

Association of Recombination to worker honey bees genome structure and behavior

By: Victoria Acosta, Shanyia Chandler, Annie Choi, Morgan Karnes and Kaitlyn Marston

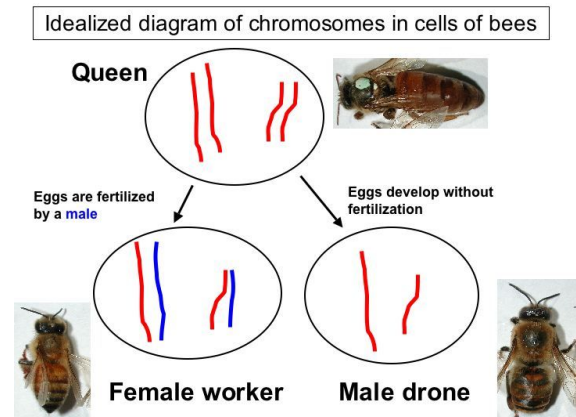
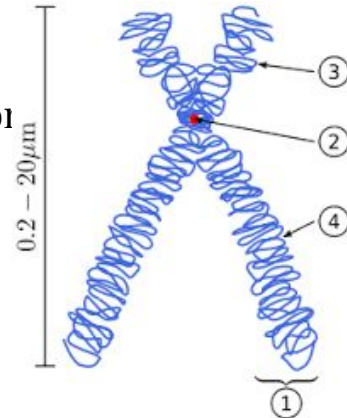
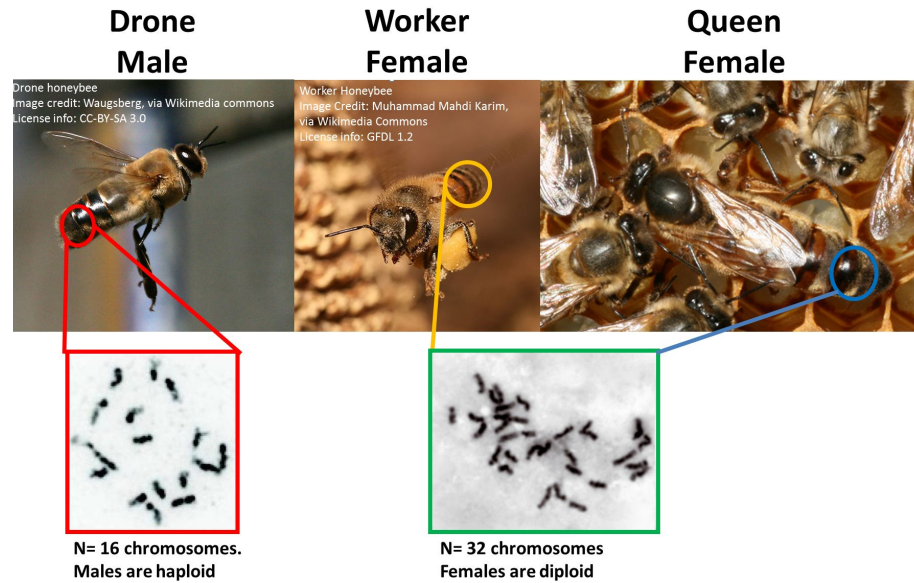
Quick Overview:

Topics include Reproduction and Genetics:

1. Review of basic genetic topics
2. Introduction into the paper
3. Methods
4. Results
5. Conclusion
6. Discussion

Chromosomes

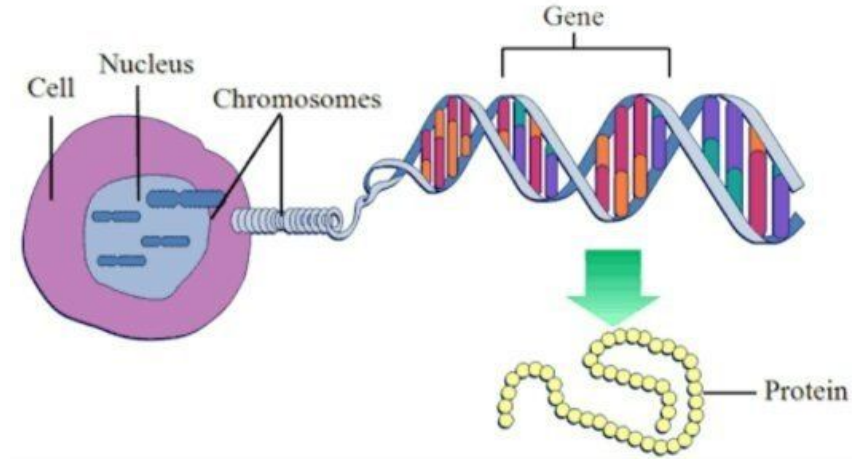
- Made of tightly wound DNA
- Worker bees and Queens:
 - Two sets totaling 32 chromosomes
- Drones:
 - One set totaling 16 chromosomes



Background

Genes

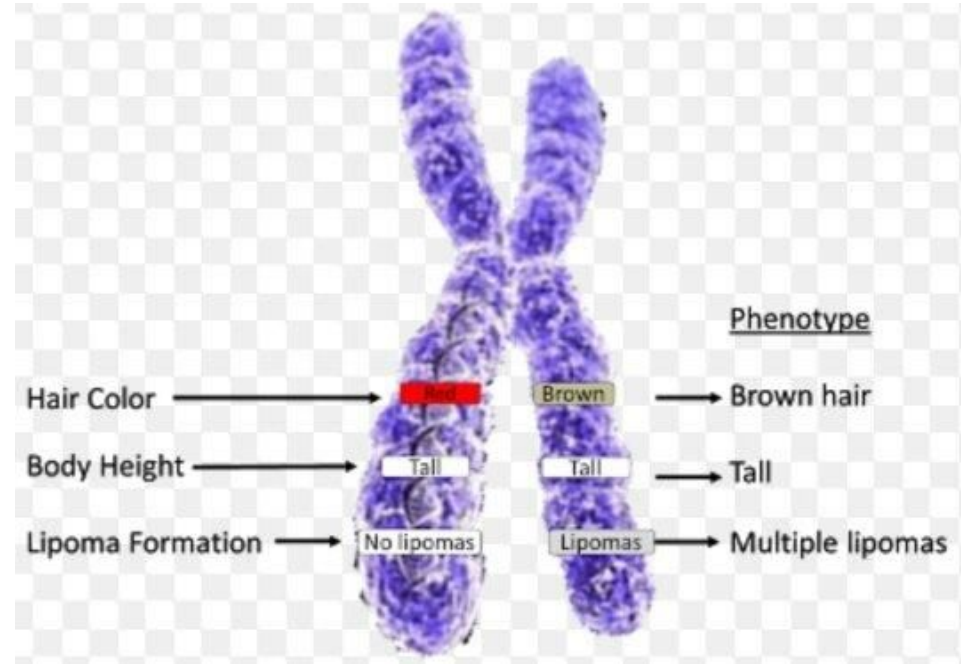
- DNA that encodes for RNA and proteins
- DNA composition:
 - Double stranded
 - Sugar and Phosphate backbone
 - 4 nucleotide bases:
 - Adenine-Thymine
 - Guanine-Cytosine
- RNA
 - Single Stranded
 - Different sugar
 - Instead of Thymine you have Uracil
- RNA translates into codons, which form amino acids to create proteins



