IN THE SUPERIOR COURT
FOR THE DISTRICT OF COLUMBIA

Civil Division

MARK Z. JACOBSON, Ph.D.,	)	
Plaintiff,	)	
v.	)	**Civil Action No. 2017 CA 006685 B**
	)	**Hon. Elizabeth Carroll Wingo**
CHRISTOPHER T. M. CLACK, Ph.D.,	)	**Next Court Date: December 29, 2017**
	)	**Event: Initial Conference**
***et al.*,**	)	
Defendants.	)	

DEFENDANT NATIONAL ACADEMY OF SCIENCES' MEMORANDUM IN SUPPORT OF ITS SPECIAL MOTION TO DISMISS PURSUANT TO THE D.C. ANTI-SLAPP ACT OR, IN THE ALTERNATIVE, MOTION TO DISMISS PURSUANT TO RULE 12(b)(6)

To silence those who disagree with him, plaintiff has sued the National Academy of Sciences ("the Academy") for defamation, breach of contract, and promissory estoppel solely because the Academy, in the Proceedings of the National Academy of Sciences ("PNAS"), published a paper by Dr. Clack and other scientists who challenged the methodologies and assumptions that plaintiff used in a paper on climate change and the feasibility of replacing high-polluting energy sources, that was previously published in PNAS. This case falls squarely within the letter and spirit of the District of Columbia Anti-SLAPP statute,[1] and should be dismissed promptly.

The goal of the Anti-SLAPP statute is to ensure that defendants are not intimidated or prevented by abusive lawsuits from engaging in public policy debates. By demanding retraction and $10 million, plaintiff seeks to censor the Academy for providing a forum for robust

[1] D.C. Code §§ 16-5501 et seq. SLAPP stands for "strategic lawsuits against public participation."

scientific debate on one of the foremost issues of public concern today, and chill the critical exchange of ideas essential to scientific progress and the public interest.

The Anti-SLAPP statute was enacted to prevent litigation intended to suppress the exercise of constitutionally protected speech. It is evident that plaintiff's aim is to limit free expression by eliminating any challenges to his paper: plaintiff threatened a preliminary injunction against the Academy to prohibit publication of the Clack paper, Compl. Ex. 8; he threatened to send cease and desist letters to the authors of the Clack paper and the presidents of the universities that employed them, Compl. Ex. 9; he demanded that the Academy limit the length of the Clack paper, Compl. Ex. 16; and he requested that there be no response to the rebuttal that the Academy afforded to him (published on the same day as the paper it was rebutting).

Plaintiff cannot state a claim as to the three counts of his complaint, which amount to little more than an unvarnished attempt to muzzle speech and end-run the First Amendment. His defamation claim must fail because he is suing over quintessential scientific debate and opinion about methodology and assumptions – exactly what the courts have consistently said should not be decided by the judiciary, but rather left to robust scientific debate and rebuttal, as PNAS has done. This Court should not be thrust into deciding whose scientific opinion is better. Plaintiff's breach of contract and promissory estoppel claims must fail because he is attempting to use non-binding editorial guidelines as a sword to censor any challenges to his work in perpetuity.

After filing suit, and nearly two years after his paper was published, plaintiff submitted Errata that address omissions in his paper which relate to two of what he claims are the three "egregious" defamatory statements.

The National Academy of Sciences respectfully asks this Court to dismiss this action in its entirety and with prejudice, pursuant to both the Anti-SLAPP statute and Rule 12(b)(6).

BACKGROUND

The Academy is a private, non-profit organization of distinguished scholars, established by an Act of Congress signed by President Lincoln.[2] The Academy is charged with providing independent, objective advice to the nation on matters related to science and technology.[3] The Academy publishes PNAS.

PNAS is a widely cited, comprehensive multidisciplinary scientific journal.[4] One of the premier international journals publishing the results of original research, PNAS is published daily online (with the full text of all papers) and weekly in print.[5]

In December, 2015, PNAS published a paper entitled, "*Low-cost solution to the grid reliability problem with 100% penetration of intermittent wind, water and solar for all purposes*," the lead author of which is the plaintiff in this case. The paper was subjected to peer review prior to publication.[6] In 2016, plaintiff's paper received the Academy's Cozzarelli Prize for scientific excellence and originality. Compl. ¶ 9.

In 2016, Dr. Clack and his co-authors submitted to PNAS a paper that challenged some of the methodologies and assumptions in plaintiff's paper. Compl. ¶ 11. The Clack paper was also subjected to peer review. PNAS sent the Clack drafts to the plaintiff for comment. Compl. Exs. 6 and 9. Plaintiff sent comments on the Clack drafts back to PNAS.[7] PNAS forwarded the

---

[2] http://www.nationalacademies.org/brochure/index.html?_ga=2.146899760.64871978714972702.
[3] *Id.* 36 U.S.C. § 150303.
[4] http://www.pnas.org/site/aboutpnas/index.xhtml.
[5] *Id.*
[6] Peer review is a process whereby independent, knowledgeable experts review a paper before it is approved for publication. *See* http://www.pnas.org/site/misc/reviewprocess.pdf and Ex. H.
[7] Compl. Exs. 8, 9 & 16.

final draft of plaintiff's comments to the Clack authors.[8] They revised their paper based on information from plaintiff. *Compare* Compl. Ex. 7 and 11. The Academy also offered plaintiff an opportunity to publish a rebuttal to the Clack paper. Compl. Exs. 6 and 17.

Dissatisfied with the revisions, plaintiff accused the Clack authors of falsehoods. Compl. Ex. 9. He demanded that the Academy not publish the Clack paper, and at one point threatened to file a preliminary injunction to prevent publication. Compl. Ex. 8. He also argued that the Clack paper should be limited to a letter of 500 words, Compl. Ex. 16, even though PNAS has in the past published other comments as research papers. *See* Ex. B.

PNAS published both the Clack paper and plaintiff's rebuttal on the same day – June 19, 2017 – online.[9] Though the guidelines generally provide 500 words for a rebuttal, plaintiff asked for and was permitted far more words. As published, his rebuttal was 1,300 words long. Ex. A; Compl. ¶ 69. The rebuttal, which included a point-by-point response to the Clack paper including 26 footnotes, opened: "The premise and all error claims by Clack . . . are demonstrably false." Ex. A. Plaintiff told the Academy that no response to his rebuttal should be allowed. Indeed, plaintiff himself has acknowledged that "PNAS published our response to Clack equally and simultaneously, giving us the last words by not allowing Clack to respond to us."[10]

When the Academy refused plaintiff's demands for retraction of the Clack paper and for $10 million, plaintiff filed his Complaint, which states: "Dr. Jacobson has acknowledged that the Jacobson Article was not clear in the actual text . . . about the hydropower assumption and

---

[8] Compl. Ex. 10.

[9] http://www.pnas.org/content/114/26/E5021.full.

[10] https://www.ecowatch.com/national-review-mark-jacobson-2454398939.html.

that there was an omission of the cost of the additional hydropower turbines," but avers that neither was material. Compl. ¶72.

Twelve days after filing the Complaint, and 21 months after the publication of his paper, plaintiff submitted to PNAS (and later posted on his website[11]) Errata, which acknowledge omissions about the assumptions he made in his 2015 paper, including two that relate to what he claims are the three allegedly most "egregious" defamatory statements in Clack's paper. Ex. C.

ARGUMENT

I. Legal Standards For Dismissal

The Anti-SLAPP Act ensures prompt dismissal of lawsuits based upon "communicating views to members of the public in connection with an issue of public interest." D.C. Code § 16-5501(1)(B). Once defendant shows the alleged acts were in furtherance of the right of advocacy on issues of public interest, the Court must dismiss the complaint with prejudice unless plaintiff can demonstrate that he is likely or probable to succeed on the merits. *See* D.C. Code § 16-5502(b) and (d).

To survive a Rule 12(b)(6) motion, a complaint "must set forth sufficient information to outline the legal elements of a viable claim for relief." *Chamberlain v. Am. Honda Fin. Corp.*, 931 A.2d 1018, 1023 (D.C. 2007) (internal quotation omitted). The "[f]actual allegations must be enough to raise a right to relief above the speculative level . . . ." *Clampitt v. Am. Univ.*, 957 A.2d 23, 29 (D.C. 2008), quoting *Bell Atl. Corp. v. Twombly*, 550 U.S. 544, 555 (2007). A complaint must contain "more than labels and conclusions" or "a formulaic recitation of the elements of a cause of action." *Murray v. Motorola, Inc.*, 982 A.2d 764, 783 n.32 (D.C. 2009).

[11] https://web.stanford.edu/group/efmh/jacobson/Articles/I/CombiningRenew/combining.html.

II. **The Complaint Asserts Claims Arising From Protected Advocacy Rights**

The Anti-SLAPP Act[12] protects against claims arising from statements or other acts in furtherance of the right of advocacy on issues of public interest. Because this case involves an issue of public interest – renewable energy and climate change – plaintiff has the burden of demonstrating *at this stage of the litigation* that he is likely to succeed on the merits of his claims. Plaintiff cannot meet this burden, and accordingly, the Academy is entitled to the expeditious dismissal of plaintiff's claims, with prejudice.

The Act defines an issue of public interest as one "related to health or safety, environmental, economic, or community well-being." D.C. Code § 16-5501(3). Acts in furtherance of the right of advocacy on issues of public interest include "any written or oral statement made" in a "public forum in connection with an issue of public interest" or "[a]ny other expression or expressive conduct that involves . . . communicating views to members of the public in connection with an issue of public interest." D.C. Code § 16-5501(1)(A) and (B). Publication of the Clack paper unquestionably involved issues of public interest: climate change, reliance on fossil fuels, the use of nuclear power, and the feasibility of large-scale renewable energy projects are hotly debated within both the academic community and society generally. The Academy's publication of the Clack paper also constituted a written statement made "in a public forum . . . to members of the public." Both plaintiff's and the Clack authors' papers appeared in the online version of PNAS, a forum where anyone with internet access may view it. *See Farah v. Esquire Magazine*, 863 F. Supp. 2d 29, 38 (D.D.C. 2012) (an online blog post qualifies as a written or oral statement made in a place open to the public). Because the Clack

---

[12] D.C. Code § 16-5501 (2013) *et seq.* The analysis and outcome would be the same under California's Anti-SLAPP Act, Cal. Code Civ. Proc. § 425.16(a). This Court has found the D.C. Act to be "an almost identical act to the California act." *Mann v. Nat'l Review, Inc.*, No. 12-CA-8263 B, 2013 D.C. Super. LEXIS 7, at *10 (D.C. Super. Ct. July 19, 2013), *rev'd in part on other grounds*, 150 A.3d 1213, 1240 (D.C. 2016).

paper was made in a public forum in connection with an issue of public interest that relates to health, safety, environmental, economic, and community well-being, the Academy has satisfied its burden of establishing a *prima facie* case, and is therefore entitled to dismissal with prejudice.

Plaintiff thus bears the heavy burden of establishing that his claims are "likely" to succeed on the merits. *See* D.C. Code § 16-5502(b). The burden imposed is the same as the exacting standard applied to preliminary injunctions: "a substantial likelihood that he will prevail on the merits."[13] Because plaintiff cannot meet his heavy burden, all claims against the Academy should be dismissed with prejudice.

III. Plaintiff Cannot Show His Defamation Claim Is Likely To Succeed On The Merits

Plaintiff's defamation claim is based on statements that are not defamatory, and must therefore be dismissed under both the Anti-SLAPP Act and Rule 12(b)(6). The statements merely involved the expression of opinions in the type of scientific debate that should be resolved in the scientific arena, not in the courts. As the courts have recognized, "[s]cientific controversies must be settled by the methods of science rather than by the methods of litigation." *Underwager v. Salter,* 22 F.3d 730, 736 (7th Cir. 1994).[14] And "[t]he remedy for this kind of academic dispute is the publication of a rebuttal, not an award of damages." *Lott v. Levitt,* 556 F.3d 564, 570 (7th Cir. 2009).[15]

Four months before the Clack paper was published, the Academy notified plaintiff that Dr. Clack and his co-authors had submitted for publication a paper challenging the conclusions

---

[13] *Ctr. for Advanced Defense Studies v. Kaalbye*, No. 2014 CA 002273 B, 2015 WL 4477660 at *3 (D.C. Super. Ct. Apr. 7, 2015) (quoting *Zirkle v. District of Columbia*, 830 A.2d 1250, 1255 (D.C. 2003) (internal citations omitted).

[14] *See also Resolute Forest Products, Inc. v. Greenpeace International,* No. 17-cv-02824, 2017 WL 4618676, at *9 (N.D. Cal. Oct. 16, 2017) ("The academy, and not the courthouse, is the appropriate place to resolve scientific disagreements of this kind.").

[15] *See also Dillworth v. Dudley*, 75 F.3d 307, 310 (7th Cir. 1996) ("[J]udges are not well equipped to resolve academic controversies, . . . and scholars have their own remedies for unfair criticism of their work – the publication of a rebuttal.").

in plaintiff's paper. *See* Compl. ¶ 12; Compl. Ex. 6. The Academy offered plaintiff the opportunity to publish a response to the Clack paper, and to coordinate the publication of the Clack paper and plaintiff's response so readers could see both papers on the same day. *See* Compl. Ex. 6. In addition, the Academy provided plaintiff with three separate drafts of the Clack paper before it was published, so plaintiff could fully evaluate and address the Clack statements in his rebuttal. *See* Compl. ¶¶ 12, 16 & 18; Compl. Exs. 6, 17.

The Academy published plaintiff's rebuttal on-line the same day as the Clack paper. *See* Ex. A.[16] The rebuttal began:

> The premise and all error claims by Clack et al. (1) in PNAS, about Jacobson et al.'s (2) report, are <u>demonstrably false</u>. We reaffirm Jacobson et al.'s conclusions. [Emphasis added].

It addressed point-by-point what plaintiff considered to be inaccurate in the Clack paper, under the headings "False Premise," "False Error Claims," "Unsubstantiated Claims About Assumptions," and "False Model Claims," and concluded:

> In sum, Clack et al.'s analysis (1) is <u>riddled with errors</u> and has no impact on Jacobson et al.'s (2) conclusions. [Emphasis added].

By simultaneously giving readers both the Clack challenges to plaintiff's paper and plaintiff's rebuttal, the Academy served as an objective forum for scientific debate, enabling readers to evaluate the authors' respective opinions. That is how scientific disputes are supposed to be evaluated and resolved.

A. <u>Plaintiff Cannot Establish The Elements Of Defamation</u>

Plaintiff alleges that, by publishing the Clack paper, the Academy published false statements "to the District of Columbia, national and international scientific community audience that reads articles published in PNAS." Compl. ¶ 85. He also contends that the paper was

[16] Plaintiff's rebuttal is publicly available at www.pnas.org/cgi/doi/10.1073/pnas.1708069114.

published "to the much larger additional D.C., national and international press readership." Compl. ¶ 76.

Under District of Columbia defamation law, plaintiff must prove:

> (1) that the defendant made a false and defamatory statement concerning the plaintiff; (2) that the defendant published the statement without privilege to a third party; (3) that the defendant's fault in publishing the statement [met the requisite standard]; and (4) either that the statement was actionable as a matter of law irrespective of special harm or that its publication caused the plaintiff special harm.

Competitive Enterprise Institute v. Mann, 150 A.3d 1213, 1240 (D.C. 2016) (citations and internal quotations omitted) (emphasis added).[17] A publication is defamatory under District of Columbia law "if it tends to injure [the] plaintiff in his trade, profession or community standing, or lower him in the estimation of the community." *Competitive Enterprise Institute v. Mann*, 150 A.3d at 1241. To be actionable, the allegedly defamatory statement "must be more than unpleasant or offensive; the language must make the plaintiff appear odious, infamous, or ridiculous." *Id.* (citations and internal quotations omitted).[18]

The elements of a defamation claim are similar in California: libel is "a false and unprivileged publication by writing . . . which exposes any person to hatred, contempt, ridicule, or obloquy, . . . or which has a tendency to injure him in his occupation." Cal. Civ. Code § 45 (emphasis added). *See also Smith v. Maldonado,* 72 Cal. App. 4th 637, 645, 85 Cal. Rptr.2d 397, 402 (1999) (defamation "involves the intentional publication of a statement of fact that is false, unprivileged, and has a natural tendency to injure or which causes special damage.") (emphasis added).

---

[17] *See also Beeton v. District of Columbia*, 779 A.2d 918, 923 (D.C. 2011).
[18] *See also Rosen v. American Israel Public Affairs Committee, Inc.*, 41 A.3d 1250, 1256 (D.C. 2012); *Community for Creative Non-Violence v. Pierce*, 814 F.2d 663, 671 (D.C. Cir. 1987); *Xereas v. Heiss*, 933 F. Supp.2d 1, 17 (D.D.C. 2013); *Machie v. Nguyen*, 824 F. Supp.2d 146, 152 (D.D.C. 2011).

Plaintiff bears an additional and especially heavy burden of proof because he is a public figure in the ongoing scientific debate over the feasibility of replacing fossil fuels. He "cannot recover [on his defamation claim] unless he proves by clear and convincing evidence that the defendant published the [allegedly] defamatory statement with actual malice, *i.e.*, with 'knowledge that it was false or with reckless disregard of whether it was false or not.'" *Masson v. New Yorker Magazine, Inc.*, 501 U.S. 496, 510 (1991), quoting *New York Times Co. v. Sullivan,* 376 U.S. 254, 279-280 (1964).[19] This standard is "a daunting one." *McFarlane v. Sheridan Square Press, Inc.*, 91 F.3d 1501, 1515 (D.C. Cir. 1996).

Plaintiff cannot demonstrate that he is likely to satisfy these requirements.

B. Plaintiff Cannot Prove The Academy Published Defamatory Statements

The First Amendment "protects public debate on matters of public concern, including scientific matters." *McMillan v. Togus Regional Office, Dep't of Veteran Affairs*, 294 F. Supp. 2d 305, 316 (E.D.N.Y. 2003), *aff'd* 120 F. App'x 849 (2d Cir. 2005). In the scientific arena, matters of public concern depend upon vigorous and unfettered debate:

> As in political controversy, "science is, above all, an adversary process. It is an arena in which ideas do battle. . . ." Our technology and lives depend on modern science. Any unnecessary intervention by the courts in the complex debate and interplay among scientists that comprises modern science can only distort and confuse. . . . Scientists should not have to conduct their studies defensively, looking over their shoulders at unnecessary costly litigations.

Id. at 317 (internal citations omitted). Thus, "[a]s a matter of constitutional principle, when the issue is whether liability may be imposed for speech expressing scientific or policy views, the question is not who is right; the First Amendment protects the expression of all ideas, good and bad." *Competitive Enterprise Institute v. Mann*, 150 A.3d at 1242.

---

[19] *See also Competitive Enterprise Institute v. Mann*, 150 A.3d at 1241-42; *Resolute Forest Products, Inc. v. Greenpeace International*, No. 17-cv-02824, 2017 WL 4618676, at *10 (N.D. Cal. Oct. 16, 2017).

Statements of opinion can only provide the basis for a defamation claim if "'they imply a provably false fact, or rely upon stated facts that are provably false.'" *Id.* Thus, "under the First Amendment [an allegedly false] statement is not actionable 'if it is plain that a speaker is expressing a subjective view, an interpretation, a theory, conjecture, or surmise, rather than claiming to be in possession of objectively verifiable facts.'" *Id.* at 1241 (citations omitted).

Applying these principles in a case involving published comments about a scientist engaged in the debate over global warming, the Court of Appeals drew the line between statements that were actionable and those that were not:

> [T]he law distinguishes between statements expressing ideas and false statements of fact. <u>To the extent statements in [defendants'] articles take issue with the soundness of [plaintiff's] methodology and conclusions – i.e., with ideas in a scientific or political debate – they are protected by the First Amendment</u>. But defamatory statements that are personal attacks on an individual's honesty and integrity and assert or imply that [plaintiff] engaged in professional misconduct and deceit to manufacture the results he desired, if false, do not enjoy constitutional protection and may be actionable.

Id. at 1242 (emphasis added). Thus, the court found statements such as "wrongdoing," "deceptions," "data manipulation," "misconduct," and "the Jerry Sandusky of climate science" to be defamatory (*Id.* at 1243), but found not defamatory the statement that Dr. Mann's work was "fraudulent," taking that word to mean "intellectually bogus and wrong." *Id.* at 1249.

The statements challenged by plaintiff fall far short of the line drawn in *Mann*. They also fall far short of the terms plaintiff uses about others: for example, he told his 13,900 twitter followers that one commentator "fakes data" (Ex. D at p. 3), that another "intentionally falsified data" (*Id.* at p. 5), and that yet another's comment on his work was a "flat out lie" (*Id.* at p. 4). Indeed, plaintiff has told his twitter followers that the Clack paper is "intentionally scientifically fraudulent with falsified data." *Id.* at p. 10.

Plaintiff has identified what he characterizes as three "egregious" false statements contained in the Clack paper. Compl. ¶ 84. Ironically, two of the three supposedly "egregious" false statements relate to omissions that plaintiff identified in the Errata he submitted to the Academy and posted on his website after suit was filed and more than 21 months after his paper was published. He has also referred to Exhibit 12 to the complaint as a list of "additional numerous falsehoods and misstatements." *Id.* None of those alleged falsehoods or misstatements involves a provably false statement about plaintiff's honesty or integrity, or imply that plaintiff engaged in professional misconduct or deceit. Rather, they simply reflect opinions of the Clack authors about the methodology and conclusions that plaintiff described in his paper, which is protected speech that cannot form the basis for a defamation claim. Moreover, both the allegedly false statements in the Clack paper and plaintiff's explanations of why he believes the statements are false involve highly esoteric scientific issues that are ill-suited to resolution by a judge or jury -- precisely the type of issues that "must be settled by the methods of science rather than by the methods of litigation." *Underwager v. Salter*, 22 F.3d at 736.

