Biomath Manuscripts

Mette Olufsen

Type of paper

- Pure math
- Applied math
- Biologists
- Engineering
- ….

The typical math paper

- Introduction
- Sections describing the mathematical innovation
- Sections illustrating and discussing mathematical method

Sciences

- Introduction (lots of references)
  - Formulate the problem/question
  - Why is it important
  - What is novel
  - What has been done previously (experimentally and mathematically) and why is your contribution important
  - What is included in this paper (outline) and what did you find [allow you to judge if you should continue reading]
- Methods (some references)
  - Data (include reference to approval of experimental protocols)
  - Model
  - Model analysis
  - Statistical methodology (could be part of model analysis)
- Results (no references)
  - Present results but do NOT judge the results
- Discussion (lots of references)
  - Discusses results and relates your findings to previous studies, and to what has been highlighted in the introduction
  - Limitations
- Conclusions
- Acknowledgements
- Appendices/Supplements
  - Data
  - Simulation results with all data
  - Code

Writing guides

- Applied Math
  - video: https://www.youtube.com/watch?v=oNgqQyF0GfY
  - and the slides: https://www.slideshare.net/masonporter/paper-writing-in-applied-mathematics-slightly-updated-slides
- Jay Humphrey (Texas A&M)
  - https://www.morganclaypool.com/doi/abs/10.2200/S00128ED1V01Y200809ENG009
- Biomedical Engineering
  - http://www.biomedicaleditor.com/active-voice.html

Authorship

- Math/Applied Math (few authors)
  - Alphabetical

- Sciences (many authors)
  - First author (the person who wrote the paper and did most of the experiments, often student or postdoc)
  - Last author (the PI)
  - Beginning – junior people, list the ones contributing the most first
  - End – senior people

  - Joint fist authorship
  - Joint last authorship

- People whose data you use may want to be co-authors

- Make sure authorship has been clarified before the paper is submitted

How to write a paper

- Make figures illustrating the story
- Write introduction
  - What is the problem to be addressed here and why is it important
- Write the results section
- Write the method section
- Write the introduction
- Write the conclusion

- Re-iterate until a cohesive story is told

- Simple but “complete” make sure that each manuscript focus on ONE story

Figure captions

- Mathematics
  - Simple

- Sciences
  - Comprehensive – you should be able to read the figure caption and understand what it illustrates WITHOUT reading the text

Figure captions (examples)

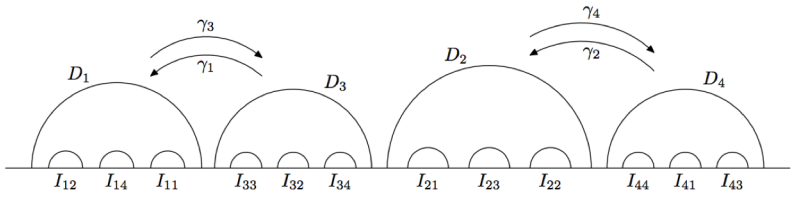


FIGURE 2. A Schottky structure with $r = 2$.

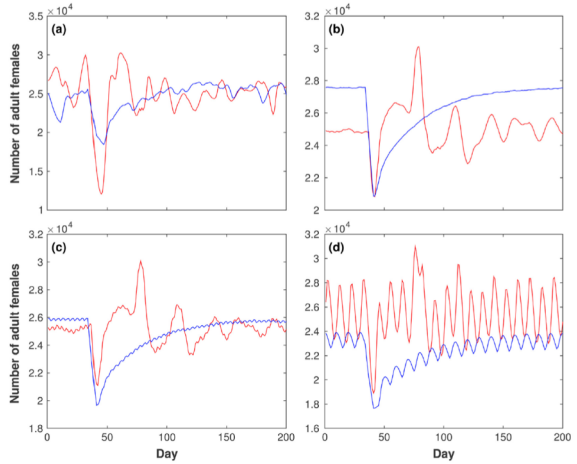


Fig. 3. Effect of a 1-week drop in temperature on the number of adult females. A sudden drop in temperature from 25.8°C (the annual average) to 20°C is simulated between days 35 and 41 inclusive, under various rainfall regimes: (a) actual daily rainfall, (b) rainfall every day, (c) rainfall every 4 d, and (d) rainfall every 10 d. The amount of daily rainfall is adjusted in the periodic regimes so that the annual total is the same across all four regimes. Plotted curves correspond to the time course of the number of adult females according to the Skeeter Buster model (red curves) and the AedesBA model (blue curves). All results shown are averages over 20 simulation runs.

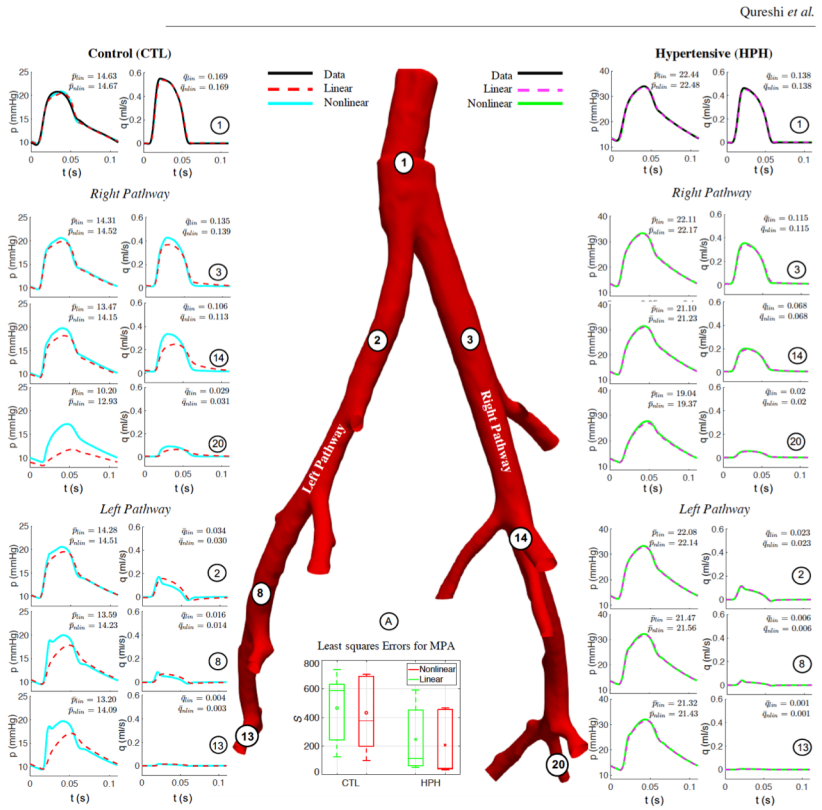


Fig. 4 Pressure and flow predictions along the pulmonary arterial network using linear (dashed line '- -') and non-linear (solid lines '–') wall models. Results are shown for a representative control (left) and hypertensive (right) mouse. The center panel (a) show the least squares error avenged across CTL and HPH groups.