Background

Alleles

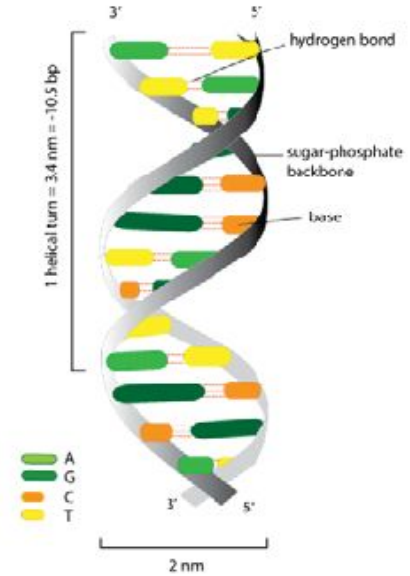
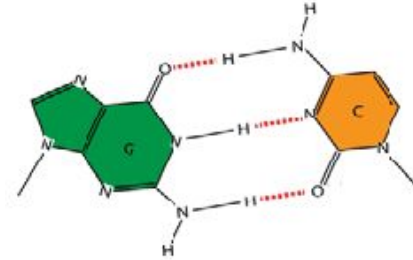
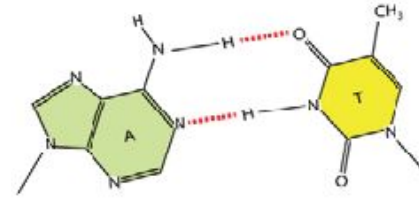
- Alternative forms of genes
- Arise from mutations
- Located at genetic loci
- Allele frequency:
 - How much a certain allele occurs



Background

GC Content

- Complementary bases
 - Adenine pairs with Thymine
 - Guanine pairs with Cytosine
- Hydrogen bonds hold bases together
 - A-T have two H-bonds
 - G-C have three H-bonds
- GC content is amount of G-C bonds within a strand of DNA
- G-C poor lower percentage, G-C rich higher percentage



Introduction

Gene Conversion

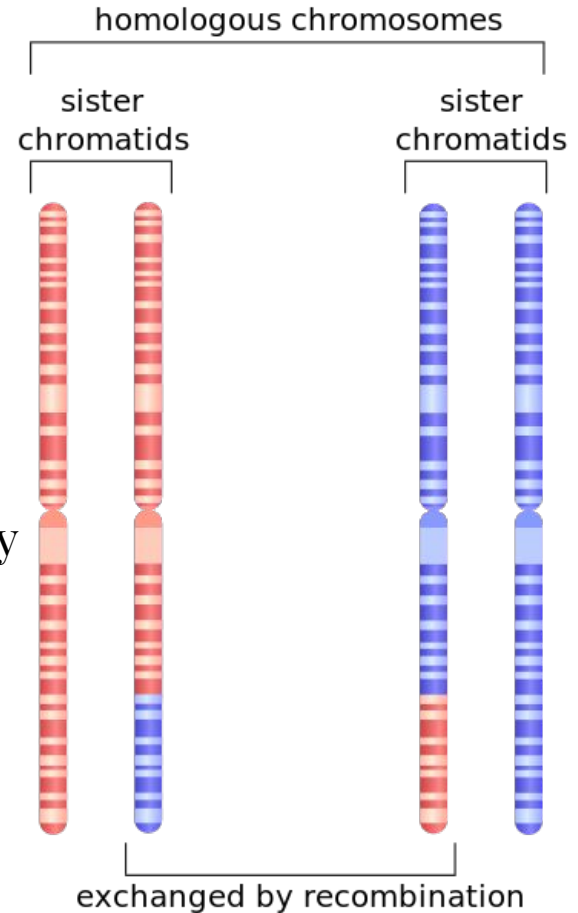
- Two forms:
 - Allelic- occurs in meiosis
 - Nonallelic/ ectopic- repair of double strand breaks from DNA damage
- Conversion refers to one allele being converted into another
- This occurs because of mismatched base pairs during homologous recombination
- Biased gene conversion (bGC) is when one allele has a higher probability of being the donor than the other
- Unbiased gene conversion alleles have same probability

Introduction

Homologous Recombination

- Repair of double-stranded breaks in DNA
- Recombination occurs in both mitosis and meiosis
- Mitosis in somatic cells repair helps prevent cancer
- Meiosis in gametes it helps to create genetic diversity

Introduction



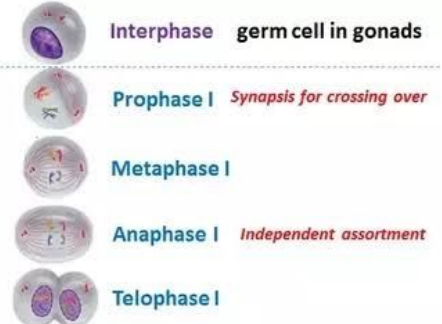
Recombination in Meiosis

- There are two stages
- Occurs in gametic cells not somatic
- Crossing over occurs in Prophase I

Meiosis: two stages

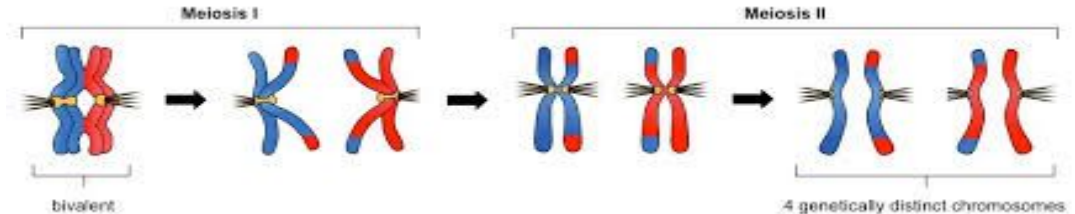
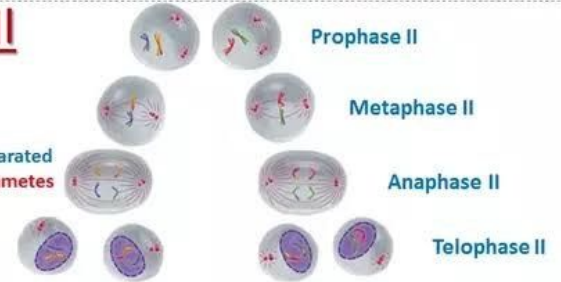
Meiosis I