1. <u>Alleged False Statement Regarding The Figures In Plaintiff's Table 1</u>

Plaintiff alleges that the following is one of the three supposedly "egregious" statements in the Clack paper that defamed him:

> [T]he flexible load used by LOADMATCH is more than double the maximum possible value from table 1 of [the Jacobson Article]. The maximum possible from table 1 of [the Jacobson Article] is given as 1,064.16 GW, whereas figure 3 of [the Jacobson Article] shows that flexible load (in green) used up to 1,944 GW (on day 912.6).

Compl. ¶¶ 42-43 and 84. Plaintiff contends this language is false and defamatory because it refers to the load figures in plaintiff's paper as "maximum" load figures, instead of "average" load figures. Compl. ¶¶ 42-49. He contends that he told the Academy the load figures were "average" figures before the Clack paper was published. Compl. ¶ 42.

Even assuming that the Table 1 figures are "average" figures, the statement was not defamatory. The statement did not "make the plaintiff appear odious, infamous, or ridiculous," did not constitute a personal attack on plaintiff's "honesty and integrity," and did not imply that he had engaged in "professional misconduct and deceit." *Competitive Enterprise Institute v. Mann*, 150 A.3d at 1241-42, At most, the Clack authors' statement involved an interpretation of a table in plaintiff's paper. This is not defamation.

Moreover, before the Clack paper was published, the Academy provided the Clack authors with plaintiff's written explanation of his position on the "maximum" vs. "average" issue, so they could determine whether they agreed with his position. Compl. Exs. 9 (Item 25) and 10. And, simultaneous with publishing the Clack paper, the Academy published plaintiff's rebuttal, which explained why plaintiff believed the numbers should be characterized as "average" instead of "maximum." Ex. A at p. 2. The Academy therefore provided its readers with both plaintiff's and the Clack authors' positions on how the data should be interpreted, which is how scientific disagreements are supposed to be addressed and resolved.

2. <u>Alleged False Statement Regarding Plaintiff's Hydropower Discharge Rates</u>

Plaintiff alleges that the underlined phrase in the following language from the Supporting Information attached to the Clack paper is the second "particularly egregious" statement that defamed him:

> The hydroelectric production profiles depicted throughout the dispatch figures reported in both the paper and its supplemental information routinely show hydroelectric output far exceeding the maximum installed capacity as well. . . . <u>This error is so substantial that we hope there is another explanation for the large amounts of hydropower depicted in these figures</u>.

Compl. Exhibit 11 at SI p. 2; Compl. ¶¶ 50, 54-55 (emphasis added). Plaintiff contends that the Clack paper made an "intentionally misleading" statement when it said "we hope there is another explanation" (Compl. ¶ 51), because plaintiff told Clack and the Academy that he had achieved the high discharge capacity estimates reflected in his paper by assuming additional turbines would be added to existing hydroelectric dams (Compl. ¶¶ 51, 52). There are at least four reasons why the Clack statement cannot support a defamation claim.

First, the Clack statement does not assert or imply any fact at all, much less a false fact that makes "plaintiff appear odious, infamous, or ridiculous." [20]

Second, based on information provided by plaintiff, the Clack authors revised their paper to state that plaintiff had assumed additional turbines would be added to the existing hydroelectric plants, and explained why they believed that was not a reasonable assumption to make. *See* Compl. Ex 11 at p. 6723 (plaintiff relied on "adding turbines" at "current reservoirs without consideration of hydrological constraints or the need for additional supporting infrastructure") & SI p. 2 ("[T]he authors assumed they could build capacity in hydroelectric plants for free . . . . [T]he extra piping needed to supply water to these turbines would cause considerable engineering issues . . . .").

Third, plaintiff acknowledges that his own paper did not disclose his "additional turbines" assumption, Compl. ¶ 52, and he recently asked the Academy to publish errata to correct that omission. Ex. C.

Finally, in the rebuttal published simultaneously with the Clack paper, plaintiff explained both the "additional turbines" assumption that had not been disclosed in his original paper, and why he believed the turbines could be added at reasonable cost. Ex. A at p. 1-2.

---

[20] *Competitive Enterprise Institute v. Mann,* 150 A.3d at 1241; *Rosen v. American Israel Public Affairs Committee, Inc.*, 41 A.3d at 1256; *Community for Creative Non-Violence v. Pierce*, 814 F.2d at 671; *Xereas v. Heiss*, 933 F. Supp.2d at 17; *Machie v. Nguyen*, 824 F. Supp.2d at 152.

Thus, the explanation for the large amounts of hydropower production projected in plaintiff's paper was provided to PNAS readers by both the Clack paper and by plaintiff's rebuttal.

3. Alleged False Statement In Figure 3 Of The Clack Paper

Plaintiff alleges that a drawing -- the following "Figure 3" -- from the Clack paper is the third "particularly egregious" statement that defamed him:

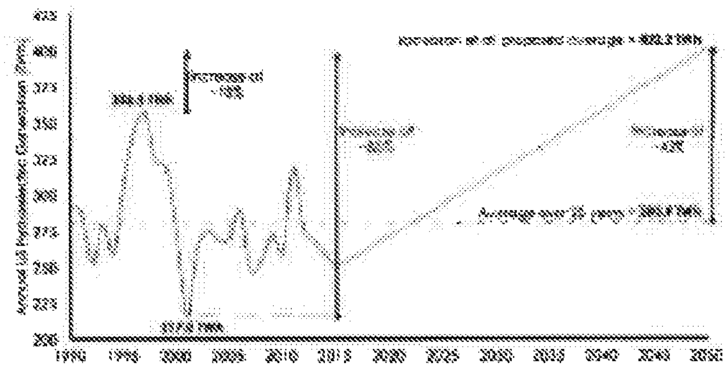


Fig. 3. Historical and proposed hydroelectric generation per year. The historical data (www.eia.gov/todayinenergy/detail.php?id=2650) show generation averaging 280.9 TWh/yr; generation proposed in ref. 11 is 402.2 TWh, 13% higher than the 25-y historical maximum of 356.5 TWh (1997) and 85% higher than the historical minimum of 217 TWh (2001).

Compl. ¶ 84. Plaintiff claims this figure is misleading because it compares the United States' historical annual production of hydroelectric power to the future annual hydroelectric power that plaintiff's paper projected, which he says included both the United States' projected power production and additional projected power to be imported from Canada. Compl. ¶ 63.[21]

But nothing in Figure 3 constitutes a false statement of fact. Plaintiff does not allege that either set of numbers is wrong. Instead, he claims that comparing the two sets of numbers is like comparing apples and oranges, because one set involved only United States power generation while the other allegedly included some Canadian power generation as well. *See* Compl. ¶ 62.

[21] Plaintiff contends that, "[b]y failing to subtract off the 45 TWh of Canadian imported hydropower out of 402.2 TWh of total hydropower, the Clack Authors misled readers into thinking the Jacobson Authors assumed an unreasonably high annually-averaged hydropower output." Compl. ¶ 63.

But that would present a simple disagreement over scientific methodology, which cannot be defamation. *E.g. Competitive Enterprise Institute v. Mann*, 150 A.3d at 1242.

Moreover, even though Figure 3 was in pre-publication drafts that plaintiff received and spent months flyspecking, he insists that he did not realize that the future projections (taken from his own paper) should not have been compared to historical U.S. production until after the Clack paper was published. Compl. ¶ 62. If plaintiff himself did not notice the allegedly "egregious" problem with Figure 3 when he was scrutinizing the paper,[22] it could not possibly have made him appear odious, infamous or ridiculous to PNAS readers.

Finally, as in the case of the "additional turbines" assumption he omitted to mention in his paper, plaintiff has now found it necessary to explain his "Canadian power" assumption in the errata he submitted after filing suit. *See* Ex. C (where plaintiff "clarifies" footnote 4 of his paper to explain he included the Canadian power).

4. The Additional Alleged Falsehoods Listed In Plaintiff's Exhibit 12

In addition to the three, so-called "egregious" statements discussed above, plaintiff alleges that Exhibit 12 to his complaint identifies numerous other "false and misleading statements" in the Clack paper. Compl. ¶ 84. Incorporating Exhibit 12 into his complaint violates the letter and spirit of Rule 8, which requires a short and concise statement of the claims. The Academy is nonetheless compelled to respond to the assertions set out in Exhibit 12.

Exhibit 12 is a "Line-by-Line Response" to the entire Clack paper, which parses through and debates the validity of almost every sentence from the paper in excruciating scientific detail.

[22] The lists of alleged errors that he submitted to the Academy not only failed to mention the problem that he now claims exists in Figure 3, they actually asserted that there was an entirely different explanation for the projections depicted in Clack's Figure 3. *See*, *e.g.*, Compl. Ex. 8 (Item 24) ("Figure 3 of Clack et al. falsely shows that the Jacobson et al. (2015b) hydropower output in 2050 is 402.2 TWh. In fact, they forgot to multiply by the transmission and distribution efficiency, a mean of 0.925, so it is really 372.035 TWh, much closer to current output.").

Nothing in Exhibit 12 shows that the Clack paper contained any statement that made plaintiff "appear odious, infamous, or ridiculous." And nothing in Exhibit 12 shows the Clack paper made personal attacks on plaintiff's honesty and integrity, or asserted he had engaged in professional misconduct or deceit to manufacture results.

Rather, Exhibit 12 merely shows that plaintiff disagrees with the scientific analyses and opinions that the Clack authors expressed in their paper. Plaintiff's complaints about the paper are therefore matters that can and must be addressed through scientific debate. They do not provide the basis for a defamation claim. *See, e.g., Competitive Enterprise Institute v. Mann*, 150 A.3d at 1242 ("As a matter of constitutional principle, when the issue is whether liability may be imposed for speech expressing scientific or policy views, the question is not who is right; the First Amendment protects the expression of all ideas, good and bad."); *Resolute Forest Products, Inc. v. Greenpeace International,* No. 17-cv-02824, 2017 WL 4618676, at *9 (N.D. Cal. Oct. 16, 2017) ("The academy, and not the courthouse, is the appropriate place to resolve scientific disagreements of this kind."); *Underwager v. Salter,* 22 F.3d 730, 736 (7th Cir. 1994) ("Scientific controversies must be settled by the methods of science rather than by the methods of litigation.").[23]

C. Plaintiff Is A Public Figure And Cannot Prove That The Academy Acted With Actual Malice

Plaintiff's defamation claim must also be dismissed because he is a public figure and cannot show that the Academy acted with actual malice when it published the Clack paper.

1. Plaintiff Is A Public Figure In The Scientific Debate Over The Feasibility Of Replacing Fossil Fuels

[23] If the Court wishes to review the litany of allegedly "false and misleading statements" described in Plaintiff's "Line-by-Line Response" to the Clack paper, a brief explanation of why each one cannot possibly support a defamation claim is set forth in Exhibit E.

An individual may be a public figure for a limited purpose where he "injects himself or is drawn into a particular public controversy and thereby becomes a public figure for a limited range of issues." *Gertz v. Welch,* 418 U.S. 323, 351 (1974). Plaintiff is a public figure for purposes of the ongoing public and scientific dispute over the need for and feasibility of replacing fossil fuels. Plaintiff's complaint asserts that he is "a renowned scientist on global warming and air pollution and the development of large-scale clean, renewable energy solutions for those problems," Compl. ¶ 1, who has published two books, published 152 peer-reviewed scientific papers, and been cited more than 11,000 times in the peer-reviewed literature. *Id.*[24]

Thus, plaintiff has injected himself into the public and scientific controversy over fossil fuels and alternative energy sources, and is a public figure for purposes of this case.

2. Plaintiff Cannot Show That The Academy Acted With Actual Malice

As a public figure, plaintiff cannot recover on a defamation claim unless he proves "by clear and convincing evidence" that the Academy published a defamatory statement "with actual malice, *i.e.*, with 'knowledge that it was false or with reckless disregard of whether it was false or not.'" *Masson v. New Yorker Magazine, Inc.,* 501 U.S. at 510, quoting *New York Times Co. v. Sullivan,* 376 U.S. at 279-80. Plaintiff cannot meet that standard, and the defamation claim should be dismissed.

The Clack paper was authored by twenty-one scientists and, pursuant to the Academy's practices, was subjected to peer review before it was published. The Academy therefore had

---

[24] Moreover, plaintiff's sixty-four page on-line *curriculum vitae* shows that his involvement in the public debate over the use of fossil fuels and alternative energy sources is not limited to the scientific journals: he has made over 480 presentations, and served on more than 35 panels. He has given more than 55 interviews to producers of television shows and documentaries. Plaintiff also has a Twitter account with almost 14,000 followers. Plaintiff's curriculum vitae is posted at http://web.stanford.edu/group/efmh/jacobson/vita/.

assurance that the paper's analyses and statements were supported by over twenty scientists, and had satisfied peer review.

Moreover, six weeks before the Clack paper was published, the Academy provided its authors with the May 5, 2017 version of plaintiff's "Line-By-Line Responses" to the paper. *See* Compl. Ex. 10. And the Clack authors made some changes to their original draft in response to what they received from plaintiff, including the addition of language to address plaintiff's assumption about adding turbines to existing hydroelectric dams. *See* Part III.B(2), *supra*. Thus, before the Clack paper was published, the Academy also had assurance that the authors were aware of plaintiff's criticisms of their paper and had made the changes they thought were necessary.

Finally, as previously discussed, the Academy gave plaintiff sixteen weeks' advance notice before publishing the Clack paper, multiple pre-publication drafts of the paper, and the opportunity to publish a rebuttal to the paper. And the Academy published plaintiff's rebuttal the same day it published the Clack paper. The Academy therefore ensured that plaintiff was given ample opportunity to identify any error he believed existed in the Clack paper, and that the PNAS readers would receive plaintiff's views simultaneously with the Clack authors' views.

This conduct was not malicious. It was, instead, the process used in the search for the scientific truth. *E.g.*, *McMillan v. Togus Regional Office*, 294 F. Supp. 2d at 317 ("'[S]cience is, above all, an adversary process. It is an arena in which ideas do battle . . . .'"). The Academy merely provided the forum for plaintiff and the Clack authors to debate their respective views.

IV. Plaintiff's Breach of Contract And Promissory Estoppel Claims Fail As A Matter Of Law

Like the defamation claim, plaintiff's contract and promissory estoppel claims "amount [ ] to nothing more than a backdoor attempt" to stifle protected expression.[25] Both the breach of contract and promissory estoppel claims fail because they are predicated on two unfounded propositions: (1) that the Academy's editorial guidelines are somehow binding on the Academy and not subject to modification or the discretion of the Academy; and (2) that plaintiff has the right to demand that his interpretation of the editorial guidelines be strictly applied to any paper challenging his work. Further, both counts also depend on the dubious proposition that plaintiff submitted his paper to PNAS, one of the most widely-cited scientific publications in the world, and part of the respected and prestigious Academy, because of the PNAS editorial guidelines, and that his selection constituted consideration for the implied agreement that PNAS would strictly apply plaintiff's interpretation of the editorial guidelines to other authors in the future. Compl. ¶¶ 95, 100, 101. Plaintiff further alleges that the Academy breached this "contract" with plaintiff when it published the Clack paper. Compl. ¶ 96. But he can point to nothing whatsoever to suggest that his interpretation of the editorial guidelines is binding on PNAS, or that he or any other author may impose his or her views of the guidelines – that is the purview of the publisher, PNAS.

A. Plaintiff's Breach Of Contract Claim Fails Because He Cannot Demonstrate The Existence Of An Implied Contract With The Academy

Plaintiff cannot show the existence of an implied contract with the Academy.[26]

---

[25] *See Compuserve Corp. v. Moody's Investors Services*, 499 F.3d 520, 531 (6th Cir. 2007).

[26] A breach of contract claim includes four elements: "(1) a valid contract between the parties; (2) an obligation or duty arising out of the contract; (3) a breach of that duty; and (4) damages caused by breach." *Tsintolas Realty Co. v. Mendez*, 984 A.2d 181, 187 (D.C. 2009). Plaintiff's claim also fails under California law because he cannot identify a contract or enforceable promise, a requirement of such a claim. *Lewis v. YouTube, LLC*, 244 Cal. App. 4th 118, 124,

The only written documents exchanged between plaintiff and the Academy are the Academy's author conflict of interest and comments form (the "Author COI and Comments Form") (Ex. F), and the PNAS License to Publish (Ex. G). By their express terms, these forms did not require PNAS to abide by the editorial guidelines. The Academy could not breach a written contract that did not obligate it to do anything.

In the absence of a written contract, plaintiff asserts that the Academy's editorial guidelines somehow created a separate implied contract. But general guidelines do not provide a basis for a breach of contract claim. *See Jurin v. Google Inc.*, 768 F. Supp. 2d 1064 (E.D. Cal. 2011). In *Jurin*, plaintiff brought a breach of contract claim against Google, alleging Google breached its contract by failing to investigate trademark infringement as required by Google's AdWords policy. The court found the parties did not have a contract regarding investigation of trademark infringement, noting that the plaintiff had:

> not referred to any written agreement between itself and Defendant above and beyond the AdWords policy. Plaintiff alleges no facts to support its contention that this policy was a contract between Plaintiff and Defendant, and not just a general policy statement on Defendant's website. <u>A broadly stated promise to abide by its own policy does not hold Defendant to a contract</u>.

Id. at 1073 (emphasis added). *See Beverage Distrib., Inc. v. Olympia Brewing Co.*, 440 F.2d 21, 29 (9th Cir. 1971) (" [a] gratuitous and unsolicited statement of policy or of intention which receives the concurrence of the party to whom it is addresses, does not constitute a contract").[27]

Plaintiff does not and cannot point to any facts showing the Academy's intent to be bound by editorial guidelines. *See Ponder v. Chase Home Fin., LLC*, 666 F. Supp. 2d 45, 48–49

---

197 Cal. Rptr. 3d 219, 224 (2015), *review denied* (Apr. 13, 2016) (citing *Oasis West Realty, LLC v. Goldman,* 51 Cal. 4th 811, 821, 124 Cal. Rptr. 3d 256, 250 P.3d 1115 (2011).

[27] *See also Dyer v. Northwest Airlines Corps.*, 334 F. Supp. 2d 1196, 1200 (D.N.D. 2004) ("[B]road statements of company policy do not generally give rise to contract claims."); *In re Northwest Airlines Privacy Litig.*, No. Civ. 04-126 (PAM/JSM), 2004 WL 1278459, at *6 (D. Miss. June 6, 2004) ("general statements of policy are not contractual.").

(D.D.C. 2009), citing *Vereen v. Clayborne*, 623 A.2d 1190, 1193 (D.C. 1993) ("An implied-in-fact contract . . . is inferred from the conduct of the parties in the milieu in which they dealt.").[28] Instead, he offers only the conclusory allegation that he believed that the Academy would "adhere to its publication policies for all publications." Compl. ¶ 95.

Moreover, plaintiff does not point to a single written or oral communication, or any action by the Academy that objectively manifests the Academy's intent to be contractually bound by its editorial guidelines. Nor could he. According to plaintiff's theory, the Academy allegedly transferred to plaintiff the right to dictate the form, authorship, and content of any paper critical of his work. Instead of any facts to support such an abdication of editorial control, plaintiff merely alleges that he was aware of the Academy's editorial guidelines, that his selection of PNAS for publication of his paper constituted consideration for an implied agreement between plaintiff and the Academy, and that the Academy would "adhere to its publication policies for all publications." Compl. ¶ 95. But plaintiff's awareness of the Academy's editorial guidelines falls far short of facts showing that the Academy intended for its discretionary editorial guidelines to become a binding agreement that plaintiff could unilaterally enforce for all future PNAS publications.

Perhaps recognizing that the Academy's own guidelines do not create the contract he seeks to enforce, plaintiff alleges that, because PNAS is a member of the Committee on Publication Ethics ("COPE"), it is required "to investigate every single claim of fabrication both before and after publication of an article" according to COPE's general guidelines. Compl. ¶ 97. Given that the Academy's own guidelines do not create a contract, it is specious to suggest that

---

[28] The party alleging an implied contract bears the burden of establishing each of the essential elements of breach of an implied contract. *Grunseth v. Marriott Corp.*, 872 F.Supp. 1069, 1073 (D.D.C. 1995); *Ponder v. Chase Home Fin., LLC*, 666 F. Supp. 2d 45, 48 (D.D.C. 2009).

an enforceable contract can be created by the guidelines of an organization of which PNAS is a member. In any event, the COPE guidelines make clear that they are merely general guidelines. As explained on its website, "COPE provides advice to editors and publishers on all aspects of publication ethics." The COPE Code of Conduct provides that "[b]est practices for editors would include . . . being *guided* by the COPE flowcharts . . . in cases of suspected misconduct,"[29] and "Best Practice Guidelines are more aspirational."[30] COPE does not impose any requirements on its members for investigating claims of fabrication, and provides no basis for plaintiff to enforce those guidelines in any event.