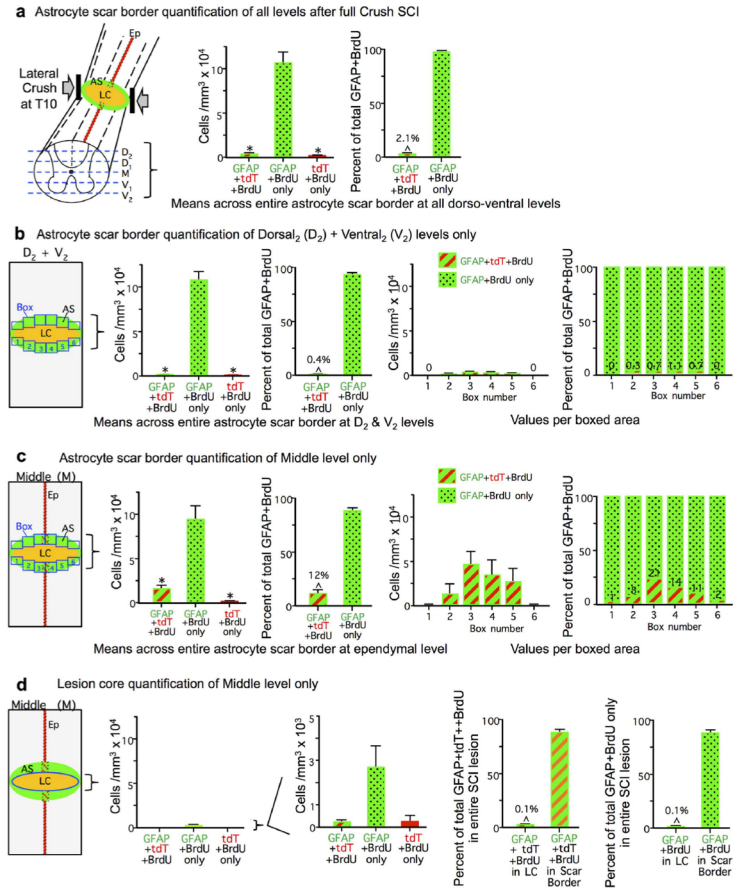


Figure 4. Quantification of BrdU positive cells co-labeled with either GFAP or Foxj1-tdT, or both in different regions of the lesion at 2 weeks after SCI. (**a**) Schematic of 5 dorso-ventral levels quantified and graphs showing the total number of cells per tissue volume and mean percent of BrdU labeled cells that are co-labeled with GFAP or Foxj1-tdT or both across the entire astrocyte scar (AS). Total values were derived by averaging counts from 6 boxes across the entire transverse spinal cord at each of the 5 levels as shown in (**b**,**c**). (**b**) Schematic of 6 counting boxes evaluated across the transverse spinal cord at dorsal and ventral levels D_2 and V_2. Graphs show the mean number of cells per volume and mean percent of BrdU labeled cells that are co-labeled with GFAP or Foxj1-tdT or both across the entire transverse plane, as well number or percent of cells per box. (**c**) Schematic of 6 counting boxes evaluated across the transverse spinal cord at the middle (M) level of the ependymal layer (Ep). Graphs show the mean number of cells per volume and mean percent of BrdU labeled cells that are co-labeled with GFAP or Foxj1-tdT or both across the entire transverse plane, as well number or percent of cells per box. (**d**) Schematic of lesion core (LC) at the middle (M) level containing the ependymal layer (Ep). Graphs show the total number per volume of BrdU labeled cells that are co-labeled with GFAP or Foxj1-tdT or both across the entire transverse plane, as well the percent of such cells in the entire SCI lesion that are located in either the lesion core or astrocyte scar. n = 6 per group, *p < 0.001 versus GFAP + BrdU only (ANOVA with Newman-Keuls), ^p < 0.001 (t-test).

Tables

Table 5 Bounds for optimization.

Parameter	β	p_1	γ	(r_1, r_2, c_1)
Lower bound	β_0	$\beta_0/2\pi$	1	0.05
Upper bound	$2.5\beta_0$	$2\beta_0$	2π	2.5

Table 6 Optimized values for individual mice in each group. Parameters were computed back from optimized values of β, p_1, γ, r_1, r_2 and c_1. Last row present the P-value, computed using the *ttest* in the Matlab, for all parameters inferred using the linear and nonlinear wall models. A P-value< 0.05 indicates a significant change in parameter value due to HPH.

	Linear								**Nonlinear**							
Control	β	R_T	$C_T \times 10^{-3}$	$I_R \times 10^{-2}$	$C_p/C_T(\%)$	$\tau \times 10^{-2}$	a (%)	S	p_1	γ	R_T	$C_T \times 10^{-3}$	$I_R \times 10^{-2}$	$C_p/C_T(\%)$	$\tau \times 10^{-2}$	a (%)	S
1	67.5	90.7	3.09	29.6	79.1	28.0	2.9	199	28.6	2.6	94.1	2.45	31.7	64.2	23.0	8.6	147
2	61.9	28.0	3.64	30.4	84.6	10.2	4.5	640	41.0	2.2	34.7	2.61	31.8	88.0	9.10	12.4	654
3	36.1	22.4	3.32	40.3	69.8	7.40	8.4	587	15.3	2.5	34.4	2.61	33.2	72.2	8.20	6.0	376
4	56.3	46.9	3.79	29.8	82.9	17.8	3.0	369	23.9	3.0	58.9	2.22	33.8	82.6	13.1	11.0	348
5	49.9	52.3	4.10	31.7	81.3	21.4	4.1	606	19.8	4.7	65.1	2.94	33.4	73.0	19.2	13.5	702
6	52.1	65.6	2.86	36.0	74.5	18.8	5.4	125	22.1	4.0	78.7	2.03	38.8	67.4	16.0	12.3	99
7	70.4	102.9	3.01	32.3	82.3	31.0	5.8	737	47.8	3.5	99.8	2.65	32.5	85.8	26.5	11.6	698
HPH	β	R_T	$C_T \times 10^{-3}$	$I_R \times 10^{-2}$	$C_p/C_T(\%)$	$\tau \times 10^{-2}$	a (%)	S	p_1	γ	R_T	$C_T \times 10^{-3}$	$I_R \times 10^{-2}$	$C_p/C_T(\%)$	$\tau \times 10^{-2}$	a (%)	S
1	151.3	146.2	0.67	56.5	62.8	9.80	15.4	49	102.7	1.6	147.5	0.57	58.4	81.9	8.40	12.9	32
2	86.3	94.3	1.72	36.3	73.2	16.2	10.2	593	36.6	2.6	100.9	1.52	39.3	76.3	15.3	13.0	450
3	121.0	155.6	1.23	33.3	73.5	19.1	7.0	407	82.2	2.1	158.7	1.02	35.7	84.6	16.2	6.9	467
4	117.4	142.2	1.16	42.9	71.0	16.5	14.4	63	49.8	2.2	146.4	0.94	44.3	75.8	13.7	13.4	41
5	108.2	130.9	1.11	48.0	66.6	14.5	15.0	113	73.5	2.3	133.5	0.85	49.2	77.5	11.4	13.0	42
P-value $\approx$	0.0001	0.0009	0.000	0.02	0.008	0.35*	0.0008	0.15*	0.004	0.035	0.0007	0.000	0.007	0.51*	0.33*	0.52*	0.23*

β (mmHg), R_T (mmHg s/ml), C_T (ml/mmHg), I_R, $C_p/C_T(\%)$,a (dimensionless), p_1 (mmHg), γ (dimensionless), S: least square error.
*Differences in these parameter values due to HPH is statistically significant.

