
interdisciplinary approach, 181
interferon, 45, 60
International Human Genome Sequencing Consortium, 227
intron, 51, 201, 203, 275
intuition, 18–19
intuitive approximation, 22
iPS (induced pluripotent stem) cells, 261–262, 275
Ishino, Toshimuzi, 234
Itakura, Keiichi, 56–57
IVF (in vitro fertilization), 154, 155, 157, 260

J
Jacob, François, 66, 149
Jansen, Ruud, 234–235
Jarrett, William, 189
Jeffreys, Alec, 214–215
Jinek, Martin, 239
Johanson, Donald, 52
John Innes Centre, 248
Johns Hopkins University, 213
Journal of Molecular Evolution, 235
Journal of Negative Results in Biomedicine, 66
journals. See scientific journals

K
karyotype, 217, 275
Kawakami, Thomas, 189
Kelvin, Lord, 205
Kendrew, John, 106, 172
King, Mary-Claire, 49
King’s College, London, 171–172, 177–178
knockout mice, 130
Köhler, Georges, 43–44
Koprowski, Hilary, 45
Koshland, Dan, 138–139
Kuhn, Thomas, 191
Kurzweil, Ray, 9
Kyoto University, 261

L
laboratory, running of, 129
laboratory notebooks, as property of institution, 136
Lamarckism, 245, 255–256
Lancet, 94, 98, 100
Lander, Eric, 149, 156–157, 227, 241–242
Latour, Bruno, 23
Leakey, Richard, 52
Leeds University, 174
Leibniz, Gottfried Wilhelm, 52
leukemia, 189
Levan, Albert, 210
Levene, Phoebus, 168, 171
Levi-Montalcini, Rita, 136
Lévi-Strauss, Claude, 3
licensing, 39, 43, 53, 55, 58, 85
Lilly, 57
linkage, 212, 275
linkage map, 211, 211–213, 215
The Logic of Scientific Discovery (Popper), 15
luck, 166, 170
Luria, Alexander, 136
Lyon, Mary, 250
Lysenko, Trofim, 243, 245–247, 256
Lysenkoism, 245
human embryo research, 154
human genome sequencing, 218–219, 226
Information Exchange Groups (IEGs), 111
open access, 115
patenting DNA sequences, 46–47, 49
proportion of grant applications funded, 129
recombinant DNA experiments, 151, 157
R01 grant, 132, 218–219
National Research Council, 136
National Research Development Corporation (NRDC), 44–45
Nature
acceptance rate, 118
advertisements in, 88, 89
authors, average number of, 180 as brand, 99
CRISPR paper, rejection of, 235
criticism of, 118
editorial decision process, 96
expedited papers, 172
external peer review, 70
germline editing letter, 156
in hierarchy of scientific journals, 91, 94–96, 97
impact factor, 94, 114
influence of, 114
interferon paper, 45
monoclonal antibody paper, 43
multiple journals, 98
objection to Information Exchange Groups (IEGs), 111
paywalls, 112
purchase by Springer-Verlag, 98
refusal of Montagnier paper, 72
Reproducibility Project, 80
reverse transcription papers, 186
Springer Nature, 77, 100, 120
Watson–Crick paper (1953), 177, 177
ncRNA, 275
Neanderthal DNA, 37, 230, 321
negative results, publication of, 66–67, 83
New England Journal of Medicine, 33, 102–103, 189
in hierarchy of journals, 94
income, 112–113
objection to e-BioMed, 113
Newton, Isaac, 52, 136
Newton effect, 96–97
next-generation sequencing, 224, 276
NHGRI (National Human Genome Research Institute), 36
NIAID (National Institute of Allergy and Infectious Diseases), 157
NIH. See National Institutes of Health (NIH)
NINDS (National Institute of Neurological Diseases and Stroke), 46
Nobel, Alfred, 52
Nobel Prize, 105–108. See also specific recipients
α-helix in protein structure, 172
apportioning credit, 178–179
control of gene expression, 66, 248
CRISPR, 10, 239
DNA sequencing, 216
DNA structure discovery, 178
electron charge, 135
for industrial research, 54
iPS cells, 261
monoclonal antibodies, 44
nuclear transfer, 257
prion discovery, 194, 196–197
protein folding, 196
restriction enzymes, 213
ribozyme discovery, 198, 200
senior authors, 178
sequencing Neanderthal DNA, 230
simultaneous or related discoveries, recognition of, 52
splicing and interrupted genes, 201
Northwestern University, 236
NRDC (National Research Development Corporation), 44–45
nuclear transfer, 155, 157, 257–261
nucleosome, 73, 276
nucleotides, 168, 276
O
objective of science, 29
objectivity, 7
obsession in scientists, 61–62
Office of Research Integrity (ORI), 147–148
Office of Scientific Integrity, 147
oil drop experiment, 135
omics, 36–37, 267, 276
online journals
citations in, 96
speed of publication, 120
supplemental material, links to, 120
online publication, 109–124
peer review, 110–111, 113, 118–120, 124
supplemental information, inclusion of, 120, 123
online publishing
gatekeeper function, 121–122
speed of publication, 103
open access, 113, 115–120, 122–123
open data, 33–34
ORI (Office of Research Integrity), 147–148
Ortega y Gasset, José, 96–97
oversampling, 226–227