Finally, plaintiff cannot demonstrate that the Academy intended to be bound by its general guidelines, given that it did not apply strictly those guidelines to plaintiff's own rebuttal to the Clack paper. And the Academy has, in fact, published comments as research papers, thereby both undermining plaintiff's reliance and demonstrating that the Academy was not – and had no intent to be -- contractually bound by its guidelines. Moreover, even apart from the Academy's practices, plaintiff's breach of contract claim would be barred by the statute of frauds because there is no signed writing reflecting an intent to be bound. D.C. Code § 28–3502 (1996).

B. Plaintiff's Promissory Estoppel Claim Fails Because The Academy Did Not Promise To Strictly Enforce Its Editorial Guidelines

Plaintiff also cannot establish a claim for promissory estoppel. To establish a claim for promissory estoppel, a plaintiff must show (1) a promise; (2) that the promise reasonably induced reliance on it; and (3) that the promisee relied on the promise to his or her detriment.

---

[29] *See* COPE Code of Conduct and Best Practice Guidelines for Journal Editors, page 3 (emphasis added), available at: https://publicationethics.org/about.

[30] *Id.* at page 1.

Simard v. Resolution Trust Corp., 639 A.2d 540, 552 (D.C. 1994).[31] "[R]eliance on an indefinite promise is not reasonable." *In re U.S. Office Prods. Co. Secs. Litig.*, 251 F. Supp. 2d 77, 97 (D.D.C. 2003). And "though a promise need not be as specific and definite as a contract, it must still be a promise with definite terms on which the promisor would expect the promisee to rely." *Id.*, citing *Bender v. Design Store Corp.*, 404 A.2d 194, 196 (D.C. 1979). Plaintiff's promissory estoppel claim fails because there is no promise on which plaintiff could have reasonably relied: as discussed above, the Academy did not make any promises regarding its editorial guidelines.

Further, plaintiff cannot show reliance. Plaintiff alleges that he was induced by the editorial guidelines "to submit the Jacobson Article for publication in PNAS rather than any other competing scientific journal." Compl. ¶ 101. This allegation is facially pretextual. PNAS is one of the world's most-cited scientific journals in publication today.[32] PNAS is part of the prestigious National Academy of Sciences.[33] Common sense would suggest that the reputation and readership of PNAS is a far greater inducement to an author than the Academy's general editorial guidelines. Moreover, plaintiff cannot show reliance because publication of the Clack paper as a research paper is consistent with prior Academy practice. *See, e.g.*, Ex. B. Plaintiff cannot show reliance on a promise when the Academy's prior actions demonstrate affirmatively that no such promise existed.

CONCLUSION

As demonstrated above, this is a case where the important purpose behind the Anti-SLAPP Act – to ensure defendants are not intimidated or prevented from engaging in public policy debates by abusive lawsuits – must be vindicated by dismissal of plaintiff's claims in their

[31] The elements of promissory estoppel are the same under California law; plaintiff fails to state a claim under California law as well. *Granadino v. Wells Fargo Bank, N.A.*, 236 Cal. App. 4th 411, 416, 186 Cal. Rptr. 3d 408, 413 (2015), as modified (Apr. 29, 2015).
[32] http://www.nasonline.org/publications/pnas.
[33] *See* PNAS Marketing Brochure, *available at*: http://www.pnas.org/site/aboutpnas/index.xhtml.

entirety with prejudice. The three counts – intended to stifle expression on a critical issue in America today – are merely a transparent attempt to circumvent the First Amendment, and plaintiff cannot prevail on the merits. Pursuant to the Anti-SLAPP Act and Rule 12(b)6), the Academy respectfully asks this Court to dismiss plaintiff's claims in their entirety and with prejudice.

Dated: November 27, 2017

Respectfully submitted,

/s/ Joseph P. Esposito

Joseph P. Esposito (D.C. Bar No. 375507)
William E. Potts, Jr. (D.C. Bar No. 945824)
HUNTON & WILLIAMS LLP
2200 Pennsylvania Avenue, N.W.
Washington, D.C. 20037
Tel. No. (202) 955-1500
jesposito@hunton.com
wpotts@hunton.com

Counsel for Defendant National Academy of Sciences

Of Counsel:

Audrey Byrd Mosley (D.C. Bar No. 263376)
General Counsel
Marc S. Gold
Deputy General Counsel
NATIONAL ACADEMY OF SCIENCES
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

Eric H. Feiler
HUNTON & WILLIAMS LLP
Riverfront Plaza, East Tower
951 East Byrd Street
Richmond, VA 23219

EXHIBIT A

LETTER

The United States can keep the grid stable at low cost with 100% clean, renewable energy in all sectors despite inaccurate claims

Mark Z. Jacobson[a,1], Mark A. Delucchi[b], Mary A. Cameron[a], and Bethany A. Frew[a]

The premise and all error claims by Clack et al. (1) in PNAS, about Jacobson et al.'s (2) report, are demonstrably false. We reaffirm Jacobson et al.'s conclusions.

False Premise

Clack et al.'s (1) premise that deep decarbonization studies conclude that using nuclear, carbon capture and storage (CCS), and bioenergy reduces costs relative to "other pathways," such as Jacobson et al.'s (2) 100% pathway, is false.

First Clack et al. (1) imply that Jacobson et al.'s (2) report is an outlier for excluding nuclear and CCS. To the contrary, Jacobson et al. are in the mainstream, as grid stability studies finding low-cost up-to-100% clean, renewable solutions without nuclear or CCS are the majority (3–16).

Second, the Intergovernmental Panel on Climate Change (IPCC) (17) contradicts Clack et al.'s (1) claim that including nuclear or CCS reduces costs (7.6.1.1): "...high shares of variable RE [renewable energy] power...may not be ideally complemented by nuclear, CCS,..." and (7.8.2) "Without support from governments, investments in new nuclear power plants are currently generally not economically attractive within liberalized markets,..." Similarly, Freed et al. (18) state, "...there is virtually no history of nuclear construction under the economic and institutional circumstances that prevail throughout much of Europe and the United States," and Cooper (19), who compared decarbonization scenarios, concluded, "Neither fossil fuels with CCS or nuclear power enters the least-cost, low-carbon portfolio."

Third, unlike Jacobson et al. (2), the IPCC, National Oceanic and Atmospheric Administration, National Renewable Energy Laboratory, and International Energy Agency have never performed or reviewed a cost analysis of grid stability under deep decarbonization. For example, MacDonald et al.'s (20) grid-stability analysis considered only electricity, which is only ~20% of total energy, thus far from deep decarbonization. Furthermore, deep-decarbonization studies cited by Clack et al. (1) have never analyzed grid stability. Jacobson et al. (2) obtained grid stability for 100% wind, water, and solar power across all energy sectors, and thus simulated complete energy decarbonization.

Fourth, Clack et al.'s (1) objectives, scope, and evaluation criteria are narrower than Jacobson et al.'s (2), allowing Clack et al. (1) to include nuclear, CCS, and biofuels without accounting for their true costs or risks. Jacobson et al. (2, 21) sought to reduce health, climate, and energy reliability costs, catastrophic risk, and land requirements while increasing jobs. Clack et al. (1) focus only on carbon. By ignoring air pollution, the authors ignore bioenergy, CCS, and even nuclear health costs (22); by ignoring land use they ignore bioenergy feasibility; by ignoring risk and delays, they ignore nuclear feasibility, biasing their conclusions.

Fifth, Clack et al. (1) contend that Jacobson et al. (2) place "constraints" on technology options. In contrast, Jacobson et al. include many technologies and processes not in Clack et al.'s (1) models. For example, Jacobson et al. (2) include, but MacDonald et al. (20) exclude, concentrated solar power (CSP), tidal, wave, geothermal, solar heat, any storage (CSP, pumped-hydro, hydropower, water, ice, rocks, hydrogen), demand-response, competition among wind turbines for kinetic energy, electrification of all energy sectors, calculations of load decrease upon electrification, and so forth. Model time steps in MacDonald et al. (20) are also 120-times longer than in Jacobson et al. (2).

False Error Claims

Clack et al. (1) claim wrongly that Jacobson et al. (2) assume a maximum hydropower output of 145.26 GW, even though table S2 in Jacobson et al. shows 87.48 GW. Clack et al. (1) then claim incorrectly that the 1,300 GW drawn in figure 4B of Jacobson et al. (2) is wrong because it exceeds 87.48 GW, not recognizing that 1,300 GW is instantaneous and 87.48 GW, a maximum possible annual average [table S2, footnote 4 in Jacobson et al. (2) and the available LOADMATCH code]. The value of 1,300 GW is correct, because turbines were

[a]Department of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305; and [b]Institute of Transportation Studies, University of California, Berkeley, CA 94720

Author contributions: M.Z.J. and M.A.D. designed research; M.Z.J., M.A.D., M.A.C., and B.A.F. performed research; M.Z.J. and M.A.D. contributed new reagents/analytic tools; M.Z.J. and M.A.D. analyzed data; and M.Z.J., M.A.D., M.A.C., and B.A.F. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. Email: jacobson@stanford.edu.

assumed added to existing reservoirs to increase their peak instantaneous discharge rate without increasing their annual energy consumption, a solution not previously considered. Increasing peak instantaneous discharge rate was not a "modeling mistake" but an assumption consistent with Jacobson et al.'s (2) table S2, footnote 4, and LOADMATCH, and written to Clack on February 29, 2016.

Jacobson et al. (2) only neglect the cost of additional turbines, generators, and transformers needed to increase the maximum discharge rate. Such estimated cost for a 1000-MW plant (23) plus wider penstocks is ~\$385 (325–450)/kW, or ~14% of hydropower capital cost. When multiplied by the additional turbines and hydropower's fraction of total energy, the additional infrastructure costs ~3% of the entire wind, water, and solar power system and thus doesn't impact Jacobson et al.'s (2) conclusions. Increasing CSP's—instead of hydropower's—peak discharge rate also works.

In their figure 3, Clack et al. (1) then claim mistakenly that Jacobson et al.'s (2) annual hydropower energy output is 402 TWh/yr and too high, when it is actually 372 TWh/yr because they missed transmission and distribution losses. This is less than half the possible United States hydropower output today and well within reason.

Clack et al. (1) next claim wrongly that in Jacobson et al.'s (2) table 1, loads are "maximum possible" loads, even though the text clearly indicates they are annual-average loads. The word "maximum" is never used. Clack et al. (1) compound this misrepresentation by claiming flexible loads in Jacobson et al.'s (2) time figures are twice "maximum possible" loads, even though Jacobson et al. clearly state that the annual loads are distributed in time.

Unsubstantiated Claims About Assumptions. Clack et al. (1) assert that underground thermal energy storage (UTES) can't be expanded nationally, but we disagree. UTES is a form of district heating, which is already used worldwide (e.g., 60% of Denmark); UTES is technologically mature and inexpensive; moreover, hot-water storage or heat pumps can substitute for UTES. Similarly, molten salt can substitute for phase change materials in CSP storage.

Clack et al. (1) further criticize Jacobson et al.'s (2) hydrogen scale-up, but this is easier than Clack et al.'s (1) proposed nuclear or CCS scale-up. Clack et al. (1) also question whether aviation can adopt hydrogen, but a 1,500-km range, four-seat hydrogen fuel cell plane already exists, several companies are now designing electric-only planes for up to 1,500 km, and Jacobson et al. (21) propose aircraft conversion only by 2035–2040.

Clack et al. (1) question whether industrial demand is flexible, yet the National Academy of Sciences (24) review they cite states, "Demand response can be a lucrative enterprise for industrial customers."

Clack et al. (1) criticize Jacobson et al.'s (2) use of a 1.5–4.5% discount rate, even though that figure is a well-referenced social discount rate for a social cost analysis of an intergenerational project (21).

Clack et al. (1) state misleadingly that Jacobson et al.'s (2) storage capacity is twice United States electricity capacity, failing to acknowledge that Jacobson et al.'s (2) report treats all energy, which is five times electricity, not just electricity, and in Jacobson et al. (2), storage is only two-fifth of all energy. Furthermore, in Jacobson et al.'s (2) report, storage is mostly heat.

Clack et al. (1) claim the average installed wind density is 3 W/m^2, but fail to admit this includes land for future project expansion and double counts land where projects overlap. Furthermore, real data from 12 European and Australian farms give 9.4 W/m^2.

Clack et al. (1) claim that Jacobson (22) didn't rely on consensus data for CO_2 lifecycle estimates, although Jacobson's nuclear estimate was 9–70 g-CO_2/kWh, within the IPCC's (17) range, 4–110 g-CO_2/kWh.

Clack et al. (1) claim falsely that Jacobson (22) didn't include a planning-to-operation time for offshore wind, even though ref. 22 states 2–5 y.

Clack et al. (1) criticize Jacobson (22) for considering weapons proliferation and other nuclear risks, although the IPCC (17) agrees (Executive Summary): "Barriers to and risks associated with an increasing use of nuclear energy include operational risks and the associated safety concerns, uranium mining risks, financial and regulatory risks, unresolved waste management issues, nuclear weapons proliferation concerns, . . .(robust evidence, high agreement)."

False Model Claims. Clack et al. (1) claim falsely that the gas, aerosol, transport, radiation-general circulation mesoscale, and ocean model (GATOR-GCMOM) "has never been adequately evaluated," despite it taking part in 11 published multimodel intercomparisons and 20 published evaluations against wind, solar, and other data; despite Zhang's (25) evaluation that GATOR-GCMOM is "the first fully-coupled online model in the history that accounts for all major feedbacks among major atmospheric processes based on first principles"; and despite hundreds of processes in it still not in any other model (26).

Clack et al. (1) contend that LOADMATCH is not transparent, even though LOADMATCH has been publicly available since Jacobson et al.'s (2) publication.

Clack et al. (1) criticize LOADMATCH for not treating power flows, and claim that Jacobson et al.'s (2) transmission costs are "rough." However Clack et al. (1) do not show such costs are unreasonable or acknowledge Jacobson et al.'s (2) high-voltage direct current cost per kilometer (21) are far more rigorous than MacDonald et al.'s (20).

Finally, Clack et al. (1) falsely claim that LOADMATCH has perfect foresight, thus is deterministic. However, LOADMATCH has zero foresight, knowing nothing about load or supply the next time step. It is prognostic, requiring trial and error, not an optimization model.

In sum, Clack et al.'s (1) analysis is riddled with errors and has no impact on Jacobson et al.'s (2) conclusions.

1 Clack CTM, et al. (2017) Evaluation of a proposal for reliable low-cost grid power with 100% wind, water, and solar. *Proc Natl Acad Sci USA* 114:6722–6727.
2 Jacobson MZ, Delucchi MA, Cameron MA, Frew BA (2015) Low-cost solution to the grid reliability problem with 100% penetration of intermittent wind, water, and solar for all purposes. *Proc Natl Acad Sci USA* 112:15060–15065.
3 Mason IG, Page SC, Williamson AG (2010) A 100% renewable electricity generation system for New Zealand utilising hydro, wind, geothermal and biomass resources. *Energy Policy* 38:3973–3984.
4 Hart EK, Jacobson MZ (2011) A Monte Carlo approach to generator portfolio planning and carbon emissions assessments of systems with large penetrations of variable renewables. *Renew Energy* 36:2278–2286.
5 Connolly D, Lund H, Mathiesen BV, Leahy M (2011) The first step towards a 100% renewable energy-system for Ireland. *Appl Energy* 88:502–507.

6 Connolly D, Mathiesen BV (2014) A technical and economic analysis of one potential pathway to a 100% renewable energy system. *Intl J Sustainable Energy Planning & Management* 1:7–28.
7 Connolly D, Lund H, Mathiesen BV (2016) Smart energy Europe: The technical and economic impact of one potential 100% renewable energy scenario for the European Union. *Renew Sustain Energy Rev* 60:1634–1653.
8 Mathiesen BV, Lund H, Karlsson K (2011) 100% Renewable energy systems, climate mitigation and economic growth. *Appl Energy* 88:488–501.
9 Mathiesen BV, et al. (2015) Energy systems for coherent 100% renewable energy and transport solutions. *Appl Energy* 145:139–154.
10 Elliston B, MacGill I, Diesendorf M (2013) Least cost 100% renewable electricity scenarios in the Australian National Electricity Market. *Energy Policy* 59:270–282.
11 Elliston B, MacGill I, Diesendorf M (2014) Comparing least cost scenarios for 100% renewable electricity with low emission fossil fuel scenarios in the Australian National Electricity Market. *Renew Energy* 66:196–204.
12 Rasmussen MG, Andresen GB, Greiner M (2012) Storage and balancing synergies in a fully or highly renewable pan-European power system. *Energy Policy* 51:642–651.
13 Budischak C, et al. (2013) Cost-minimized combinations of wind power, solar power and electrochemical storage, powering the grid up to 99.9% of the time. *J Power Sources* 225:60–74.
14 Steinke F, Wolfrum P, Hoffmann C (2013) Grid vs. storage in a 100% renewable Europe. *Renew Energy* 50:826–832.
15 Becker S, et al. (2014) Features of a fully renewable U.S. electricity-system: Optimized mixes of wind and solar PV and transmission grid extensions. *Energy* 72:443–458.
16 Bogdanov D, Breyer C (2016) North-east Asian super grid for 100% renewable energy supply: Optimal mix of energy technologies for electricity, gas, and heat supply options. *Energy Convers Manage* 112:176–190.
17 IPCC; Bruckner T, et al. (2014) Energy systems. *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, eds Edenhofer O, et al. (Cambridge Univ Press, Cambridge, UK).
18 Freed J, Allen T, Nordhaus T, Lovering J (2017) Is nuclear too innovative? Third Way, Available at https://medium.com/third-way/is-nuclear-too-innovative-a14fb4fef41a#.qag59xnk0. Accessed Feburary 28, 2017.
19 Cooper M (2016) The economic and institutional foundations of the Paris Agreement on climate change: The political economy of roadmaps to a sustainable electricity future. Available at dx.doi.org/10.2139/ssrn.2722880. Accessed May 12, 2017.
20 MacDonald AE, et al. (2016) Future cost competitive electricity systems and their impact on US CO_2 emissions. *Nat Clim Chang* 6:526–531.
21 Jacobson MZ, et al. (2015) 100% clean and renewable wind, water, and sunlight (WWS) all-sector energy roadmaps for the 50 United States. *Energy Environ Sci* 8:2093–2117.
22 Jacobson MZ (2009) Review of solutions to global warming, air pollution, and energy security. *Energy Environ Sci* 2:148–173.
23 International Renewable Energy Agency (2012) Renewable energy technologies: Cost analysis series, Vol. 1. Available at www.irena.org/documentdownloads/publications/re_technologies_cost_analysis-hydropower.pdf. Accessed February 10, 2014.
24 National Academy of Sciences, National Academy of Engineering, National Research Council (2010) *Real Prospects for Energy Efficiency in the United States* (National Academies Press, Washington, DC), p 251.
25 Zhang Y (2008) Online coupled meteorological and chemistry models: History, current status, and outlook. *Atmos Chem Phys* 8:2895–2932.
26 Jacobson MZ (2012) History of, Processes in, and Numerical Techniques in GATOR-GCMOM. Available at web.stanford.edu/group/efmh/jacobson/GATOR/GATOR-GCMOMHist.pdf. Accessed May 11, 2017.

EXHIBIT B

Proc. Natl. Acad. Sci. USA
Vol. 93, p. 14314, December 1996
Chemistry

Comments on the article "Persistent confusion of total entropy and chemical system entropy in chemical thermodynamics" [(1996) *Proc. Natl. Acad. Sci. USA* 93, 7452–7453]

JOHN ROSS

Department of Chemistry, Stanford University, Stanford, CA 94305

Contributed by John Ross, September 4, 1996

A comment is necessary on the article entitled "Persistent confusion of total entropy and chemical system entropy in chemical thermodynamics" by Gregorio Weber, which appeared in these PROCEEDINGS (1). The article purports to show that in all prior treatises and texts on thermodynamics the temperature variation of the Gibbs free energy of a single substance, and therefore also the van't Hoff equation for the temperature variation of the Gibbs free energy change of a reaction, have been misinterpreted.

Weber originally published his views in *J. Phys. Chem.* (2). His work there was criticized in refs. 3 and 4; he replied in ref. 5 but ignored most of the criticism. The purpose here is to point explicitly to Weber's errors in a simple way; the valid and pertinent arguments in refs. 3 and 4 are not repeated.

Consider a one-variable system. The Gibbs free energy is

$$G = H - TS; \qquad [1]$$

for differential changes we have

$$dG = Vdp - SdT. \qquad [2]$$

Symbols without subscripts refer to the system. For a process at constant pressure but not constant temperature

$$(\partial G/\partial T)_p = -S; \qquad [3]$$

in Eq. **3** S is the entropy of the system, not the entropy of the system and the entropy of the surroundings, as Weber claims.

For a process at constant temperature and constant pressure

$$(dG)_{T,p} = 0, \qquad [4]$$

where we have added on dG the notation that T and p are constant. Next we return to Eq. **1** and derive for a process at constant temperature

$$(dG)_T = dH - TdS. \qquad [5]$$

If in addition we hold the pressure constant, then

$$dH = đQ_p \qquad [6]$$

and the entropy change in the surroundings is given by

$$-dH = -đQ_p = T_{surr}dS_{surr}. \qquad [7]$$

Since both T and p are kept constant we may write

$$T = T_{surr} \qquad [8]$$

and

$$(dG)_{T,p} = -T[dS + dS_{surr}]. \qquad [9]$$

From Eq. **4** we see that Eq. **9** equals zero always. Another way of seeing that result comes from the fact that for a one-component system a process at constant T and p is reversible; hence

$$dS = -dS_{surr}, \qquad [10]$$

and Eq. **9** is zero. Eq. **9** is the same as equation **3** in ref. 1.

