AUTOSOMAL DNA COUSINS? NOT SO FAST!

STEPHEN M. BAIRD, M.D.

April 3, 2017

FAMILY TREE OF ALL LIFE

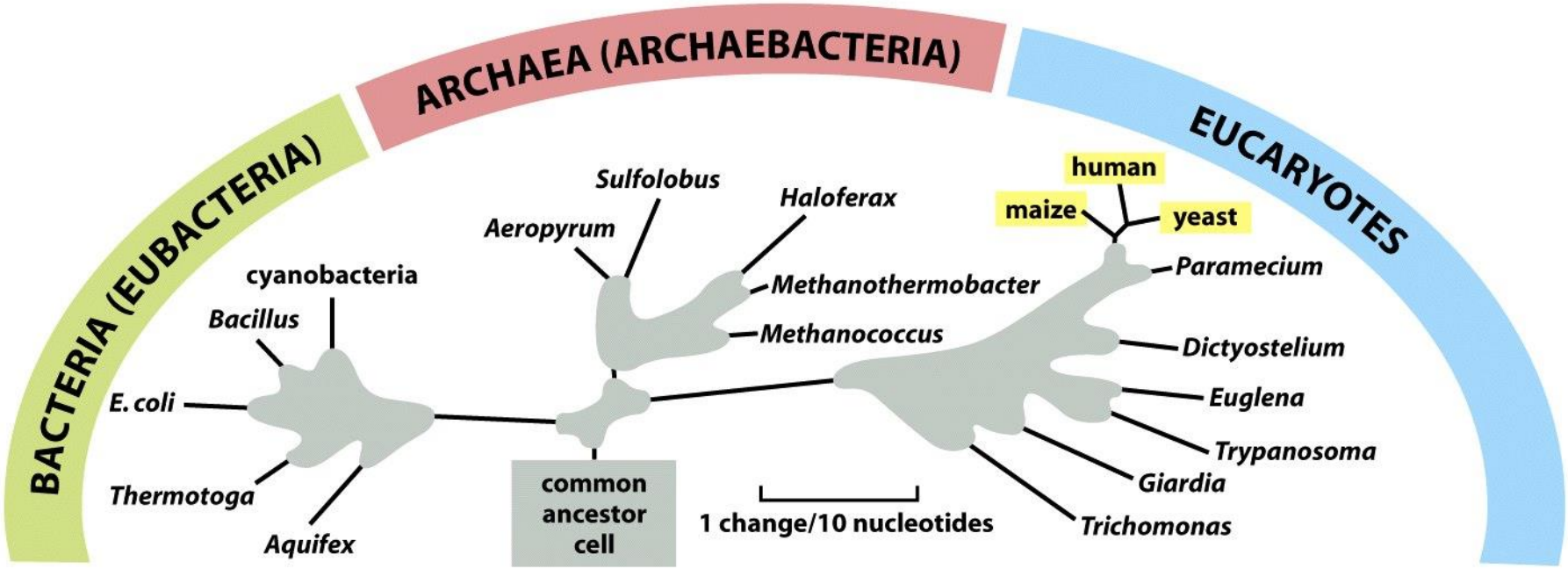


Figure 1-21 Molecular Biology of the Cell 5/e (© Garland Science 2008)

Time span: about 3.5 billion years from the original common ancestor cell to the present.
To some degree, we are cousins with every living thing on Earth today.

These are all of our cousins, so far. 2.3 million named species out of an estimated 8.7 million. 15 thousand are discovered every year. About 5% have had their DNA completely sequenced.

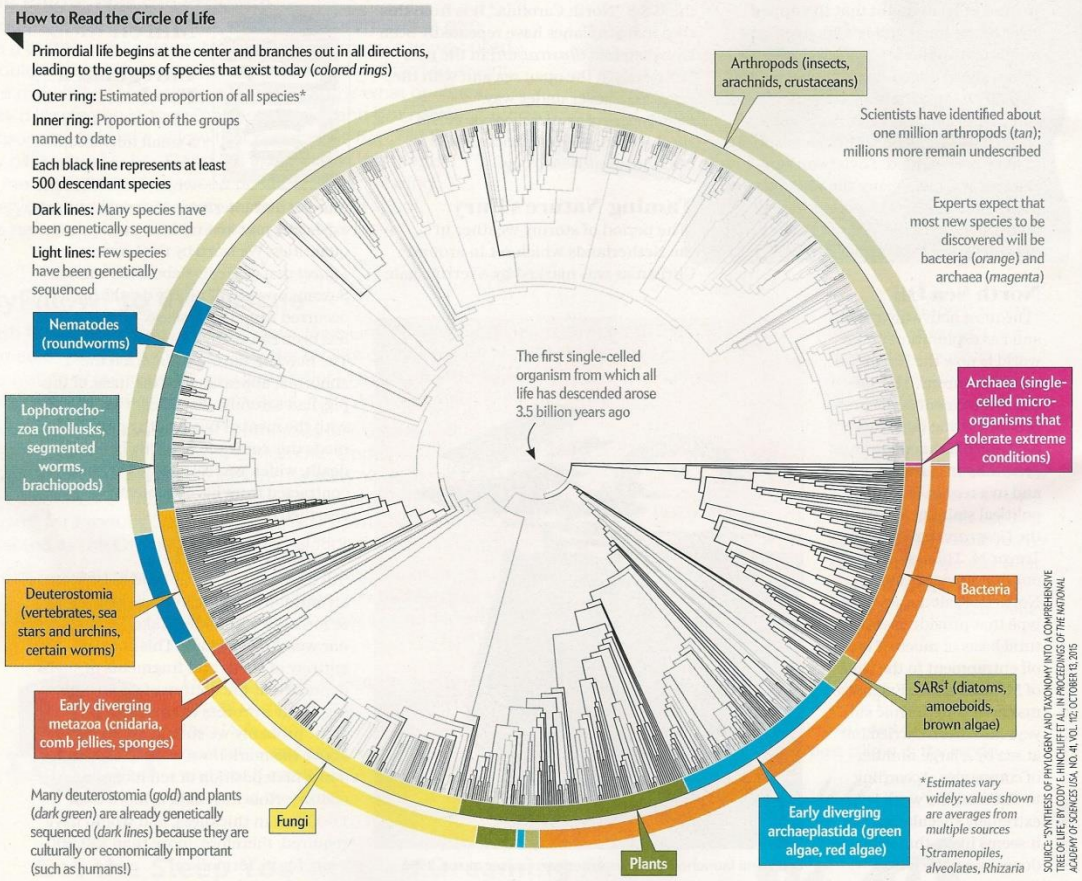
SCIENTIFIC AMERICAN ONLINE
FOR MORE ON RAPIDLY EXPANDING TREES OF BACTERIAL LIFE, GO TO SCIENTIFICAMERICAN.COM/MAR2016/GRAPHIC-SCIENCE

The Circle of Life

Lineages of all known species on earth are finally pieced together

Since Charles Darwin's day, biologists have depicted how new organisms evolve from old ones by adding branches to numerous trees that represent portions of the animal, plant and microbial kingdoms. Researchers from a dozen institutions recently completed a three-year effort to combine tens of thousands of trees into one diagram, most readable as a circle (*below*). The lines inside the circle represent all 2.3 million species that have been named. Biologists have genetic sequences for only about 5 percent of them, however; as more are finished, the relationships within and across groups of species may change. Experts estimate that up to 8.7 million species may inhabit the planet (about 15,000 are discovered every year). "We expect the circle to broaden," says Karen Cranston, a computational evolutionary biologist at Duke University.

Anyone can propose updates to the database (OpenTreeOf Life.org). Greater detail could improve understanding of evolution and help scientists invent drugs, make crops more productive and better control infectious diseases. —*Mark Fischetti*



Graphic by Stephen Smith

TOTAL NUMBER OF GENOMES SEQUENCED, 2014

Organism	Completed sequences
Bacteria, (archaea)	17,420 (362)
Fungi	356
Insects	98
Plants	150
Vertebrates (mammals)	235 (80)
Marine invertebrates	16
Nematodes (worms)	17
Drosophila (fly)	21

SHARED DNA SEQUENCES AMONG DISTANT RELATIVES

Figure 1-22 Molecular Biology of the Cell 5/e (© Garland Science 2008)

"Railroad ties" between sequences indicate identical bases. E. Coli and Methanococcus are eubacteria and archaebacteria, respectively.

MECHANISMS FOR CHANGING GENES

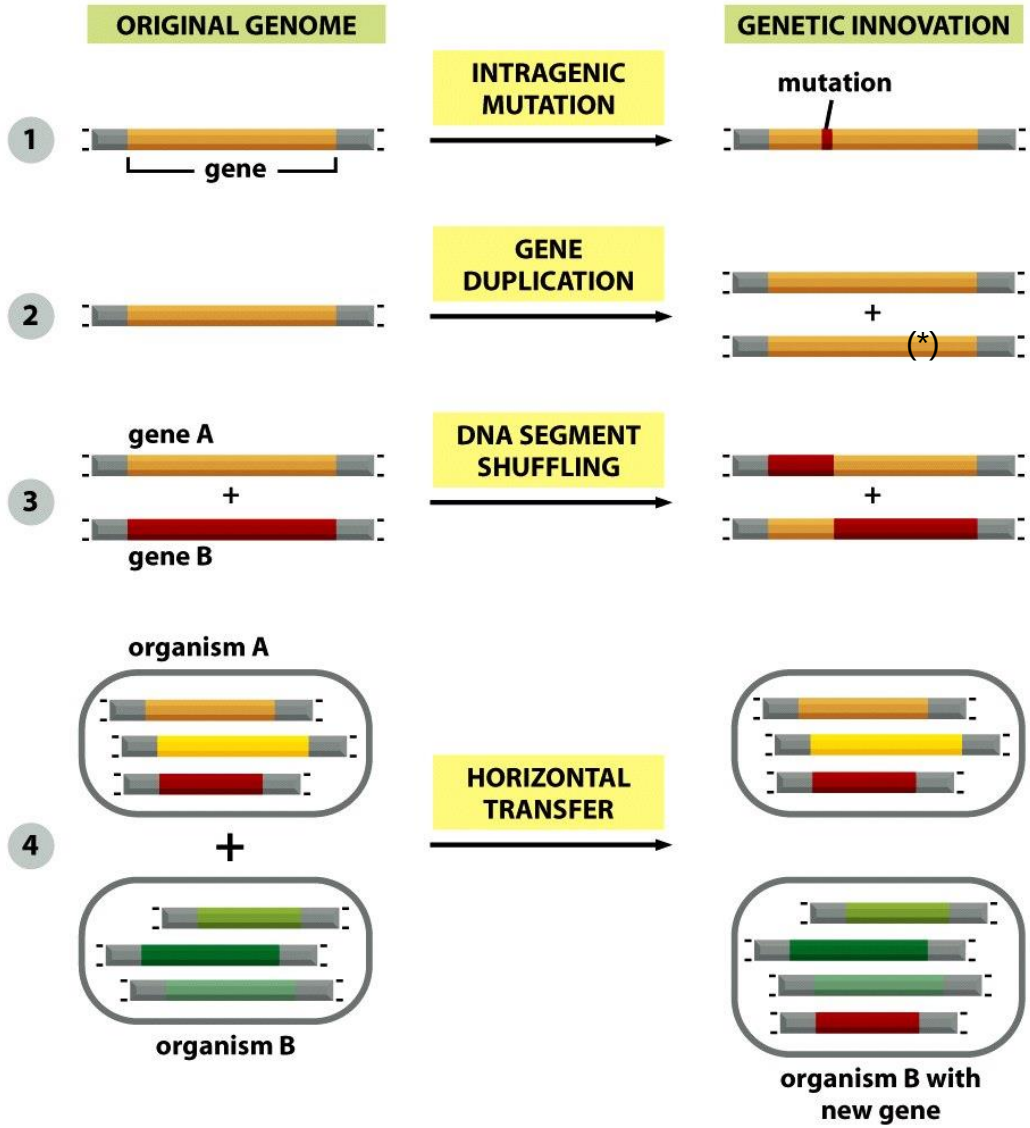


Figure 1-23 Molecular Biology of the Cell 5/e (© Garland Science 2008)

A “snp”

Hemoglobins, clotting factors

Red and green visual pigments, X chromosome

Antibiotic resistance in bacteria

New species?

