The Biogeography of Speciation: new genomic insights about reinforcement using biological collections

Stephen A. Sefick

2016-12-02

Stephen A. Sefick The Monagaraphy of Specialism, new renorm 2016-12-02 1 / 48

Outline

- Introduction
- 2 Methods
- Significance
- Broader Impacts
- Questions

Topic

- Introduction
- 2 Methods
- Significance
- Broader Impacts
- Questions

Speciation

- Speciation
 - Speciation results in biodiversity

- Speciation
 - Speciation results in biodiversity
 - Biodiversity declining

- Speciation
 - Speciation results in biodiversity
 - Biodiversity declining
 - Important to study how biodiversity is produced

- Speciation
 - Speciation results in biodiversity
 - Biodiversity declining
 - Important to study how biodiversity is produced

Biogeography important to speciation process

- Speciation
 - Speciation results in biodiversity
 - Biodiversity declining
 - Important to study how biodiversity is produced
- Biogeography important to speciation process
- Genomics

- Speciation
 - Speciation results in biodiversity
 - Biodiversity declining
 - Important to study how biodiversity is produced
- Biogeography important to speciation process
- Genomics
 - since 2006 near exponential increase in genomic data

- Speciation
 - Speciation results in biodiversity
 - Biodiversity declining
 - Important to study how biodiversity is produced
- Biogeography important to speciation process
- Genomics
 - since 2006 near exponential increase in genomic data
 - Recently computationaly feasable to identify genetic variants

- Speciation
 - Speciation results in biodiversity
 - Biodiversity declining
 - Important to study how biodiversity is produced
- Biogeography important to speciation process
- Genomics
 - since 2006 near exponential increase in genomic data
 - Recently computationaly feasable to identify genetic variants
 - simple

- Speciation
 - Speciation results in biodiversity
 - Biodiversity declining
 - Important to study how biodiversity is produced
- Biogeography important to speciation process
- Genomics
 - since 2006 near exponential increase in genomic data
 - Recently computationaly feasable to identify genetic variants
 - simple
 - complex

- Speciation
 - Speciation results in biodiversity
 - Biodiversity declining
 - Important to study how biodiversity is produced
- Biogeography important to speciation process
- Genomics
 - since 2006 near exponential increase in genomic data
 - Recently computationaly feasable to identify genetic variants
 - simple
 - complex
- Transform our understanding of the relationship of biogeography with speciation

4 D > 4 B > 4 B > 4 B > 9 Q C

5 / 48

Stephen A. Sefick The Biogeography of Speciation: new genom 2016-12-02

Allopatric Speciation

- Allopatric Speciation
 - interupted species range (i.e., stream)

- Allopatric Speciation
 - interupted species range (i.e., stream)
 - decreased migration and gene flow

- Allopatric Speciation
 - interupted species range (i.e., stream)
 - · decreased migration and gene flow
 - 2 incipient species

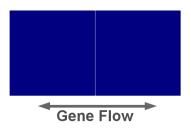
6 / 48

Stephen A. Sefick The Dispropriate of Speciation new renon 2016-12-02



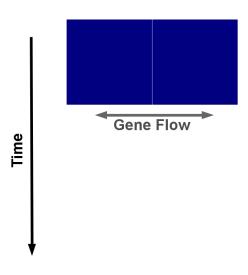
Stephen A. Sefick The Biogeography of Speciation: new

7 / 48

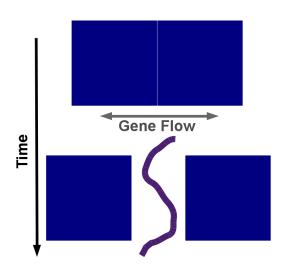


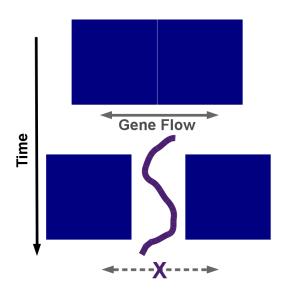
Stephen A. Sefick The Biogeography of Speciation: new genom