At this point Weber claims that Eq. **3**, with S interpreted by him to be the entropy of the system and the entropy of the surroundings (his equation **4**), follows from Eq. **9** (his equation **3**), but he gives no derivation. This is impossible, since the constraints on Eq. **9** are constant T and p, but the constraints on Eq. **3** are constant p. Furthermore, Eq. **9** always equals zero, whereas Eq. **3** is not zero for any T variation. S in Eq. **3** is the entropy of the system, not the entropy of the system plus that of the surroundings, as Weber claims.

Weber failed to notice the different constraints on Eq. **9** (his equation **3**) and on Eq. **3** (his equation **4**); further, he failed to notice that Eq. **9** (his equation **3**) is always zero. Hence all that follows in Weber's article is incorrect.

Another argument can be made against Weber's interpretation of Eq. **3**.

Consider Weber's suggested relation for a one-variable system

$$\left(\frac{\partial G}{\partial T}\right)_p = -[S_{system} + S_{surr}],$$

which is equation **4** in Weber's article in the PROCEEDINGS. G is the Gibbs free energy of the systems. For consistency we choose G to be per mole of the system, S_{system} to be the entropy per mole of the system, and S_{surr} to be the entropy per mole of the surroundings. Suppose the surroundings are made of $N_2(g)$, which has a given value for its entropy per mole; if we change the surroundings to be water, then the entropy per mole of H_2O has another value. Thus if we integrate the above equation at constant pressure

$$G(T_2, p) - G(T_1, p) = -\int_{T_1}^{T_2} [S_{system} + S_{surr}]dT,$$

then G is no longer a state function—that is, a function dependent on the state of the system only. The integral depends on the specific material constituting the surroundings. Hence, given $G(T_1, p)$, the value of $G(T_2, p)$ depends not only on T_2, p of the system but also on the specific material of the surroundings. This conclusion is absurd, and Weber's arguments cannot be correct.

This work was supported in part by the Department of Energy, Basic Energy Sciences, Engineering Program.

1. Weber, G (1996) *Proc. Natl. Acad. Sci. USA* **93**, 7452–7453.
2. Weber, G. (1995) *J. Phys. Chem.* **99**, 1052–1059.
3. Holtzer, A. (1995) *J. Phys. Chem.* **99**, 13048–13049.
4. Ragone, R., Colonna, G. & Ambrosone, L. (1995) *J. Phys. Chem.* **99**, 13050.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 7452–7453, July 1996
Chemistry

Persistent confusion of total entropy and chemical system entropy in chemical thermodynamics

(van't Hoff equation/model dependent thermodynamics)

GREGORIO WEBER

University of Illinois School of Chemical Sciences, 600 South Mathews Avenue, Urbana, IL 61801

Contributed by Gregorio Weber, February 29, 1996

ABSTRACT **The change in free energy with temperature at constant pressure of a chemical reaction is determined by the sum (dS) of changes in entropy of the system of reagents, dS_i, and the additional entropy change of the surroundings, dS_H, that results from the enthalpy change, dH. A faulty identification of the total entropy change on reaction with dS_i has been responsible for the attribution of general validity to the expressions $(d\Delta G/dT)_P = -\Delta S_i$ and $d(\Delta G/T)/d(1/T) = \Delta H$, which are found in most textbooks and in innumerable papers.**

The free energy associated to each reagent entering in chemical reaction under constant temperature and pressure is given by the familiar Gibbs function

$$G = H - TS_i. \qquad [1]$$

Eq. 1 is a statement of the heat content of the reagent, part of which is actual and determined by its intrinsic entropy S_i, and partly potential, determined by H, the energy locked in the bonds that partake in the chemical reaction. In an isothermal reaction

$$dG = dH - TdS_i, \qquad [2]$$

the actual change in the composition of the system does not provide a measure of the enthalpy change dH, which can only be appraised by the corresponding change in heat content of the surroundings of magnitude $dQ_H = -dH$. Similarly, TdS_i provides for an additional change in the heat content of the surroundings of magnitude dQ_i. Calorimetric measurements do not separate these two interdependent but distinct heat changes and only their sum $dQ = -(dQ_H + dQ_i)$ is measurable. Neglecting the very small contribution to the external work, pdV, which is the case in practically all reactions in solution, the change in free energy of a reagent in an isothermal reaction may be expressed as a change in entropy dS of its surroundings, the sum of changes dS_H and dS_i

$$dG = -T(dS_H + dS_i) = -TdS. \qquad [3]$$

As such, Eq. 3 satisfies Planck's criterion (1, *) that the decrease in free energy in a spontaneous process is determined by the change in entropy of *all* the bodies in which the heat content changes in the reaction, a condition essentially required by the second law (2). It follows from Eq. 3 that the free energy change of the reagent with temperature at constant pressure equals

$$(dG/dT)_p = -S. \qquad [4]$$

From Eq. 1, we have

$$\left(\frac{dG}{dT}\right)_p = \frac{dH}{dT} - T\frac{dS_i}{dT} - S_i. \qquad [5]$$

Eq. 4 refers the changes in free energy to the entropy of system and surroundings, as demanded by the second law of thermodynamics, whereas Eq. 5 expresses the free energy change in terms of properties intrinsic to the system. If we neglect this difference and simply write Eq. 5 in the form

$$\left(\frac{dG}{dT}\right)_p = \frac{dH}{dT} - T\frac{dS}{dT} - S, \qquad [5']$$

we risk concluding from Eqs. 4 and 5′ that

$$(dG/dT)_P = -S_i \qquad [6]$$

and inadvertently also that $S_H + S_i = S_i$. Yet the simple mistake embodied in Eq. 6 has been exhibited by textbooks and is current in the literature for the last 70 and perhaps more years. That dS_i does not represent the total entropy change is brought out by the many cases in which the standard change in S_i in the reaction is negative. These instances do not contradict the second law because of the larger entropy increase of the surroundings that follows the decrease in H. In fact, an immediate corollary of the second law is that: If the entropy of any part of a system decreases in the course of an isothermal process, there must be a concomitant and larger increase in another part of the same system.

The erroneous identification of the entropy change of the chemical system dS_i with the total entropy change $dS_H + dS_i$ leads immediately to a fixed relation of the changes in enthalpy and entropy with temperature: If $dG/dT = -S_i$, it follows from Eq. 5′ that

$$dH/dT = T(dS_i/dT) \qquad [7]$$

is valid for each reagent in a chemical reaction, and then for all reactions

$$d\Delta H/dT = T(d\Delta S_i/dT), \qquad [8]$$

where ΔH and ΔS_i are the standard changes in enthalpy and intrinsic entropy in the reaction, respectively.

It also follows that

$$\frac{1}{T}\frac{d\Delta H}{d(1/T)} = \frac{d\Delta S_i}{d(1/T)}. \qquad [8']$$

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

*Planck establishes the distinction between the entropy of the environment, ϕ_o, and that of the reagents, ϕ, states the application of the second law to reactions as $d(\phi + \phi_o) \geq 0$, and consistently uses ϕ in the same way as I use S_i. He notes that $dU = dH - pdV$ often exceeds $Td\phi$, thus justifying Berthelot's rule. Unfortunately, he did not explicitly state that in these cases one can have $d\phi \leq 0$, which should have precluded the systematic erroneous identification of $S = \phi + \phi_o$ with $S_i = \phi$.

Starting with the Gibbs relation $\Delta G = \Delta H - T\Delta S_i$ and dividing by T,

$$\Delta G/T = \Delta H/T - \Delta S_i \quad [9]$$

and

$$\frac{d(\Delta G/T)}{d(1/T)} = \Delta H + \frac{1}{T}\frac{d\Delta H}{d(1/T)} - \frac{d\Delta S_i}{d(1/T)} \quad [10]$$

but, according to Eq. **8′**, because the second and third term cancel each other

$$\frac{d(\Delta G/T)}{d(1/T)} = \Delta H \quad [11]$$

The last is the celebrated van't Hoff equation that has been used for so long in the determination of enthalpy changes in chemical reactions that all papers and most textbooks consider it unnecessary to enter into a detailed discussion of its derivation. Actual examples of the mistaken derivation of Eq. **11** encountered in classical textbooks of chemical thermodynamics are given elsewhere (3). No textbook that I have consulted writes explicitly Eq. **10**, which could have prompted somebody to ask what happens when $\Delta H = 0$ and $\Delta S_i > 0$. Instead they omit all differences in notation necessary to distinguish S_i from S of Eq. **3**, and then from Eqs. **4** and **5′** derive Eq. **8′** and with this in hand declare the universal worth of the van t'Hoff relation of Eq. **11**. The most important consequence of the dismissal of the relations in Eqs. **8** and **8′** is that in a chemical reaction, there are no fixed relations between the standard changes in enthalpy and entropy deducible from the laws of thermodynamics. Consequently, calculation of the effects of pressure and temperature on the chemical equilibria require suitable thermodynamic models that describe the relations of H and S expected for each reagent. While success of these ad hoc models will not confer on the results the certainty wrongly attributed in the past to the standard changes in enthalpy and entropy determined by the van't Hoff equation, they provide us with an opportunity to establish meaningful relations between macroscopic thermodynamics and microscopic chemistry.

It may be asked why the van't Hoff equation has proven useful in spite of its spurious derivation. It is simply that most chemical reactions, as already recognized by Berthelot (4) in 1897, are driven by the change in enthalpy. Any error owing to the neglect of the much smaller entropy contribution would pass unnoticed because we have no other means of independently evaluating either ΔH or ΔS_i. The consideration of the association of protein subunits, which are typical entropy driven reactions (5, 6), led me to observe the extremely different results obtained from Eqs. **10** and **11** when specific relations of H and S of the reagents were assumed (7, 8) to permit modeling the reaction. Eventually, I was able to trace the origin of Eq. **11** to the erroneous Eqs. **8** and **8′**, and the original confusion between S_i and $S = S_H + S_i$. This long-standing mistake has clearly arisen from inability to distinguish between the variational conditions that define the chemical equilibrium, originally stated by Gibbs (9), and the application of the laws of thermodynamics to actual processes like those that result from finite changes in temperature or pressure. It provides a striking example of Truesdell's dictum (10): In thermodynamics, "Confusion of the nature of the equilibrium of a large class of bodies with the effect of processes undergone by members of a small class of bodies is nearly universal."

I thank Drs. H. G. Drickamer, D. M. Jameson, G. U. Nienhaus, and G. D. Reinhart for their comments and suggestions. This work was supported by U.S. Public Health Service Grant GM11223.

1. Planck, M. (1922) *Treatise on Thermodynamics* (Dover, New York), pp. 103–119.
2. Serrin, J. (1977) *Arch. Rational Mech. Anal.* **70**, 355–371.
3. Weber, G. (1995) *J. Phys. Chem.* **99**, 13051.
4. Berthelot, M. (1897) *Thermochimie* (Gauthier–Villars, Paris), Vol. 1, p. 13.
5. Lauffer, M. (1975) *Entropy Driven Reactions in Biology* (Springer, New York).
6. Silva, J. L. & Weber, G. (1993) *Annu. Rev. Phys. Chem.* **44**, 89–113.
7. Weber, G. (1993) *J. Phys. Chem.* **97**, 7105–7115.
8. Weber, G. (1993) *J. Phys. Chem.* **99**, 1052–1059.
9. Gibbs, J. W. (1993) *Scientific Papers I: Thermodynamics* (Ox Bow, Woodbridge, CT), pp. 56–62.
10. Truesdell C. (1984) *Rational Thermodynamics* (Springer, New York), 2nd Ed., p. 31.

No evidence of nanodiamonds in Younger–Dryas sediments to support an impact event

Tyrone L. Daulton[a,1], Nicholas Pinter[b], and Andrew C. Scott[c]

[a]Department of Physics and Center for Materials Innovation, Washington University in St. Louis, St. Louis, MO 63130; [b]Department of Geology, Southern Illinois University, Carbondale, IL 62901; and [c]Department of Earth Sciences, Royal Holloway University of London, Egham, Surrey TW20 OEX, United Kingdom

Edited by Mark H. Thiemens, University of California at San Diego, La Jolla, CA, and approved July 27, 2010 (received for review March 24, 2010)

The causes of the late Pleistocene megafaunal extinctions in North America, disappearance of Clovis paleoindian lithic technology, and abrupt Younger–Dryas (YD) climate reversal of the last deglacial warming in the Northern Hemisphere remain an enigma. A controversial hypothesis proposes that one or more cometary airbursts/impacts barraged North America ≈12,900 cal yr B.P. and caused these events. Most evidence supporting this hypothesis has been discredited except for reports of nanodiamonds (including the rare hexagonal polytype) in Bølling–Ållerod-YD-boundary sediments. The hexagonal polytype of diamond, lonsdaleite, is of particular interest because it is often associated with shock pressures related to impacts where it has been found to occur naturally. Unfortunately, previous reports of YD-boundary nanodiamonds have left many unanswered questions regarding the nature and occurrence of the nanodiamonds. Therefore, we examined carbon-rich materials isolated from sediments dated 15,818 cal yr B.P. to present (including the Bølling–Ållerod-YD boundary). No nanodiamonds were found in our study. Instead, graphene- and graphene/graphane-oxide aggregates are ubiquitous in all specimens examined. We demonstrate that previous studies misidentified graphene/graphane-oxide aggregates as hexagonal diamond and likely misidentified graphene as cubic diamond. Our results cast doubt upon one of the last widely discussed pieces of evidence supporting the YD impact hypothesis.

archaeology | paleoclimate | Quaternary extinctions | carbon spherules | fungal sclerotia

During the end of the last glacial period in the Northern Hemisphere near 12,900 cal yr B.P., deglacial warming of the Bølling–Ållerod interstadial ceased abruptly (1) and glacial conditions were restored for a ≈1,300-yr interval known as the Younger–Dryas (YD) stadial (2–5). Conclusion of the YD was abrupt with temperature increasing to present-day Holocene-interglacial levels within a few decades (2, 6). In North America at least 17 genera of megafauna (e.g., mammoths, mastodons, giant short-faced bears, saber-tooth tigers, and numerous other megafauna species) became extinct near the YD onset, although the details of the chronology are in question (7, 8). North America's earliest known human populations arrived and dispersed prior to the YD, and their Clovis lithic technology disappeared from the sedimentary record near the onset of the YD (9, 10). Although the YD climate reversal as well as the geologically abrupt Pleistocene megafauna extinctions and disappearance of the Clovis culture are not disputed, their causes are matters of intense debate.

Most paleoclimatologists believe the YD stadial resulted when a massive volume (~9,500 km^3) of fresh water from the proglacial Lake Agassiz released into the northern Atlantic and disabled the thermohaline circulation (4, 11). However, the YD has also been attributed to cessation in the El Niño Southern oscillation caused by changes in Earth's orbital configuration (12). Although climate change undoubtedly applied stress to the megafauna and human populations, it remains unclear if that alone was sufficient to catastrophically impact animal populations. Widespread hunting of megafauna by Clovis paleoindians and disease have also been proposed.

A recent hypothesis suggests that an extraterrestrial body, either a fragmented chondritic meteorite or comet, detonated as airburst(s) and/or impact(s) over North America, igniting continent-wide wildfires and injecting a large mass of dust into the atmosphere (13–16). The energy deposited by the bolides is speculated to have induced partial melting of the Laurentide ice sheet, disabling the North Atlantic thermohaline circulation and initiating the YD stadial. The combined environment impact would have adversely affected animal populations. In numerous sites throughout North America, the Bølling–Ållerod-YD sedimentary boundary is characterized by a black organic-containing layer (often termed "black mat") (9, 17), whose base is near the boundary and is interpreted by impact proponents as a hemisphere-wide deposit from wildfires (13–16). Elevated concentrations (with respect to overlying and underlying sediments) of impact markers are reported in YD black mats from 10 Clovis-age archaeological sites and 15 Carolina bays on the Atlantic Coastal Plain: specifically, iridium, nickel, magnetic microspherules, fullerenes enriched in trapped ^3He, charcoal/soot, carbon spherules, glass-like carbon, and nanodiamonds (13–16). Among these reported markers is the rare hexagonal (2H) polytype of diamond, lonsdaleite (16). Lonsdaleite is often associated with shock pressures related to impacts where it has been found to occur naturally; see refs. 18–21. Many of the impact markers reported in YD black mats have been widely discredited (22–27) with the exception of the enigmatic nanodiamonds.

Results and Discussion

To investigate the presence and nature of Bølling–Ållerod-YD-boundary nanodiamonds, we microcharacterized the carbon allotropes in carbonaceous materials from the base of YD black mats and other dated sources using transmission electron microscopy (TEM). The carbonaceous phases in carbon spherules and microcharcoal isolated from YD black mats are identical to those in spherules and glassy carbon isolated from older sediments as well as obtained from a modern forest fire. All specimens were predominantly C and contained the same dominant minerals: amorphous carbon (a-C), graphene, graphene/graphane, and graphite (all but the former displaying varying degrees of disorder). Neither cubic nor hexagonal diamond was identified in any of the samples. Trace amounts of submicrometer to nanometer-sized minerals were observed, including Fe and Cu oxides as well as Ti-, Si-, and/or Ca-rich grains.

The dominant crystalline carbonaceous phase in carbon spherules, microcharcoal, and glassy carbon was graphene in the form

GEOLOGY

Author contributions: T.L.D., N.P., and A.C.S. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: tdaulton@physics.wustl.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1003904107/-/DCSupplemental.

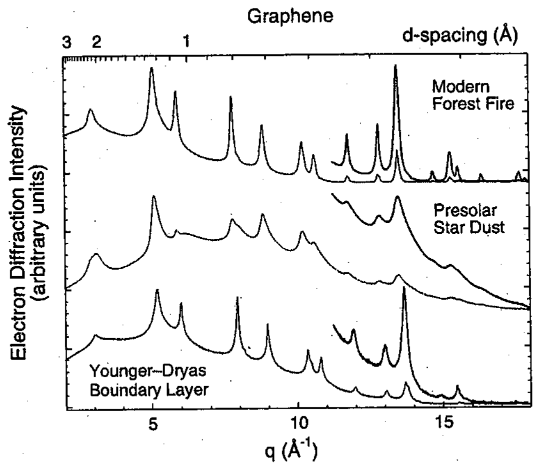


Fig. 1. Radial mean of electron diffraction intensity for graphene (thin lines) of modern origin (charred fungal Sclerotium), of supernova origin [cores of graphitic spherules isolated from the carbonaceous chondrite meteorite Murchison (30)], and from the onset of the Younger–Dryas (carbon spherules isolated from Santa Rosa Island, CA). Diffraction patterns were also acquired with longer exposure times to measure faint, high-angle peaks (thick lines). Graphene in microcharcoal from Murray Springs, AZ (~12,900 cal yr B.P.) and in carbon spherules as well as glassy carbon from Santa Cruz Island, CA (15,498–16,209 cal yr B.P.) are nearly identical to above.

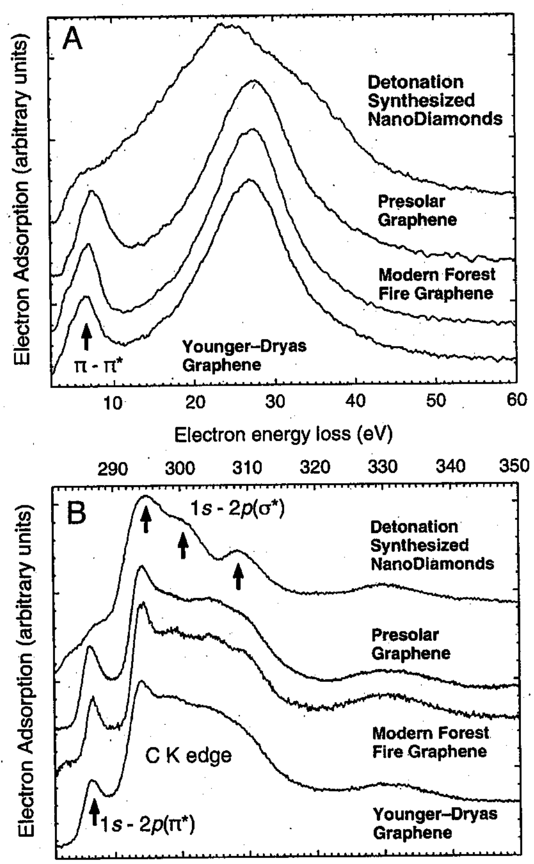


Fig. 2. Comparison of low-loss (*A*) and core-loss (*B*) electron energy loss spectra acquired from specimens overhanging holes in the support film (EELS collection half angle $2\beta = 6.34 \pm 0.06$ mrad). Shown are detonation synthesized nanodiamonds as well as graphene of supernova origin [Murchison meteorite (30)], of modern origin (charred fungal Sclerotium), and from the onset of the Younger–Dryas (Santa Rosa Island, CA). See Fig. 1 caption for expanded details. Arrows indicate peaks associated with π-π^* and 1*s*-2*p*(π^*) sp^2 graphitic transitions as well as 1*s*-2*p*(σ^*) sp^3 diamond transitions.

of polycrystalline aggregates (Figs. 1–3 and Table 1). Graphene is a two-dimensional, single-atom-thick planar molecule with sp^2-bonded carbon (1.42 ± 0.1 Å bond length) in a hexagonal arrangement of 2.46 ± 0.02 Å edge length (28, 29). In the form of a polycrystalline aggregate, as first observed in the cores of many circumstellar graphite spherules isolated from chondritic meteorites (30), graphene sheets are randomly oriented and lack any correlation. When periodically stacked normal to their plane (e.g., *AB*, *AA*, or *ABC* stacking), graphene sheets form various graphite polytype structures or turbostratic graphite if the stacking is disordered.