FAMILY TREE OF HEMOGLOBINS

One chain

Two chains, now with separate mutations

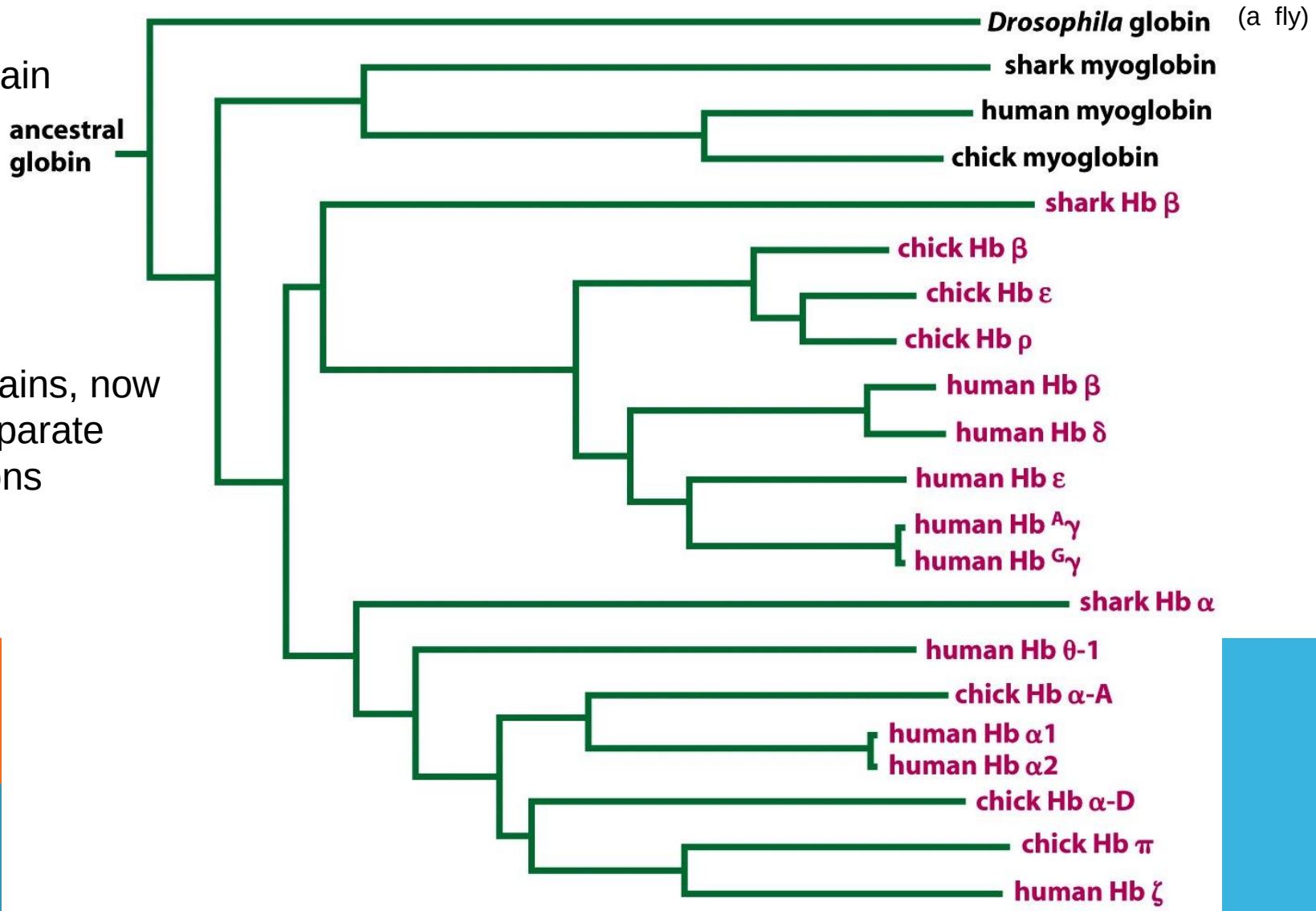


Figure 1-26 Molecular Biology of the Cell 5/e (© Garland Science 2008)

DIAGRAM OF EUKARYOTIC CELL (TRUE NUCLEUS)

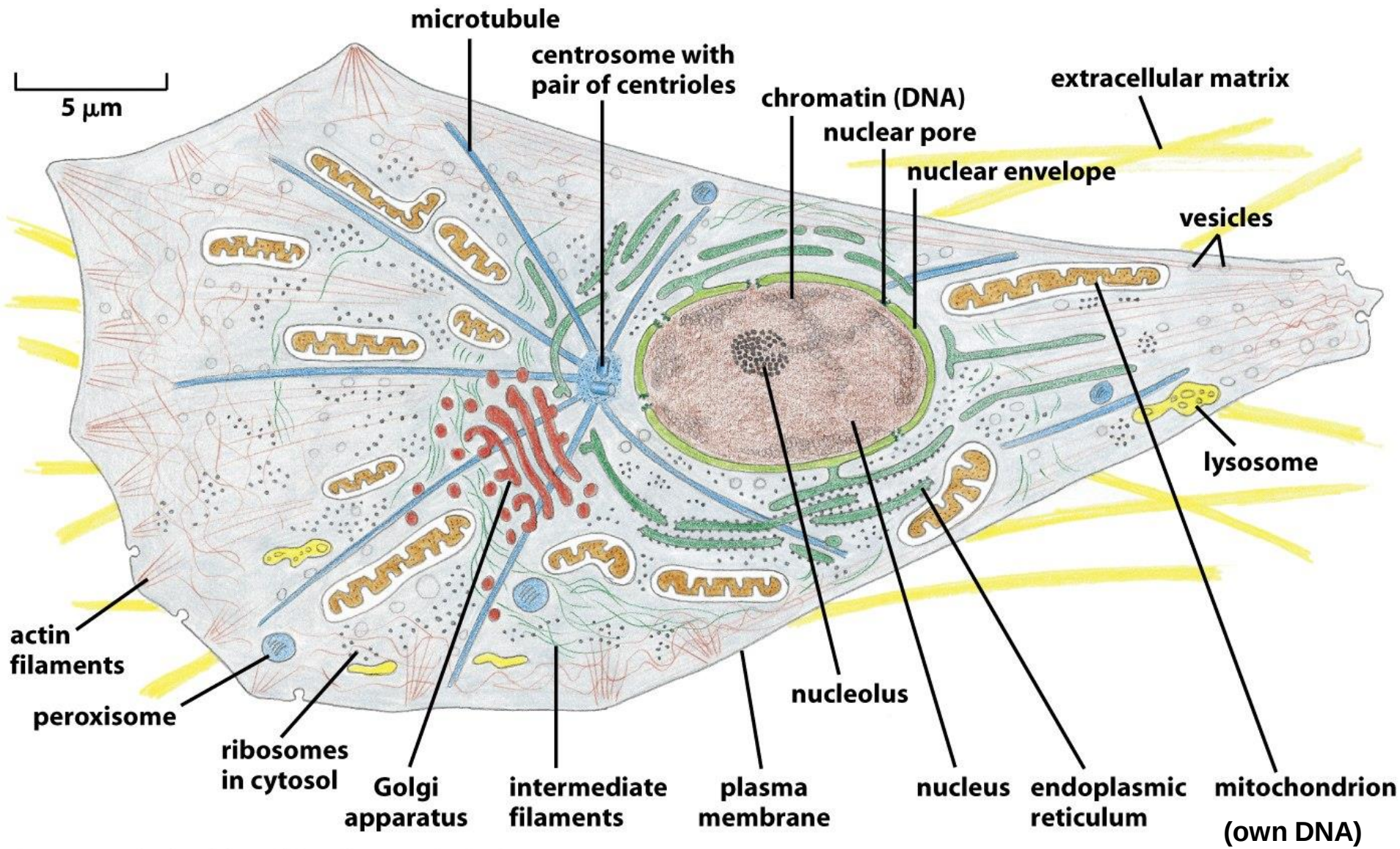


Figure 1-30 Molecular Biology of the Cell 5/e (© Garland Science 2008)