Equations

- Methods section
- Appendix
- Figure

A more complex model was used to create simulated populations of septic patients and to simulate a clinical trial of anti-TNF therapy (15). In this paper, we report a further enhancement to our mathematical model of acute inflammation and shock, based on some known underlying cellular and molecular mechanisms of inflammation. Our model represents the simultaneous interactions of several different pathways. Specifically, we show that it can account for the temporal changes in the concentrations of three selected cytokines and nitric oxide by-products in mice, for disparate initial insults involving bacterial lipopolysaccharide (LPS, endotoxin), surgical trauma, and hemorrhage.

MATERIALS AND METHODS

Experimental procedures

Mice—All animal experiments were approved by the Institutional Animal Care and Use Committee of the University of Pittsburgh. The experiments were performed in adherence to the National Institutes of Health Guidelines on the Use of Laboratory Animals. All studies were carried out in C57Bl/6 mice (6–10 weeks old; Charles River Laboratories, Charles River, ME).

Endotoxemia protocol—Mice received either LPS (from *E. coli* O111:B4, 3, 6, or 12 mg/kg intraperitoneally; Sigma Chemical Co., St. Louis, MO) or saline control. At various time points following this injection, the mice (four to eight separate mice per time point) were euthanized, and their sera were obtained for measurement of various analytes (see below). All of the mice survived this high dose of LPS until the final time point (24 h following injection of LPS).

Surgical trauma and hemorrhagic shock protocols—For surgical trauma and hemorrhagic shock treatment, mice were anesthetized, and both femoral arteries were surgically prepared and cannulated. For hemorrhagic shock, the mice were then subjected to withdrawal of blood with a mean arterial pressure (MAP) maintained at 25 mmHg for 2.5 h with continuous monitoring of blood pressure as described previously (16). The normal MAP in mice is approximately 100 mmHg. In the resuscitated hemorrhage groups, the mice were resuscitated over 10 min with their remaining shed blood plus two times the maximal shed blood amount in lactated Ringer solution via the arterial catheter. For trauma, only the surgical preparation was conducted. In some cases, LPS was administered intraperitoneally to mice undergoing hemorrhagic shock. Animals were euthanized by exsanguination, and their serum analyzed as described below.

Analysis of cytokines and NO_2^-/NO_3^-—The following cytokines were measured using commercially available ELISA kits (R&D Systems, Minneapolis, MN): TNF, IL-10, and IL-6. Nitric oxide was measured as NO_2^-/NO_3^- by the nitrate reductase method using a commercially available kit (Cayman Chemical, Ann Arbor, MI) (17). Aspartate aminotransferase (AST) was measured using a commercially available kit (Vitros Chemistry™; Ortho-Clinical Diagnostics, Raritan, NJ) according to manufacturer's instructions.

Mathematical model of acute inflammation

We constructed a mathematical model of acute inflammation that incorporates key cellular and molecular components of the acute inflammatory response (see Results, Table 1, and Appendix). The mathematical model consists of a system of 15 ordinary differential equations that describe the time course of these components. Included in the model equations are two systemic variables that represent mean arterial blood pressure and global tissue dysfunction and damage. "Global tissue damage/dysfunction" describes the overall health of the organism because the hallmark of the pathology accompanying sepsis and hemorrhagic shock is the eventual, sequential failure of multiple organs. Given the complexity of simulating individual organs, we approximated this process by treating it as a gradual, ongoing process occurring in the whole body and driven by inflammation. Thus, unrecoverable tissue damage/dysfunction served as a surrogate for death, whereas damage/dysfunction that tended to return to baseline over a several-day period was a proxy for survival. In the model, pathogen-derived products, trauma, and hemorrhage are initiators of inflammation. (We note that hemorrhage is caused by injury that disrupts the integrity of blood vessels, and thus the two processes must be included in an accurate simulation).

Each equation was constructed from known interactions among model components as documented in the existing scientific literature. In deriving the mathematical model, we balanced biological realism with simplicity. Our goal was to find a fixed set of parameters that would qualitatively reproduce many known scenarios of inflammation found in the literature, correctly describe our data, and serve as a platform for eventually testing novel predictions experimentally.

The model and parameters were specified in three stages. In the preliminary stage, the model was constructed so it could reproduce qualitatively several different scenarios reported in the literature. In this stage, direct values of parameters such as cytokine half-lives were used when available. In the second stage, the model was matched to our experimental data by adjusting some of the parameters using our knowledge of the biological mechanisms together with the dynamics of the model to attain desired time course shapes. In the third stage, the parameters were optimized using a stochastic gradient descent algorithm that was implemented in software of Immunetrics, Inc. (Pittsburgh, PA).

The units of all the quantities in the model were specified so that for the experimental protocols tested they ranged from zero to one. This was done for computational convenience and has no implications for the underlying biology.

Appendix

process represented by this particular arrow. In addition, are two systemic variables representing blood pressure and tissue damage/dysfunction. The effect of trauma is expressed as an exponential decay of influence after an initial insult. It represents possible released cellular material that can trigger inflammation. The function fB(B) is phenomenological and represents the stimulatory effect of a decrease in blood pressure on triggering a stress response. This occurs in a variety of ways, and we are currently examining how to model it more exactly as part of ongoing projects in larger animals. Small deviations from normality probably are of little consequence, whereas larger deviations are proportionally more harmful (and rapidly so) as the ability of the organism to compensate decreases. This is what explains the phenomenological fourth-power exponent. For numerical ease, the variables of the equations are defined in abstract units of concentration. The actual units are restored with a linear scaling factor when compared to experimental data. The time unit is fixed at hours. The differential equations were solved numerically using the XPPAUT freeware written by Dr. G. B. Ermentrout (University of Pittsburgh, Department of Mathematics; *www.math.pitt.edu/~phase*) as well as proprietary software of Immunetrics, Inc. The fitting process included quantitative data, but the overall behavior of the model had also to be compatible with various qualitative scenarios extracted from the literature. These scenarios are listed in Table 2. Failure to comply with any of those qualitative behaviors resulted in discarding a given model.