Segregation
Homologous pairs are separated
reducing chromosome number by half



Meiosis II

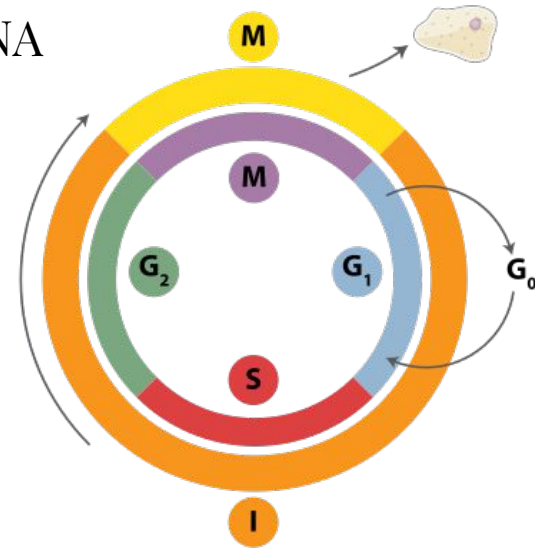
Sister chromatids are separated
producing four haploid gametes



Introduction

Recombination in Mitosis

- In mitosis it occurs before start of M stage and after DNA replication in the S and G₂ phase
- More sister chromatids are present
- Sister chromatids are used because they are identical copy of chromosome
- There are several pathways

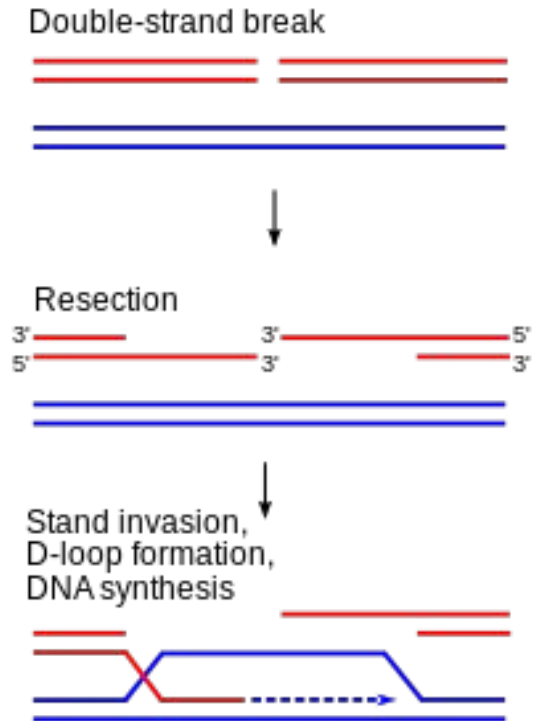


Introduction

Beginning steps in Double-Strand Break Repair and Synthesis-Dependent Strand Annealing Pathways

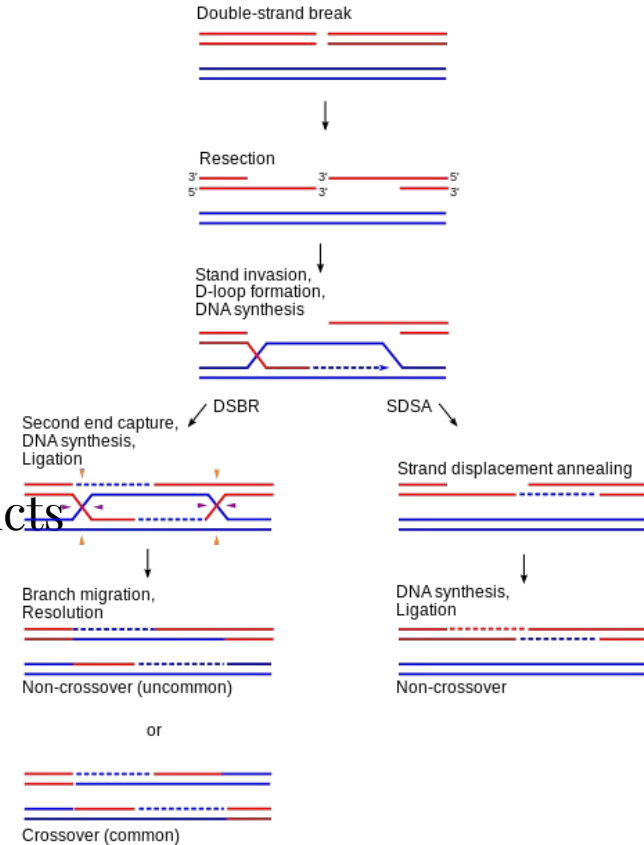
- MRX complex recruits proteins and enzymes to cut back 5' end of DNA break
- Other proteins then bind to 3' end overhang to form a nucleoprotein filament
- This filament invades the sister chromatid and searches for similar sequence
- This forms a displacement-loop
- DNA polymerase then synthesizes new DNA

Introduction



Double-Strand Break Repair Pathway

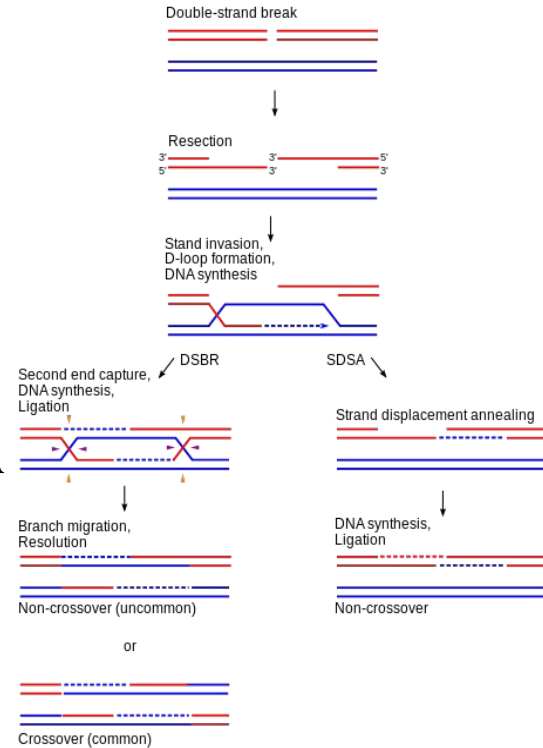
- Also known as Double Holliday Junction Model
- Can occur in both meiosis and mitosis
- The other 3' end forms a Holliday junction
- Nicking endonuclease forms recombination products
- Crossing over usually occurs
- Determined by how Holliday junction is cut