BECOMING A MITOCHONDRION, THE CELL'S ENERGY FACTORY

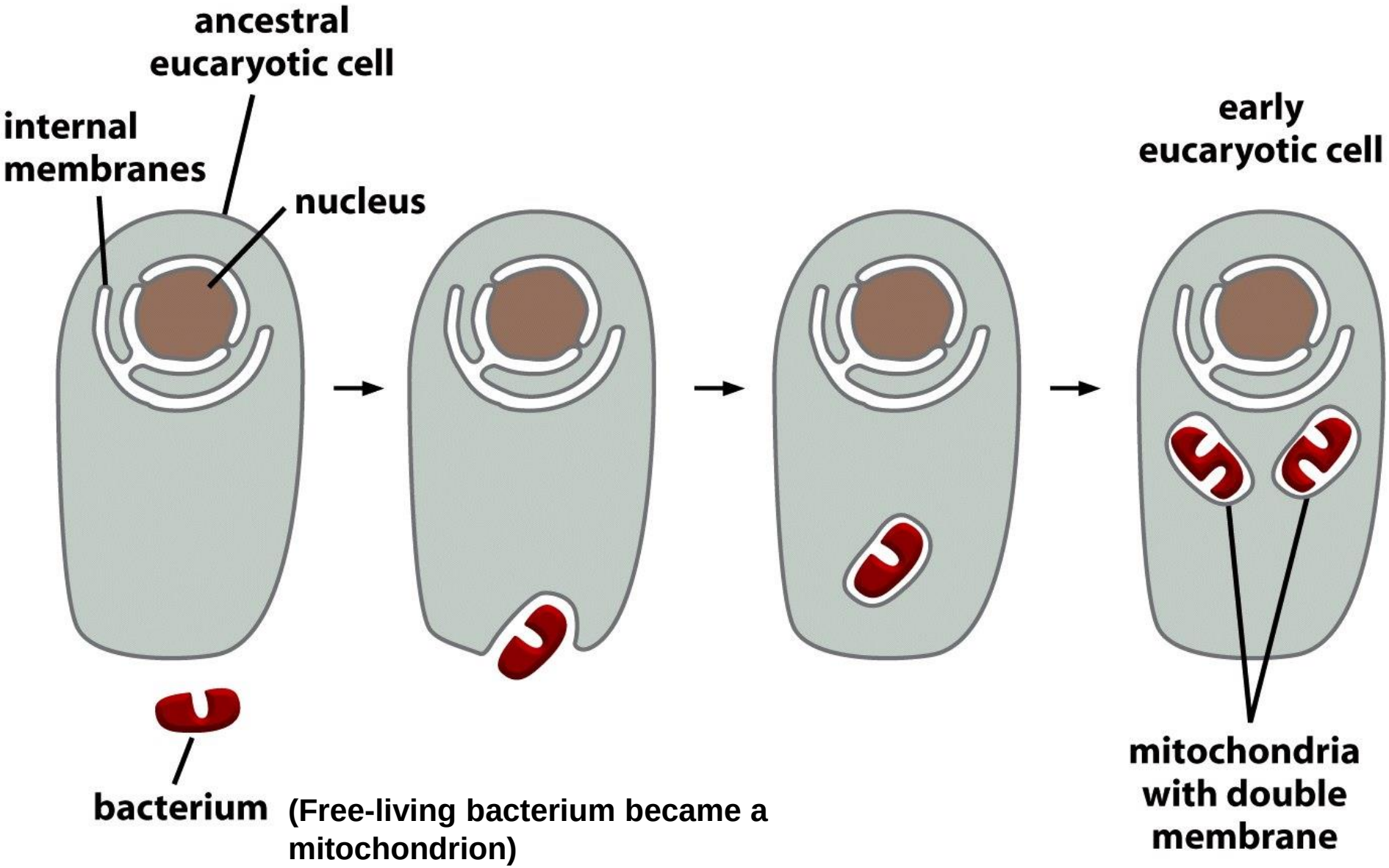
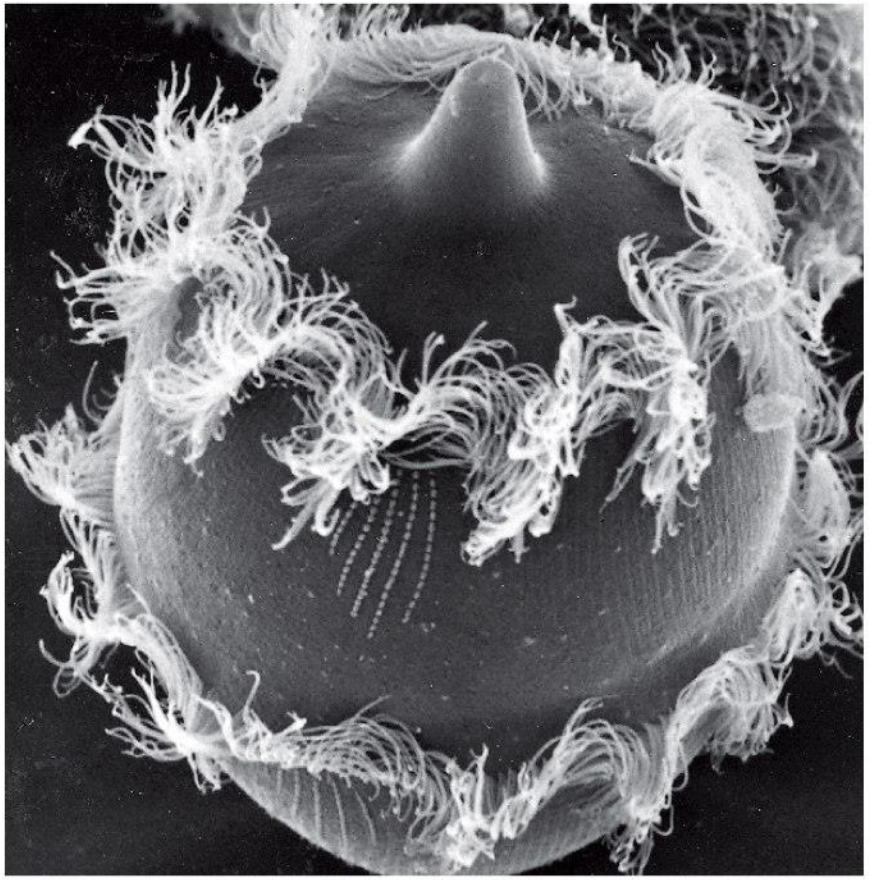


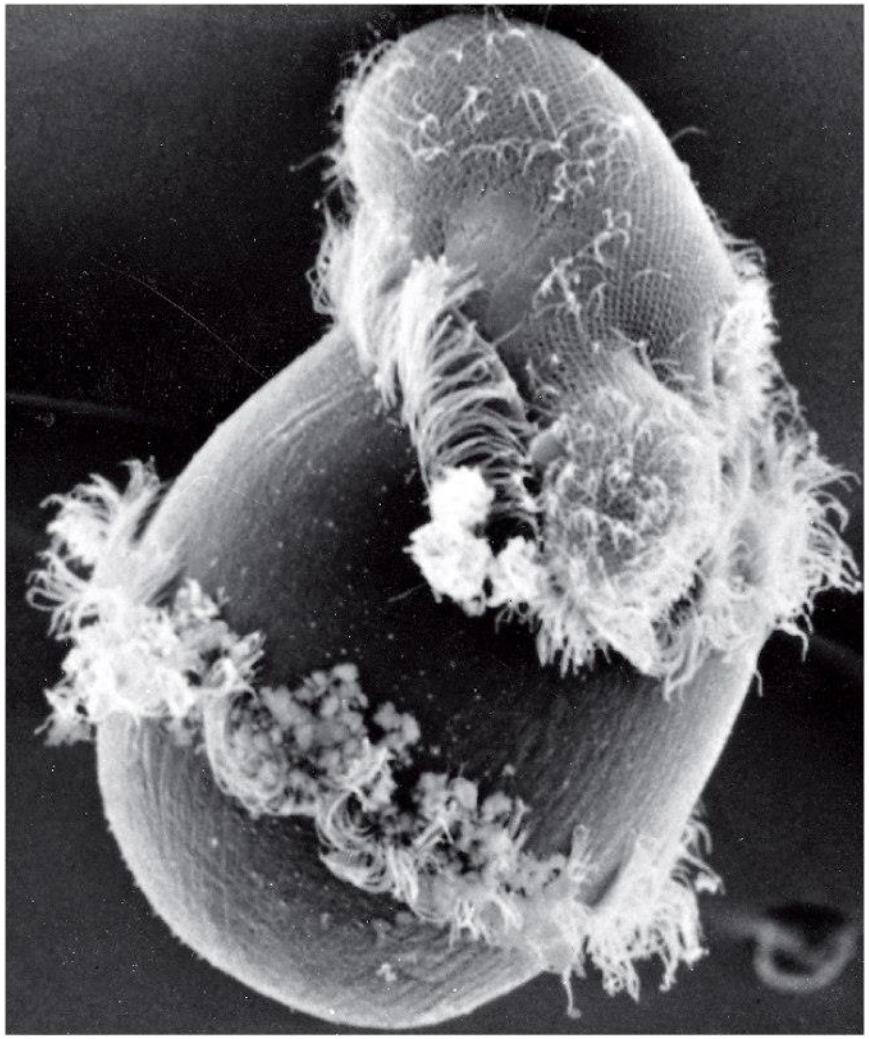
Figure 1-34 Molecular Biology of the Cell 5/e (© Garland Science 2008)

BECOMING A MITOCHONDRION TODAY - THIS PROCESS IS STILL GOING ON



(A)

100 μm

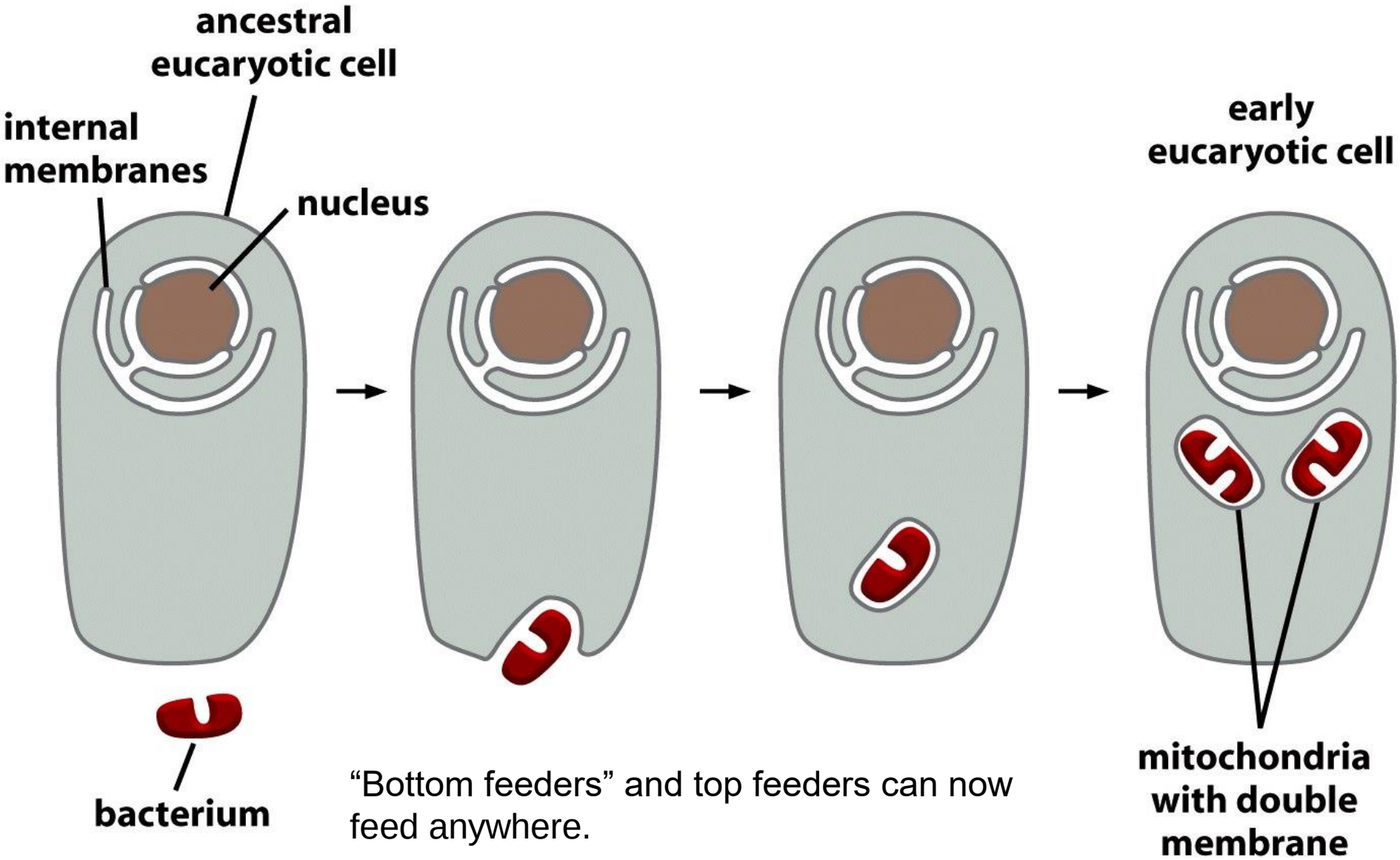


(B)

Cell on bottom eating cell on top

Figure 1-32 Molecular Biology of the Cell 5/e (© Garland Science 2008)

BECOMING A MITOCHONDRION, SYMBIOSIS - ONE ORGANISM ANEROBIC; ONE AEROBIC



"Bottom feeders" and top feeders can now feed anywhere.