Equations

$$M_A' = \left[\left(k_{MLPS}\frac{LPS(t)^2}{1+(LPS(t)/x_{MLPS})^2} + k_{MD}\frac{D^4}{x_{MD}^4+D^4}\right) \times \left(\frac{TNF^2}{x_{MTNF}^2+TNF^2} + k_{M6}\frac{IL6^2}{x_{M6}^2+IL6^2}\right) + k_{MTR}TR(t) + k_{MB}f_B(B)\right] \frac{1}{1+((IL10+CA)/x_{M10})^2} M_R - k_{MA}M_A$$

$$N_R' = -\left(k_{NLPS}\frac{LPS(t)}{1+LPS(t)/x_{NLPS}} + k_{NTNF}\frac{TNF}{1+TNF/x_{NTNF}} + k_{N6}\frac{IL6^2}{1+(IL6/x_{N6})^2} + k_{ND}\frac{D^2}{1+(D/x_{ND})^2} + k_{NB}f_B(B) + k_{NTR}TR(t)\right) \times \frac{1}{1+((IL10+CA)/x_{N10})^2} N_R - k_{NR}(N_R - S_N)$$

$$M_R' = -\left[\left(k_{MLPS}\frac{LPS(t)^2}{1+(LPS(t)/x_{MLPS})^2} + k_{MD}\frac{D^4}{x_{MD}^4+D^4}\right) \times \left(\frac{TNF^2}{x_{MTNF}^2+TNF^2} + k_{M6}\frac{IL6^2}{x_{M6}^2+IL6^2}\right) + k_{MTR}TR(t) + k_{MB}f_B(B)\right] \frac{1}{1+((IL10+CA)/x_{M10})^2} M_R - k_{MR}(M_R - S_M)$$

$$N_A' = \left(k_{NLPS}\frac{LPS(t)}{1+LPS(t)/x_{NLPS}} + k_{NTNF}\frac{TNF}{1+TNF/x_{NTNF}} + k_{N6}\frac{IL6^2}{1+(IL6/x_{N6})^2} + k_{ND}\frac{D^2}{1+(D/x_{ND})^2} + k_{NB}f_B(B) + k_{NTR}TR(t)\right) \times \frac{1}{1+((IL10+CA)/x_{N10})^2} N_R - k_{NA}N_A$$

$$iNOSd' = (k_{INOSN}N_A + k_{INOSM}M_A + k_{INOSEC} \left(\frac{TNF^2}{1+(TNF/x_{INOSTNF})^2} + k_{INOS6}\frac{IL6^2}{1+(IL6/x_{INOS6})^2}\right)) \times \frac{1}{1+(IL10/x_{INOS10})^2}\frac{1}{1+(NO/x_{INOSNO})^4} - k_{INOSd}iNOSd$$

$$iNOS' = k_{iNOS}(iNOSd - iNOS),$$

$$eNOS' = k_{ENOSEC}\frac{1}{1+TNF/x_{ENOSTNF}}\frac{1}{1+LPS(t)/x_{ENOSLPS}}\frac{1}{1+(TR(t)/x_{ENOSTR})^4} - k_{ENOS}eNOS$$

$$NO_3' = k_{NO3}(NO - NO_3)$$

$$TNF' = (k_{TNFN}N_A + k_{TNFN}M_A) \frac{1}{1+((IL10+CA)/x_{TNF10})^2}\frac{IL6}{1+(IL6/x_{TNF6})^3} - k_{TNF}TNF$$

$$IL6' = \left(k_{6N}N_A + M_A\right)\left(k_{6M} + k_{6TNF}\frac{TNF^2}{x_{6TNF}^2+TNF^2} + k_{6NO}\frac{NO^2}{x_{6NO}^2+NO^2}\right) \frac{1}{1+((CA+IL10)/x_{610})^2} + k_6(S_6 - IL6)$$

$$IL12' = k_{12M}M_A\frac{1}{1+(IL10/x_{1210})^2} - k_{12}IL12$$

$$CA' = k_{CATR}A(t) - k_{CA}CA$$

$$IL10' = (k_{10N}N_A + M_A(1 + k_{10A}A(t)) \left(k_{10MR} + k_{10TNF}\frac{TNF^4}{x_{10TNF}^4+TNF^4} + k_{106}\frac{IL6^4}{x_{106}^4+IL6^4}\right) \left((1-k_{10R})\frac{1}{1+(IL12/x_{1012})^4} + k_{10R}\right) - k_{10}(IL10 - S_{10})$$

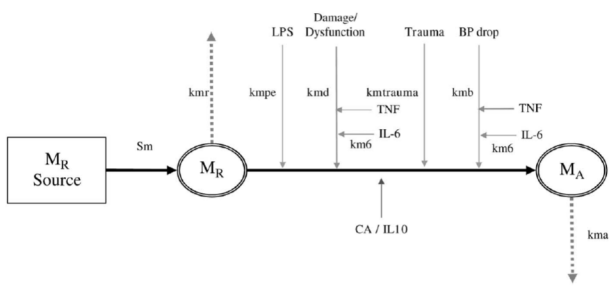
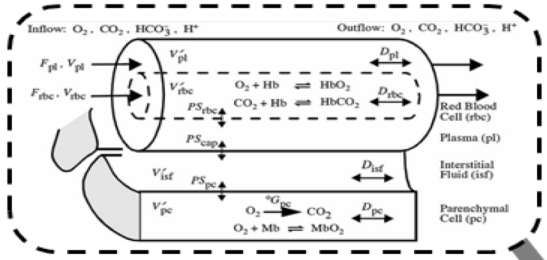
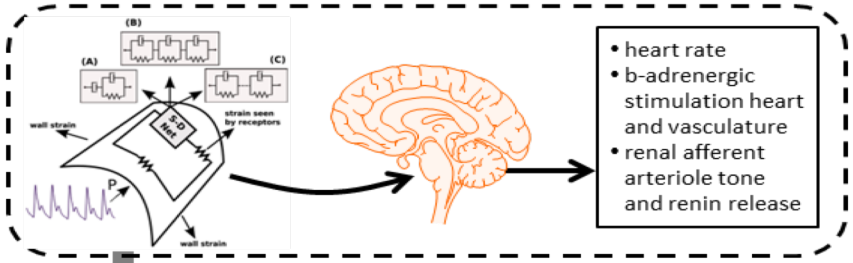


FIG. A1. **A simplified version of macrophage dynamics**. In the model used herein, resting macrophages (M_R) are activated by a number of physiologic processes, including endotoxin (PE), damage/dysfunction, trauma and hypotension (blood pressure [BP] drop). This recruitment process can be up-regulated (green lines) in the presence of tumor necrosis factor TNF and interleukin (IL)-6, whereas IL-10 and other antiinflammatory (CA) molecules down-regulate (red line) these activating influences. Both resting and activated macrophages (M_A) "die" at their respective rates (gray dotted line). Each process is supported by a literature search.

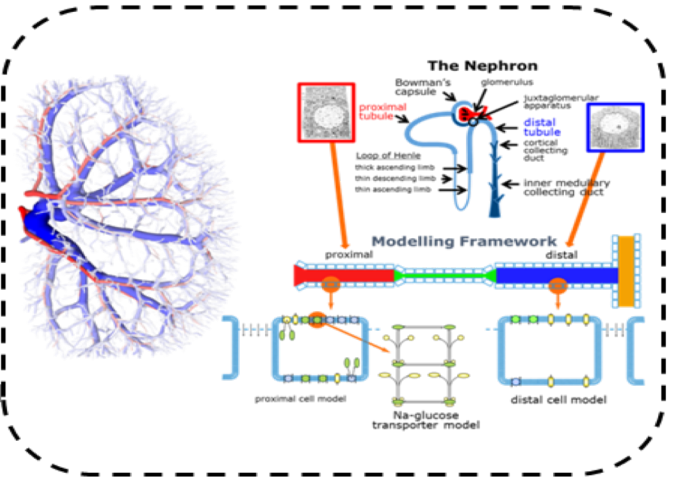
Lung gas exchange (**Project 3**)