Core-loss absorption edges were measured using TEM electron energy loss spectroscopy (EELS) to determine the elements present, quantify their concentration, and characterize their local bonding. Elemental maps acquired using EELS spectrum imaging of core-loss edges demonstrate that terrestrial graphene aggregates contain predominantly C, but also heterogeneous distributions of O at upward of ≈6 at. % indicating partial and inhomogeneous oxidation. Discrete tens-of-nanometer diameter Ca- and O-rich grains were sometimes observed embedded within graphene aggregates. The near-edge structure of EELS core-loss spectra is highly dependent on the local bonding of an element and is sensitive to valence, crystal field splitting (e.g., atom coordination and low- or high-spin configurations), spin-orbit interactions, atomic Coulomb repulsion, and exchange effects. Low-loss EELS spectra are also dependent on local bonding. Further confirming the grains identified as (sp^2-bonded) graphene are not diamond, both the low- and core-loss EELS spectra of graphene are distinct from that of sp^3-bonded diamond (Fig. 2).

For all specimens examined, some graphene aggregates exhibited diffraction lines that were asymmetric (indicative of texturing) and doubled (Figs. 3 and 4). The extra set of reflections showed varying degrees of diffuseness (i.e., disorder). For some aggregates, their second set of reflections could be isolated from

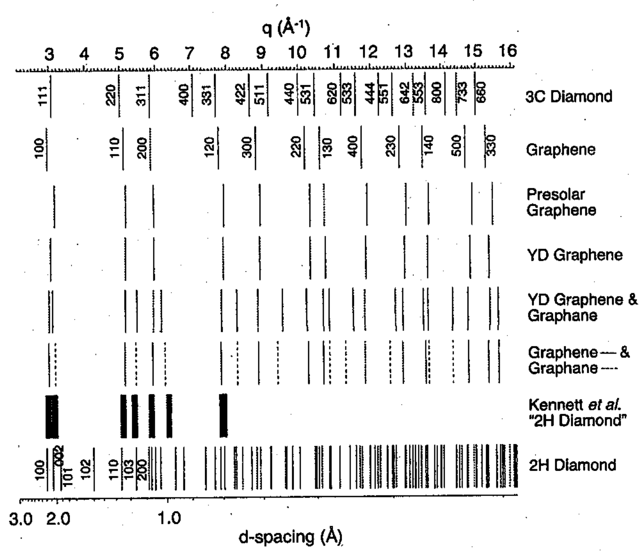


Fig. 3. Electron diffraction peaks calculated for 3C cubic diamond, 2H hexagonal diamond, graphene, and graphene/graphane mixture compared to those measured from presolar graphene as well as graphene and graphene/graphane aggregates in carbon spherules isolated from Santa Rosa Island, CA. In the electron diffraction patterns identified as hexagonal diamond by Kennett et al. (16), asymmetrically doubled diffraction lines are clearly evident in their Fig. S2*B* as well as discernible in both their Fig. 2*F* and Fig. S2*A*. Peaks measured from the doubled diffraction lines in Fig. S2*B* of Kennett et al. (16) are shown (we calibrated the reported {100} reflection to 2.189 Å, and the line widths represent the error in our measurement). Kennett et al. (16) identified their Fig. S2*C* as hexagonal diamond; however, it is also consistent with diffraction normal to the graphite basal plane; graphite {100} (2.139 Å) is close to that of lonsdaleite {100} (2.182 Å).

Table 1. Electron diffraction planar spacings

3C Diamond		Graphene/graphane—oxide					
Indices	Calculated, Å	Indices	Calculated graphene, Å	Presolar graphene, Å*	Younger–Dryas graphene, Å*,†	Younger–Dryas graphane, Å*,†,‡	Calculated graphane, Å
111	2.053	100	2.130	2.033 (6)	2.076 (4)	2.004 (7)	2.021
220	1.257	110	1.230	1.230 (2)	1.222 (2)	1.158 (2)	1.167
311	1.072	200	1.065	1.069 (3)	1.061 (2)	0.991 (3)	1.010
400	0.889						
331	0.816	120	0.805	0.807 (2)	0.798 (1)	0.754 (2)	0.764
422	0.726	300	0.710	0.709 (2)	0.705 (1)	0.657 (1)	0.674
511/333	0.684						
440	0.629	220	0.615	0.616 (1)	0.609 (1)	0.575 (2)	0.583
531	0.601	130	0.591	0.593 (1)	0.584 (1)	0.547 (5)	0.561
620	0.562						
533	0.542	400	0.533	0.534 (1)	0.525 (1)	0.496 (2)	0.506
444	0.513						
551/711	0.498	230	0.489	0.489 (1)	0.479 (1)	0.460 (1)	0.464
642	0.475						
553/731	0.463	140	0.465	0.465 (1)	0.455 (1)	0.436 (6)	0.441
800	0.445						
733	0.434	500	0.426	0.427 (1)	0.418 (1)	0.400 (1)	0.404
660/822	0.419	330	0.410	0.410 (1)	0.397 (1)	0.381 (1)	0.389

*Detonation synthesized nanodiamonds were used to calibrate diffraction camera length of microscope. Values in parentheses are the measurement error (in the least significant digit) based on standard error of replicate measurements and the error in camera length calibration (±0.2%).
†Hexagonal edge length varies slightly from grain to grain.
‡Measured from graphene/graphane aggregates.

the first using a small selected-area aperture (Fig. 4), and EELS showed those regions were also predominantly C with varying O concentrations, demonstrating the presence of a modified form of graphene. The modified graphene exhibited a $5.1 \pm 0.3\%$ (see Table 1) contraction in hexagonal edge length, although the contraction varied somewhat from aggregate to aggregate. This is consistent with the previously theorized but only recently synthesized hydrogenated form of graphene, termed graphane (29), which exhibits a 5% reduction in edge length resulting from buckling of C bonds out of the plane of the C sheet due to H bonding on sheet faces. The third-most abundant crystalline carbonaceous phase was graphite with various degrees of graphene-sheet stacking disorder.

Careful analysis must be exercised in identifying diamond polytypes in carbonaceous specimens. Graphene diffraction lines closely resemble those of cubic diamond for small Bragg angle (Fig. 3 and Table 1). The first five diffraction spacings of graphene approximately mirror those of cubic diamond with the notable exception that graphene lacks analogous diffraction intensity to that from {400} diamond planes. Differences between the two structures become more pronounced at larger Bragg angles. It is possible that graphene aggregates, which are ubiquitous in carbon spherules, microcharcoal and glassy carbon, were misidentified as cubic diamond in previous studies of the YD-boundary deposits (13, 15, 16) because no {400} reflections (or EELS spectra) were reported and diffraction lines were measured only over a small range of Bragg angles (15, 16).

Also reported in YD-boundary deposits (15, 16) is the *proposed n*-diamond (or fcc carbon) modification of diamond; see (31, 32). No *n* diamond was observed in any of our specimens.

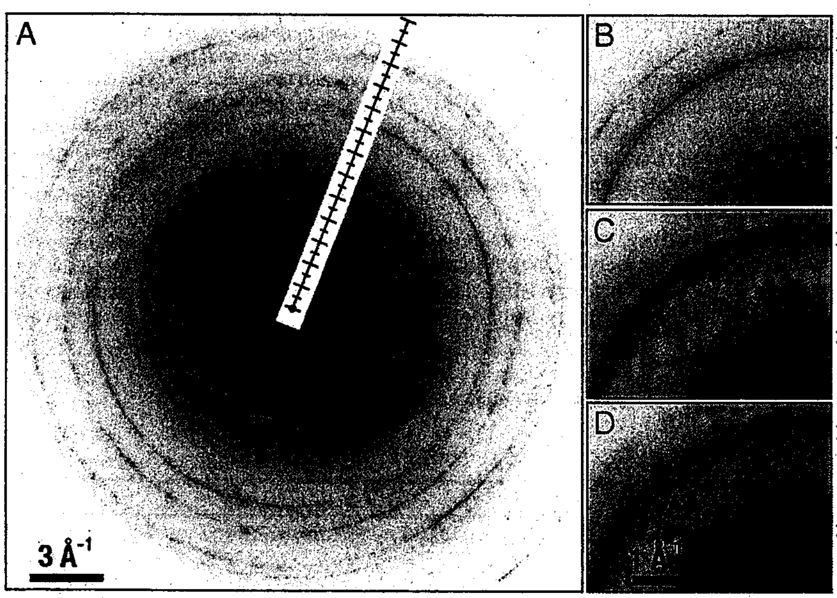


Fig. 4. Typical electron diffraction pattern from graphene/graphane aggregates exhibits (*A*) asymmetrical electron diffraction rings indicative of texturing (i.e., nonhomogeneous dispersion) of two phases. Different regions within the *same* aggregate exhibit diffraction rings from (*B*) only graphene (solid triangles), (*C*) both graphene and graphane, or (*D*) only graphane (open triangles). To aid in visual clarity, the diffraction patterns are displayed in reverse contrast.

In the YD Arlington specimen, we identified an aggregate of nanocrystalline Cu, and in several of our specimens we identified nanocrystalline Cu_2O. These minerals can be misidentified as several of the proposed polytypes/modifications of diamond. Copper (space group 225, $a = 3.6149$ Å) has the same diffraction lines as n diamond with planar spacings (i.e., Cu {111} at 2.087 Å) differing by <1.7% from diamond. Further, Cu nanocrystals mounted on standard a-C coated Cu TEM grids can be misidentified as pure C by TEM energy dispersive X-ray spectroscopy. Copper oxide, Cu_2O (space group 201, $a = 4.267$), has diffraction lines very similar to those of the *proposed* i-carbon modified diamond structure [occasionally reported synthesized along with n diamond (31)].

Although diffraction lines of 2H hexagonal diamond are distinctly different from those of the major carbonaceous phases identified in YD-boundary carbons (see Table 1), previous researchers (16) misidentified graphene/graphane aggregates as 2H hexagonal diamond in YD black mats. Diffraction patterns reported as 2H diamond by Kennett et al. (16) are clearly missing many relatively intense 2H diamond reflections, e.g., the (101) and (102) (18–20). Furthermore, close inspection of those patterns reveals asymmetric doubled diffraction lines that match closely to those of graphene/graphane aggregates and are inconsistent with 2H diamond (Fig. 3).

Conclusion

Our work emphasizes that rigorous analysis of electron diffraction patterns must be performed together with appropriate elemental quantification or other supplemental structural analysis (e.g., EELS or Raman spectroscopy) for the proper identification of diamond polytypes in carbonaceous materials. Unfortunately, many TEM studies of reported nanometer- to submicrometer-sized diamond polytypes in the literature do not perform such a definitive microanalysis. We demonstrate that previous studies of YD-boundary sediments (15, 16) clearly misidentified graphene/graphane aggregates, shown here to be ubiquitous in several types of carbon-rich materials from sediments (dated from before the YD to the present), as 2H diamond and may have misidentified graphene as cubic diamond. Importantly, we observe no nanodiamonds in any Bølling–Ållerod-YD-boundary sediments examined in this study. It is possible nanodiamonds occur inhomogeneously and only in some of the YD-boundary carbons and hence are not observed in our study. However, Kennett et al. (16) state that "lonsdaleite crystals at Arlington co-occur with carbon spherules and other diamond polymorphs…" and describe the occurrence of nanodiamonds in carbon spherules with "…a TEM study revealed conspicuous subrounded, spherical, and octahedral crystalline particles (2-300 nm) distributed in their carbonaceous matrices…. Analysis of the particles by electron diffraction shows reflections consistent with cubic diamonds…."

The usefulness of cubic nanodiamonds as impact markers in sediments remains unclear because processes other than impact might account for them. Diffraction, supported by EELS and Raman spectroscopy, identified submicrometer (and perhaps nanometer-sized) cubic diamond in an indeterminate population of carbon spherules isolated in upper soils from various sites in Germany and Belgium (33). No links to impact structures have been established, and the origins of these diamonds remain unclear. The reported presence of 2H hexagonal diamond in Bølling–Ållerod-YD-boundary sediments (16) represented the strongest evidence suggesting possible shock processing and an YD impact event. However, we demonstrate that nondiamond carbonaceous minerals were misidentified as 2H diamond in the previous study (16). The observation that morphologically similar carbon spherules occur throughout late Pleistocene to modern sediments (27) together with our results that YD-boundary carbons lack diamonds (particularly the 2H polytype) and are mineralogically similar to older as well as modern spherules casts significant doubt upon the YD impact hypothesis.

Materials and Methods

Microcharcoal aggregates (<63 μm size separate) were isolated from the base of black mat sediment layer at the same locality and stratum (Murray Springs, AZ) reported to contain cubic nanodiamonds (15). Although we did not carbon-date our Murray Springs specimen, Haynes et al. (34) have obtained consistent dates for the base of the black mat throughout the Murray Springs site of 10,260 ± 430 B.P. (calibrated: 11,358–12,546 cal yr B.P., 1σ range), 11,000 ± 100 B.P. (calibrated: 12,737–13,061 cal yr B.P., 1σ range), 10,410± 190 B.P. (calibrated: 12,005–12,559 cal yr B.P., 1σ range), respectively. Therefore, we assume that our specimen dates to be of the same range and are representative of the Bølling–Ållerod-YD boundary. Carbon spherules were isolated from the same locality (Arlington Canyon, Santa Rosa Island, CA) reported to contain hexagonal nanodiamonds (16). Kennett et al. (14, 16) dated the entire basal 5 m of the Arlington YD sequence within the 1σ range 13,100–12,830 cal yr B.P. and reported nanodiamonds in the deepest meter-thick layer (16). In contrast, we obtained calibrated radiocarbon dates spanning >5,000 years over that same 5-m sequence. From that sequence, we examined two specimens for nanodiamonds (from the lowest meter) dating to 12,766–13,044 and 13,379–13,560 cal yr B.P. (1σ range). On neighboring Santa Cruz Island, carbon spherules and glassy carbon were collected from Bølling–Ållerod sediments dated at 15,498–16,209 cal yr B.P. (1σ range) in Sauces Canyon. Our radiocarbon ^{14}C dates were analyzed by the University of California at Irvine Accelerator Mass Spectrometry lab and were calibrated using *Calib* v6.0 calibration software (35, 36); see Table 2. Charred fungal sclerotia were collected from a modern (2006) low-intensity fire at Thursley Bog in Surrey, southern England, and were strikingly similar in morphological form to Pleistocene- as well as Holocene-age carbon spherules (27). All carbonaceous grains were extracted from the collected sediments using a combination of dilute (10%) hydrogen peroxide to break down the clays, followed by digestion in hydrofluoric acid; for further details, see ref. 27.

Several particles (i.e., carbon spherules, microcharcoal, or glassy carbon aggregates) from each specimen were crushed between quartz disks (without solvent), and the fine powder was mounted directly on holey a-C film-coated TEM grids by gently placing the TEM support film in physical contact with the powder. Between 262 and 545 submicrometer-sized particles per TEM mount (2,200 particles total) were individually characterized by TEM selected-area electron diffraction to identify their structure. Elemental and supplemental structural analysis of representative grains from the different structural types was performed using EELS. Specimens were analyzed using a 200-kV JEOL JEM-2100F field emission scanning transmission electron microscope (at Washington University), equipped with high-resolution pole piece and a Schottky field emission gun. The instrument is equipped with a Gatan Tridiem imaging filter capable of performing EELS, EELS spectrum imaging, and electron energy-fil-

Table 2. ^{14}C radiocarbon and calibrated ages

Laboratory/ specimen number	Height (m above datum)	Material dated	^{14}C age, yr	Calibrated age, cal yr B.P.: median*	Calibrated age, cal yr BP: 1-sigma*	Calibrated age, cal yr BP: 2-sigma*
Santa Cruz Island Sauces Canyon						
UCIAMS 46051/SCI-07-P4	3.37	Organics from centrifuge	13,080 ± 30	15,818	15,498–16,209	15,221–16,402
Santa Rosa Island Middle Arlington Canyon						
UCIAMS 66950/SRI-09028	Near basal	Charred wood	11,020 ± 25	12,890	12,766–13,044	12,718–13,079
UCIAMS 66951/SRI-09-29c	Near basal	Charcoal	11,625 ± 25	13,456	13,379–13,560	13,341–13,619

*Calibrated using *Calib* v6.0 Radiocarbon Calibration software (35, 36).

tered imaging. EELS spectra were collected in the diffraction mode of the transmission electron microscope (i.e., image coupling to the EELS spectrometer) and were corrected for dark current and channel-to-channel gain variation of the CCD detector array. To quantify elemental compositions, EELS partial cross-sections were calculated from Hartree–Slater models.

ACKNOWLEDGMENTS. We thank T. J. Bernatowicz and K. Croat for presolar graphite specimens as well as valuable discussions. This work was partially funded by the Center for Materials Innovation at Washington University, National Geographic Society, National Science Foundation (EAR-0746015), and the Royal Holloway strategy fund.

1. Björck S, et al. (1996) Synchronized terrestrial-atmospheric deglacial records around the North Atlantic. *Science* 274:1155–1160.
2. Alley RB, et al. (1993) Abrupt increase in Greenland snow accumulation at the end of the Younger Dryas event. *Nature* 362:527–529.
3. Alley RB, Clark PU (1999) The deglaciation of the Northern Hemisphere: A global perspective. *Annu Rev Earth Planet Sci* 27:149–182.
4. Broecker WS (2003) Does the trigger for abrupt climate change reside in the ocean or in the atmosphere? *Science* 300:1519–1522.
5. Muscheler R, et al. (2008) Tree rings and ice cores reveal ^{14}C calibration uncertainties during the Younger Dryas. *Nat Geosci* 1:263–267.
6. Severinghaus JP, Sowers T, Brook EJ, Alley RB, Bender ML (1998) Timing of abrupt climate changes at the end of the Younger Dryas interval from thermally fractionated gases in polar ice. *Nature* 391:141–146.
7. Grayson DK (2007) Deciphering North American Pleistocene extinctions. *J Anthropol Res* 63:185–213.
8. Faith JT, Surovell TA (2009) Synchronous extinction of North America's Pleistocene mammals. *Proc Natl Acad Sci USA* 106:20641–20645.
9. Haynes CV, Jr (2008) Younger Dryas "black mats" and the Rancholabrean termination in North America. *Proc Natl Acad Sci USA* 105:6520–6525.
10. Buchanan B, Collard M, Edinborough K (2008) Paleoindian demography and the extraterrestrial impact hypothesis. *Proc Natl Acad Sci USA* 105:11651–11654.
11. Teller JT, Leverington DW, Mann JD (2002) Freshwater outbursts to the oceans from glacial Lake Agassiz and their role in climate change during the last deglaciation. *Quaternary Sci Rev* 21:879–887.
12. Clement AC, Cane MA, Seager R (2001) An orbitally driven tropical source for abrupt climate change. *J Climate* 14:2369–2375.
13. Firestone RB, et al. (2007) Evidence for an extraterrestrial impact 12,900 years ago that contributed to the megafaunal extinctions and the Younger Dryas cooling. *Proc Natl Acad Sci USA* 104:16016–16021.
14. Kennett DJ, et al. (2008) Wildfire and abrupt ecosystem disruption on California's Northern Channel Islands at the Ållerød–Younger Dryas boundary (13.0–12.9 ka). *Quaternary Sci Rev* 27:2530–2545.
15. Kennett DJ, et al. (2009) Nanodiamonds in the Younger Dryas boundary sediment layer. *Science* 323:94.
16. Kennett DJ, et al. (2009) Shock-synthesized hexagonal diamonds in Younger Dryas boundary sediments. *Proc Natl Acad Sci USA* 106:12623–12628.
17. Quade J, Forester RM, Pratt WL, Carter C (1998) Black mats, spring-fed streams, and late-glacial-age recharge in the Southern Great Basin. *Quaternary Res* 49:129–148.
18. Bundy FP, Kasper JS (1967) Hexagonal diamond—A new form of carbon. *J Chem Phys* 46:3437–3446.
19. Hanneman RE, Strong HM, Bundy FP (1967) Hexagonal diamonds in meteorites: Implications. *Science* 155:995–997.
20. Frondel C, Marvin UB (1967) Lonsdaleite, a hexagonal polymorph of diamond. *Nature* 214:587–589.
21. Erlich EI, Hausel WD (2002) *Diamond Deposits: Origin, Exploration, and History of Discovery* (Society for Mining, Metallurgy, and Exploration, Littleton, CO).
22. Marlon JR, et al. (2009) Wildfire response to abrupt climate change in North America. *Proc Natl Acad Sci USA* 106:2519–2524.
23. Paquay FS, et al. (2009) Absence of geochemical evidence for an impact event at the Bølling–Allerød/Younger Dryas transition. *Proc Natl Acad Sci USA* 106:21505–21510.
24. Surovell TA, et al. (2009) An independent evaluation of the Younger Dryas extraterrestrial impact hypothesis. *Proc Natl Acad Sci USA* 106:18155–18158.
25. Kaiser K, et al. (2009) Palaeopedological marker horizons in northern central Europe: Characteristics of Lateglacial Usselo and Finow soils. *Boreas* 38:591–609.
26. Haynes CV, Jr, et al. (2010) The Murray Springs Clovis site, Pleistocene extinction, and the question of extraterrestrial impact. *Proc Natl Acad Sci USA* 107:4010–4015.
27. Scott AC, et al. (2010) Fungus, not comet or catastrophe, accounts for carbonaceous spherules in the Younger Dryas 'impact layer'. *Geophys Res Lett* 37:L14302, 10.1029/2010GL043345.
28. Geim AK, Novoselov KS (2007) The rise of graphene. *Nat Mater* 6:183–191.
29. Elias DC, et al. (2009) Control of graphene's properties by reversible hydrogenation: Evidence for graphane. *Science* 323:610–613.
30. Bernatowicz TJ, et al. (1996) Constraints on stellar grain formation from presolar graphite in the Murchison Meteorite. *Astrophys J* 472:760–782.
31. Yamada K, Sawaoka AB (1994) Very small spherical crystals of distorted diamond found in a detonation product of explosive/graphite mixtures and their formation mechanism. *Carbon* 32:665–673.
32. Konyashin I, et al. (2006) A new hard allotropic form of carbon: Dream or reality? *Int J Refract Met H* 24:17–23.
33. Yang ZQ, et al. (2008) TEM and Raman characterisation of diamond micro- and nanostructures in carbon spherules from upper soils. *Diam Relat Mater* 17:937–943.
34. Haynes CV, Jr (2007) *Murray Springs: A Clovis Site with Multiple Activity Areas in the San Pedro Valley, Arizona*, eds CV Haynes, Jr and BB Huckell (Univ Arizona Press, Tucson), pp 229–239.
35. Stuiver M, Reimer PJ (1993) Extended ^{14}C data base and revised CALIB 3.0 ^{14}C age calibration program. *Radiocarbon* 35:215–230.
36. Reimer PJ, et al. (2009) INTCAL09 and MARINE09 radiocarbon age calibration curves, 0–50,000 years cal BP. *Radiocarbon* 51:1111–1150.