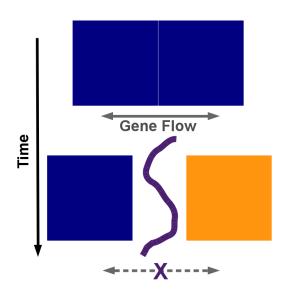
8 / 48











- 2 main types of speciation in a biogeographic context
 - Allopatric Speciation
 - interupted species range (i.e., stream)
 - decreased migration and gene flow
 - 2 incipient species

Sympatric Speciation

Stephen A. Sefick The Biogeography of Speciatio

- 2 main types of speciation in a biogeographic context
 - Allopatric Speciation
 - interupted species range (i.e., stream)
 - · decreased migration and gene flow
 - 2 incipient species

- Sympatric Speciation
 - no species range interuption

2016-12-02

13 / 48

Stephen A. Sefick The Biogeography of Speciation: new genom

- 2 main types of speciation in a biogeographic context
 - Allopatric Speciation
 - interupted species range (i.e., stream)
 - · decreased migration and gene flow
 - 2 incipient species

- Sympatric Speciation
 - no species range interuption
 - proceeds with gene flow

13 / 48

Stephen A. Sefick The Blogeography of Speciation, new genom 2016-12-02

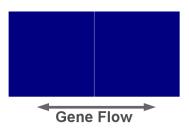
- 2 main types of speciation in a biogeographic context
 - Allopatric Speciation
 - interupted species range (i.e., stream)
 - · decreased migration and gene flow
 - 2 incipient species

- Sympatric Speciation
 - no species range interuption
 - proceeds with gene flow
 - 2 incipient species

Stephen A. Sefick The Biogeography of Speciation 1

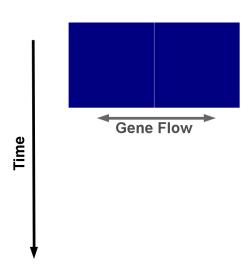


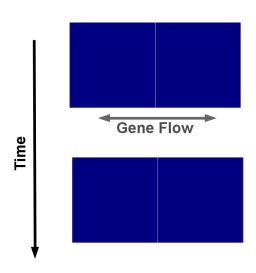
Stephen A. Sefick The Singergraphy of Specialism new paper 2016-12-02 14 / 48



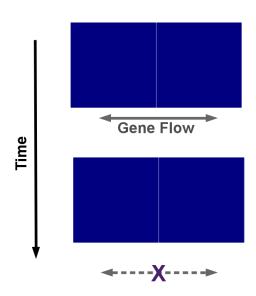
15 / 48

Stephen A. Sefick The Glogopy phy of Speciations (co. 1909) 2016-12-02



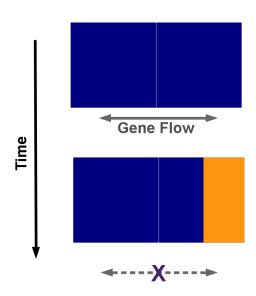






18 / 48

Stephen A. Sefick The Biogeography of Speciation, new genom 2016-12-02



- Allopatric Speciation
 - interupted species range (i.e., stream)
 - decreased migration and gene flow
 - · 2 incipient species
- Sympatric Speciation
 - no geographic interuption
 - proceeds with geneflow
 - 2 incipient species

2016-12-02

20 / 48

Stephen A. Sefick The Biogeography of Speciation: new genom

Types of speciation

- Allopatric Speciation
 - interupted species range (i.e., stream)
 - decreased migration and gene flow
 - 2 incipient species
- Sympatric Speciation
 - no geographic interuption
 - proceeds with geneflow
 - 2 incipient species
 - Reinforcement recently shown to be important

20 / 48

Stephen A. Sefick The Biogeography of Speciations new genom 2016-12-02

Stephen A. Sefick The Stephen purpose of Spectrum (1997) 1 2016-12-02 21 / 48

2 previously allopatric populations come into contact



22 / 48

Stephen A. Sefick To Siegos reply of Speciation (1997) 2016-12-02

- 2 previously allopatric populations come into contact
- Speciation process not complete



22 / 48

Stephen A. Sefick To Siegos reply of Speciation (1997) 2016-12-02

- 2 previously allopatric populations come into contact
- Speciation process not complete
- There is selection against hybrids