Introduction

Synthesis-Dependent Strand Annealing Pathway

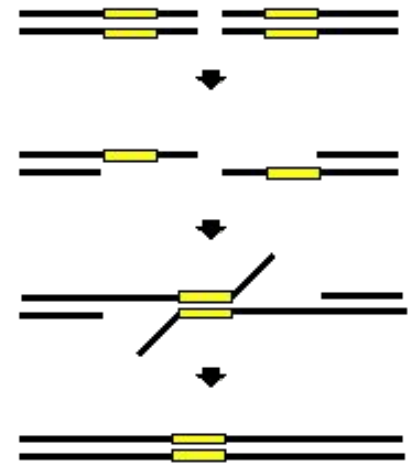
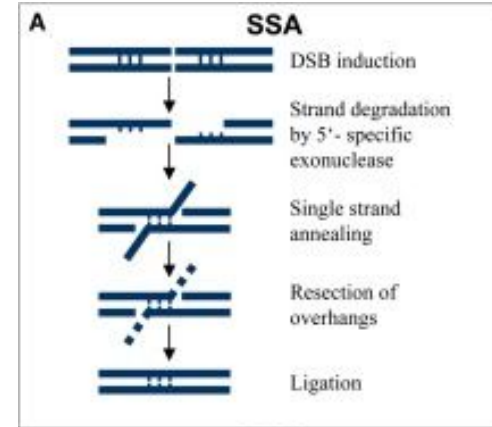
- Can occur in meiosis and mitosis
- The invading strand extends along sister chromatid
- Releases as Holliday Junction slides a process known as branching migration
- The new 3' end anneals to 3' overhang in damaged DNA
- Any flaps are removed and DNA ligation occurs
- Crossing over does not occur



Introduction

Single-Strand Annealing Pathway

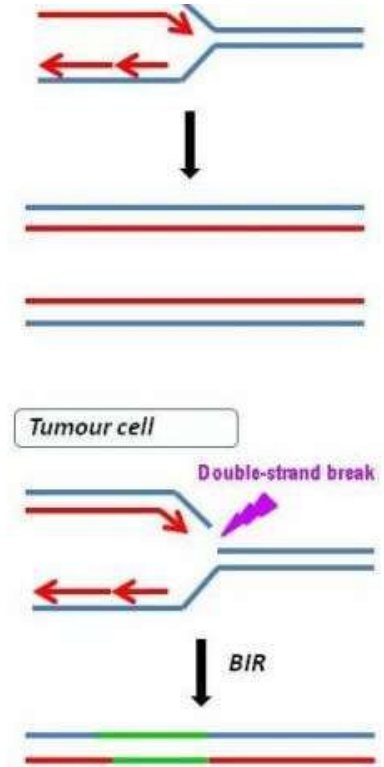
- Repairs double-strand breaks between repeat sequences
- DNA duplex are cut back around the site of breakage
- Resulting 3' overhangs align and anneal



Introduction

Break-Induced Replication Pathway

- Fixes breaks at the DNA replication fork
- Process unknown
- There are a few proposed mechanisms
- Believe to start with strand invasion
- Theories differ in steps after strand invasion
 - Migration of displacement loop and other phases
- Can help maintain telomere lengths



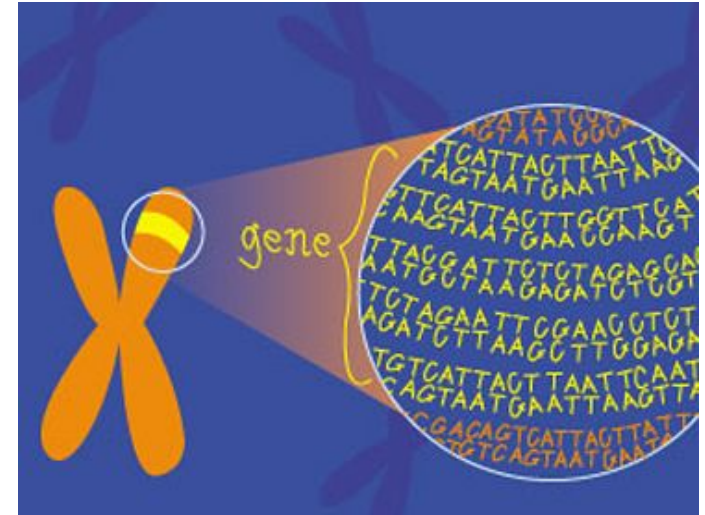
Introduction

Introduction Session



What is being tested?

- Conducted analysis that combined population genetic, molecular evolution, and transcriptomic studies of the honey bee to determine the connection between recombination, GC-content, and social behavior.



Datasets of Worker Bees

1, Sanger data

2. SNP data

3. Baylor Data

- *A. mellifera*
 - *A. cerana*
 - *A. dorsata*
 - *A. florea*

Methods

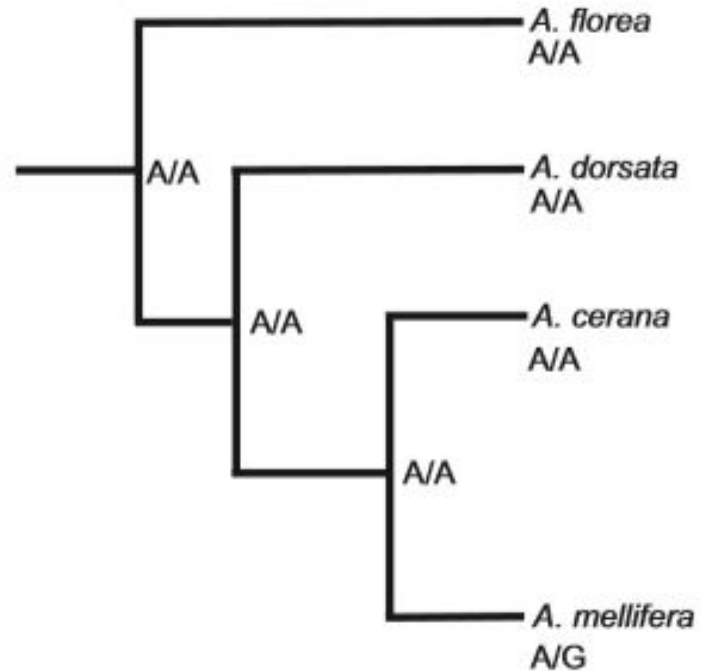


Fig. S1. Phylogeny of *Apis* spp. sequenced. Phylogenetic relationships among the four *Apis* spp. sequenced are based on a recent molecular phylogeny.