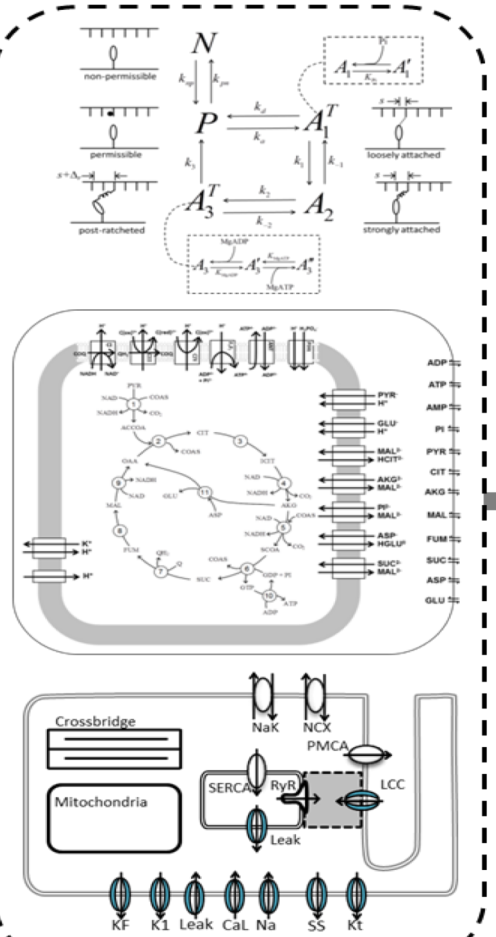
Strain sensing and autonomic reflexes (**Project 1**)

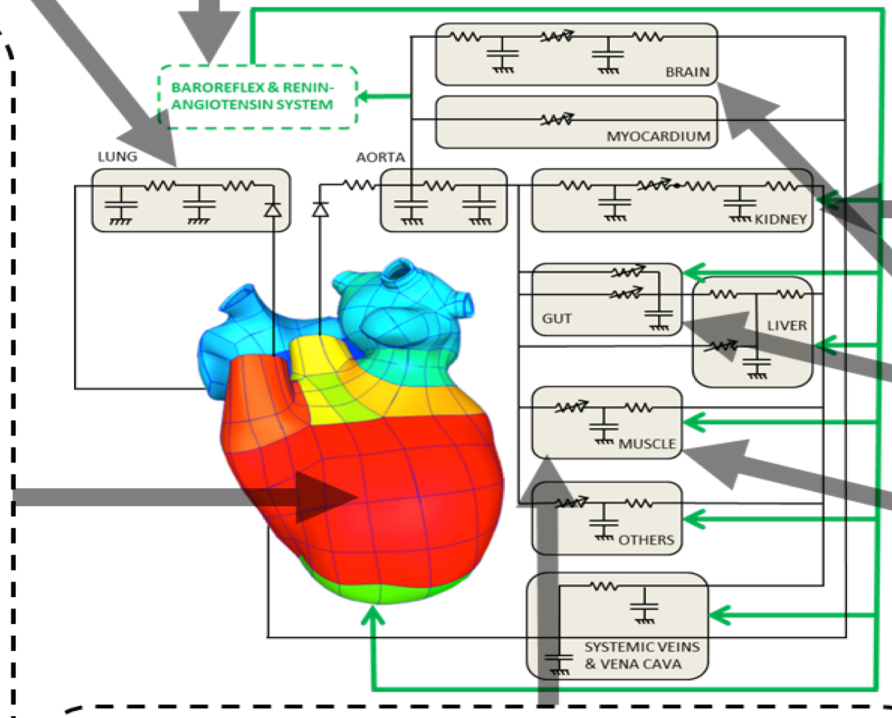


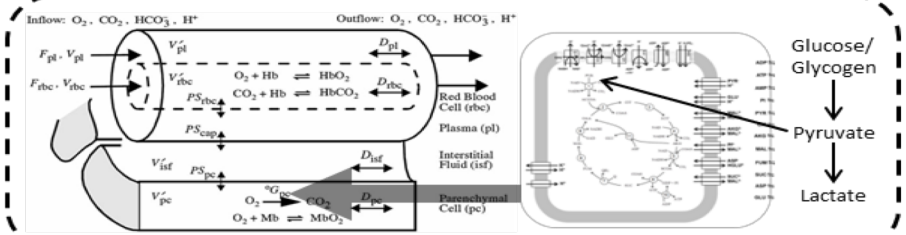
Renal blood flow , urine formation, and renin release (**Project 1**)



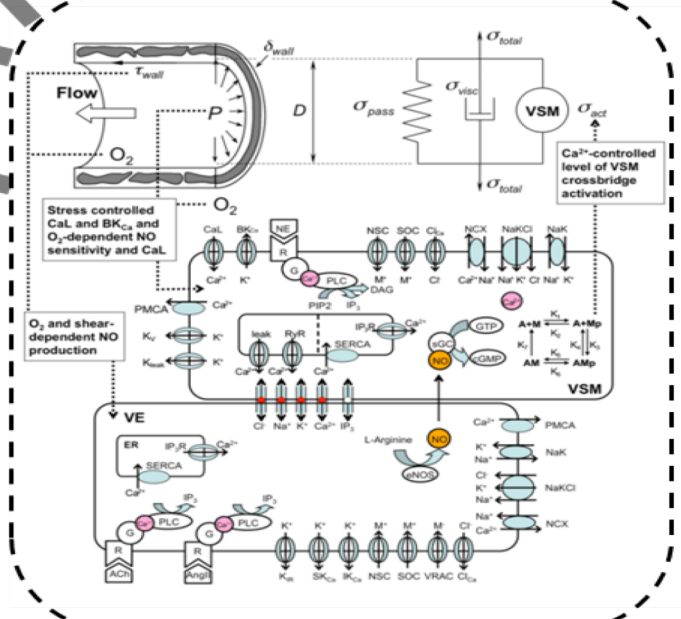
Myocardial metabolism , mechanics, and remodeling (**Projects 2 and 3**)







Muscle gas exchange and metabolism (**Project 3**)



Control of vascular resistance (**Project 1 and 3**)

Data and Code

- **Work supported by NIH and NSF, code and data should be made available at least upon request.**
- **Does not have to be distributed until work has been completed.**

Writing guide

By J. Humphrey, Texas A&M

Contents

2.2 REMOVING REDUNDANCIES AND UNNECESSARY WORDS

Now that we have a feel for an overall approach to writing well, let us begin to address specific aspects of "critical editing." Recall that effective technical writing is first and foremost clear and concise, which for obvious reasons is better written "Recall that effective technical writing is clear and concise." One way to ensure such characteristics in our writing is to *remove redundancies and unnecessary words*, sentence by sentence. Let us consider a few specific examples below (note: the original version is on the left and the corrected version is on the right, hence it is best to cover the right side first and consider how you might improve each example before looking at the suggested change):

The cells were cultured for a period of three weeks.	The cells were cultured for three weeks.
The temperature of the chamber remained between 35 and 39°C.	The chamber remained between 35 and 39°C.
The associated mechanisms are not known at this time.	The associated mechanisms are not known. (or, . . .remain unknown.)
The experiments were performed over a period of 10 hours.	The experiments were performed for 10 hours.
The new transducer is much smaller in size, which simplifies the design.	The new transducer is smaller, which simplifies the design.
The temperature increased at a rate of 3°/min.	The temperature increased at 3°/min.
The signal is lost below a threshold level of 10 Hz.	The signal is lost below a threshold of 10 Hz.