Figure 1-34 Molecular Biology of the Cell 5/e (© Garland Science 2008)

MITOCHONDRIAL STRUCTURE - LINES OF MEMBRANES FOR ELECTRON AND ENERGY TRANSPORT. DNA NOT SHOWN

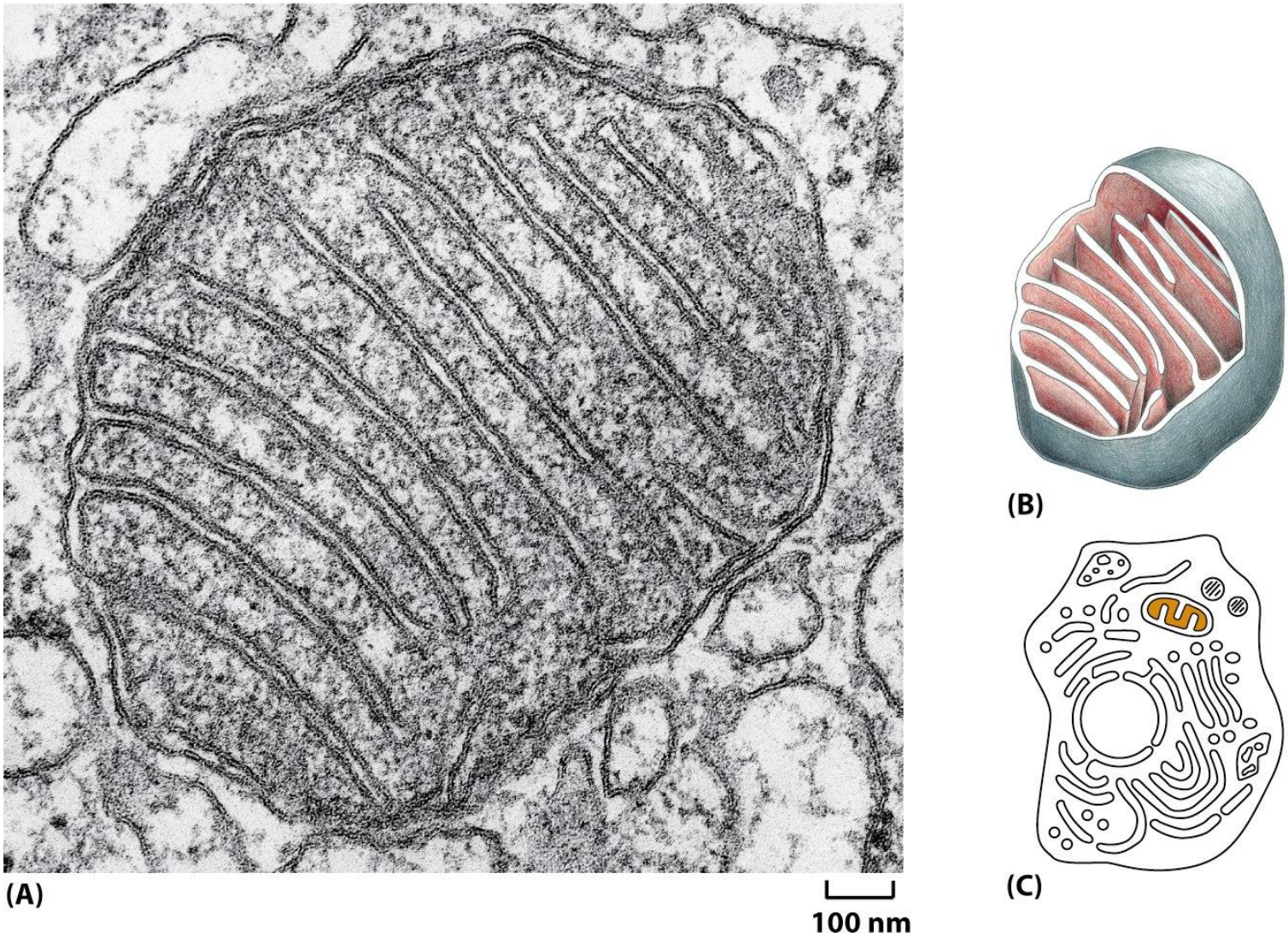


Figure 1-33 Molecular Biology of the Cell 5/e (© Garland Science 2008)

TWO "COUSINS" SHARING THE SAME GENE (KIT) WITH THE SAME MUTATION: THE GENE IS A TYROSINE KINASE THAT FUNCTIONS IN NORMAL DEVELOPMENT AND IS ALSO RESPONSIBLE FOR SOME MALIGNANCIES

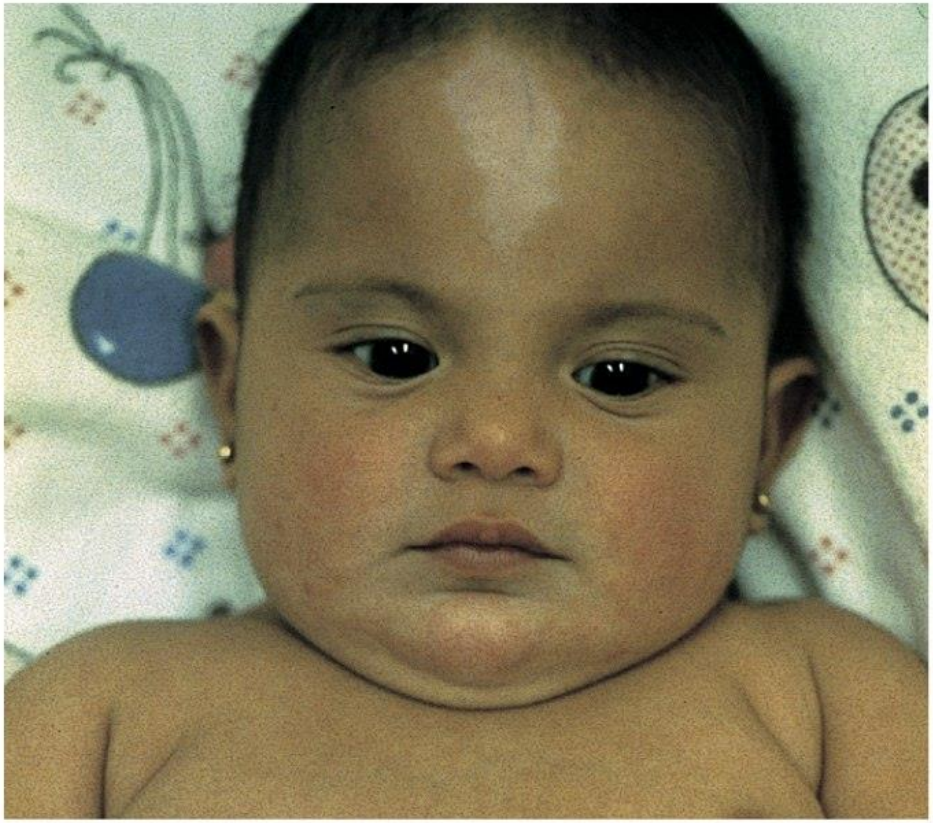
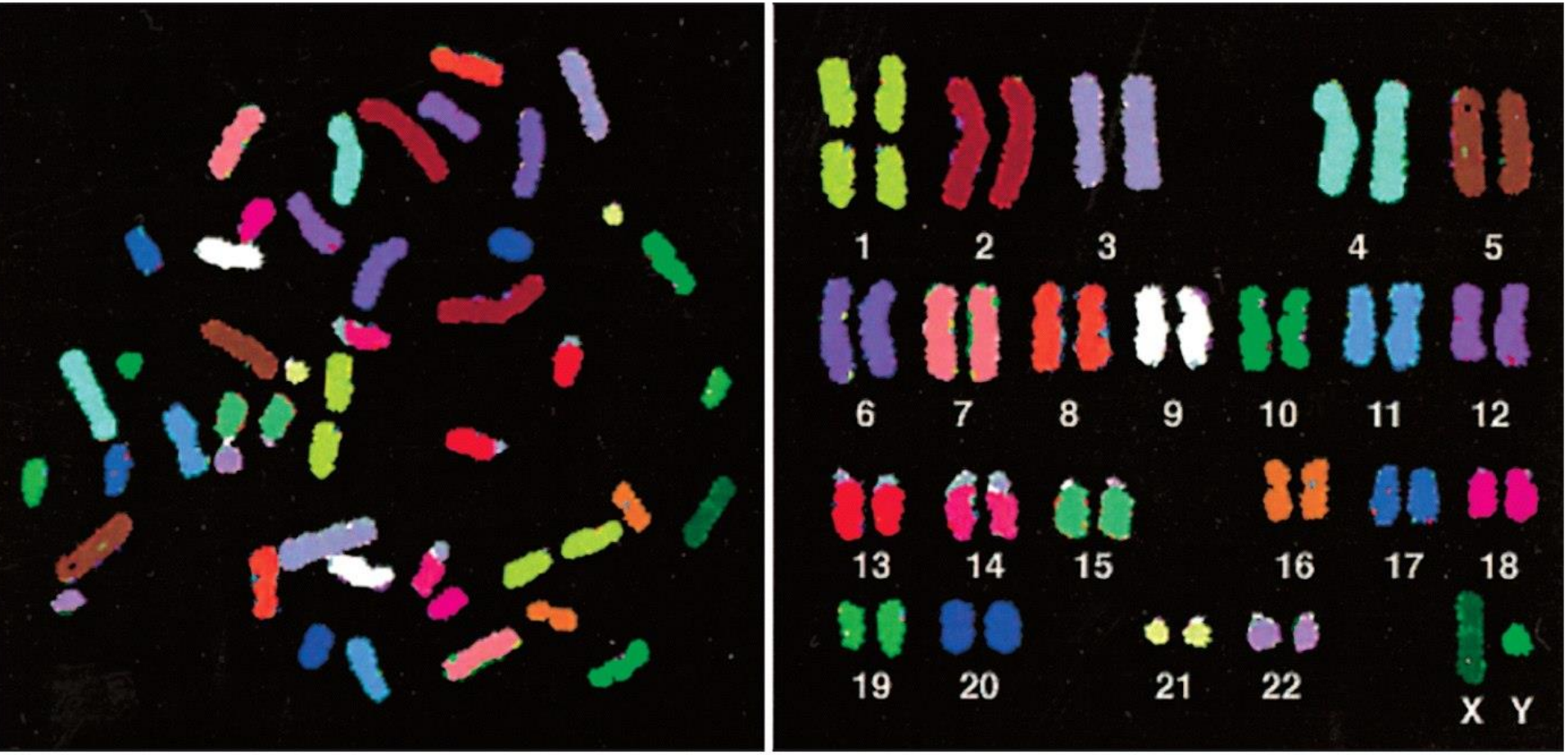


Figure 1-53 Molecular Biology of the Cell 5/e (© Garland Science 2008)

“PAINTED” HUMAN CHROMOSOMES, DIFFERENT GROUPS OF GENES CAN BE STAINED DIFFERENT COLORS



(A) (B) 10 μm

All chromosomes except X and Y are “autosomes”

Figure 4-10 Molecular Biology of the Cell 5/e (© Garland Science 2008)