P
Pääbo, Svante, 230–231
Painter, Theophilus, 210
palindrome, 106
Pannetta, Leon, 158
panspermia, directed, 266
paper, scientific. See scientific paper
paradigm shift, 191, 193, 197, 202, 219
patents, 39–52
adrenaline, 40, 53–54
authorship, 56
BRCA1 and BRCA2, 49–50
cDNA, 50, 51
challenges to, 49
CRISPR, 47, 233, 236, 241
DNA sequences, 40, 45–51, 48, 51
effect on science, 50–51
exclusion of products of nature, 40
gene synthesis, 59
lawsuits concerning, 60–61
as legitimate objective in academic research, 62
licensing, 39, 43, 55, 58
on living organisms, 40–41, 45
monoclonal antibodies, 43–45
novelty of invention, 39
Pseudomonas putida, 40–41
recombinant DNA, 42–43, 45, 55
Supreme Court rulings, 41
timing of application filing, 42, 44–45, 51
usefulness criterion, 39, 46
Pauli, Wolfgang, 79
Pauling, Linus, 172, 175
paywalls, 112
PCR (polymerase chain reaction), 54
PDB, 86, 276
PDGF, 33
peer review
drug companies, 103
fraud detection, 141
of grants, 93, 132–133
online publication, 110–111, 113, 118–120, 124
of papers, 4, 68–74, 86, 95
PLoS, 113
quality control by, 9
retractions, 139
signed reviews, 113
spectrum of rigor, 119
penicillin, discovery, 45
Pennica, Dianne, 72
Pergamon Press, 90
Perkins, Thomas, 55
personal communication, 111
Perutz, Max, 106, 172
PhD candidate, 16, 61, 127
philosophers of science, 22–23
Philosophical Transactions, 89
photograph, 51, 173–175, 174
Physical Review, 69
physics
arXiv system, 110–111
big data/big science, 9, 18
biology compared, 8–9, 265, 266, 267–271
CERN particle accelerator, 226
discoveries in, 265, 266
disdain for other fields of science, 181
emergence in, 269
experimental science, 8
gene behavior, 167
number of authors on scientific papers, 34, 35
partnerships between academia and private companies, 53
quantum mechanics, 268, 270–271
reductionism, 267–271
P.I. (principal investigator), 128, 131, 276
plagiarism, 137, 145

© 2023 by Benjamin Lewin. Published by Cold Spring Harbor Laboratory Press. All rights reserved.
INDEX