Phasing and Recombination

Phase v2

- Determine differences between haplotype from unphased diploid genotypes from the Sanger sequencing
 - Outputs: R software and DNAsp

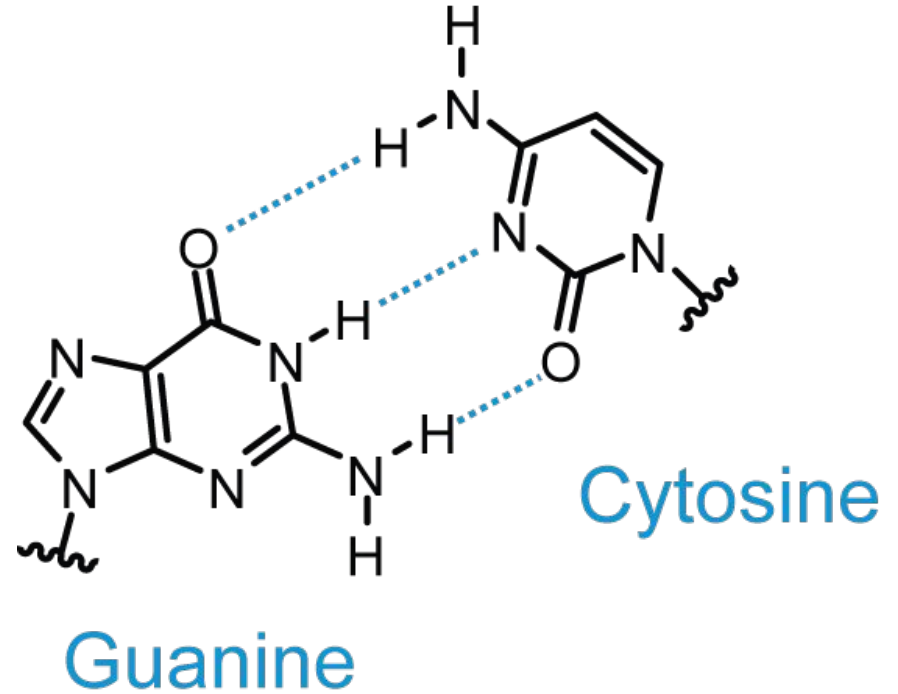
Reducing errors

- Through redundancy: Forward and reverse primers
- Eliminated population structure, size change and artificial selection

Methods

GC Content Classification

- The percentage of nitrogen bases that are of GC content
 - Identification of rich or poor
- Consistent with other studies
 - GC₃ mean content of 33 %



Methods

Ancestral State Construction

Used maximum Parsimony

- *A. mellifera*
 1. *A. cerana*
 2. *A. florea*
 3. *A. dorsata*

High frequency of strong GC from mutations

- Fixation rate : ancestral base pair for transitions
- Polymorphism rate: the reverse transition

Methods

Diversity and Divergence

Measuring GC₃ with DNAsp

- Synonymous diversity (π_s)
- Synonymous (K_s) and Nonsynonymous (K_a) divergence

Perl Scripts

- Estimated GC content

Methods



Gene Ontology and GC

Orthologs of Predicted honey bees

- Found by reciprocal blastp as the best match

Comparisons of GO Groups using:

1. *A. mellifera*
2. *Nasonia vitripennis*
3. *Acromyrmex echinator*
4. *Pogonomyrmex barbatus*



1



2



3



4
© 1999 John Abbott / naturepl.com

Methods

General Statistical Methods

Nonparametric Tests

- Wilson Rank Sum Test

Statistical Tests

- R software



Methods

Association Between Biased Gene Conversion and Recombination/GC

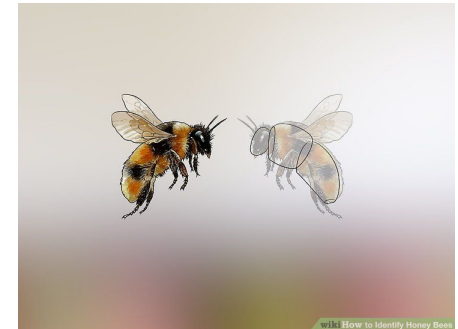
- Polymorphism Data:
 - High and low groups based on recombination rate
- Fixation data:
 - Strong and weak synonymous fixations
 - Ancestral weak and strong synonymous sites for each gene



Methods

Analysis of GC Content and Recombination across four Hymenopteran Genomes

- Calculated the GC content of the coding sequence for every predicted gene in each species
 - Median absolute deviation (nonparametric estimate of spread)
- Recombination rate on GC content was examined with GO groups
 - Restricted Maximum Likelihood (REML)



Methods

Analysis of Microarray Studies

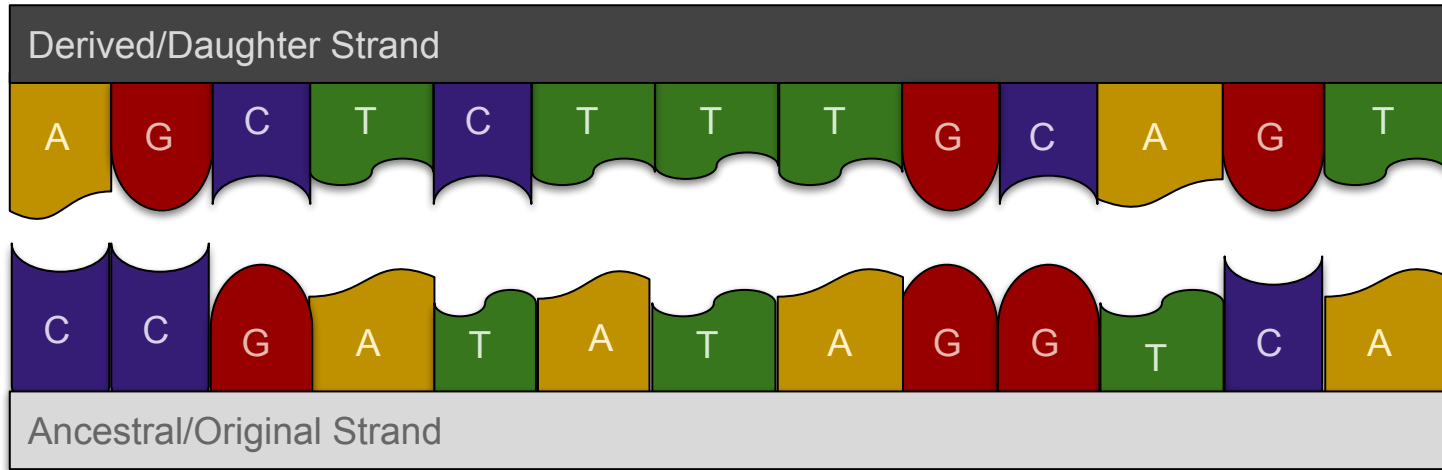
- Two datasets used to examine expressed genes in the bee brain:
 - Test 1: compared gene expression of queens and workers
 - Test 2: compared gene expression in drones and workers, and nurses and foragers
- Analyzed only array probes that were identical to a predicted gene in official gene set
- Caste related genes had significantly higher levels of expression