HUMAN CHROMOSOME DIAGRAM, DIFFERENT DNA SEOUENCES ARE DARK OR LIGHT

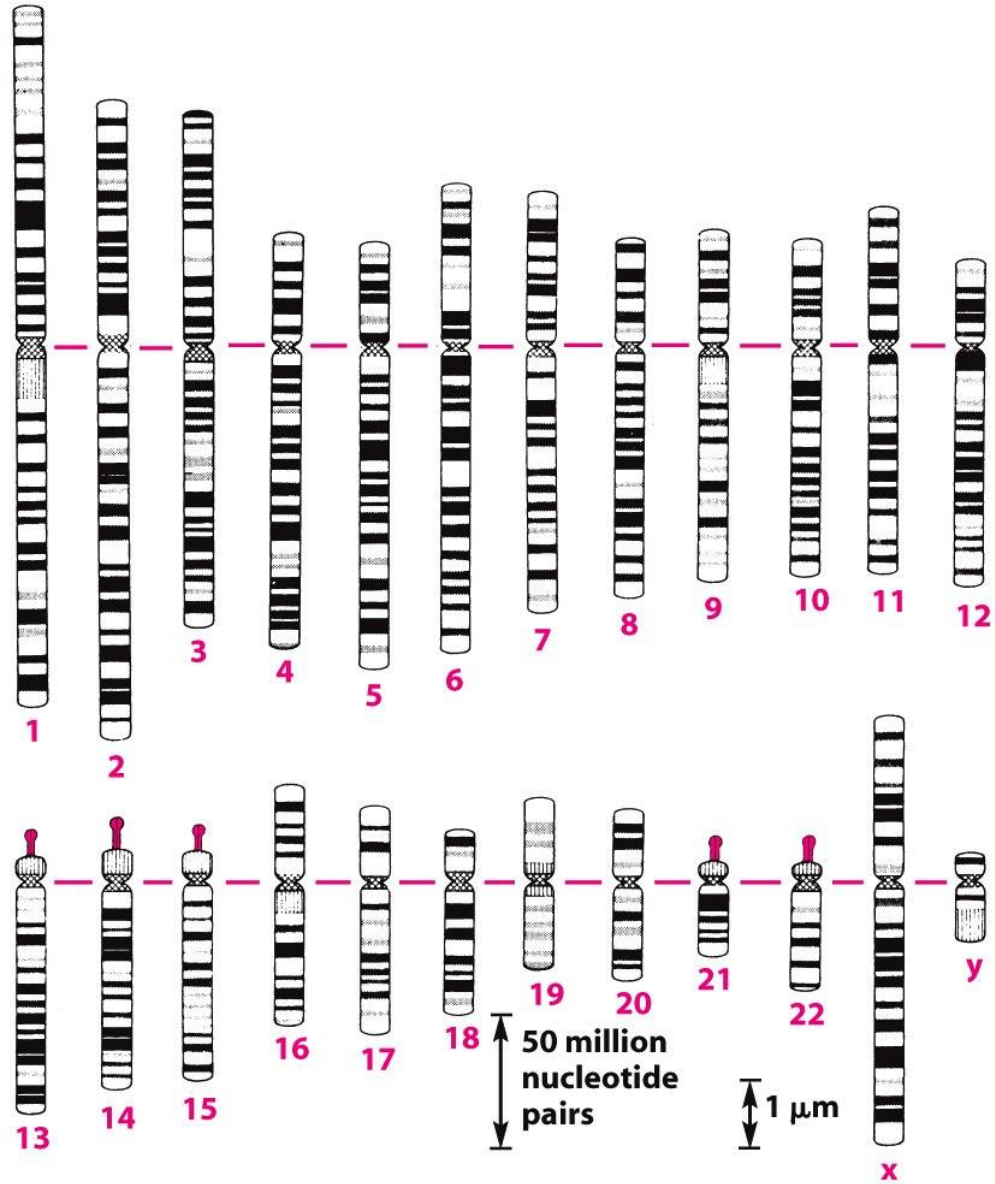
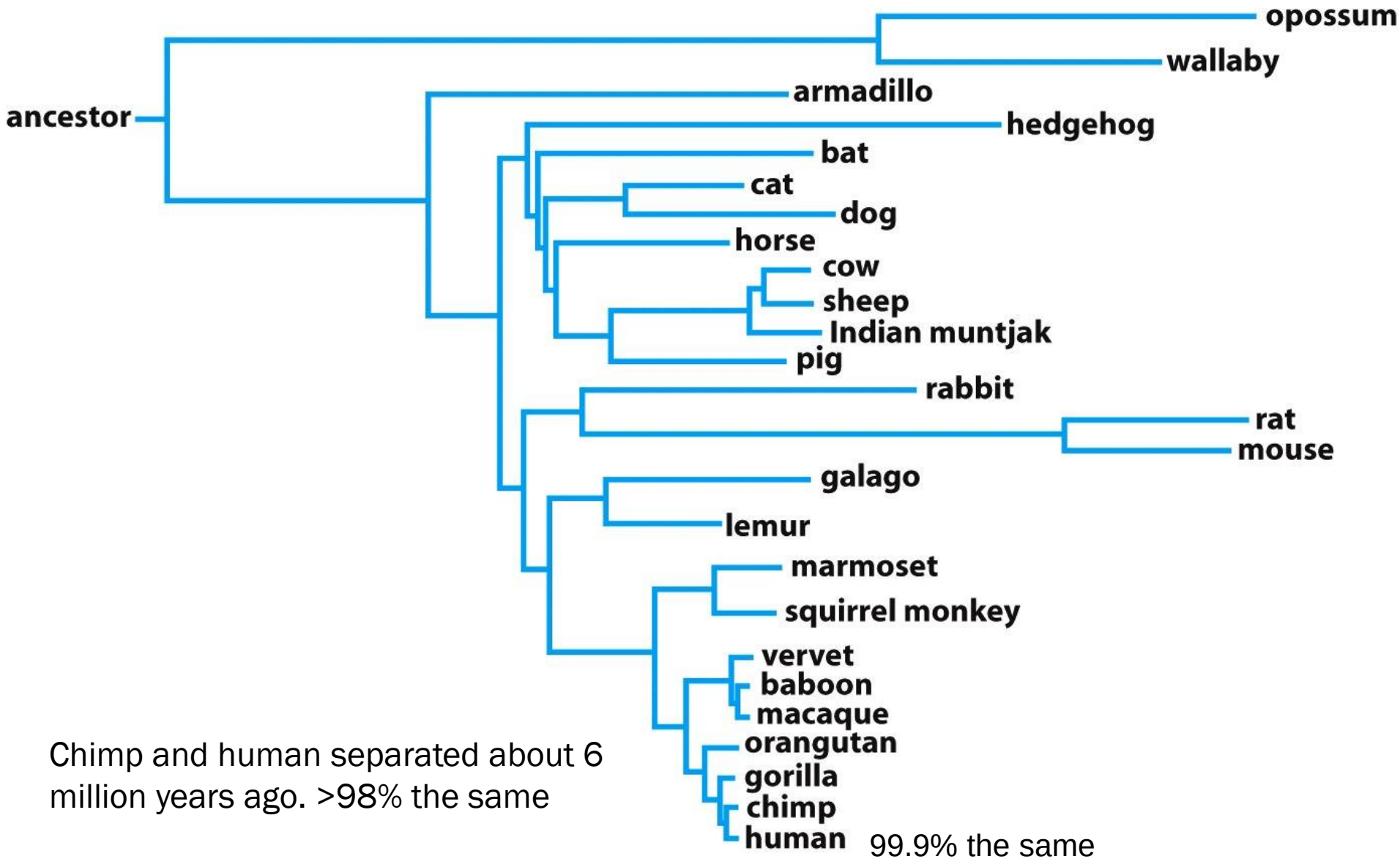


Figure 4-11 Molecular Biology of the Cell 5/e (© Garland Science 2008)

All but X and Y are called autosomes

Three billion total nucleotide pairs. All autosomes can pair with each other and exchange genes. One autosome from mom, one from dad.

FAMILY TREE OF DISTANT RELATIVES (COUSINS)



Chimp and human separated about 6 million years ago. >98% the same

99.9% the same

Figure 4-77 Molecular Biology of the Cell 5/e (© Garland Science 2008)

FAMILY TREE OF CLOSER RELATIVES

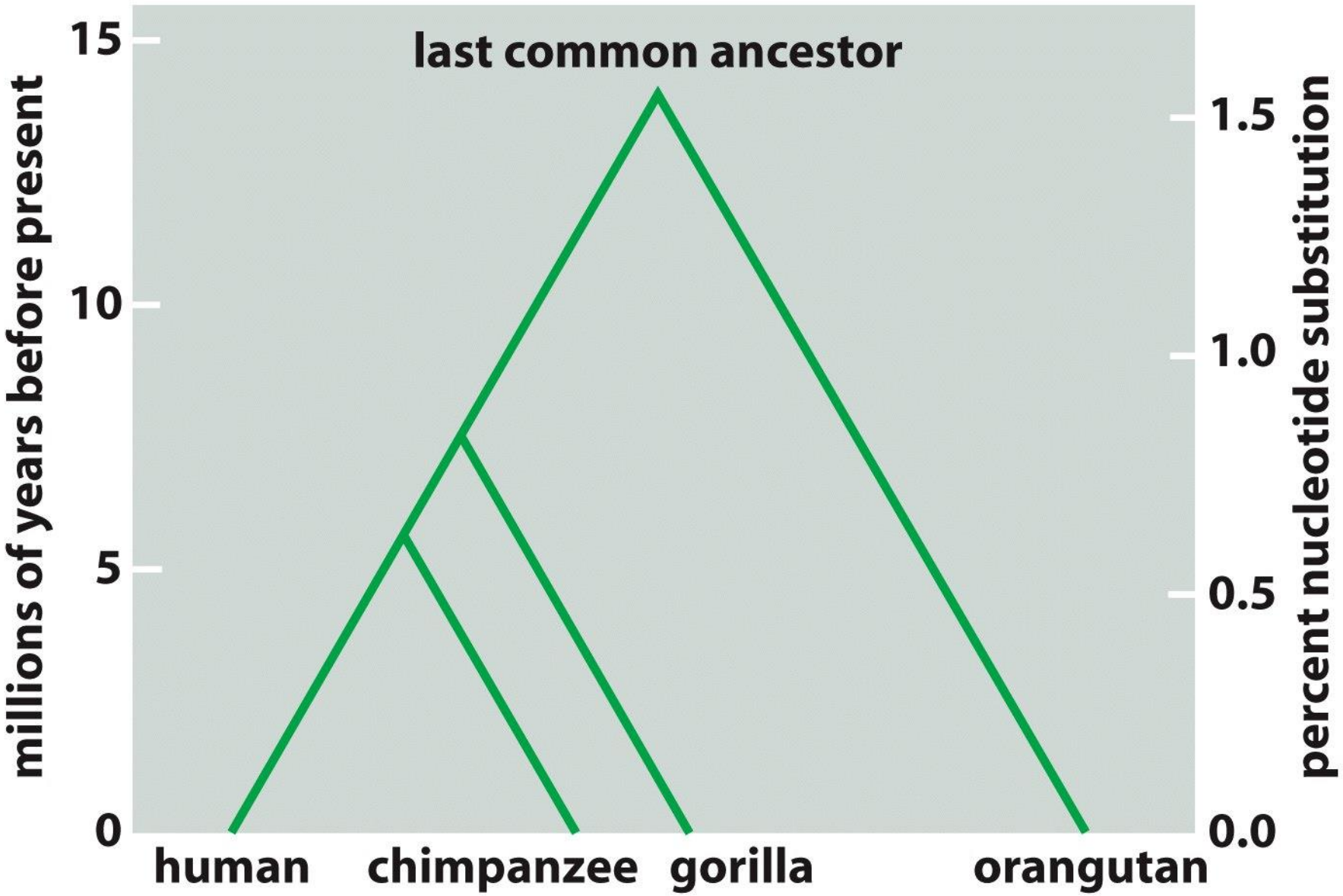


Figure 4-75 Molecular Biology of the Cell 5/e (© Garland Science 2008)