Concise wording

is used to develop	develops
is dependent on	depends on
We propose to use the combination of	We propose to combine
results in the simplification of	simplifies
It is interesting to note that	Note that (or, omit)
due to the fact that	because
in order to	to
in spite of the fact that	despite
as a result of	Omit
appears to be	seems
experienced a peak at	peaked at
in the event that	if
was found to be	was
a number of	many (or, various)
may be a mechanism responsible for	may cause
It is well-known that	Omit
for a long period of time	for a long period
is described in detail in	is detailed in
in the absence of	without
It is not uncommon that	It is common that
The finding is not inconsistent with	The finding is consistent with

Concise wording

Because the structure is assumed to remain circular, . . .	Assuming the structure remains circular, . . .
This will enable us to develop a better understanding of. . .	This will enable us to understand better. . .
This finding is the opposite of that reported by. . .	This finding is opposite that reported by. . .
The model is capable of describing. . .	The model can describe. . .
Table 1 is a list of all findings. . .	Table 1 lists all findings. . .
The next section is a brief description of the experimental methods.	The next section briefly describes the experimental methods.
The faculty advisor was the supervisor of both the undergraduate and the graduate students.	The faculty advisor supervised both the undergraduate and the graduate students.
The temperature readings will be dependent upon the contact stress. . .	The temperature readings will depend upon the contact stress. . .
Our laboratory technician also serves as the budget manager.	Our laboratory technician also manages the budgets.
The following example is an illustration of the basic concepts of. . .	The following example illustrates the basic concepts of. . .

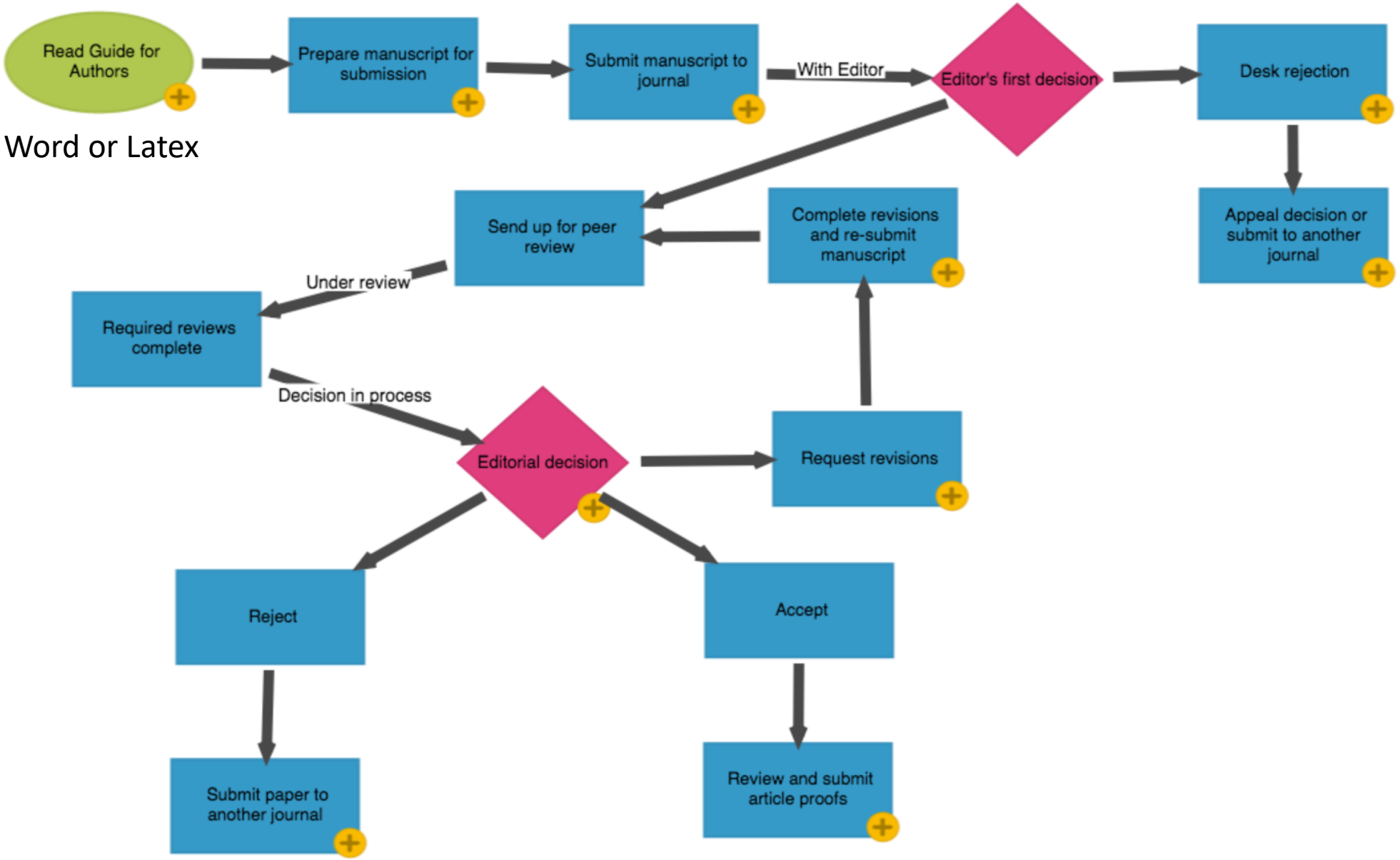
2.3 ACTIVE VOICE, FIRST PERSON, AND DIFFERENT TENSES

2.3.1 Voice

In *active voice*, the subject of the sentence performs the action indicated by the verb. Conversely, in *passive voice*, the subject of the sentence receives the action of the verb. The simple example below distinguishes between passive and active voice:

Passive voice: The data were analyzed by him using an ANOVA.[3]
Active voice: He analyzed the data using an ANOVA.

The specimen is connected to the device through a custom cannula.	The specimen connects to the device through a custom cannula.
The output signal is fed into a signal conditioner.	The output signal feeds into a signal conditioner.
In the next section, the underlying theory is given.	The next section gives the underlying theory.
In our current research, attention is directed to finding the mechanism.	Our current research directs attention to finding the mechanism.
The theory is dependent on five basic postulates.	The theory depends on five basic postulates.
X was used to create a surface-confined computational mesh.	X created a surface-confined computational mesh.
Increasing evidence has implicated the importance of. . .	Increasing evidence implicates the importance of. . .
Three different sectioning planes were used to form. . .	Three different sectioning planes formed. . .
Experimental noise is increased when unshielded cables are used.	Experimental noise increases with the use of unshielded cables.
A reader's attention is increased by the liberal use of figures and schematic drawings.	A reader's attention increases with the liberal use of figures and schematic drawings.



Word or Latex

Review Process

- Major revision (1-2)
- Rejection with resubmission
- Minor revision (typically only once)
- Accepted with minor revision

How to choose a journal

- What is the story you would like to tell?
- Introduction of a new model -> Applied Math/Biomath journal
  - Journal of Mathematical Biology
  - Journal of Theoretical Biology
  - Mathematical Biosciences
  - Siam Journal of Applied Math
- Use the model to study a specific question
  - Nature
  - Science
  - NEJM
  - PLoS Computational Biology
  - Am J Physiology
  - ….