Planck, Max, 127, 271
Plan S, 115, 117
plants, response to temperature by, 245–248, 247
plasmid, 42, 150–151, 276
PLoS (Public Library of Science), 99–100, 113–115, 118, 145
PLoS Biology, 113–114
PLoS Medicine, 113–114
pluripotent stem cells, 257, 260–262, 276
politics, 98, 149–159, 182, 257
Pollack, Bob, 150
Polyani, Michael, 191
polymerase chain reaction (PCR), 54
Pons, Stanley, 82
Popper, Karl, 15, 23, 31, 193, 266
Porton Down facility, 80
postdoctoral fellowship, 16, 58, 61, 127, 276
postmodernism, 24
preclinical studies, reproducibility of, 81
predatory journals, 119–120
predatory pricing
by drug companies, 50
by large commercial publishers, 114
prematurity, 164, 178
preprints, 110–111, 119–122, 276
press
COVID misinformation, 153
inadequacy of reporting on science, 153
releasing information to, 101–104
principal investigator, 128, 131, 276
priority
CRISPR, 233, 241
definition, 276
establishing, 104
online publication, 121
online thread, 123
patents, 39, 49, 51–52
publication order, 187
rush to obtain, 81
scientific papers, 69, 71
Proceedings of the National Academy of Sciences (PNAS), 88, 91, 95, 187, 201, 235, 237, 253
protein
infectious, 193–197, 195, 199–200, 267
sequencing, 223
protein structure, 19–21, 20
alternative structures, 194–195, 195
cell environment, effect of, 195
databases, 85–86
folding, 195–196
proteome, 21
proteomics, 36
PrP, 194–195, 195, 276
Prusiner, Stanley, 193–194
Pseudomonas putida, 40–41
pseudoscience, 5, 244–245, 266
psychology, reproducibility problem in, 82
Ptashne, Mark, 249–250
publication. See also scientific paper
author payments for, 116–118
conflict of interest, 77
double, 102
electronic, 109–124
filters en route to publication and publicity, 71
of negative results, 66–67, 83
number of, 87–88, 88
online, 76, 109–124
order of papers in, 187
by press conference, 104
of retraction, 84
speed of, 88–89, 91, 103, 107, 120
supplemental data, 76
publishing
biotech’s patent and publish model, 58, 60
commercial, 90
groups of journals, 98–99, 99
large commercial publishers, 97–101, 110, 112, 114–115, 121, 123
predatory pricing, 114
publish or perish, 87–108
signal:noise ratio in scientific literature, 91, 94, 101
by societies, 90, 99, 111
timeline of scientific, 112
by university presses, 90, 99
PubMed Central, 111–112
pure research/science, 53, 58, 60

Q
quantum mechanics, 268, 270–271
question-driven science, 16–17, 31, 204

© 2023 by Benjamin Lewin. Published by Cold Spring Harbor Laboratory Press. All rights reserved.
INDEX

R
recombinant DNA. See also biotechnology cloning DNA, 40–42, 41, 149–150
definition, 277
history of, 149–154
insulin produced from, 55
NIH, 157
patents, 42–43, 45, 55
public concern over, 151–154
rules concerning research, 59, 79–80, 151–153
recombination frequency, 212–213
reductionism, 11, 205, 244, 265–272
Reflections on the Decline of Science in England (Babbage), 136
regeneration, 257–259, 259
Reimers, Niels, 42–43
relativism, 92
Rennie, Drummond, 27
reproducibility, 79–86
discrepancies, 144, 147
incentive for, 81
patent condition compared, 40
of preclinical studies, 81
problems with, 81–83
of published work, 4, 9
retractions due to failure, 138
study on, 80–81
unifying concept in science, 79
unique ingredients, use of, 83
reproducibility crisis, 82
Reproducibility Project, 80–81
research article/paper. See scientific paper
research parasites, 34
restriction enzyme, 212–214, 219, 233, 235–236, 277
retraction
as blemish on record, 84
fraud, 138–147, 140
mass, 100
number of authors per paper, decline with, 139, 140
Retraction Watch database, 138–139
retrovirus, 187–191, 190, 277
reverse transcription, 185, 187, 193, 202–203, 277
reward system of science, 52, 67, 81
RFLPs (restriction fragment length polymorphisms), 213, 215
ribozymes, 198, 198–200
Riggs, Arthur, 55, 57
RNA (ribonucleic acid), 169
crRNA, 237
definition, 277
genome sequences encoding, 221–222
heterogeneous RNA (hnRNA), 202–203, 275
messenger RNA (mRNA), 201–203, 275
ncRNA, 275
ribozymes, 198, 198–200
splicing, 201–204
translation into protein, 184
RNA interference, 233, 236, 238
RNA sequencing (RNA-seq), 224–225, 277
RNA virus, 185, 185, 187–188
RNA World Hypothesis, 199
Roberts, Rich, 201–203
Rockefeller University, 188, 239
Rous, Peyton, 188–189
Rous sarcoma virus, 188
Royal Society of London, 69, 89–90
running a laboratory, 129
Rutherford, Ernest, 34, 181