22 / 48

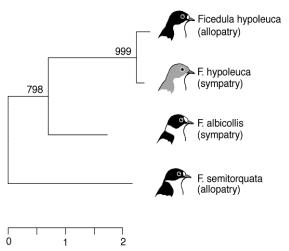
Stephen A. Sefick The Close graphy of Speciation, new genom 2016-12-02

- 2 previously allopatric populations come into contact
- Speciation process not complete
- There is selection against hybrids
 - Reproductive character displacement

22 / 48

Stephen A. Sefick The Glogago phy of Speciations new genom 2016-12-02

Flycatchers



Stre et al. 1997

Stephen A. Sefick The Sugarguphy of Spacetime new genon 2016-12-02 23 / 48

- 2 previously allopatric populations come into contact
- Speciation process not complete
- There is selection against hybrids
 - Reproductive character displacement

24 / 48

Stephen A. Sefick The Biogeography of Spesiations new perion 2016-12-02

- 2 previously allopatric populations come into contact
- Speciation process not complete
- There is selection against hybrids
 - Reproductive character displacement
 - Ecological character displacement

24 / 48

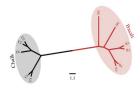
Stephen A. Sefick The Close graphy of Speciation, new genom 2016-12-02

Naked mole rat









Li et al. 2015

- 2 previously allopatric populations come into contact
- Speciation process not complete
- There is selection against hybrids
 - Reproductive character displacement
 - Ecological character displacement

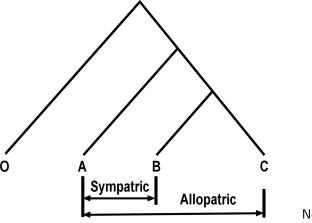
26 / 48

Stephen A. Sefick The Close graphy of Speciation, new genom 2016-12-02

- 2 previously allopatric populations come into contact
- Speciation process not complete
- There is selection against hybrids
 - Reproductive character displacement
 - Ecological character displacement
- 2 incipient species result

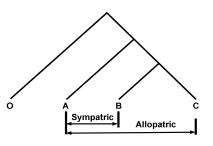
26 / 48

Stephen A. Sefick The Biogeography of Speciation, new genom 2016-12-02



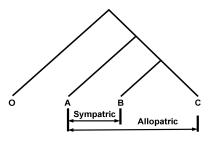
Noor 1997

 Species B more often different from A than C



Stephen A. Sefick The Stephengaphy of Specialism new yearon 2016-12-02

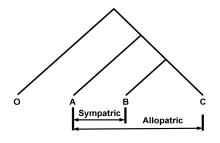
- Species B more often different from A than C
- Originally for reproductive isolation



28 / 48

Stephen A. Sefick The Signography of Speciation new renon 2016-12-02

- Species B more often different from A than C
- Originally for reproductive isolation
- Logic of test can be applied to genetic variants



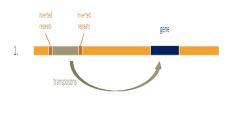
28 / 48

Stephen A. Sefick The Biogeography of Speciation, new genom 2016-12-02

• What exactly are genetic variants?

Stephen A. Sefick 12 Storger and 12

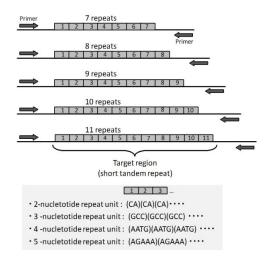
Genetic variants: Transposable elements





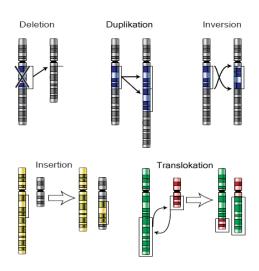
Stephen A. Sefick The Biogeography of Speciation: new genom

Genetic variants: Short tandem repeats



↓□▶ ⟨□▶ ⟨□▶ ⟨□▶ ⟨□▶ □ ♥)

Genetic variants: Structural Variants



Stephen A. Sefick

Genetic variants: Single nucleotide polymorphisms



33 / 48

Stephen A. Sefick The Biogeography of Speciation new genom 2016-12-02

Objectives and Hypotheses

 Objective: Use a diverse set of taxa with genomic data in online biological collections to investigate the relationship of genetic variants with biogeography.