Supporting Information

Daulton et al. 10.1073/pnas.1003904107

Table S1. Similarities between fcc-Cu and Cu_2O with the proposed *n*-diamond and *i*-carbon structures

"*n* diamond"						fcc-Cu space group 225 $a = 3.6149$ Å		Cu_2O space group 201 $a = 4.266$ Å		"*i* carbon"			
Indices	Calculated, Å*	Reported, Å (1)	Reported, Å (2)	Reported, Å (3, 4)	Reported, Å (5)	Indices	Calculated, Å	Indices	Calculated, Å (6)	Reported, Å (1)	Reported, Å (2)	Reported, Å (3)	Reported, Å (4)
111	2.053	2.08	2.10	2.06	2.067	111	2.087	110	3.020	3.04	3.02	3.03	3.04
200	1.778	1.80	1.81	1.78	1.791	200	1.808	111	2.463	2.55	2.48	2.49	2.42
220	1.257	1.28	1.28	1.26	1.261	220	1.278	200	2.133	2.12	2.12	2.13	2.08
311	1.072	1.09	1.09	1.07	1.078	311	1.090	211	1.742	1.82	1.77	1.78	1.70
222	1.027	1.05	1.05	1.04	1.032	222	1.044	220	1.510	1.51	1.51	1.59	1.49
400	0.889	0.903		0.898	0.892	400	0.904	311	1.286	1.30	1.29	1.29	1.26
331	0.816	0.824	0.829	0.818	0.817	331	0.830	222	1.232				1.19
420	0.795	0.808	0.808	0.796	0.796	420	0.808	400	1.067	1.10		1.09	
422	0.726	0.735		0.726	0.727	422	0.738	331	0.978			1.05	
511/333	0.684	0.673		0.683 (4)	0.686	511/333	0.695	420	0.954			0.916	
440	0.629				0.630	440	0.639	422	0.870	0.906		0.847	
531	0.601				0.601	531	0.611	511	0.821	0.835			
600/420	0.593					600/420	0.603						
620	0.562					620	0.571						
533	0.542					533	0.551						
622	0.536					622	0.545						
444	0.513					444	0.522						
551/711	0.498					551/711	0.506						

*3C diamond reflections including kinematically forbidden reflections.

1. Yamada K, Sawaoka AB (1994) Very small spherical crystals of distorted diamond found in a detonation product of explosive/graphite mixtures and their formation mechanism. *Carbon* 32:665–673.
2. Burkhard G, et al. (1994) Carbon phase transition by dynamic shock compression of a copper/graphite powder mixture. *Jpn J Appl Phys* 33:L876–L879.
3. Djerdj I, et al. (2006) Transmission electron microscopy study of carbon nanophases produced by ion beam implantation. *Mat Sci Eng C-Bio S* 26:1202–1206.
4. Hirai H, Kondo K-I (1991) Modified phases of diamond formed under shock compression and rapid quenching. *Science* 253:773–774.
5. Konyashin I, et al. (2006) A new hard allotropic form of carbon: Dream or reality? *Int J Refract Met H* 24:17–23.
6. Golden TD et al. (1996) Electrochemical deposition of Copper(I) oxide films. *Chem Mater* 8:2499–2504.

Table S2. Similarities between graphene with the proposed *p*-diamond structure

Graphene				"*p* diamond"		
Indices	Calculated, Å	Presolar, Å*	Younger–Dryas, Å* (1)	Indices	Reported, Å (2)	Reported, Å (3, 4)
100	2.130	2.033 (6)	2.076 (4)	111	2.08	2.08
110	1.230	1.230 (2)	1.222 (2)	221	1.20	1.21
200	1.065	1.069 (3)	1.061 (2)	222	1.04	1.05
120	0.805	0.807 (2)	0.798 (1)	412	0.79	0.791
300	0.710	0.709 (2)	0.705 (1)	413	0.70	0.703
				333	0.69	0.672
220	0.615	0.616 (1)	0.609 (1)	334	0.62	
130	0.591	0.593 (1)	0.584 (1)	703	0.59	
400	0.533	0.534 (1)	0.525 (1)			
230	0.489	0.489 (1)	0.479 (1)			
140	0.465	0.465 (1)	0.455 (1)			
500	0.426	0.427 (1)	0.418 (1)			
330	0.410	0.410 (1)	0.397 (1)			

*Detonation synthesized nanodiamonds were used to calibrate diffraction camera length of microscope. Values in parenthesis are the measurement error (in the least significant digit) based on standard error of replicate measurements and the error in camera length calibration (±0.2%).

1. Hexagonal edge length varies slightly from grain to grain.
2. Kleiman J, et al. (1984) Shock compression and flash heating of graphite/metal mixtures at temperatures up to 3200 K and pressures up to 25 GPa. *J Appl Phys* 56:1440–1454.
3. Yamada K, et al. (1999) Nanostructure and formation mechanism of proto diamond shock-synthesized from graphite. *Carbon* 37:275–280.
4. Yamada K, Tanabe Y, Sawaoka, AB (2000) Allotropes of carbon shock synthesized at pressures up to 15 GPa. *Philos Mag A* 80:1811–1828.

PNAS

Shock-synthesized hexagonal diamonds in Younger Dryas boundary sediments

Douglas J. Kennett[a,1], James P. Kennett[b], Allen West[c], G. James West[d], Ted E. Bunch[e], Brendan J. Culleton[a], Jon M. Erlandson[a,f], Shane S. Que Hee[g], John R. Johnson[h], Chris Mercer[i,j], Feng Shen[k], Marilee Sellers[e], Thomas W. Stafford, Jr.[l], Adrienne Stich[m], James C. Weaver[n], James H. Wittke[e], and Wendy S. Wolbach[m]

[a]Department of Anthropology and [f]Museum of Natural and Cultural History, University of Oregon, Eugene, OR 97403; [b]Department of Earth Science and Marine Science Institute and [j]Materials Department, University of California, Santa Barbara, CA 93106; [c]GeoScience Consulting, Dewey, AZ 86327; [d]Department of Anthropology, University of California, Davis, CA 95616; [e]Department of Geology and Imaging and Histology Core Facility, Northern Arizona University, Flagstaff, AZ 86011; [g]Department of Environmental Health Sciences, University of California, Los Angeles, CA 90095-1772; [h]Santa Barbara Museum of Natural History, 2559 Puesta del Sol, Santa Barbara, CA 93105; [i]National Institute for Materials Science, Tsukuba, Ibaraki 305-0047, Japan; [k]FEI Company, 5350 Northeast Dawson Creek Drive, Hillsboro, OR 97124; [l]Stafford Research, Inc., 200 Acadia Avenue, Lafayette, CO 80026; [m]Department of Chemistry, DePaul University, Chicago, IL 60604; and [n]Department of Chemical and Environmental Engineering, University of California, Riverside, CA 92521

Communicated by Steven M. Stanley, University of Hawaii, Honolulu, HI, June 11, 2009 (received for review March 3, 2009)

The long-standing controversy regarding the late Pleistocene megafaunal extinctions in North America has been invigorated by a hypothesis implicating a cosmic impact at the Ållerød-Younger Dryas boundary or YDB (≈12,900 ± 100 cal BP or 10,900 ± 100 ^{14}C years). Abrupt ecosystem disruption caused by this event may have triggered the megafaunal extinctions, along with reductions in other animal populations, including humans. The hypothesis remains controversial due to absence of shocked minerals, tektites, and impact craters. Here, we report the presence of shock-synthesized hexagonal nanodiamonds (lonsdaleite) in YDB sediments dating to ≈12,950 ± 50 cal BP at Arlington Canyon, Santa Rosa Island, California. Lonsdaleite is known on Earth only in meteorites and impact craters, and its presence strongly supports a cosmic impact event, further strengthened by its co-occurrence with other nanometer-sized diamond polymorphs (n-diamonds and cubics). These shock-synthesized diamonds are also associated with proxies indicating major biomass burning (charcoal, carbon spherules, and soot). This biomass burning at the Younger Dryas (YD) onset is regional in extent, based on evidence from adjacent Santa Barbara Basin and coeval with broader continent-wide biomass burning. Biomass burning also coincides with abrupt sediment mass wasting and ecological disruption and the last known occurrence of pygmy mammoths (*Mammuthus exilis*) on the Channel Islands, correlating with broader animal extinctions throughout North America. The only previously known co-occurrence of nanodiamonds, soot, and extinction is the Cretaceous-Tertiary (K/T) impact layer. These data are consistent with abrupt ecosystem change and megafaunal extinction possibly triggered by a cosmic impact over North America at ≈12,900 ± 100 cal BP.

Arlington Canyon | biomass burning | cosmic impact | hexagonal nanodiamonds | megafaunal extinctions

Thirty-five mammal and 19 bird genera became extinct in North America near the end of the Pleistocene (1, 2). Other animal populations were severely reduced or suffered massive range restrictions (3). Most of these animals were large, although smaller animals also underwent major biogeographic changes, and some also became extinct (1, 4). Detailed extinction and biogeographic histories for many of these genera are poorly known, but at least 16 genera and several additional species became extinct abruptly over broad parts of North America close to 13,000 years ago (5, 6). In situ bones of the most common large genera on the late Pleistocene landscape [e.g., *Equus* (horses), *Camelops* (camels), and *Mammuthus* (mammoths)] occur widely in North American sedimentary sequences up to, but never above, the base of a distinctive organic-rich black sedimentary layer (5). This biostratigraphic marker dates to ≈12.9 ± 0.1 ka (10,900 ± 100 ^{14}C years)* and indicates that some of the most common genera disappeared synchronously and broadly over North America.

A long-standing debate has culminated in a polarized standoff regarding the hypothesized role of climatic change versus human predation for the demise of these animals and is "further from resolution than it has been in its 200-year history (1)." The debate continues because the available empirical evidence supports neither hypothesis very well. Climatic and vegetation shifts certainly contributed to biogeographic changes in North America during the late Pleistocene, but the animal genera involved were highly adapted to frequent climatic oscillations of equal or greater magnitude during the late Quaternary (2). The absence of kill sites for most of these genera and paucity for others is also inconsistent with the alternative hypothesis; that humans rapidly drove all of these animals into extinction. Given the technology involved, along with several other significant problems (7, 8), it is highly unlikely that humans rapidly triggered the extinction of so many genera on a continental scale, although human-caused extinction is well supported in the prehistoric record of vulnerable island animal populations (e.g., 7, 9, 10). More sophisticated models combining environmental and human induced causes (e.g., 11) are potentially viable for explaining singular mammal extinctions (e.g., *Mammuthus*), but fall short of explaining the full taxonomic depth and ecological breadth of the latest Pleistocene extinctions.

This debate has recently been challenged by a hypothesis that ecosystem disruption and widespread extinctions were triggered by multiple airbursts/impacts (by comet or carbonaceous chondrite) in North America at 12.9 ± 0.1 ka (12, 13). Although controversial when first proposed, a major cosmic impact at the Cretaceous-Tertiary (K/T) boundary is now widely accepted as the cause of one of the largest known mass extinctions (14). The connection of impact-to-extinction and the presence of several of the same impact proxies in this widespread 12,900-year-old sedimentary layer provide an empirical basis for the Younger Dryas boundary (YDB) impact

ANTHROPOLOGY

GEOLOGY

Author contributions: D.J.K., J.P.K., T.E.B., J.M.E., S.S.Q.H., F.S., and W.S.W. designed research; D.J.K., J.P.K., A.W., G.J.W., T.E.B., B.J.C., J.M.E., J.R.J., C.M., F.S., M.S., A.S., J.C.W., J.H.W., and W.S.W. performed research; S.S.Q.H. and W.S.W. contributed new reagents/analytic tools; D.J.K., J.P.K., A.W., G.J.W., T.E.B., J.M.E., C.M., F.S., T.W.S., J.C.W., J.H.W., and W.S.W. analyzed data; and D.J.K. and J.P.K. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. E-mail: dkennett@uoregon.edu.

*Ages in this paper are expressed in thousands of calendar years before present (ka). Radiocarbon ages will be identified and clearly marked "^{14}C years."

This article contains supporting information online at www.pnas.org/cgi/content/full/0906374106/DCSupplemental.

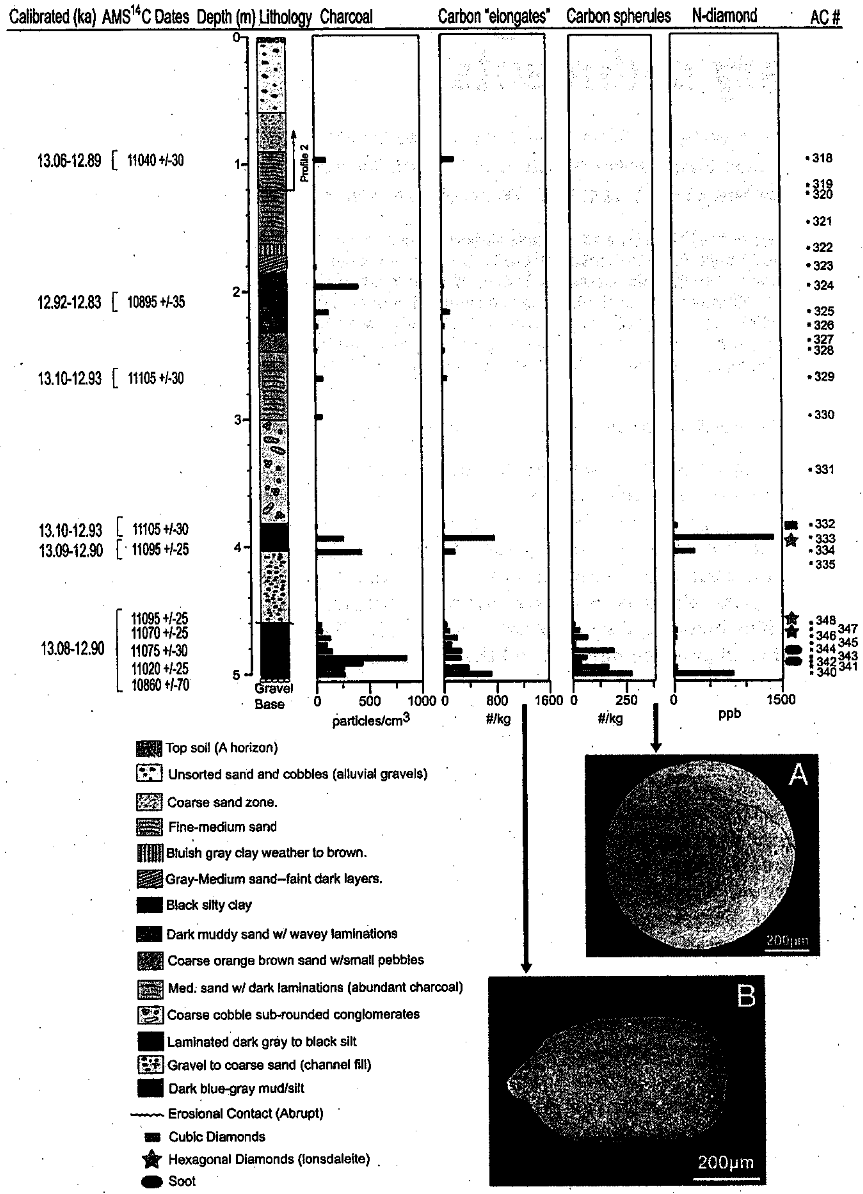


Fig. 1. Stratigraphic section of AC-003 profile (Arlington Canyon, Santa Rosa Island, CA) showing lithology, calibrated and uncalibrated radiocarbon dates, stratigraphic distribution of carbon material (charcoal, elongates, spherules, and soot), and concentrations/locations of 3 different diamond polymorphs (n-diamonds, hexagonal, and cubics). SEM images represent carbon spherules (*A*) and carbon elongates (*B*) from the lower dark sedimentary layers, both of which contain n-diamonds. * designates sediment samples analyzed in this study.

hypothesis. Massive North American animal extinctions could have resulted from the direct effects of these airbursts/impacts (shockwaves, heat, wildfires) and subsequent cascading ecological changes associated with landscape and biotic disruption, interruption of North Atlantic thermohaline circulation, abrupt climate change, and human predation on remnant animal populations. Skepticism about the YDB impact hypothesis has resulted from the absence in the YDB of some widely accepted impact markers that are present in the K/T and other documented impacts (e.g., shocked minerals, breccias, tektites, and a visible impact crater) (15). However, such markers also appear absent from an observed and widely accepted cosmic impact event, the Tunguska airburst over Siberia in 1908 (16).

Results

Here, we present evidence for shock-synthesized hexagonal nanodiamonds (lonsdaleite) in YDB sediments in North America. These diamonds occur at Arlington Canyon on Santa Rosa Island (California), which, at 12.9 ka, was joined with 3 other Northern Channel Islands to form one landmass, Santarosae (Fig. S1*B*) (17). The diamonds occur in a discrete layer that is contemporary with, and similar to, the organic-rich sedimentary layers described by Haynes across North America (see Fig. S1*A*) (5). The best known geological exposure (AC-003) of this dark sedimentary unit is 1.35 km from the modern coastline on the west side of the canyon (Universal Transverse Mercator: 10S 0762524/3764532), where a 44-cm-thick, organic-rich, dark blue-gray, silty mud (black layer) rests directly on a gravel deposit at 5 m (Fig. 1) (18). This layer is capped with a coarse cobble lag deposit (≈60 cm thick) and a second less dark layer of gray to black laminated sandy silt. The rest of the overlying sequence consists of alluvial sands and gravels. Accelerator mass spectrometry (AMS) ^{14}C dates from upper and lower parts of the sequence are statistically similar,

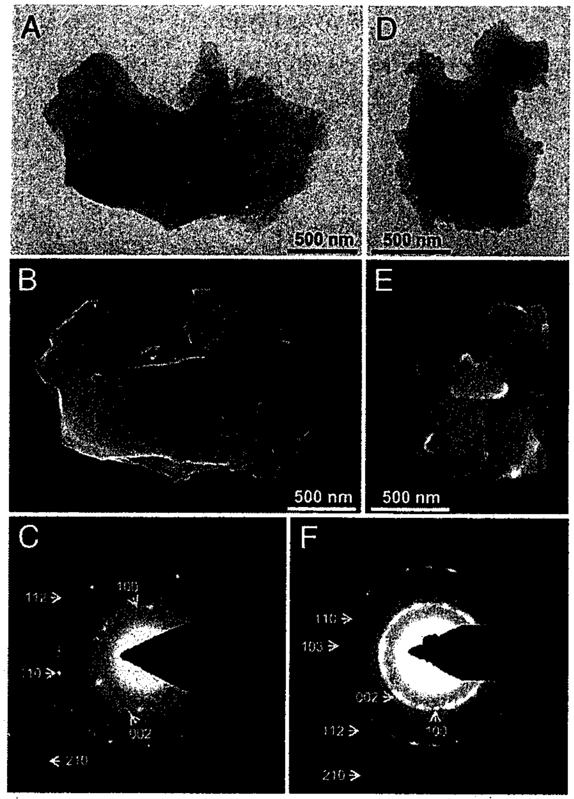


Fig. 2. TEM photomicrographs, SEM images, and diffraction patterns of stacked diamond clusters. TEM photomicrographs (*A* and *D*), SEM images (*B* and *E*), and diffraction patterns (*C* and *F*) of 2 stacked diamond clusters composed primarily of lonsdaleite crystals from the YDB (12.95 ± 0.05 ka) sedimentary layer in Arlington Canyon (AC-003). Both lonsdaleite clusters are from AC-003: 4.59–4.64 m (AC#348). See Fig. 1 for location and Figs. S2–S4 for additional images and information.