S
Salk, Jonas, 39
Sanger, Fred, 216, 223
Sanger Center, 48
Schekman, Randy, 118
Schrödinger, Erwin, 11, 65, 167
Science
acceptance rate, 118
Celera publication, 85
CRISPR papers, 236, 238–239
criticism of, 118
Dan Koshland, 138
editorial decision process, 96
external peer review, 70
germline editing letter, 156
in hierarchy of scientific journals, 91, 94–96, 97
impact factor, 94, 114
influence of, 114
objection to Information Exchange Groups (IEGs), 111
Reproducibility Project, 80
somatostatin paper, 56
speed of publication, 88
Science, the Endless Frontier (Bush), 159
Science Citation Index (SCI), 93
Scientific attitude, 5, 8–9, 23, 29, 78, 200
scientific journals
  barrier to starting, 118
  as brands, 98–99
  charges for article publication, 116–118
  gatekeeper role of, 121–122
  groups, 98–99, 99
  hierarchy, 91, 94–96, 97, 105
  impact factor, 94, 99–100, 114, 118
  number of, 87, 96, 99–101
  online, 96
  origin and history, 89–91, 90
  predatory, 119–120
  price, 90, 100–101, 116, 119
  rejection of CRISPR papers, 235, 238
  reviews and commentaries, 107
  subscriptions, 95, 110, 113–118
scientific method, 15, 17, 22–23, 177
scientific paper
  accuracy, 138–139
  artificial intelligence (AI) examination of submitted, 145
  assessment of, 17
  authorship, 75–76
  biotech industry, 58–60
  changes after acceptance, 71
citations, 92–94
  competition for "hot" papers, 91
criticism, 70, 72–74, 76
database as primary source, 85–86
  disruptive papers, 27
  expedited, 172
  filters in route to publication and publicity, 71
  format, 68, 69
  history, 89–90, 90
  importance measured by citations, 93–94
  as key unit of science, 4
  methods papers, 93
  myth of, 65–78, 69
  number of authors, 34, 35, 75, 78, 87–88, 88, 139, 140, 180
  preprints, 110–111, 119–122
  priority, 69, 71
  in PubMed database, 87, 92
  rejection of, 75–76, 89, 91–92, 96, 120–121, 141, 143, 146
  reproducibility of experiments reported, 80–81
  responsibility of authors, 26–27
retraction, 83–84, 100, 138–147, 140
review of, 4, 68–74, 86, 95, 120, 122, 141
  supplemental data, 76
  threads as alternative format, 123–124
  transition to submitting data to database, 86
  as unit of assessment, 109
  unpublished data cited in, 76
  written in first person, 68
  written in third person, 67
scientist
career, 127–129
  illiteracy, 6–7
  mindset, 9, 11, 15, 25
  obsession, 52, 61–62
  origin of term, 22–23
  specialization, 5, 65, 78
  training, 128, 132
scrapie, 194, 197
Secher, David, 45
Second Law of Thermodynamics, 6, 92
self-correcting mechanism of science
  big data/big science, 33, 231
  database data validation, 86
  discrepancies, 144, 147
  fraud, response to, 144, 146–147
  negative results, importance of, 67
  number of human genes, 229
  reproducibility and, 4, 9, 80, 82