Methods

Methods Session



Biased Gene Conversion



Results

Biased Gene Conversion

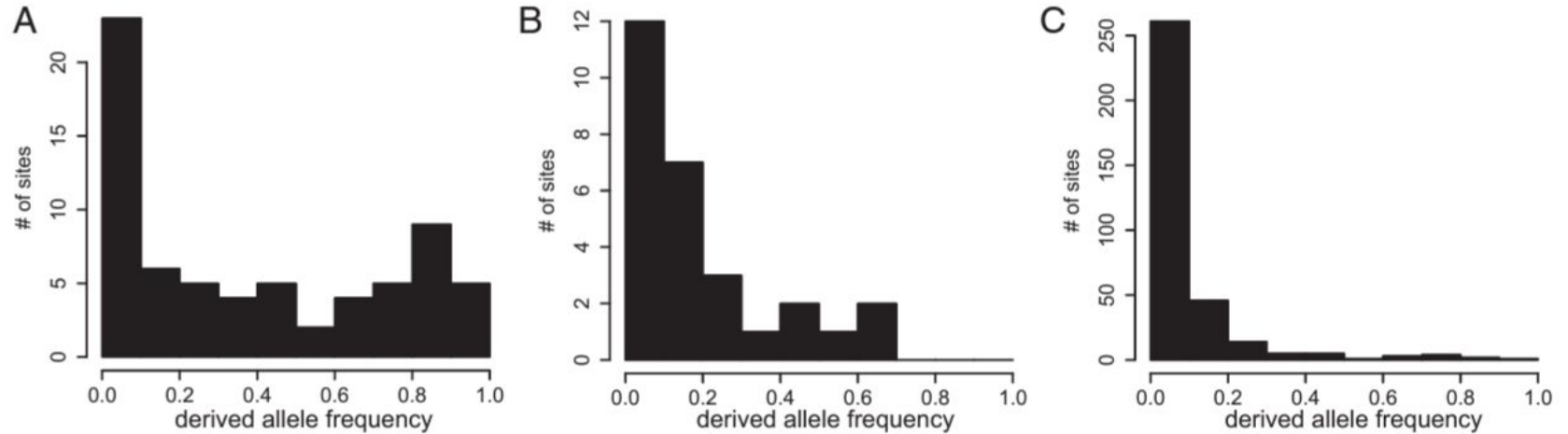


Fig. 1. Derived allele frequency spectra are consistent with bGC. (A) W→S mutations that increase GC content have higher frequencies relative to (B) mutations with no effect on GC content and (C) S→W mutations that decrease GC content ($n = 23$ genes).

Results

Biased Gene Conversion

- Recombination rate highly associated with frequency of G/C mutations
 - High-Recombination Regions: S alleles/W site outnumbered W alleles/S site by 2.5x
 - Low-Recombination Regions: W alleles/S site outnumbered S alleles/W site by 3.3x
- Recombination rate highly associated with fixation rate of G/C mutations
- Strong association between population genetic estimates of recombination rate and GC content at a fine scale
 - GC-rich genes ($\geq 38\%$) had 10.9x higher recombination rates compared with GC-poor genes
 - GC-rich genes exhibited strong association with a bias toward fixation of derived G/C mutations

Results

Biased Gene Conversion

Predictions observed in analysis:

1. An increase in the derived allele frequency of strong mutations
2. An increase in the fixation rate of derived strong mutations
3. The bias favoring strong derived mutations was in synonymous, intronic, and intergenic sites
4. The bias favoring strong derived mutations was observed in areas with high recombination rates

Results

GC Content and Behavior in Honey Bees

- Compared predicted functions of GC-rich genes using Gene Ontology (GO) database
- Correlation between GO groups (of genes) associated with neurobiology and behavior were GC-rich
- Genes associated with worker behavior can quickly evolve due to the relationship between GC content and genetic diversity and divergence

Results

GC Content and Behavior in Honey Bees

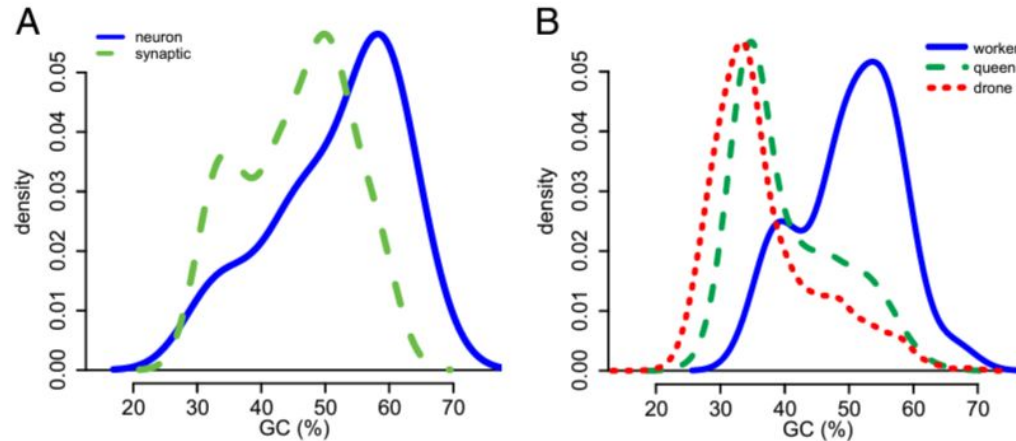


Fig. 2. We generated a smoothed scaled density plot for various groups of genes based on the GC content of coding sequences. (A) Genes from GO “synaptic transmission” (green dashes, $n = 42$ genes) and “neuron fate commitment” (blue solid, $n = 13$) illustrate how genes associated with neurobiology and behavior are often GC-rich. (B) Genes with significantly higher expression in the brains of workers (blue, $n = 701$), queens (green, $n = 389$), and drones (red, $n = 1,296$). Worker-related genes have the highest GC content.

Results

GC Content and Behavior in Honey Bees

- Genes associated with worker behavior and worker division of labor
 - Predominantly GC-rich
 - Higher rates of molecular evolution
 - Expected in regions of high rates of recombination
- “High levels of genetic diversity provide the fuel for evolution, and high recombination rates render the actions of natural selection on these mutations more efficient.”