SHARED GENES AMONG "COUSINS"

```
                    gorilla  CAA
                              Q
human   GTGCCCATCCAAAAAGTCCAAGATGACACCAAAACCCTCATCAAGACAATTGTCACCAGG
        ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
chimp   GTGCCCATCCAAAAAGTCCAGGATGACACCAAAACCCTCATCAAGACAATTGTCACCAGG
protein  V  P  I  Q  K  V  Q  D  D  T  K  T  L  I  K  T  I  V  T  R

                                                      K
human   ATCAATGACATTTCACACACGCAGTCAGTCTCCTCCAAACAGAAAGTCACCGGTTTGGAC
        ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
chimp   ATCAATGACATTTCACACACGCAGTCAGTCTCCTCCAAACAGAAGGTCACCGGTTTGGAC
protein  I  N  D  I  S  H  T  O  S  V  S  S  K  Q  K  V  T  G  L  D
                                                gorilla  AAG

                    gorilla  CCC
                              P
human   TTCATTCCTGGGCTCCACCCCATCCTGACCTTATCCAAGATGGACCAGACACTGGCAGTC
        ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
chimp   TTCATTCCTGGGCTCCACCCTATCCTGACCTTATCCAAGATGGACCAGACACTGGCAGTC
protein  F  I  P  G  L  H  P  I  L  T  L  S  K  M  D  Q  T  L  A  V

                                                  V
human   TACCAACAGATCCTCACCAGTATGCCTTCCAGAAACGTGATCCAAATATCCAACGACCTG
        |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
chimp   TACCAACAGATCCTCACCAGTATGCCTTCCAGAAACATGATCCAAATATCCAACGACCTG
protein  Y  Q  Q  I  L  T  S  M  P  S  R  N  M  I  Q  I  S  N  D  L
                                          gorilla  ATG

                         D
human   GAGAACCTCCGGGATCTTCTTCAGGTGCTGGCCTTCTCTAAGAGCTGCCACTTGCCCTGG
        |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
chimp   GAGAACCTCCGGGACCTTCTTCAGGTGCTGGCCTTCTCTAAGAGCTGCCACTTGCCCTGG
protein  E  N  L  R  D  L  L  H  V  L  A  F  S  K  S  C  H  L  P  W
                gorilla  GAC
```

Figure 4-76 Molecular Biology of the Cell 5/e (© Garland Science 2008)

NEUTRAL VERSUS OTHER MUTATIONS

In some languages, some words that mean the same thing are spelled differently:

"Grey" and "gray" mean the same thing as do "color" and "colour." Tyre, tire, jail, gaol. These are neutral mutations.

However, between different languages (species) words that have been separated for a long time usually vary more in spelling

In German, "grau" means "gray."

In Italian, "tavola" means "table."

By comparing large groups of such words linguists can estimate the length of time that two languages have been separated.

SHARED DNA SEQUENCES AMONG DISTANT RELATIVES

```
GTTCCGGGGGGAGTATGGTTGCAAAGCTGAAACTTAAAGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGAAACCTCACCC  human
|    |  |||||||| ||| ||||  |||||||||||||||||||| |||  | |||| || | | ||||||||||||| |||||| || ||||| | ||  | || |||
GCCGCCTGGGGAGTACGGTCGCAAGACTGAAACTTAAAGGAATTGGCGGGGGAGCACTACAACGGGTGGAGCCTGCGGTTTAATTGGATTCAACGCCGGGCATCTTACCA  Methanococcus
 |||||||||||||||| ||||||  | ||||| ||| |||||| ||||||  | | ||||  |||||||| || ||||||||| ||| ||||||   |    ||||||
ACCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAATGAATTGACGGGGGCCCGC.ACAAGCGGTGGAGCATGTGGTTTAATTCGATGCAACGCGAAGAACCTTACCT  E. coli
     |  |||||||| ||  |||| | | ||||| ||| ||||||||||  |  | | || ||  ||||||| || || |||||| ||  |||| ||   ||||| |||
GTTCCGGGGGGAGTATGGTTGCAAAGCTGAAACTTAAAGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGAAACCTCACCC  human
```

Figure 1-22 Molecular Biology of the Cell 5/e (© Garland Science 2008)

"Railroad ties" between sequences indicate identical bases. E. Coli, Methanococcus, and humans have been separated longer than humans and chimpanzees. Human siblings, cousins, second cousins, etc. have been separated for even shorter times.

MEIOSIS - HOW GENES GET SHUFFLED DURING GAMETOGENESIS (SPERM AND EGG)

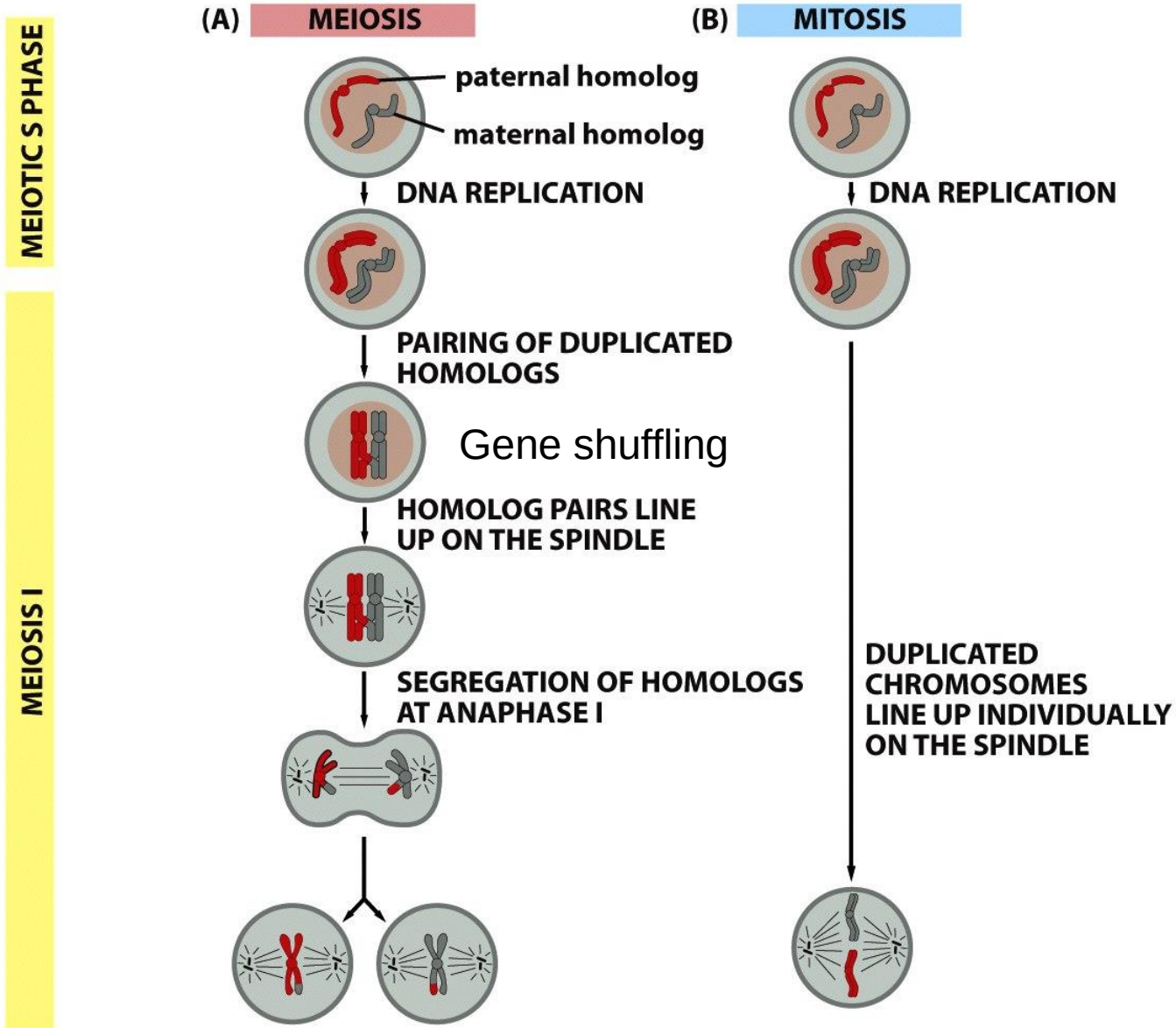


Figure 17-47 part 1 of 2 Molecular Biology of the Cell 5/e (© Garland Science 2008)

MEIOSIS – HOW GENES GET SHUFFLED

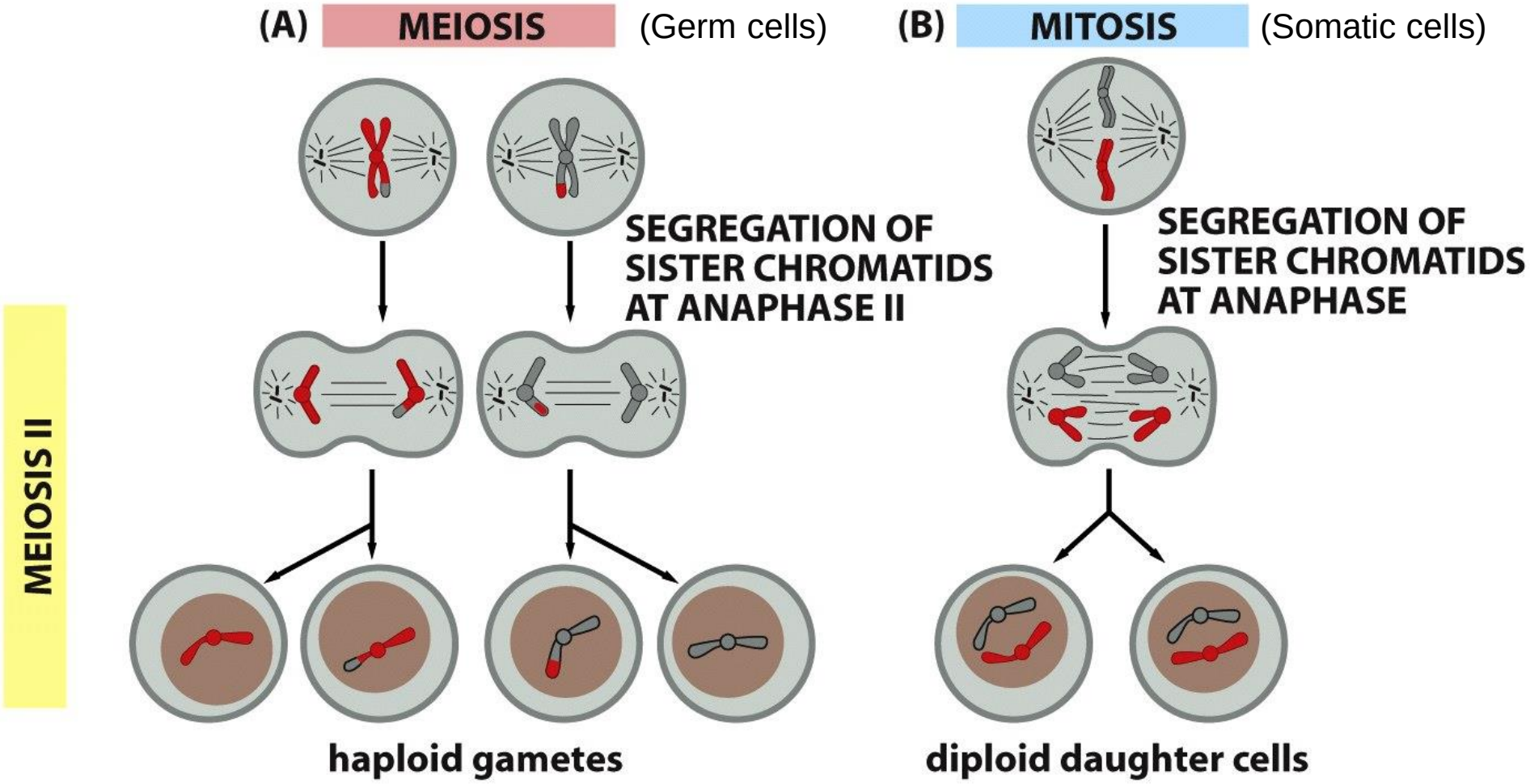


Figure 17-47 part 2 of 2 Molecular Biology of the Cell 5/e (© Garland Science 2008)

WHAT DOES SHUFFLING DO FOR YOU?