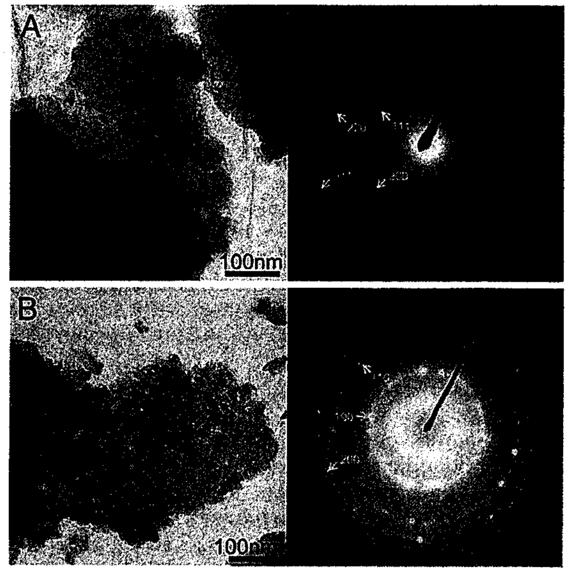


Fig. 3. Paired TEM photomicrographs (*Left*) and diffraction patterns (*Right*) of other diamond polymorphs in the YDB sedimentary layer in Arlington Canyon (AC-003). (*A*) n-diamonds embedded in carbonaceous matrix of a carbon spherule (AC#341); (*B*) a cubic diamond cluster in a carbon elongate (AC # 332). See Fig. 1 for stratigraphic position of samples.

suggesting rapid accumulation of fluvial deposits shortly after ~12.95 ± 0.05 ka (18).

Diamonds with hexagonal crystallographic modification were observed only within the lower 2 dark sedimentary layers between 3.92–4.69 m below the modern ground surface. In these stratigraphic units, we discovered both single crystals and clusters of lonsdaleite ranging in size from 20 to 1,800 nm that were either inside or adhered to elongate carbon particles common in these lowest deposits (Fig. 2 and Figs. S2–S6). Transmission electron microscopy (TEM) work demonstrates that these diamonds are mono- and polycrystalline, with some displaying individual lamellae spaced at ≈20–30 nm (Fig. S2). Electron diffraction analyses at multiple locations across these crystals confirm the hexagonal diamond polymorph with reflections corresponding to lattice planar spacings of 2.18, 1.26, 1.09, and 0.826 Å. Tabular (flake-like) morphology and stacking faults in these lonsdaleite crystals are consistent with impact-related shock transformation, in which graphite is the likely parent material (19).

The lonsdaleite in 12.95 ± 0.05 ka sediments at Arlington provides evidence for the presence in the YDB of a shock-synthesized mineral widely considered to be a cosmic impact marker. Controlled experiments indicate that graphite is transformed into a mixture of cubic and hexagonal diamond at ≈15 GPa (2 million psi) at temperatures between 1,000–1,700 °C followed by rapid quenching (19). Lonsdaleite has never been found associated with mantle-derived kimberlitic diamonds (19) and has only been found on Earth inside meteorites (20) and associated with impact craters [e.g., Popigai (21), Reis (22, 23), (Fig. S7)]. These observations make the presence of lonsdaleite an excellent shock indicator (19) and, by extension, a cosmic impact proxy when found in sediments.

The lonsdaleite crystals at Arlington co-occur with carbon spherules and other diamond polymorphs known to be concentrated in 12.9 ± 0.1 ka sediments at multiple locations across North America (13). Carbon spherules (400–1,500 μm) were found in the Arlington sequence only between 4.6–5 m below the ground surface in the lowest dark stratum (black layer) (Fig. 1). Scanning electron microscope (SEM) work shows that the carbon spherules have a characteristic reticulate interior (Fig. S6 *E* and *F*), and a TEM study revealed conspicuous subrounded, spherical, and octahedral crystalline particles (2–300 nm) distributed in their carbonaceous matrices (Fig. 3). Analysis of the particles by electron diffraction shows reflections consistent with cubic diamonds (2.06, 1.26, and 1.07 Å), as well as "forbidden" reflections at 1.78, 1.04, and 0.796 Å, indicative of a metastable "new-diamond" polymorph or n-diamond (24), which has been created in experimental TNT explosions (25), and is known from meteorites (26), but not known to be associated with mantle-derived diamonds. The forbidden reflections are attributed to flaws in the normal cubic diamond lattice (24), perhaps due to rapid quenching consistent with anoxic conditions associated with the shockwave. N-diamonds are abundant in YDB sediments across North America (13), occurring at Arlington in concentrations of ≤1,340 ppb (Fig. 1 and Table S1), equivalent to >1 billion diamonds per cm^3, a concentration comparable to those of cubic diamonds at some K/T boundary localities (3,200 ppb) (27). At Arlington, there is no evidence of upslope diamondiferous sources (e.g., older impact craters) that might have contributed reworked diamonds, and their presence cannot be due to volcanism, because high temperatures (420–570 °C) under oxidizing conditions destroy diamonds (26, 27).

N-diamonds also occur within the matrices of the carbon "elongates" associated with lonsdaleite crystals between 3.92–4.69 m (see Fig. 1). Carbon elongates differ from the carbon spherules in having an irregular array of walls and voids, whereas

ANTHROPOLOGY

GEOLOGY

carbon spherule interiors display a well-organized honeycomb (reticulated) pattern. Both types are composed entirely of glass-like amorphous carbon indicative of high-temperature formation. The general shape of elongates ranges from angular (hexagonal in cross-section) to subrounded. The elongates were found throughout the sequence, but are concentrated in the lower organic-rich layers below 3.83 m, where they are generally more highly vitrified, hard, and brittle (Fig. S5 and Fig. S6). N-diamonds were found in these elongates only below 3.83 m.

During this TEM survey, clusters of stable cubic diamonds (≈1,000 in total) were found with carbon elongates between 3.83 and 3.86 m. These diamonds appear more angular than the associated n-diamonds and have diffraction patterns typical of cubic diamonds (2.06, 1.26, 1.08 Å; Fig. 3*B*). Similar cubic diamonds have been identified in YDB sediments at Bull Creek in northwest Oklahoma and Murray Springs, Arizona (13). This diamond polymorph is also found in the ejecta layer of the K/T boundary impact (27, 28) and forms at higher temperatures and pressures than n-diamond, but the conditions of formation for all 3 polymorphs fall outside the range of Earth's known surficial processes. The co-occurrence of lonsdaleite and cubic diamonds with n-diamond at Arlington bolsters the interpretation that the n-diamond polymorph, common in YDB sediments elsewhere in North America (13), is impact-related. However, investigations are needed to determine if the lonsdaleite formed in the high pressure-temperature environment of a surface impact or were transported inside the impactor. Hexagonal diamonds are known to occur in meteorites (e.g., ref. 29; see also Fig. S7), but may not occur in sufficient abundance to account for their presence in the YDB layer. Also, because up to ≈90% of an impactor's mass may be vaporized upon impact, it seems less likely that the abundant hexagonal diamonds at Arlington arrived with the impactor and more likely that they formed upon impact. Due to their plate-like morphology, the hexagonal diamonds most likely formed via shock transformation of graphite, which does not appear to occur locally where they were discovered. However, because nanometer-sized diamonds were distributed widely through the atmosphere during the K/T impact event, it is plausible that the hexagonal diamonds arrived at Arlington after forming in a distant impact into carbon-rich target rocks. Because cosmic isotopic ratios of carbon ($\delta^{13}C$) and nitrogen ($\delta^{15}N$) often have very different values from terrestrial ratios (30), we plan to measure $\delta^{13}C$ and $\delta^{15}N$ content of these hexagonal diamonds to help determine their origin.

Discussion

Shock-synthesized hexagonal diamonds and other nanometer-sized diamond polymorphs also co-occur with high concentrations of charcoal and other forms of particulate carbon (carbon spherules and elongates) that are indicative of major biomass burning at ≈12.95 ± 0.05 ka. In addition, grape-clustered microscopic soot was found (sample AC343, 2,500 ± 250 ppm; Fig. S8 and see Fig. 1 for location) in the lowest deposits at Arlington. Soot is rare in the geological record, but is well-documented in the K/T boundary layer (65 Ma) and other cosmic impact-related deposits (31, 32). Soot is produced in flaming combustion at high temperatures consistent with intense wildfires and the reduction of conifer forests common on these islands at 12.95 ± 0.05 ka (18). The near consistency of AMS ^{14}C ages (at 12.95 ± 0.05 ka) throughout the lower 4 m of the sequence suggests that these intense fires denuded the landscape and promoted sediment mass wasting (18) that rapidly buried the diamonds and soot. The close association of shock-synthesized diamonds with a wide-range of wildfire proxies is consistent with the hypothesis that they were ignited by an intense radiation flux associated with a cosmic impact. The

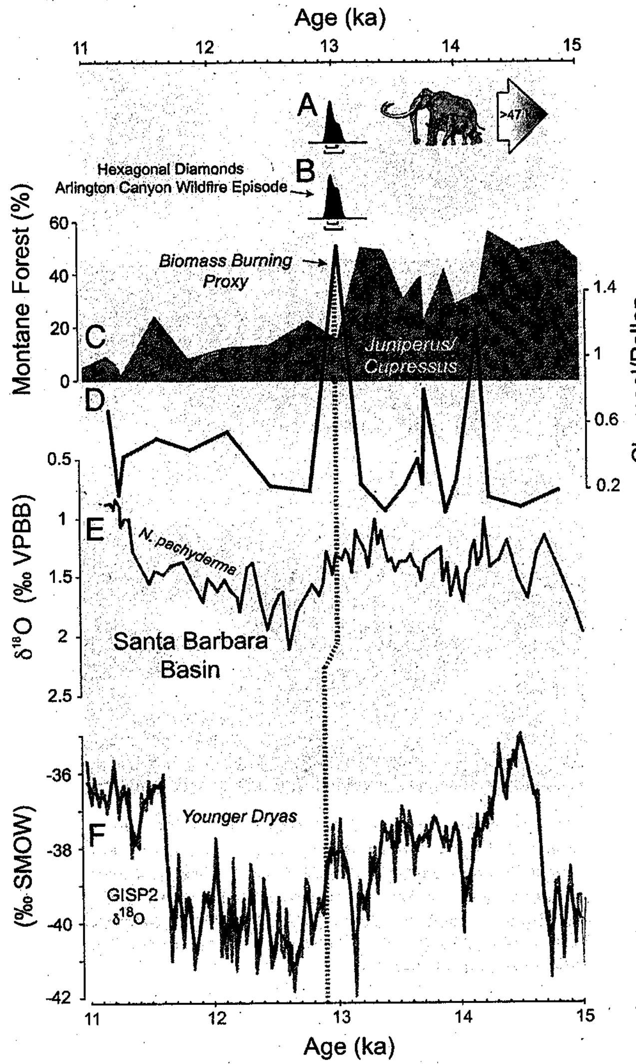


Fig. 4. Summary of climate change, vegetational shifts, and *Mammuthus exilis* (Pygmy Mammoth) extinction at the YDB shown relative to the age of hexagonal diamonds in Arlington Canyon and regional biomass burning proxies. Comparison of the known terminal age of *M. exilis* (*A*) and age of Arlington Canyon hexagonal diamonds and major YDB wildfire episode (*B*) with other proxies in Santa Barbara Basin between 15 and 11 ka; (*C*) percent abundance of montane forest types (*Juniperus-Cupressus*) (29); (*D*) ratio of pollen to charcoal concentration in Santa Barbara Basin (34); (*E*) $\delta^{18}O$ record of *Neogloboquadrina pachyderma* (35); (*F*) Comparative Greenland Ice Sheet (GISP2) $\delta^{18}O$ record (36) with heavy gauge lines representing splines of raw data. Santa Barbara Basin chronology is from ref. 35. Gray dashed line marks the initiation of the Younger Dryas cooling event in Greenland and Santa Barbara Basin, an event considered synchronous between these records. Drawing shows the size of *M. exilis* compared with *M. columbi* of the North American continent (illustration by Rusty Van Rossman).

presence of millions of diamonds embedded inside carbon spherules and the rapid burial of these deposits is inconsistent with the alternative hypothesis that micrometeoritic materials accumulated on a stable surface over an extended interval (33).

Biomass burning in Arlington Canyon is synchronous with wildfires elsewhere on the Channel Islands at ≈12.95 ± 0.05 ka (18) and contemporary with the largest of several peak burning events in the last 25,000 years recorded in nearby Santa Barbara Basin Ocean Drilling Program [Site 893 (34); Fig. 4 and see Fig. S1 for core location]. Increase in ballast minerals and higher sedimentation rates in the Santa Barbara Basin at 12.95 ± 0.05

ka (35) are also consistent with sediment mass wasting of a denuded landscape in the wake of extensive biomass burning on the adjacent islands and mainland (18). This well-dated marine sequence also records a shift to Younger Dryas (YD) cooling (36, 37) and relatively abrupt vegetational change from closed montane forest to more open habitats dominated by grasslands, chaparral, and dispersed oak stands, the dominant regional vegetation throughout the YD and Holocene (34, 38).

The unique combination of high-resolution terrestrial and marine sedimentary records from the Northern Channel Islands and adjacent Santa Barbara Basin indicate regional wildfire and abrupt ecological disruption. This disruption coincides with the last known occurrence of pygmy mammoths on Santarosae (*Mammuthus exilis*, ≈12.95 ± 0.05 ka, Fig. 4*A*) (39) and the beginning of a gap of 600–800 years in the archaeological record (18, 40). The distinctive dark layer in Arlington Canyon correlates in time with the base of other dark sedimentary layers distributed across North America that contain nanometer-sized diamonds and evidence for biomass burning (12, 13). This biomass burning appears coeval with evidence from the Greenland ice sheet for an abrupt increase in hemispheric wildfires at the beginning of the YD (41, 42, contra 43). The base of the black layer (YDB) in North America marks a major biostratigraphic change; the remains of extinct megafaunal taxa occur directly below and never above this readily discernable layer (5). This observation is consistent with radiocarbon evidence indicating abrupt extinction of mammoths and 15 other North American animal genera (5, 6). The vast majority of the North American megafaunal taxa abruptly vanished from the North American continent at the onset of the YD at 12.9 ± 0.1 ka, as marked by a major stratigraphic boundary layer that is rich in nanodiamonds (13). The uniqueness of this megafaunal extinction is highlighted by the evolutionary history of the horse, which had lineages continuously present in North America since their appearance in the early Cenozoic at about 55 Ma (44) except after 12.9 ± 0.1 ka when *Equus* suddenly disappeared from the continent (5). Before the YDB, the last known major comprehensive and abrupt megafaunal extinction recorded in North America occurred at the K/T boundary (65 Ma). Both the K/T and YDB events are uniquely similar through their association of megafaunal extinction with a distinct layer that is rich in nanodiamonds and soot. The presence of shock-synthesized hexagonal and other nanometer-sized diamonds in YDB sediments in association with soot and other wildfire indicators is consistent with a cosmic impact event at 12.9 ± 0.1 ka, and the hypothesis that the Earth crossed paths with a swarm of comets or carbonaceous chondrites producing airshocks and/or surface impacts that contributed to abrupt ecosystem disruption and megafaunal extinctions in North America.

Methods

The carbon particles (spherules, elongates, and charcoal) were separated by flotation and hand-picking. For analysis, standard techniques were used for TEM and SEM imaging, including energy dispersive x-ray spectroscopy (EDS) and electron energy loss spectroscopy (EELS). Selected particles (10–15) were collectively crushed and dispersed on a TEM grid for diamond analytical work (see *SI Methods*), and the allotropes were identified by using standard crystallographic diffraction patterns. Soot was separated and identified by using standard techniques (31).

ACKNOWLEDGMENTS. We appreciate the comments of 2 anonymous reviewers, resulting in valuable improvements to the manuscript. Early support from the Office of Research at the University of California, Santa Barbara and the University of Oregon is greatly appreciated. We thank Channel Islands National Park for logistical support; Don Morris for assistance in the field; John and Jennifer Donovan and other staff at the Lorry I. Lokey Nanotechnology Laboratories at University of Oregon, and the FEI Company for their generous support and technical assistance. This work also made use of Materials Research Laboratory Central Facilities supported by the Materials Research Science and Engineering Centers Program of the National Science Foundation under Award DMR05 and the Scanning Electron Microscope facilities of D. Morse at University of California, Santa Barbara. This research was funded by the National Science Foundation (ATM-0713769; also GRFP-2006022778 to B.J.C.) and Bray (D.J.K.) and Knight (J.M.E.) fellowships at the University of Oregon.

1. Grayson DK (2007) Deciphering North American Pleistocene extinctions. *J Anthropol Res* 63:185–213.
2. Martin PS (2005) *Twilight of the Mammoths: Ice Age Extinctions and the Rewilding of America* (University of California Press, Berkeley, CA).
3. Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol Biol Evol* 22:1185–1192.
4. Grayson DK (1987) The biogeographic history of small mammals in the Great Basin: Observations on the last 20,000 years. *J Mammal* 68:359–375.
5. Haynes CV, Jr (2008) Younger Dryas "black mats" and the Rancholabrean termination in North America. *Proc Natl Acad Sci USA* 105:6520–6525.
6. Fiedel S (2008) in *American Megafaunal Extinctions at the End of the Pleistocene*, ed Haynes G (Springer, The Netherlands), pp 21–37.
7. Grayson DK, Meltzer DJ (2003) A requiem for North American overkill. *J Archaeol Sci* 30:585–593.
8. Yule JV, Jensen CXJ, Joseph A, Goode J (2009) The puzzle of North America's Late Pleistocene megafaunal extinction patterns: Test of new explanation yields unexpected results. *Ecol Mod* 220:533–544.
9. Anderson A (1989) Mechanics of overkill in the extinction of New Zealand moas. *J Archaeol Res* 16:137–151.
10. Steadman DW (1995) Prehistoric extinctions of Pacific island birds: Biodiversity meets zooarchaeology. *Science* 267:1123–1130.
11. Barnosky AD, Koch PL, Feranec RS, Wing SL, Shabel AB (2004) Assessing the causes of Late Pleistocene extinctions on the continents. *Science* 306:70–75.
12. Firestone RB, et al. (2007) Evidence for an extraterrestrial impact 12,900 years ago that contributed to the megafaunal extinctions and Younger Dryas cooling. *Proc Natl Acad Sci USA* 104:16016–16021.
13. Kennett DJ, et al. (2009) Nanodiamonds in Younger Dryas boundary sediment layer. *Science* 323:94.
14. Alvarez LW, Alvarez W, Asaro F, Michel HV (1980) Extraterrestrial cause for the Cretaceous-Tertiary extinction. *Science* 208:1095–1108.
15. Köeberl C (2007) The geochemistry and cosmochemistry of impacts. *Treatise on Geochemistry*, ed Davis A (Elsevier, Amsterdam, The Netherlands), Vol 1, pp 1–52.
16. Longo G, Serra R, Cecchina S, Galli M (1994) Search for microremnants of the Tunguska cosmic body. *Planet Space Sci* 42:163–177.
17. Orr PC (1968) *Prehistory of Santa Rosa Island* (Santa Barbara Museum of Natural History, Santa Barbara, CA).
18. Kennett DJ, et al. (2008) Wildfire and abrupt ecosystem disruption on California's Northern Channel Islands at the Ållerød-Younger Dryas boundary (13.0–12.9 ka). *Quaternary Sci Rev* 27:2528–2543.
19. DeCarli PS, Bowden E, Jones AP, Price GD (2002) in *Catastrophic Events and Mass Extinctions: Impacts and Beyond*, eds Koeberl C, MacLeod K (Geological Society of America Special Paper 356, Boulder, CO), pp 595–605.
20. Hanneman RE, Strong HM, Bundy FP (1967) Hexagonal diamonds in meteorites: Implications. *Science* 155:995–997.
21. Köeberl C, et al. (1997) Diamonds from the Popigai impact structure, Russia. *Geology* 25:967–970.
22. Oleinik GS, Valter AA, Erjomenko GK (2003) The structure of high pressure lonsdaleite diamond grains from the impactites of the Belilovka (Zapadnaja) astrobleme (Ukraine). *34th Annual Lunar and Planetary Science Conference, March 17–21, 2003* (League City, Texas), 1561 (abstr).
23. Langenhorst F, Shafranovsky G, Masaitis VL (1998) A comparative study of impact diamonds from the Popigai, Ries, Sudbury, and Lappajärvi craters. *Meteorit Planet Sci* 33:A90 (abstract 5083).
24. Wen B, et al. (2007) Synthesis and crystal structure of n-diamond. *Int Mater Rev* 52:131–151.
25. Yamada K, Sawaoka AB (1994) Very small spherical crystals of distorted diamond found in a detonation product of explosive/graphite mixtures and their formation mechanism. *Carbon* 324:665–673.
26. Grady MM, Lee MR, Arden JW, Pillinger CT (1995) Multiple diamond components in Acfer 182. *Earth Planet Sc Lett* 136:677–692.
27. Hough RM, et al. (1995). Diamonds from the iridium-rich K-T boundary layer at Arroyo el Mimbral, Tamaulipas, Mexico. *Geology* 25:1019–1022.
28. Carlisle BC, Braman RB (1991) Nanometre-size diamonds in the Cretaceous/Tertiary boundary clay of Alberta. *Nature* 352:708–709.
29. Jenniskens P, et al. (2008) The impact and recovery of asteroid 2008 TC3. *Nature* 458:485–488.
30. Gilmour I, et al. (1992) Terrestrial carbon and nitrogen isotopic ratios from Cretaceous-Tertiary boundary nanodiamonds. *Science* 258:1624–1626.
31. Wolbach WS, Lewis RS, Anders E (1985) Cretaceous extinctions: Evidence for wildfires and search for meteoritic impact. *Science* 230:167–170.
32. Wolbach WS (1990) Carbon across the Cretaceous-Tertiary boundary. PhD thesis, (Univ of Chicago, Chicago).