- Who is the audience
- Where are similar manuscripts published
- Talk to the editors

A rejected pay may not be bad, but could have reached the wrong editor

Choose the handling editor carefully

Suggest reviewers

Example manuscripts

RESEARCH ARTICLE SUMMARY

HIV TRANSMISSION

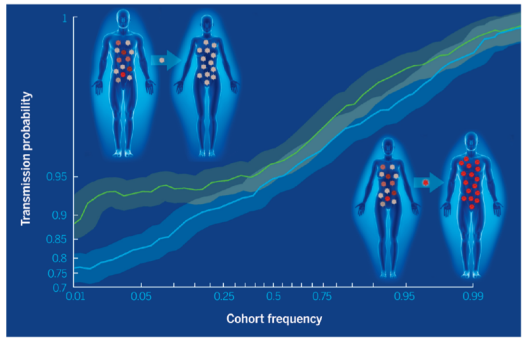
Selection bias at the heterosexual HIV-1 transmission bottleneck

Jonathan M. Carlson,* Malinda Schaefer, Daniela C. Monaco, Rebecca Batorsky, Daniel T. Claiborne, Jessica Prince, Martin J. Deymier, Zachary S. Ende, Nichole R. Klatt, Charles E. DeZiel, Tien-Ho Lin, Jian Peng, Aaron M. Seese, Roger Shapiro, John Frater, Thumbi Ndung'u, Jianming Tang, Paul Goepfert, Jill Gilmour, Matt A. Price, William Kilembe, David Heckerman, Philip J.R. Goulder, Todd M. Allen, Susan Allen, Eric Hunter*

INTRODUCTION: Heterosexual HIV-1 transmission is an inefficient process with rates reported at <1% per unprotected sexual exposure. When transmission occurs, systemic infection is typically established by a single genetic variant, taken from the swarm of genetically distinct viruses circulating in the donor. Whether that founder virus represents a chance event or was systematically favored is unclear. Our work has tested a central hypothesis that founder virus selection is biased toward certain genetic characteristics.

RATIONALE: If HIV-1 transmission involves selection for viruses with certain favorable characteristics, then such advantages should emerge as statistical biases when viewed across many viral loci in many transmitting partners. We therefore identified 137 Zambian heterosexual transmission pairs, for whom plasma samples were available for both the donor and recipient partner soon after transmission, and compared the viral sequences obtained from each partner to identify features that predicted whether the majority amino acid observed at any particular position in the donor was transmitted. We focused attention on two features: viral genetic characteristics that correlate with viral fitness and clinical factors that influence transmission. Statistical modeling indicates that the former will be favored for transmission, while the latter will nullify this relative advantage.

RESULTS: We observed a highly significant selection bias that favors the transmission of amino acids associated with increased fitness. These features included the frequency of the amino acid in the study cohort, the relative advantage of the amino acid with respect to the stability of the protein, and features related to immune escape and compensation. This selection bias was reduced in couples with high risk of transmission. In particular, significantly less selection bias was observed in men with genital inflammation and in women (regardless of inflammation status), compared to healthy men, suggesting a more permissive environment in the female than male genital tract. Consistent with this observation, viruses transmitted to women were characterized by lower predicted fitness than those in men. The presence of amino acids favored during transmission predicted which individual virus within a donor was transmitted to their partner, while chronically infected individuals with viral populations characterized by a predominance of these amino acids were more likely to transmit to their partners.



Fitter viruses (red) are favored more in woman-to-man (bottom curve) than in man-to-woman (top curve) transmission. The probability that a majority donor amino acid variant is transmitted is a function of relative fitness, here estimated by the frequency of the variant in the Zambian population. Even residues common in the population are less likely to be transmitted to healthy men than to women, indicative of higher selection bias in woman-to-man transmission.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.1254031

CONCLUSION: These data highlight the clear selection biases that benefit fitter viruses during transmission in the context of a stochastic process. That such biases exist, and are tempered by certain risk factors, suggests that transmission is frequently characterized by many abortive transmission events in which some target cells are nonproductively infected. Moreover, for efficient transmission, some changes that favored survival in the transmitting partner are frequently discarded, resulting in overall slower evolution of HIV-1 in the population. Paradoxically, by increasing the selection bias at the transmission bottleneck, reduction of susceptibility may increase the expected fitness of breakthrough viruses that establish infection and may therefore worsen the prognosis for the newly infected partner. Conversely, preventive or therapeutic approaches that weaken the virus may reduce overall transmission rates via a mechanism that is independent from the quantity of circulating virus, and may therefore provide long-term benefits to the recipient if transmission does occur. ■

RELATED ITEMS IN SCIENCE

S. B. Joseph, R. Swanstrom, A fitness bottleneck in HIV-1 transmission. *Science* **345**, 136–137 (2014).

The list of author affiliations is available in the full article online.
*Corresponding author. E-mail: carlson@microsoft.com (J.M.C.); ehunte4@emory.edu (E.H.)
Cite this article as J. M. Carlson *et al.*, *Science* **345**, 1254031 (2014). DOI: 10.1126/science.1254031

Estimating Allele Age and Selection Coefficient from Time-Serial Data

Anna-Sapfo Malaspinas,*·[1] Orestis Malaspinas,[†,‡] Steven N. Evans,[§] and Montgomery Slatkin**
*Centre for Geogenetics, Natural History Museum of Denmark, University of Copenhagen, 1350 Copenhagen, Denmark, [†]Institut Jean le Rond d'Alembert, Université Pierre et Marie Curie, F-75252 Paris cedex 5, France, [‡]Computational Science Department, Centre Universitaire d'Informatique, Université de Genève, 1227 Carouge, Switzerland, [§]Department of Statistics, University of California, Berkeley California 94720-3860, and **Department of Integrative Biology, University of California, Berkeley, California 94720-3140

ABSTRACT Recent advances in sequencing technologies have made available an ever-increasing amount of ancient genomic data. In particular, it is now possible to target specific single nucleotide polymorphisms in several samples at different time points. Such time-series data are also available in the context of experimental or viral evolution. Time-series data should allow for a more precise inference of population genetic parameters and to test hypotheses about the recent action of natural selection. In this manuscript, we develop a likelihood method to jointly estimate the selection coefficient and the age of an allele from time-serial data. Our method can be used for allele frequencies sampled from a single diallelic locus. The transition probabilities are calculated by approximating the standard diffusion equation of the Wright–Fisher model with a one-step process. We show that our method produces unbiased estimates. The accuracy of the method is tested via simulations. Finally, the utility of the method is illustrated with an application to several loci encoding coat color in horses, a pattern that has previously been linked with domestication. Importantly, given our ability to estimate the age of the allele, it is possible to gain traction on the important problem of distinguishing selection on new mutations from selection on standing variation. In this coat color example for instance, we estimate the age of this allele, which is found to predate domestication.