  self-policing, 144, 147
  triumph of data, 5
  value system of science, 28
self-policing mechanism of science, 144, 147
self-verifying system, science as, 4
Sequence Hypothesis, 183–184
sequencing. See DNA sequencing; human genome sequencing
sex chromosomes, 210–211, 277
Sharp, Phil, 201–202
short regularly spaced repeats (SRSRs), 234
shotgun cloning, 226–227, 277
shotgun sequencing, 226–227
sickle cell anemia, 211
signal:noise ratio in scientific literature, 91, 94, 101, 119, 121, 255
Šikšnys, Virginijus, 237–239
simian sarcoma virus, 33
Singularity, 9, 21
Sinsheimer, Robert, 215
<table>
<thead>
<tr>
<th>Index</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>skepticism</strong></td>
</tr>
<tr>
<td>data viewed with, 145</td>
</tr>
<tr>
<td>dogma and, 192</td>
</tr>
<tr>
<td>organized, 5</td>
</tr>
<tr>
<td>of prions, 197</td>
</tr>
<tr>
<td>protection against falling into dogma, 204</td>
</tr>
<tr>
<td>of reverse transcription, 203</td>
</tr>
<tr>
<td>Skolnick, Mark, 49</td>
</tr>
<tr>
<td>sloppiness, 135, 137, 144, 148</td>
</tr>
<tr>
<td>Smith, Dave, 217–218</td>
</tr>
<tr>
<td>Smith, Ham, 213</td>
</tr>
<tr>
<td>Snow, C.P., 3, 6–7, 25, 257</td>
</tr>
<tr>
<td>social sciences</td>
</tr>
<tr>
<td>citation frequency of papers, 92</td>
</tr>
<tr>
<td>controls, difficulty with, 5, 8</td>
</tr>
<tr>
<td>epigenetics, 244</td>
</tr>
<tr>
<td>reproducibility problem in, 82</td>
</tr>
<tr>
<td>societies, publishing by, 90, 99, 111</td>
</tr>
<tr>
<td>sociology, skepticism about, 23–24</td>
</tr>
<tr>
<td>soft sciences, 8</td>
</tr>
<tr>
<td>somatostatin, 55–57</td>
</tr>
<tr>
<td>Sontheimer, Erik, 236, 239</td>
</tr>
<tr>
<td>Southern, Ed, 93</td>
</tr>
<tr>
<td>specialization, 5, 25–26, 180, 209</td>
</tr>
<tr>
<td>Spiegelman, Sol, 186–187</td>
</tr>
<tr>
<td>splicing, 201–204, 277</td>
</tr>
<tr>
<td>split gene, 130, 201–202, 277</td>
</tr>
<tr>
<td>Springer, 100, 112–113</td>
</tr>
<tr>
<td>Springer Nature, 77, 100, 120</td>
</tr>
<tr>
<td>Springer-Verlag, 98</td>
</tr>
<tr>
<td>SRSRs (short regularly spaced repeats), 234</td>
</tr>
<tr>
<td>Stanford University, 149</td>
</tr>
<tr>
<td>statistical fallacy, 82–83</td>
</tr>
<tr>
<td>stem cells, 257–262, 258</td>
</tr>
<tr>
<td>embryonic, 257, 259–261</td>
</tr>
<tr>
<td>pluripotent, 257, 260–262, 276</td>
</tr>
<tr>
<td>Stent, Gunther, 204–206</td>
</tr>
<tr>
<td><em>The Structure of Scientific Revolution</em> (Kuhn), 191</td>
</tr>
<tr>
<td>subscription, 95, 110, 113–118</td>
</tr>
<tr>
<td>Sulston, John, 31, 47–48, 85, 228–229</td>
</tr>
<tr>
<td>Supreme Court rulings, 41</td>
</tr>
<tr>
<td>Suram, Azim, 251</td>
</tr>
<tr>
<td>SV40, 150, 213, 277</td>
</tr>
<tr>
<td>Swanson, Bob, 55, 57</td>
</tr>
<tr>
<td>symposium proceedings, 102</td>
</tr>
</tbody>
</table>