Results

GC Content and Behavior in Honey Bees

- Is this association found in other eusocial Hymenoptera?
- Compared the GC content of different GO groups with the behavior in the honey bee, two ant species, and the solitary jewel wasp
- Detected highly significant effect of average genome-wide recombination rates on GC content across these species
- Higher GC content of behaviorally related GO groups

Results

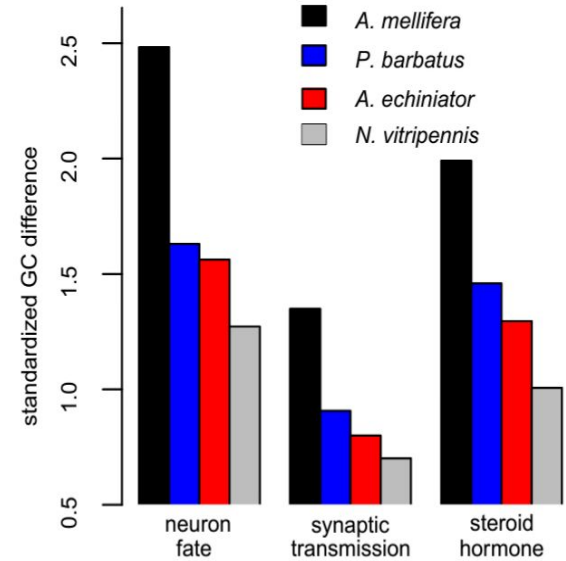


Fig. 3. A standardized measure of GC content is shown for genes across three behaviorally relevant GO groups relative to the genomic GC content in *A. mellifera*, *A. echinator*, and *P. barbatus* (all eusocial) and *N. vitripennis* (solitary). Positive values indicate that the average GC content of genes in a GO group are higher than the average GC content of all genes in a specific genome. The four species are ordered (left to right) by decreasing recombination rates. Across these three GO groups, all social species have higher GC content (relative to the species' genome average) than the solitary *N. vitripennis*, and the deviation in GC content increases as a function of recombination rate.

Results Session



Connections between recombination, GC content, molecular evolution, and social behavior

- Through analysis:
 - Recombination rate and GC content
 - Recombination rate and the rate of molecular evolution
 - GC content and the rate of molecular evolution
 - A significant enrichment of genes associated with behavior and genes with worker-biased brain gene expression in GC-rich regions



Conclusion

Gene Expression

- Worker-biased brain gene expression have 48% of more genetic diversity compared to the controls
- Genes associated with worker behavior and division of labor are located where rates of recombination and molecular evolution are the highest

Conclusion

Observations

- Bias in allele frequency of obtained G/C mutations
- Recombination and bGC act to protect GC-rich regions from A/T biased mutation rates
- Recombination rates increase the evolutionary rate of genes in worker behavior

Conclusion

Question still unanswered

- Why recombination is associated with eusociality in hymenopteran?
 - The evolution of eusociality may suggest higher rates of recombination causing an increase in colony fitness.



Discussion

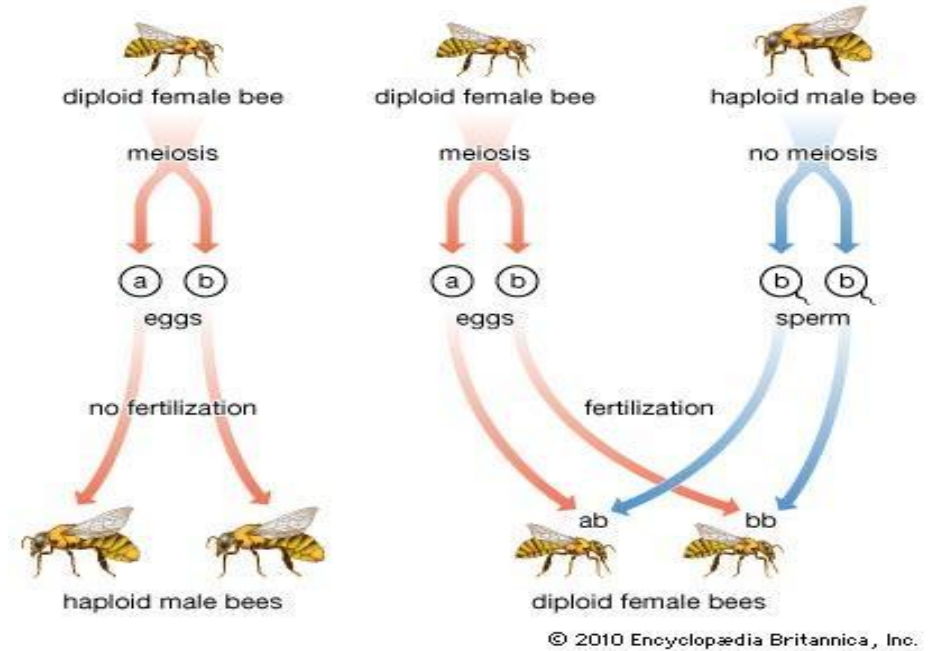
Immediate questions and reflections

Quiz questions?

Discussion

Discussion Question 1

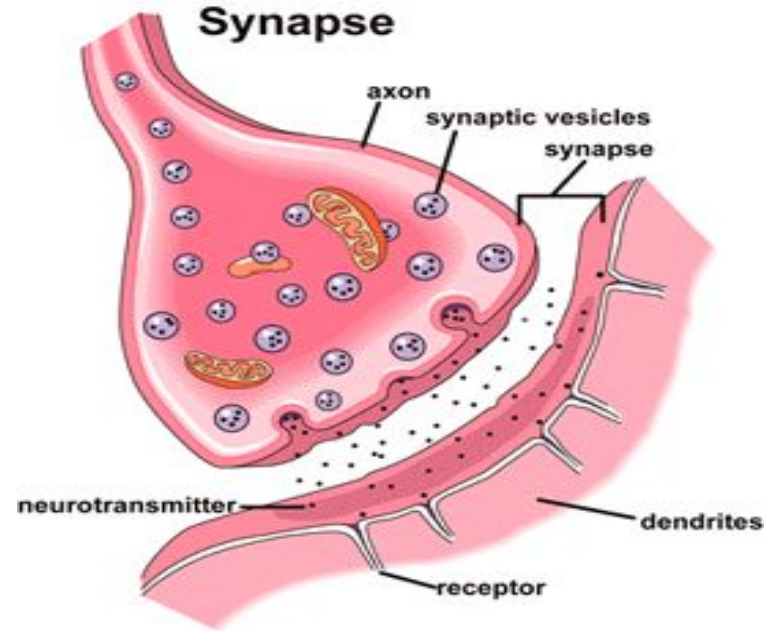
How does the honeybee genome maintain heterogeneity with strong biased gene conversion?



Discussion

Discussion Question 2

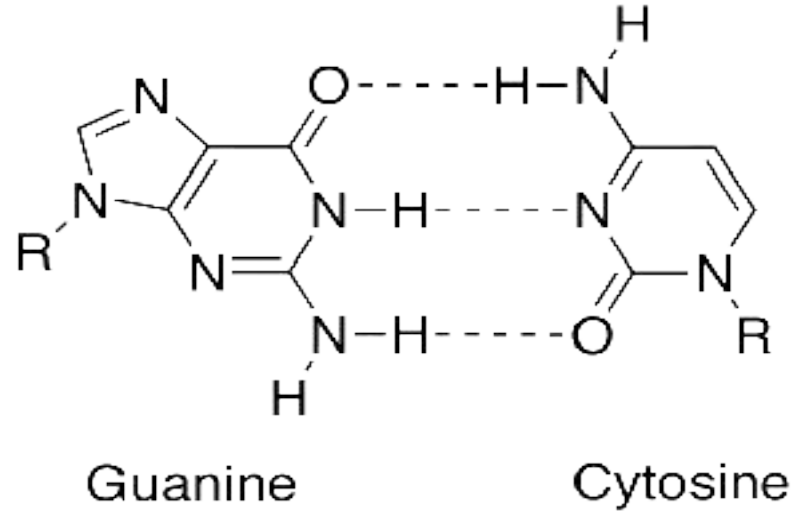
Why did the study use genes associated with neurobiology such as neuron fate commitment and synaptic transmission genes?