Gene shuffling puts genes from mother and father next to each other in new relationships. Thus the offspring, though sharing genes from both mother and father, have new mixtures that have never been tried out before in the species. In the case of immune response genes, this makes it very unlikely that a new organism such as a mutated influenza virus or smallpox, or HIV will ever be able to completely wipe out all individuals in a species.

This process also gives a baby “mommy’s eyes” and “daddy’s nose.” It is useful for babies to look something like their parents so the mothers will nurse them and the fathers won’t eat them.

CELL DIVISION: MITOTIC SPINDLE
BLUE STRUCTURES ARE CHROMOSOMES

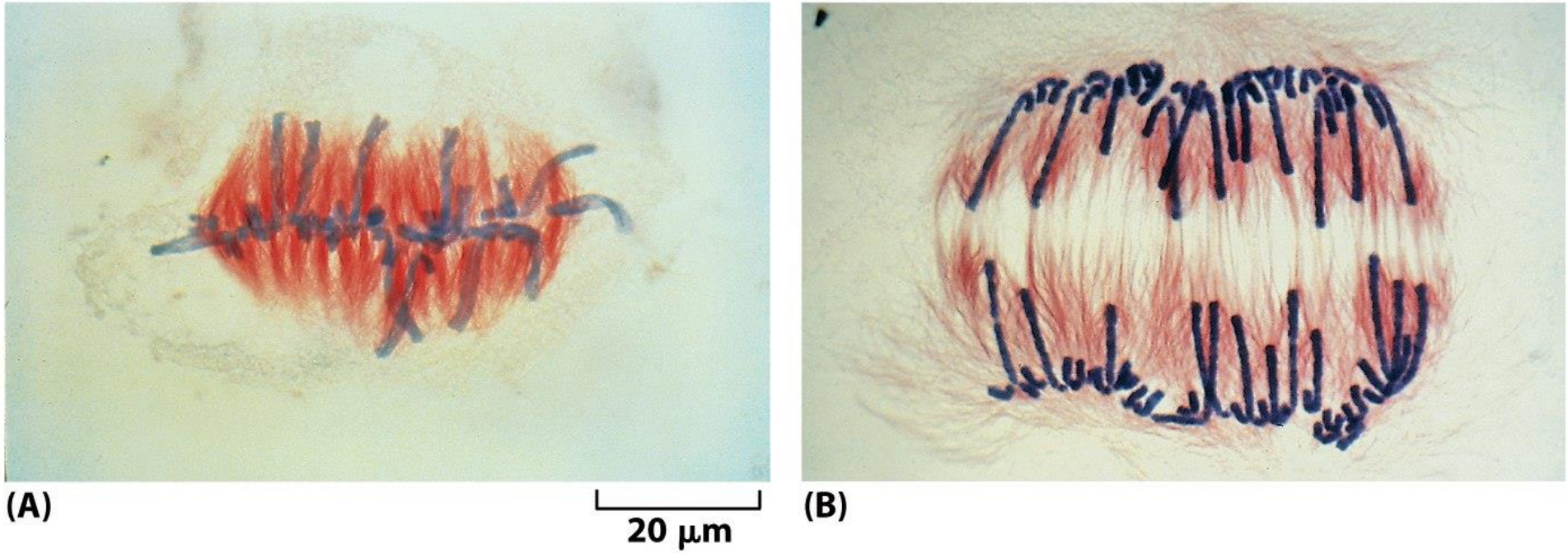
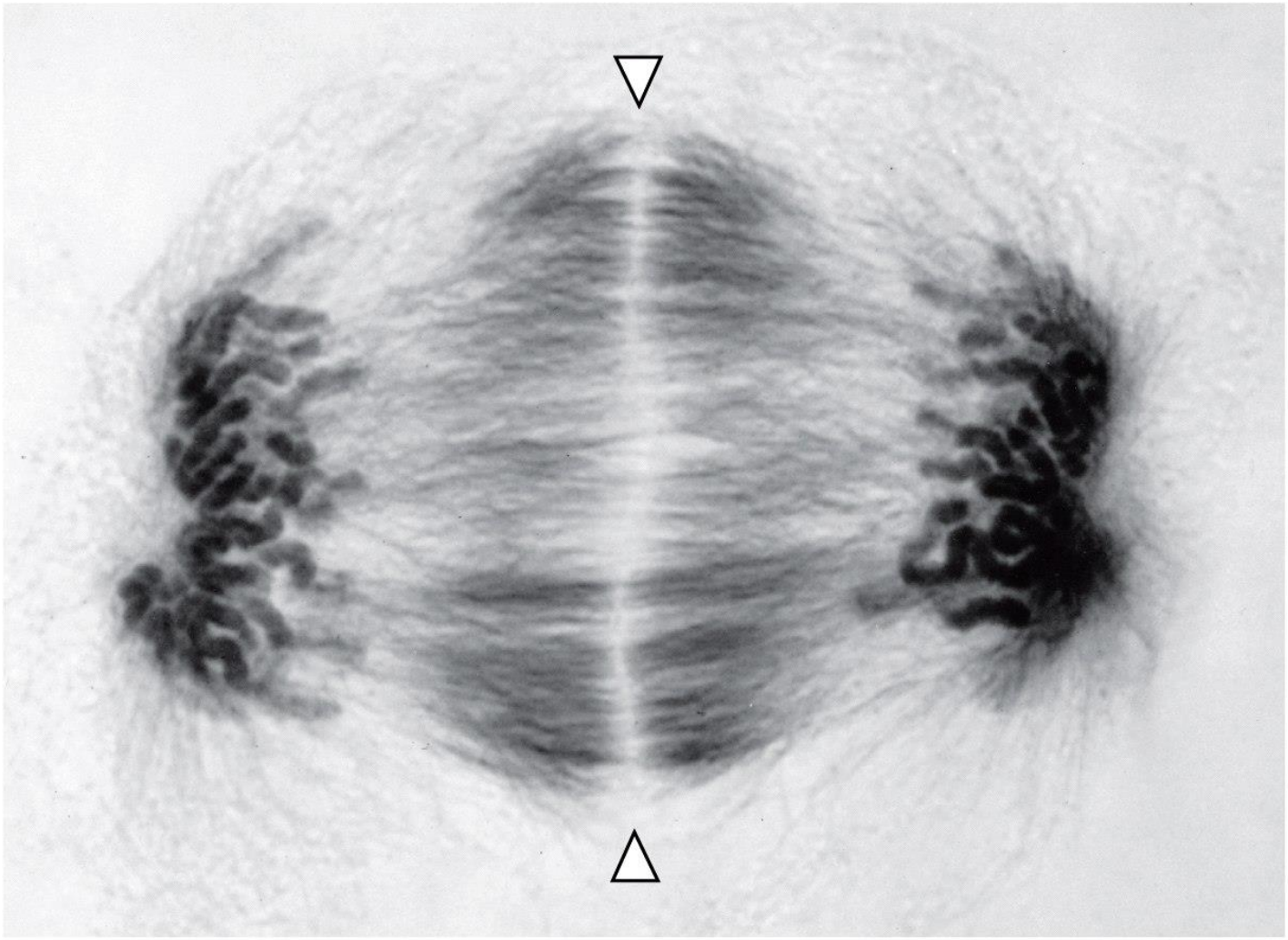


Figure 17-43 Molecular Biology of the Cell 5/e (© Garland Science 2008)

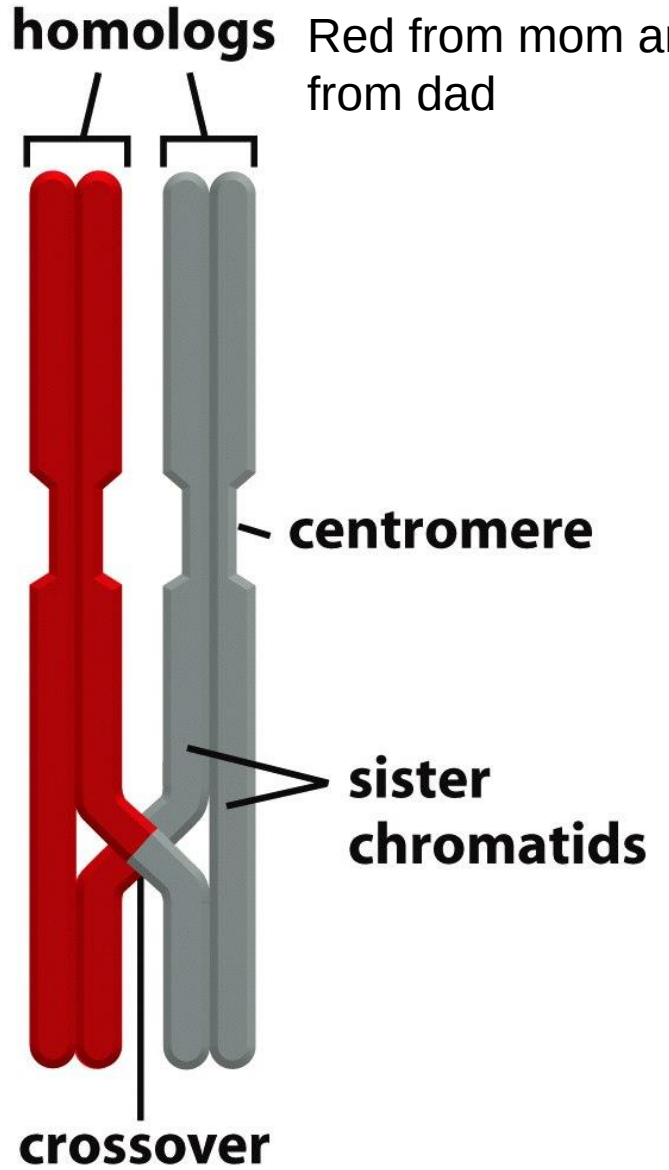
MITOTIC FIGURE APPARATUS: SPINDLE



50 μm

Figure 17-56 Molecular Biology of the Cell 5/e (© Garland Science 2008)

CROSSING OVER:
THE KEY TO UNDERSTANDING GENE SHUFFLING



Red from mom and gray from dad

Figure 17-48 Molecular Biology of the Cell 5/e (© Garland Science 2008)

Sister chromatid exchange occurs during meiosis and the generation of sperm and egg. Maternal and paternal genes are rearranged rather randomly: mother's nose and father's eyes.