34 / 48

Stephen A. Sefick The Biogeography of Speciation: new genom 2016-12-02

Objectives and Hypotheses

 Objective: Use a diverse set of taxa with genomic data in online biological collections to investigate the relationship of genetic variants with biogeography.

• H1: Complex genomic variants will show the pattern of reinforcement.

Stephen A. Sefick The Biogeography of Speciation new genom 2016-12-02

Objectives and Hypotheses

- Objective: Use a diverse set of taxa with genomic data in online biological collections to investigate the relationship of genetic variants with biogeography.
- H1: Complex genomic variants will show the pattern of reinforcement.
- H2: SNPs representative of reinforcement will be associated with functions indicative of ecological character displacement.

Stephen A. Sefick The Biogeography of Speciation: new genom 2016-12-02

Topic

- Introduction
- 2 Methods
- Significance
- 4 Broader Impacts
- Questions

Stephen A. Sefick

Conducted literature review

Stephen A. Sefick 10: 310 years play of Specialism payments 2016-12-02 37 / 48

- Conducted literature review
 - phylogeny

37 / 48

Stephen A. Sefick Dis Biographics of Specialism nav genom 2016-12-02

- Conducted literature review
 - phylogeny
 - biogeographic context

37 / 48

Stephen A. Sefick The Glogography of Speckylon new grown 2016-12-02

- Conducted literature review
 - phylogeny
 - biogeographic context
 - whole genome sequencing

37 / 48

Stephen A. Sefick The Biogeography of Speciation, new genom 2016-12-02

- Conducted literature review
 - phylogeny
 - biogeographic context
 - whole genome sequencing
- Publically avaliable data (e.g., NCBI)

37 / 48

Stephen A. Sefick The Elegeography of Speakulon new senson 2016-12-02

- Conducted literature review
 - phylogeny
 - biogeographic context
 - whole genome sequencing
- Publically avaliable data (e.g., NCBI)
- Appropriate phylogeny, biogeographic context, and whole genome sequencing

Organism	Acession Numbers
Mosquitoes (Anopheles)	NCBI: PRJNA6751 and PRJNA254046
Horses (Equus)	ENA: PRJEB7446
Butterflies (Heliconius)	ENA: ERP002440
Flycatchers (Ficula)	ENA: PRJEB7359
Dogs (Canis)	Authors Contacted
Cichlids	NCBI: PRJNA78915, PRJNA60369, PRJNA60363, and PRJNA78185

Download data (NCBI)

39 / 48

Stephen A. Sefick The Biogeography of Speciation new genom 2016-12-02

- Download data (NCBI)
- Use HPCs



- Download data (NCBI)
- Use HPCs
- GATK pipline (variant/variant filtration)

Stephen A. Sefick The Signed group of Signed group my segret 2016-12-02 39 / 48

- Download data (NCBI)
- Use HPCs
- GATK pipline (variant/variant filtration)
 - SNPs

39 / 48

Stephen A. Sefick The Siegenments of Specialism new system 2016-12-02

- Download data (NCBI)
- Use HPCs
- GATK pipline (variant/variant filtration)
 - SNPs
- Computationionally identify complex genetic variants

39 / 48

- Download data (NCBI)
- Use HPCs
- GATK pipline (variant/variant filtration)
 - SNPs
- Computationionally identify complex genetic variants
 - TEs

Stephen A. Sefick The Biogeography of Speciation: new genon

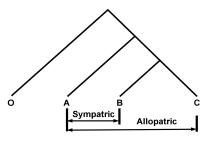
39 / 48

- Download data (NCBI)
- Use HPCs
- GATK pipline (variant/variant filtration)
 - SNPs
- Computationionally identify complex genetic variants
 - TEs
 - STRs

Stephen A. Sefick

- Download data (NCBI)
- Use HPCs
- GATK pipline (variant/variant filtration)
 - SNPs
- Computationionally identify complex genetic variants
 - TEs
 - STRs
 - SVs