Discussion

Discussion Question 3

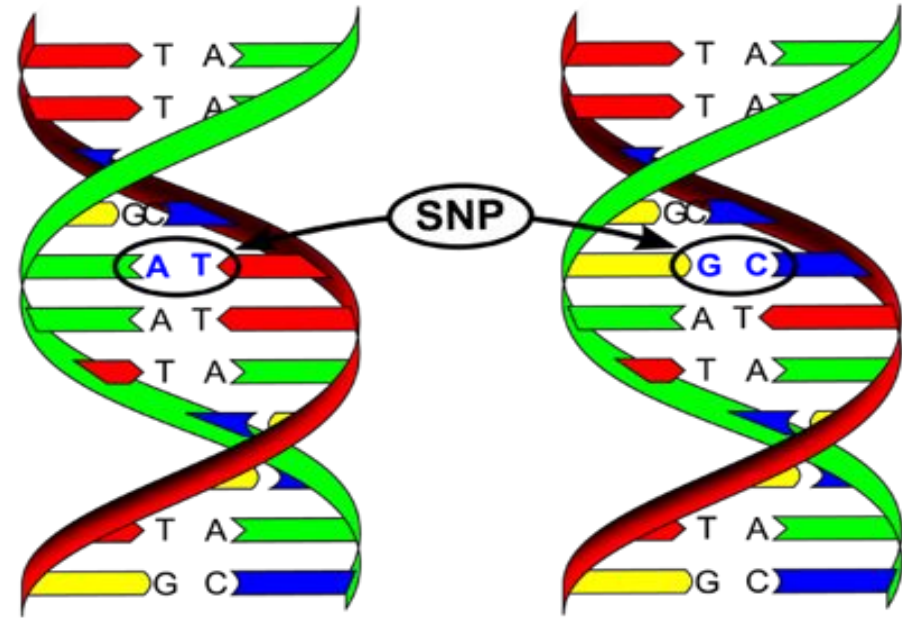
Why doesn't the genome of these bees completely turn to G/C content?



Discussion

Discussion Question 4

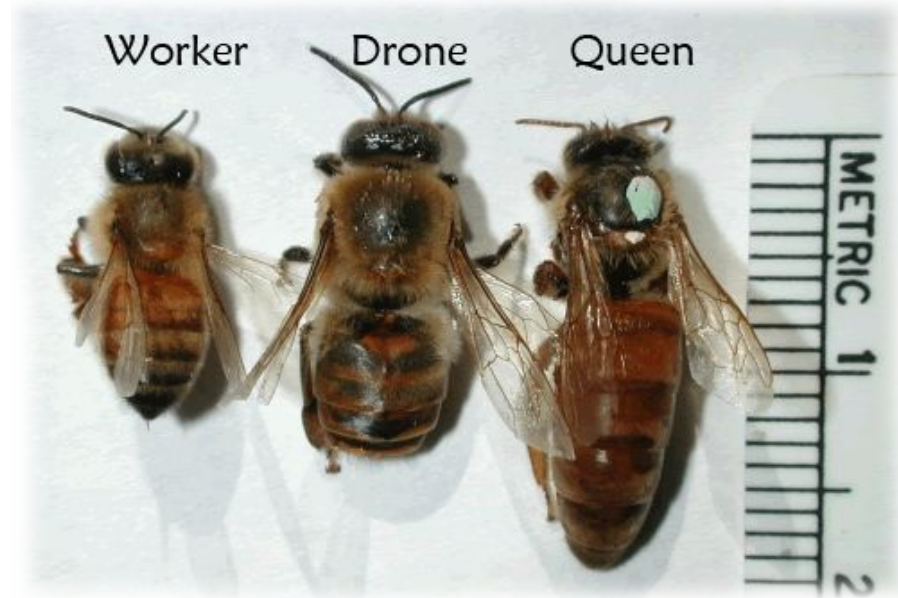
Why are polymorphisms important in this study?



Discussion

Discussion Question 5

Why do worker bees exhibit higher G/C content in their genomes compared to drones and queens?

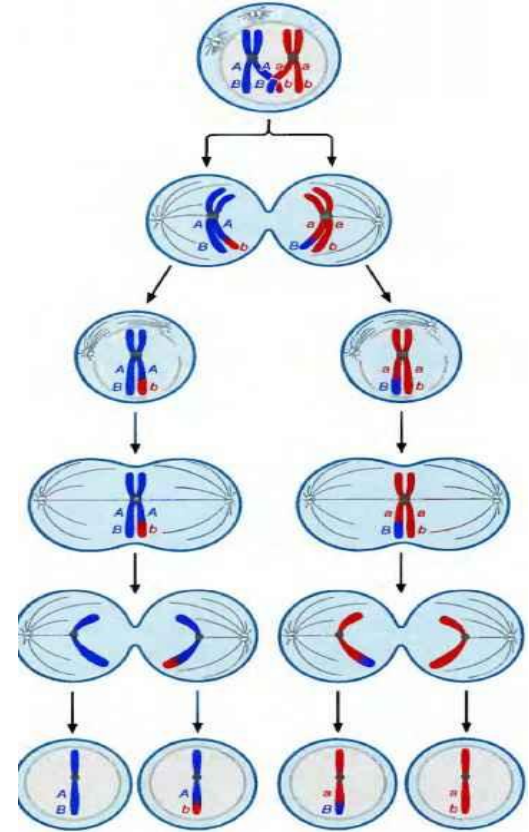


Discussion

Discussion Question 6

What are some different types of recombination?

Discussion



References

- <http://www.ub.edu/dnasp/>
- <https://www.thoughtco.com/synonymous-vs-nonsynonymous-mutations-1224600>
- Picture
https://earthjustice.org/blog/2015-april/11-amazing-reasons-to-save-the-honeybees?utm_source=crm&utm_content=grizzlytitle&utm_curation=ebrief
- <https://honeybeesuite.com/tag/pheromones/>
- <https://www.sciencedirect.com/science/article/pii/B0122270800005036>
- http://photobiology.info/Smith_DSB.html
- <https://www.nature.com/scitable/content/repair-of-dna-double-strand-breaks-by-41523>
- <http://www.bio.brandeis.edu/haberlab/jehsite/ssa.html>
- <http://www.bio.brandeis.edu/haberlab/jehsite/bir.html>
- <https://www.britannica.com/story/whats-the-difference-between-a-gene-and-an-allele>