The probability of crossing over between any two gene loci is described by the "centiMorgan, cM." One cM is a likelihood of .01 that two genes will be separated in one generation. Its size is about a million base pairs. (Genome is 3 billion base pairs)

HUMAN CHROMOSOME DIAGRAM, DIFFERENT DNA SEOUENCES ARE DARK OR LIGHT

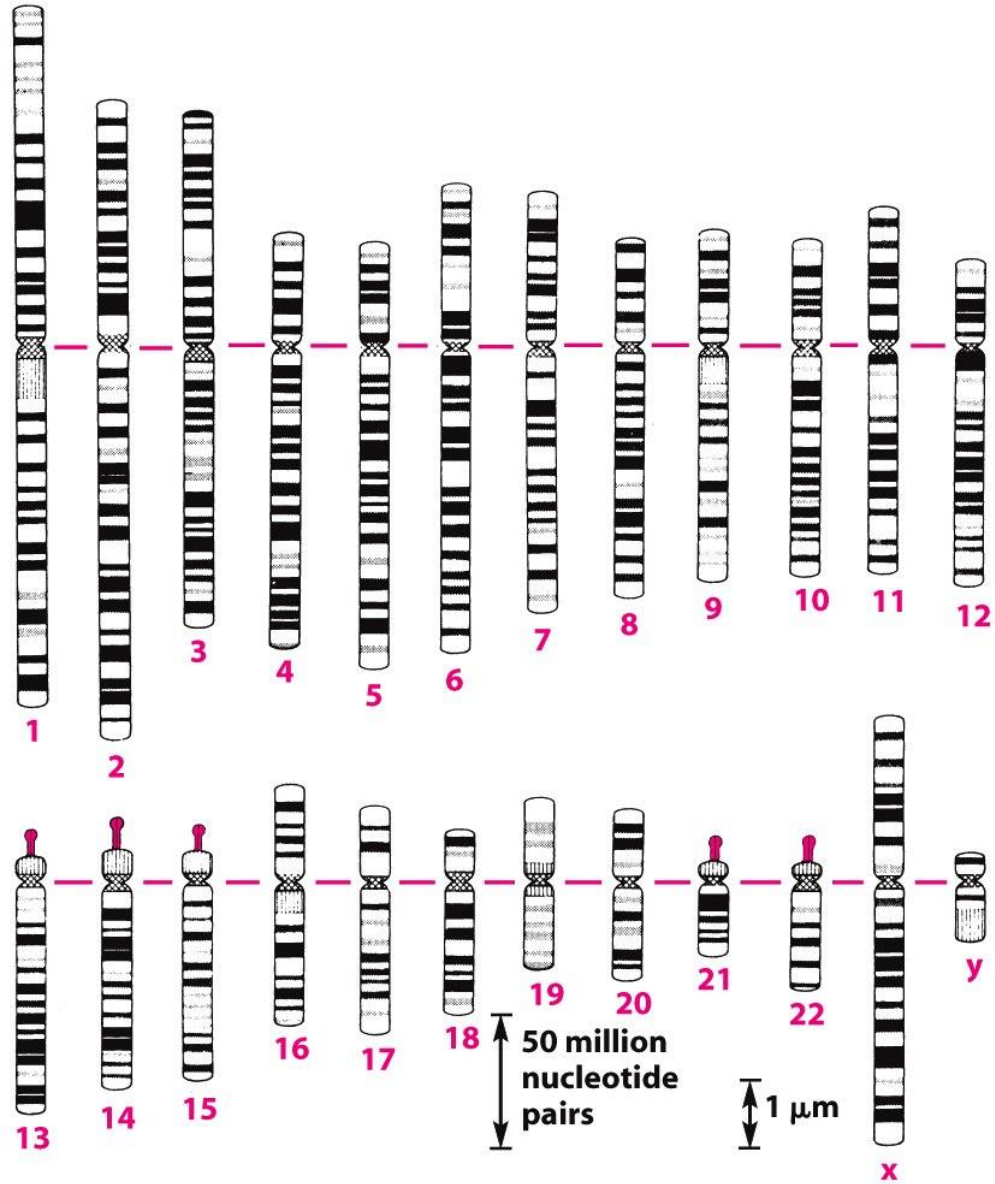


Figure 4-11 Molecular Biology of the Cell 5/e (© Garland Science 2008)

All but X and Y are called autosomes

Three billion total nucleotide pairs. ~25,000 genes. All autosomes can pair with each other and exchange genes. One autosome from mom, one from dad.

USE OF GENE SHUFFLING TO FIND RELATIVES

- A genealogical DNA test examines the nucleotides at specific locations on a person's DNA for genetic genealogy purposes. The test results are not meant to have any informative medical value and do not determine specific genetic diseases or disorders; they are intended only to give genealogical information. Genealogical DNA tests generally involve comparing the results of living individuals to historic populations.
- The Family Finder test matches you confidently with any cousin within five generations. The science is simple - shared blocks of DNA across the 22 autosomal chromosomes are matched between two people. The degree of matching yields evidence for the relationship.
- The closer the relationship the more likely that different genes will stick together because there have been fewer shuffles

HUMAN RELATIONSHIPS

YOU	1
PARENTS	2
GRANDPARENTS	4
GENERATION 10	1024
GENERATION 20	1 MILLION
GENERATION 30	1 BILLION
GENERATION 40 (1,000 years)	1 TRILLION

- Total population of Earth ~ 7.5 billion
- Current best estimates are that about 100-1,000 humans migrated from Africa to fill the rest of the world about 100,000 years (5,000 generations ago) Human generation time is about 20-25 years and getting longer.

- Inbreeding within human populations must have occurred
- That's why all humans are 99.9% the same.

CARD (GENE) SHUFFLING

LONGEST SEQUENTIAL RUN RELATED TO NUMBER OF SHUFFLES

Shuffle #	Hearts	Diamonds	Clubs	Spades
0	13	13	13	13
1	6	6	9	4
2	4	6	4(2)	3
3	3	6	4(2)	3
4	3	4	4(1)	3
8	2	2	3	1

SHARED GENE SEQUENCES AMONG RELATIVES

Sibling	2350 cM	~18-2800 cM
½ Sib	1730 cM	790-2130
1st cousin	880	80-1500
2nd cousin	245	50-760
3rd cousin	90	0-330
4th cousin	50	0-120
5th cousin	25	0-?
Anybody *	~10	0-?

* Shared sequences depend on geographical closeness of ancestors and restrictions on mating (ghettoized Jews in the European Middle Ages)

READING AND INTERPRETING YOUR REPORT

Download Your Family Finder Results

Filter Matches By:	Relation: Close and Immediate	Exact Spelling for Surname Matches	Hide 3rd Party Results
Advanced Filters:	Name:	Ancestral Surnames:	Filter
Export To File:	Excel	CSV	

	GENEALOGICAL MATCHES BASED ON DNA				KNOWN GENEALOGY	
Name	Suggested Relationship	Relationship Range	Shared cM	Longest Block	Known Relationship	Ancestral Surnames (Bolded names match your surnames)
Helmut Schwab	2nd Cousin	2nd Cousin - 3rd Cousin	187.91	17.50	Assign	
Ashley Floyd Fields	2nd Cousin	2nd Cousin - 3rd Cousin	142.16	8.16	Assign	Feldman Osterman
Alice Lieberman	2nd Cousin	2nd Cousin - 3rd Cousin	141.24	15.34	Assign	Ransahoff Sternau (Stern?) Loeb Fred

More distant relationships (4th – 5th cousins)

Daniel Gideon Krasnegor	-	5th Cousin - Distant Cousin	99.20	7.92	Assign	Botwinick (Minsk Belarus) Krasnegor (Shepetivka Ukraine) Seefer (Mlinov
Harold Katz	-	4th Cousin - Distant Cousin	99.13	11.08	Assign	
Larry Heisler	-	5th Cousin - Distant Cousin	98.83	7.81	Assign	
Sigrid Esther Belinkoff	-	5th Cousin - Distant Cousin	98.72	8.91	Assign	Peyser (Lithuania or Poland) Rubin (Belarus) Shanedling (Lithuania) Harris (Poland) Krein (Lithuania)
Bennett Greenspan	4th Cousin	3rd Cousin - 5th Cousin	98.40	12.69	Assign	Borchonoysky Braun Garrop Gorobzoff Greenspan

CARD (GENE) SHUFFLING

LONGEST SEQUENTIAL RUN RELATED TO NUMBER OF SHUFFLES

Shuffle #	Hearts	Diamonds	Clubs	Spades
0	13	13	13	13
1	6	6	9	4
2	4	6	4(2)	3
3	3	6	4(2)	3
4	3	4	4(1)	3
8	2	2	3	1

WHAT IS THE PROBABILITY THAT MY RELATIVE AND I SHARE ENOUGH DNA SEQUENCE TO BE DETECTED BY FAMILY FINDER?

Relationship	Match Probability
2nd cousins or closer	> 99%
3rd cousin	> 90%
4th cousin	> 50%
5th cousin	> 10%
6th cousin and more distant	Remote (typically less than 2%)

Lots of people from the same geographical area, Scotland, for example, share enough DNA to be said to be related at the sixth cousin level. People in the same town, county, country, continent, marry each other.

LINEAL vs GEOGRAPHICAL "COUSINS"

- Lineal cousins share named grandparents, great grandparents, etc.
- Geographic cousins share the same autosomal gene pool which repeatedly shuffles but the genes being shuffled are just like the cards in a 52-card deck
- Many lineal male cousins will share the same Y chromosome
- Many lineal male and female cousins will share the same mtDNA
- Many geographic cousins may also share Y chromosome or mtDNA haplogroups because they are common in that geographic area, i.e ***I1a*** Y chromosome haplogroup in Norway and Scotland (That's me)
- However, lineal cousins will tend to <u>share longer blocks of genes</u> than geographic cousins because there have been fewer shuffles from their most recent common ancestor

Bottom line: All humans are related at about the sixth or seventh cousin level. At about this level of relationship, if you can prove it on paper, your lineal relative will be no more related to you genetically than any other person on Earth.

INSIDE EDITION ACCURACY TESTS

23andMe, Family Tree DNA, and Ancestry DNA were all tested for accuracy on two sets of identical, female triplets and one set of identical ,male quadruplets

Triplet set 1: all 99% European but one girl was 11% French and German, another was 18%, and a third was 22 percent.

Triplet set 2: British Iles ancestry was 59%, 66%, and 70% on the three girls

Quadruplet set: all results essentially identical. The boys did not show the variation that the girls did.

How could this happen?

Small mutations in genes always occur during life, even in identical twins, triplets, etc.

Females have two X chromosomes: one from mom and one from dad. Females randomly inactivate one or the other in all somatic cells. If any of the genes scored for country of origin, are on the X, they would give different results in different females. Males have only one X.

These results might also reflect the “error bars” in the laboratory test and computer algorithms.

The cemetery of Geneva, Kansas, where two centuries of our maternal ancestors rest. We placed my sister, Janet's ashes there on 2/25/2016.

The Mabie family plot, one of the founding families of Geneva, Kansas, a town that no longer exists. Edmund Mabie and Clara May Esse were the parents of my grandmother, Alice Miller (Mabie.)
I am placing some of Janet's ashes with the Mabies.

The Esse (Ess) family plot. They were another of the founding families of Geneva. We placed some of Janet's ashes with each of her ancestral families.

Grandpa Ben Miller and Grandma Alice were the parents of our mother, Lois Miller Baird.

The gravestone of our mother, Lois Miller Baird and my sister, Janet.

Letting the Kansas wind take the rest of Janet’s ashes away.


MILLER