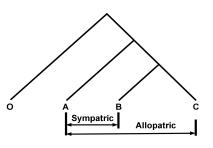
Restrictions:



40 / 48

Stephen A. Sefick The Chapen uply of Spectations new general 2016-12-02

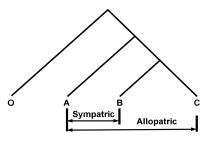
- Restrictions:
 - C Allopatric to all other closely related species



40 / 48

Stephen A. Sefick The Stephengaphy of Spectations new reason 2016-12-02

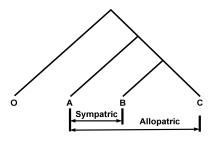
- Restrictions:
 - C Allopatric to all other closely related species
 - but B and C some overlap



40 / 48

Stephen A. Sefick The Siegeography of Specialism now groon 2016-12-02

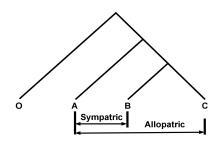
- Restrictions:
 - C Allopatric to all other closely related species
 - but B and C some overlap
 - effects of sympatry shared



40 / 48

Stephen A. Sefick The Stephen And Security of Specialism new group 2016-12-02

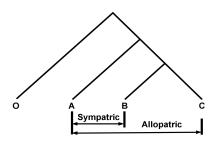
- Restrictions:
 - C Allopatric to all other closely related species
 - but B and C some overlap
 - effects of sympatry shared
- outgroup (O) to restrict search for genetic variants to those derived after most recent common ancestor



40 / 48

- Restrictions:
 - C Allopatric to all other closely related species
 - but B and C some overlap
 - effects of sympatry shared
- outgroup (O) to restrict search for genetic variants to those derived after most recent common ancestor

$$D = \frac{\sum_{n=1}^{i} Sym_i - Allo_i}{\sum_{n=1}^{i} Sym_i + Allo_i}$$

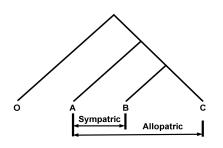


- Restrictions:
 - C Allopatric to all other closely related species
 - but B and C some overlap
 - effects of sympatry shared
- outgroup (O) to restrict search for genetic variants to those derived after most recent common ancestor

•

$$D = \frac{\sum_{n=1}^{i} Sym_i - Allo_i}{\sum_{n=1}^{i} Sym_i + Allo_i}$$

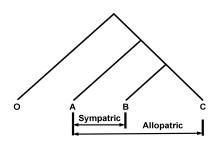
• 1 > D > -1



- Restrictions:
 - C Allopatric to all other closely related species
 - but B and C some overlap
 - effects of sympatry shared
- outgroup (O) to restrict search for genetic variants to those derived after most recent common ancestor

$$D = \frac{\sum_{n=1}^{i} Sym_i - Allo_i}{\sum_{n=1}^{i} Sym_i + Allo_i}$$

- 1 > D > -1
 - Reinforcement D > 0



Stephen A. Sefick

• H1: Complex genomic variants will show the pattern of reinforcement.

• H1: Complex genomic variants will show the pattern of reinforcement.

• Genome wide D

Stephen A. Sefick 10.5 Biography of Spreading 1997 person 2016-12-02 41 / 48

• H1: Complex genomic variants will show the pattern of reinforcement.

- Genome wide D
- Novel permutation proceedure to assess significance

41 / 48

• H1: Complex genomic variants will show the pattern of reinforcement.

- Genome wide D
- Novel permutation proceedure to assess significance
- Predictions

41 / 48

- H1: Complex genomic variants will show the pattern of reinforcement.
- Genome wide D
- Novel permutation proceedure to assess significance
- Predictions
 - Reinforcement *D* > 0

41 / 48

• H2: SNPs representative of reinforcement will be associated with functions indicative of ecological character displacement.

42 / 48

• H2: SNPs representative of reinforcement will be associated with functions indicative of ecological character displacement.