**T**

Takamine, Jokichi, 53

tandem repeats, 234, 234

techniques, advancement of science, 209

Temin, Howard, 185–187, 202

*Tetrahymena*, 198–199

tetr nucleotide theory, 168, 175

tetranucleotide theory

emergence of new, 192

experiment relationship with, 200

science advancement and, 191

threat of becoming dogma, 204

Thomson, Jamie, 260

threads, 123–124

thyroxine, 54

T4 phage, 169

training of scientists, 127–129, 277

transcription, 184–187, 277

transcription factors, 261

transcRNA (*trans*-encoded small RNA), 237, 239

translation, 184, 277

transparency, 119, 121

tribal, science as, 24

trimming, 136

truth, objective, 8

*The Two Cultures* (Snow), 6, 25

**U**

Umeå, 237, 239

uncertainty, 18, 267, 271

universalism, 4

University of Alicante, 234

University of Amsterdam, 214

University of Berkeley, 237–238

University of California, Los Angeles (UCLA), 155–156

University of California, San Francisco (UCSF), 114, 193

University of California, Santa Cruz (UCSC), 215

University of Colorado, 198

University of Edinburgh, 93, 108

University of Leicester, 214

University of London, 131

University of Oxford, 257

University of Utah, 49, 82, 130

University of Wisconsin, 185, 260

university presses, 90, 99

unpublished data, 76, 111, 123, 140, 173, 178

**V**

value system of science, 23–24

commercial attitudes, 61

features of, 4–5

© 2023 by Benjamin Lewin. Published by Cold Spring Harbor Laboratory Press. All rights reserved.
in grant funding, 134
looking for fraud, 145
patents, 39
paying for information, 47
self-correction, 28
van der Oost, John, 237
Varmus, Harold, 47, 133, 135, 154, 157, 190
Velazquez, Jose Luis Perez, 131
Venter, Craig, 46–47, 219, 226–227
venture capitalists, 55
verification, 21, 28, 32
vernalization, 245–246, 247
Vickers, Tony, 43
Villa-Komaroff, Lydia, 80
Vilnius University, 237–238
virus
  bacteriophage, 167–169, 169, 204, 212, 237
cancer-associated, 185, 188–191
discovery of split genes in, 130
retrovirus, 187–191, 190
RNA, 185, 185, 187–188
vitamin B12, 54
Voltaire, 78

W
Waddington, Conrad, 248, 249
Walter, Peter, 114
Washington University, 48
Watson, Jim
  bacteriophage DNA, 167
  Cold Spring Harbor Laboratory, 189, 201, 222–223
DNA structure, 171–173, 175, 175–178, 180, 209
  The Double Helix, 67, 172–173
genome sequencing, 209, 219, 221
Human Genome Project, 209, 222–223
  on human improvement, 149
  on insider information, 70
  on NIH RO1 grants, 219
  Nobel Prize, 106, 178
  Weinberg, Steve, 9, 268–269
  Wellcome Trust, 115, 118
  Westfall, Richard, 136
  What Is Life? (Schrödinger), 11, 167
  What Mad Pursuit (Crick), 173–174
  Whewell, William, 22
  whistleblower, 147
  Whitehead Institute, 156
  Wiedenheft, Blake, 239
  Wilkins, Maurice, 67, 171–173, 176–178, 180
  Wistar Institute, 45
  Wyngaarden, James, 222

X
X chromosome, 210–211, 250–251, 277
Xenopus laevis, 258, 259
X-inactivation, 250–251, 277
X-ray diffraction, 171, 174, 177, 180

Y
Yale, 198
Yamamoto, Keith, 114
Yamanaka, Shinya, 261
Yeats, William Butler, 134
yogurt, 235–236, 241

Z
Zhang, Feng, 156, 239–240
Zuckerman, Harriet, 105–106