TIME-series analysis is widespread in several fields, such as meteorology, economics, and physics (Hamilton 1994) with the relation being statistical models designed to deal with a time-ordered sequence of observations. Such observations are also prevalent in several areas of biology. Until recently, however, time-series molecular data were only available for time spanning a few generations in higher organisms. Therefore, in the context of population genetics, time-serial data were mostly limited to viral or experimental evolution (Wichman *et al.* 2005; Bollback and Huelsenbeck 2007; Nelson and Holmes 2007; Gresham *et al.* 2008).

However, with recent advances in DNA sequencing and DNA preparation techniques, the study of extinct and long dead organisms is now entering a new era, an era in which time-sampled measurements spanning hundreds or thousands of generations for even mammalian species may be obtained. For example, while previous studies were limited to short segments of mitochondrial DNA, whole nuclear genomes are now available from several ancient samples (Rasmussen *et al.* 2010; Reich *et al.* 2010), and it is now additionally possible to target specific DNA regions in ancient organisms (Lalueza-Fox *et al.* 2007; Ludwig *et al.* 2009; Rusk 2009). Therefore, time-serial data will become increasingly available for a whole range of organisms allowing one to test evolutionary questions using not only present day samples, but also samples from extinct populations.

The relevant theory to describe such temporal changes in allele frequency has existed since the advent of population genetics (Fisher 1922; Wright 1931). Although not very common, several statistical methods and estimators to deal with time-serial data have been developed and applied to, for example, estimate historical changes in population size (Waples 1989; Williamson and Slatkin 1999; Anderson *et al.* 2000; Drummond and Rambaut 2007). More recently,

doi: 10.1534/genetics.112.140939
Manuscript received April 7, 2012; accepted for publication June 18, 2012
Supporting information is available online at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.112.140939/-/DC1.
[1]Corresponding author: Centre for Geogenetics, Natural History Museum of Denmark, Øster Voldgade 5-7, 1350 Copenhagen K, Denmark. E-mail: anna.sapfo.malaspinas@snm.ku.dk

Evolution, 54(1), 2000, pp. 64–71

EVOLUTION OF VIRULENCE IN A HETEROGENEOUS HOST POPULATION

ROLAND R. REGOES,[1,2] MARTIN A. NOWAK,[3,4] AND SEBASTIAN BONHOEFFER[1,5,6]
[1]*Experimental Ecology, Swiss Federal Institute of Technology Zürich, ETH Zentrum NW, CH-8092 Zürich, Switzerland*
[2]*E-mail: regoes@eco.umnw.ethz.ch*
[3]*Institute for Advanced Study, Princeton, New Jersey 08540*
[4]*E-mail: nowak@ias.edu*
[5]*Friedrich-Miescher-Institute, Maulbeerstrasse 66, CH-4058 Basel, Switzerland*
[6]*E-mail: bonhoeffer@eco.umnw.ethz.ch*

Abstract.—There is a large body of theoretical studies that investigate factors that affect the evolution of virulence, that is parasite-induced host mortality. In these studies the host population is assumed to be genetically homogeneous. However, many parasites have a broad range of host types they infect, and trade-offs between the parasite virulence in different host types may exist. The aim of this paper is to study the effect of host heterogeneity on the evolution of parasite virulence. By analyzing a simple model that describes the replication of different parasite strains in a population of two different host types, we determine the optimal level of virulence in both host types and find the conditions under which strains that specialize in one host type dominate the parasite population. Furthermore, we show that intrahost evolution of the parasite during an infection may lead to stable polymorphisms and could introduce evolutionary branching in the parasite population.

Key words.—Evolution of virulence, genetic trade-offs, host heterogeneity, host-parasite coevolution, intrahost evolution, serial passage.

Received December 31, 1998. Accepted July 16, 1999.

The relationship between a parasite and its host is a story of benefits and harms. The parasite benefits from the host by living in and on it and by using host resources to reproduce. The parasite's benefit gives rise to the host's harm: The life span of the host is usually shortened by infection, and important fitness traits of the host, such as fecundity, are often negatively affected by the parasite (Ewald 1994). However, by reducing the life span or the fitness of its host, the parasite may inflict harm upon itself. The question emerges: What evolutionary forces determine the level of harm inflicted on the host? In this paper, we focus on the evolution of parasite virulence, here defined as parasite-induced host mortality.

If virulence were an independent trait, natural selection should favor parasites with low virulence. A parasite that does not kill its host has more time to exploit it and to be transmitted, thus increasing its own fitness. Therefore, in the long run the parasite should evolve to be avirulent. Nevertheless, there are many examples of intermediately or highly virulent host-parasite systems with a long coevolutionary history (Fenner and Ratcliffe 1965; Herre 1993). To account for virulent host-parasite interactions, trade-offs between virulence and other parasite traits such as infectivity, transmissibility, or reproduction rate were postulated (Anderson and May 1979, 1981, 1982, 1991; May and Anderson 1979, 1983, 1990; Levin and Pimentel 1981; Bremermann and Pickering 1983; Ewald 1983; Knolle 1989; Frank 1992; Antia et al. 1994). Furthermore, it was argued that in systems in which the parasite can super- or coinfect hosts or in which the parasites frequently generate new mutant strains within a host, a parasite strain with high virulence may have a competitive advantage over less virulent strains within the host (Hamilton 1972; Bremermann and Pickering 1983; Knolle 1989; Sasaki and Iwasa 1991). In models that take into consideration this competition between different mutants within a single host in addition to the competition for transmission between hosts, one finds the persistence of parasite mutants that are highly virulent and in some cases too virulent to survive on their own (Bonhoeffer and Nowak 1994; Nowak and May 1994; May and Nowak 1995; van Baalen and Sabelis 1995).

The genetic or phenotypic composition of the host population also influences the dynamics of a host-parasite system (Jaenike 1978, 1996; Hamilton 1980; Lively 1987; Lipsitch et al. 1995; Ebert and Hamilton 1996). Many theoretical studies investigate the interaction of parasites with a heterogeneous host population, concentrating mainly on the stability of such systems (Cramer and May 1972; Murdoch and Oaten 1975; Roughgarden and Feldman 1975; Comins and Hassel 1976; Fujii 1977; Hassel 1979; Dobson 1990; Jones et al. 1994). Frank analyzed the coevolution of polymorphic host-pathogen systems using models in which parasite virulence and host resistance are closely linked (Frank 1991, 1994).

In the present paper, we focus on the evolution of parasite virulence in a genetically heterogeneous host population. We investigate a model that describes many parasite strains in a population of two different hosts. The parasite strains differ with regard to their reproduction rate and their virulences in the two hosts, and we assume trade-offs between the parasite virulences. To our knowledge, there are no studies investigating the evolution of virulence in a genetically heterogeneous host population as a consequence of such trade-offs.

In the following two sections, we introduce our model and give equilibrium solutions. Then we address two main questions: (1) What determines whether generalist or specialist strains evolve? and (2) How does host heterogeneity influence the evolution of parasite virulence? In the last part of the paper, we extend our model to include within-host evolution of the parasite. There are many studies reporting changes in parasite virulence as a consequence of within-host evolution during an infection (Bull 1994; Ewald 1994; Ebert and Hamilton 1996; Lipsitch and Moxon 1997; Ebert 1998, 1999). We incorporate these findings, to study their effect on the