D in sliding windows

42 / 48

- H2: SNPs representative of reinforcement will be associated with functions indicative of ecological character displacement.
- D in sliding windows
- Novel permutation proceedure to assess significance

42 / 48

- H2: SNPs representative of reinforcement will be associated with functions indicative of ecological character displacement.
- D in sliding windows
- Novel permutation proceedure to assess significance
- Use Gene Ontology to assign functions where D > 0

42 / 48

- H2: SNPs representative of reinforcement will be associated with functions indicative of ecological character displacement.
- D in sliding windows
- Novel permutation proceedure to assess significance
- Use Gene Ontology to assign functions where D > 0
 - Use GO enrichment analysis

42 / 48

- H2: SNPs representative of reinforcement will be associated with functions indicative of ecological character displacement.
- D in sliding windows
- Novel permutation proceedure to assess significance
- Use Gene Ontology to assign functions where D > 0
 - Use GO enrichment analysis
- Predictions

42 / 48

- H2: SNPs representative of reinforcement will be associated with functions indicative of ecological character displacement.
- D in sliding windows
- Novel permutation proceedure to assess significance
- Use Gene Ontology to assign functions where D > 0
 - Use GO enrichment analysis
- Predictions
 - Significantly enriched GO terms ecological character displacement

42 / 48

- H2: SNPs representative of reinforcement will be associated with functions indicative of ecological character displacement.
- D in sliding windows
- Novel permutation proceedure to assess significance
- Use Gene Ontology to assign functions where D > 0
 - Use GO enrichment analysis
- Predictions
 - Significantly enriched GO terms ecological character displacement
 - e.g., Habitat preferences

42 / 48

- H2: SNPs representative of reinforcement will be associated with functions indicative of ecological character displacement.
- D in sliding windows
- Novel permutation proceedure to assess significance
- Use Gene Ontology to assign functions where D > 0
 - Use GO enrichment analysis
- Predictions
 - Significantly enriched GO terms ecological character displacement
 - e.g., Habitat preferences
- Alternatives

Stephen A. Sefick The Biogeography of Speciation: new p

- H2: SNPs representative of reinforcement will be associated with functions indicative of ecological character displacement.
- D in sliding windows
- Novel permutation proceedure to assess significance
- Use Gene Ontology to assign functions where D > 0
 - Use GO enrichment analysis
- Predictions
 - Significantly enriched GO terms ecological character displacement
 - e.g., Habitat preferences
- Alternatives
 - Significantly enriched GO terms reproductive character displacement

4□ > 4□ > 4□ > 4□ > 3

- H2: SNPs representative of reinforcement will be associated with functions indicative of ecological character displacement.
- D in sliding windows
- Novel permutation proceedure to assess significance
- Use Gene Ontology to assign functions where D > 0
 - Use GO enrichment analysis
- Predictions
 - Significantly enriched GO terms ecological character displacement
 - e.g., Habitat preferences
- Alternatives
 - Significantly enriched GO terms reproductive character displacement
 - e.g., sperm development

42 / 48

Topic

- Introduction
- Methods
- Significance
- 4 Broader Impacts
- Questions

Stephen A. Sefick

• taxonomic breadth

44 / 48

• taxonomic breadth

• genetic variants (complex variants and SNPs)

- taxonomic breadth
- genetic variants (complex variants and SNPs)
- Potential to transform our understanding of speciation's relationship with biogeography

- taxonomic breadth
- genetic variants (complex variants and SNPs)
- Potential to transform our understanding of speciation's relationship with biogeography
- Reveal impotant basic insights into how biodiversity is produced through speciation

Topic

- Introduction
- 2 Methods
- Significance
- 4 Broader Impacts
- Questions

Stephen A. Sefick

7th/8th Graders

• Speciation module AUMNH Junior Curator Camp

Stephen A. Sefick The Characteristic Stephen A. Sefick 2016-12-02 46 / 48

7th/8th Graders

- Speciation module AUMNH Junior Curator Camp
- Computer Based Bioinformatics lesson



46 / 48

7th/8th Graders

- Speciation module AUMNH Junior Curator Camp
- Computer Based Bioinformatics lesson
- Program to connect online and traditional biological collections

46 / 48

Topic

- Introduction
- 2 Methods
- Significance
- 4 Broader Impacts
- Questions

Stephen A. Sefick

Questions



Stephen A. Sefick The Stephen Stephen A. Sefick 2016-12-02 48 / 48