33. Pinter N, Ishman SE (2008) Impacts, mega-tsunami, and other extraordinary claims. *GSA Today* 18:37–38.
34. Heusser LE, Sirocko F (1997) Millennial pulsing of environmental change in southern California from the past 24 k.y.: A record of Indo-Pacific ENSO events? *Geology* 25:243–246.
35. Nederbragt AJ, Thurow JW, Bown PR (2008) Paleoproductivity, ventilation, and organic carbon burial in the Santa Barbara Basin (ODP site 893, off California) since the last glacial. *Paleoceanography* 23:1–15.
36. Hendy IL, Kennett JP, Roark EB, Ingram BL (2002) Apparent synchroneity of submillennial scale climate events between Greenland and Santa Barbara Basin, California from 30–10 ka. *Quaternary Sci Rev* 21:1167–1184.
37. Stuiver M, Grootes PM (2000) GISP2 oxygen isotope ratios. *Quaternary Res* 53:277–283.
38. Heusser L (1998) Direct correlation of millennial-scale changes in western North America vegetation and climate with changes in the California current system over the past ~60 kyr. *Paleoceanography* 19:1–15.
39. Agenbroad LD, Johnson JR, Morris D, Stafford TW, Jr (2005) in *Proceedings of the Sixth California Islands Symposium*, eds Garcelon DK, Schwemm CA (Institute for Wildlife Studies, Arcata, CA), pp 3–7.
40. Johnson JR, Stafford TW Jr, Ajie HO, Morris DP (2002) in *Proceedings of the Fifth California Islands Symposium*, eds Browne DR, Mitchell KL, Chaney HW (Santa Barbara Museum of Natural History, Santa Barbara, CA), pp 541–545.
41. Mayewski PA, et al. (1993) The atmosphere during the Younger Dryas. *Science* 261:195–197.
42. Mayewski PA, et al. (1997) Major features and forcing of high-latitude northern hemisphere atmospheric circulation using a 110,000-year-long glaciochemical series. *J Geophys Res* 102:26345–26366.
43. Marlon JR, et al. (2009) Wildfire responses to abrupt climate change in North America. *Proc Natl Acad Sci USA* 106:2519–2524.
44. MacFadden BJ (2008) *Fossil Horses: Systematics, Paleobiology, and Evolution of the Family Equidae* (Cambridge Univ Press, New York).

PNAS

Supporting Information

Kennett et al. 10.1073/pnas.0906374106

SI Methods

Incremental sediment samples from the Arlington Canyon geological section (AC-003) were dried and weighed. Sediments were then disaggregated, and "elongate" and spherical carbon particles were extracted via flotation and hand picked using a light microscope. Carbon elongates (≈10–15) and spherules from each stratum were grouped and crushed to a powder in 3.7-mL glass vials and mixed with 4–5 drops of 100% alcohol (ETOH) to suspend the carbonaceous powder. This admixture was pipeted to a 200-mesh copper TEM grid and dried. A representative random sample of grid cells (≈5%) was scanned for diamonds with JEOL 1200EX II, JEOL 1210, or FEI Titan transmission electron microscopes. A TEM was used to identify and image single and clustered crystals, and the diamond polymorphs were identified via selected area electron diffraction. Diamonds ranged in size between 2–1,500 nm, and crystallographic work was restricted to clusters of small crystals or larger single crystals capable of producing diffraction patterns. Multiple measurements were taken across large diamonds to determine purity. Low electron beam voltages (60–80 kV) were used because n-diamonds and lonsdaleite are metastable and vaporize at high temperatures. N-diamond concentrations inside carbon spherules were calculated using the volume of bulk sediment, the average diameter of the carbon spheres in each level, the percentage containing n-diamonds, the percentage by volume of n-diamonds in carbon spherules, and the average size of n-diamonds. Standard techniques were used for scanning electron microscope imaging. EDS and EELS analysis demonstrated that the nanoparticles identified as diamonds via selected area diffraction contain only carbon.

Soot samples were extracted and analyzed using standard procedures described by Wolbach (1, 2). Sediment samples were dried, weighed, and then demineralized using alternating 9 M HCl and 10 M HF/1 M HCl treatments to dissolve carbonates and silicates. The resultant carbonaceous residue was oxidized for 600 h with 0.2 M $Na_2Cr_2O_7$/2.0 M H_2SO_4 to selectively destroy organic carbon. The elemental carbon residue was dried, weighed, and examined on the SEM, where any soot particles present were identified by their characteristic aciniform ("bunch of grapes") morphology and quantified using image particle size analysis.

1. Wolbach WS, Lewis RS, Anders E (1985) Cretaceous extinctions: Evidence for wildfires and search for meteoritic impact. *Science* 230:167–170.
2. Wolbach WS (1990) Carbon across the Cretaceous-Tertiary boundary. PhD thesis, (Univ of Chicago, Chicago).
3. Haynes CV, Jr (2008) Younger Dryas "black mats" and the Rancholabrean termination in North America. *Proc Natl Acad Sci USA* 105:6520–6525..
4. Kennett DJ, et al. (2008) Wildfire and abrupt ecosystem disruption on California's Northern Channel Islands at the Ållerød-Younger Dryas boundary (13.0–12.9 ka). *Quaternary Sci Rev* 27:2528–2543.
5. Oleinik GS, Valter AA, Erjomenko GK (2003) The structure of high pressure lonsdaleite diamond grains from the impactites of the Belilovka (Zapadnaja) astrobleme (Ukraine). *34th Annual Lunar and Planetary Science Conference, March 17-21, 2003* (League City, Texas), 1561 (abstr).

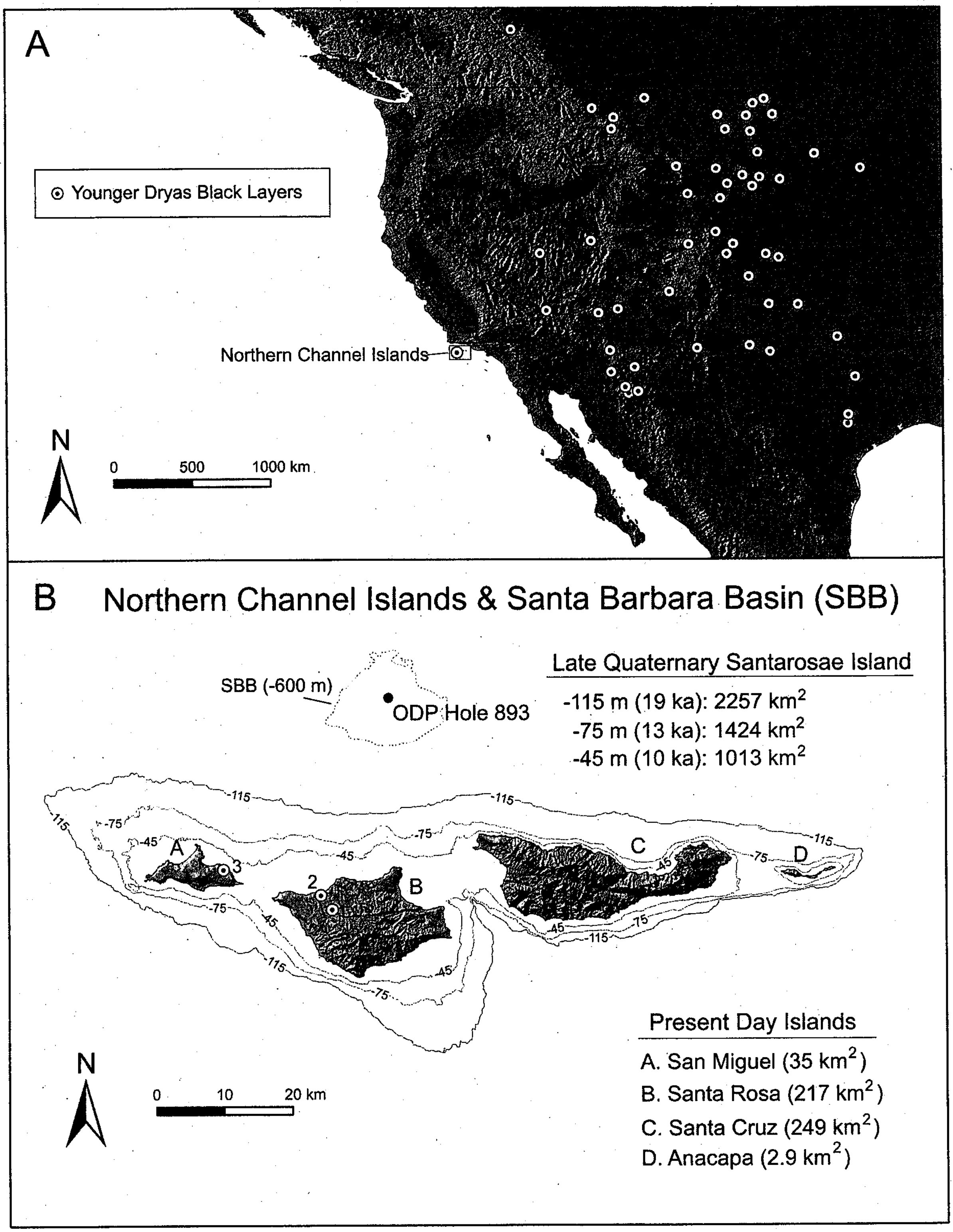


Fig. S1. Maps of western North America and Northern Channel Islands showing localities discussed in the text. (*A*) The distribution of black sedimentary layers in western North America is based on Haynes (3). (*B*) See Kennett et al. (4) for details about the bathymetry surrounding the Northern Channel Islands. Localities identified on the Northern Channel Islands are indicated by numbers: 1, Arlington Canyon (AC-003); 2, Arlington Springs Human locality (CA-SRI-173); 3, Daisy Cave. (Figure layout by Jacob Bartruff.)

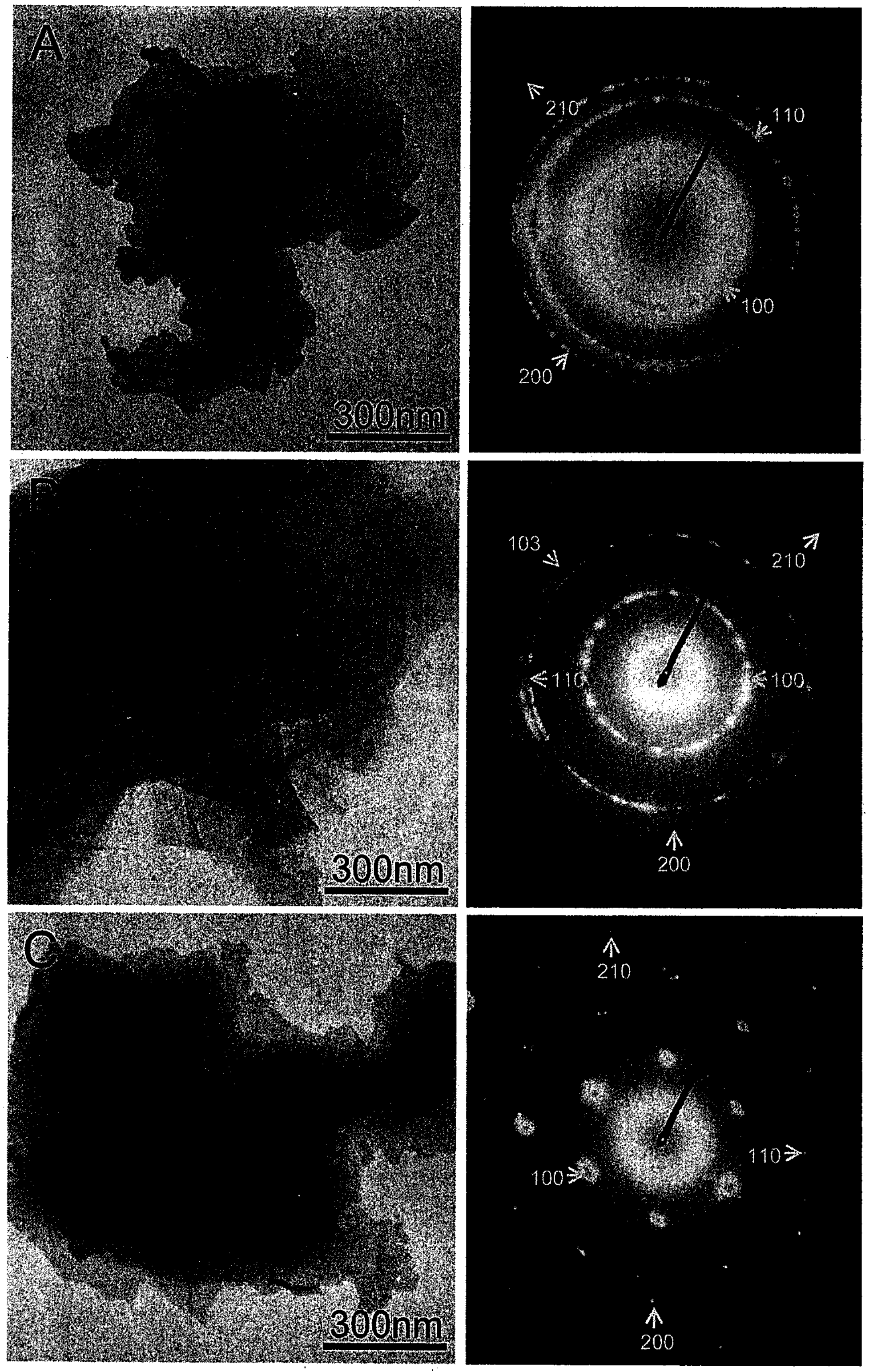


Fig. S2. Additional TEM photomicrographs (*Left*) and diffraction patterns (*Right*) of hexagonal diamond polymorphs in the YDB (12.95 ± 0.05 ka) sedimentary layer in Arlington Canyon (AC-003). (*A*) Cluster of lonsdaleite crystals and associated diffraction pattern from 4.64 and 4.69 m (AC#347); (*B*) cluster of lonsdaleite crystals and associated diffraction pattern from 4.59–4.64 m (AC#348); (*C*) close-up of single lonsdaleite crystal and associated diffraction pattern from 4.59–4.64 m (AC#348)—note lamellae. See Fig. 1 for stratigraphic position of samples.

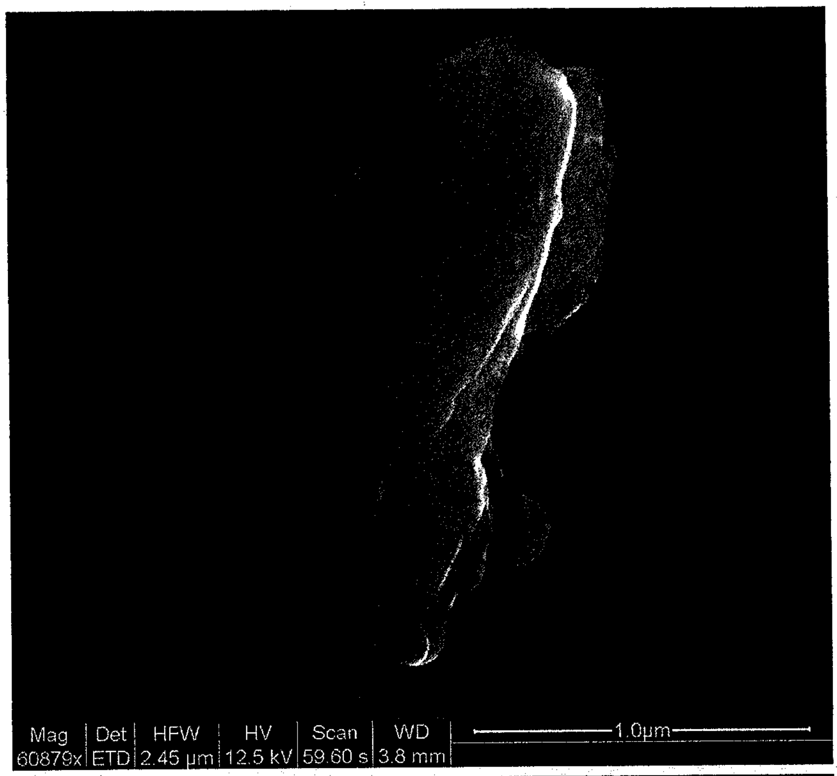


Fig. S3. Oblique SEM images of lonsdaleite crystal clusters shown in Fig. 2.

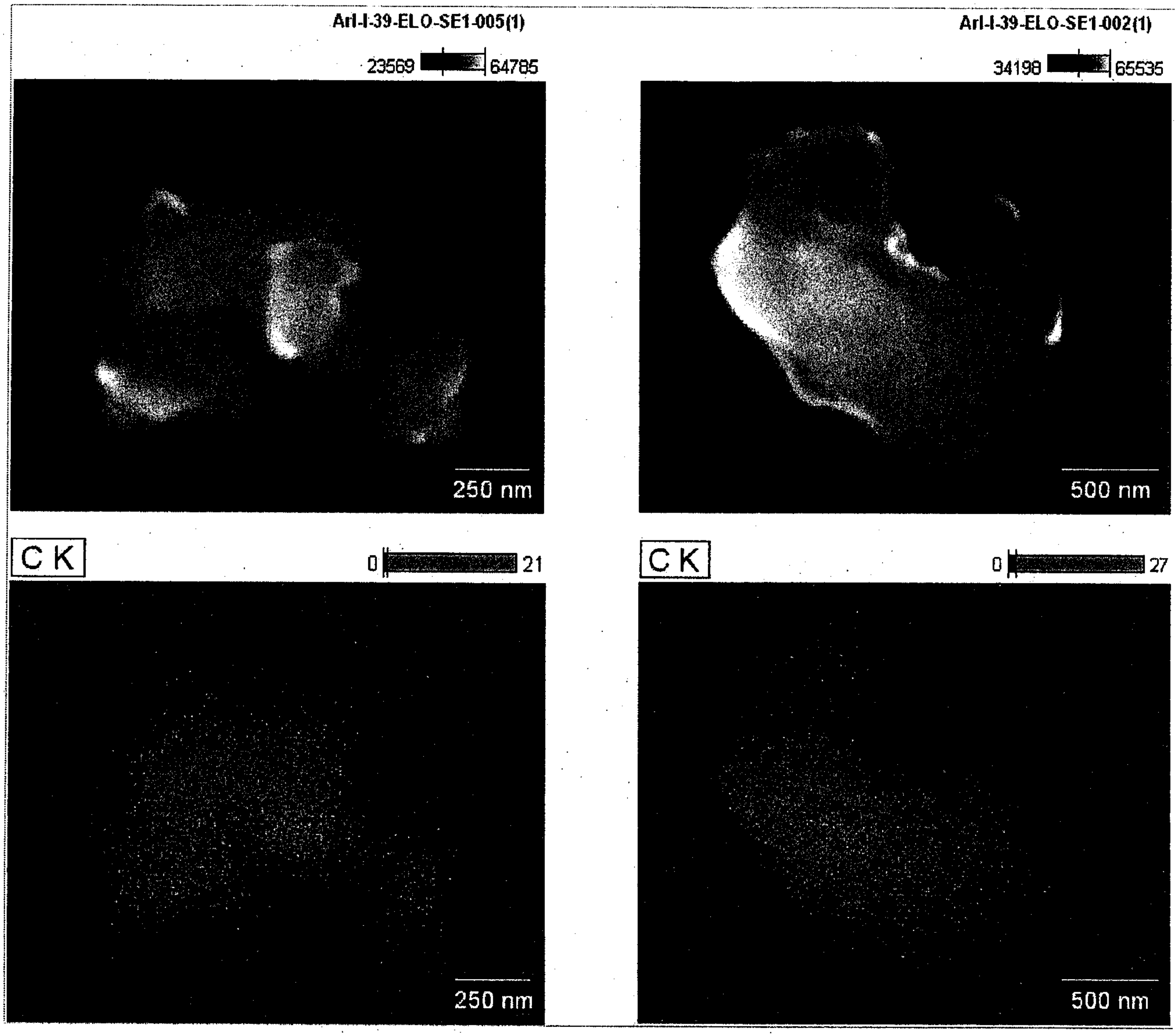


Fig. S4. Image map of EDS spectra of lonsdaleite crystal clusters shown in Fig. 2 that demonstrates dominant carbon composition.

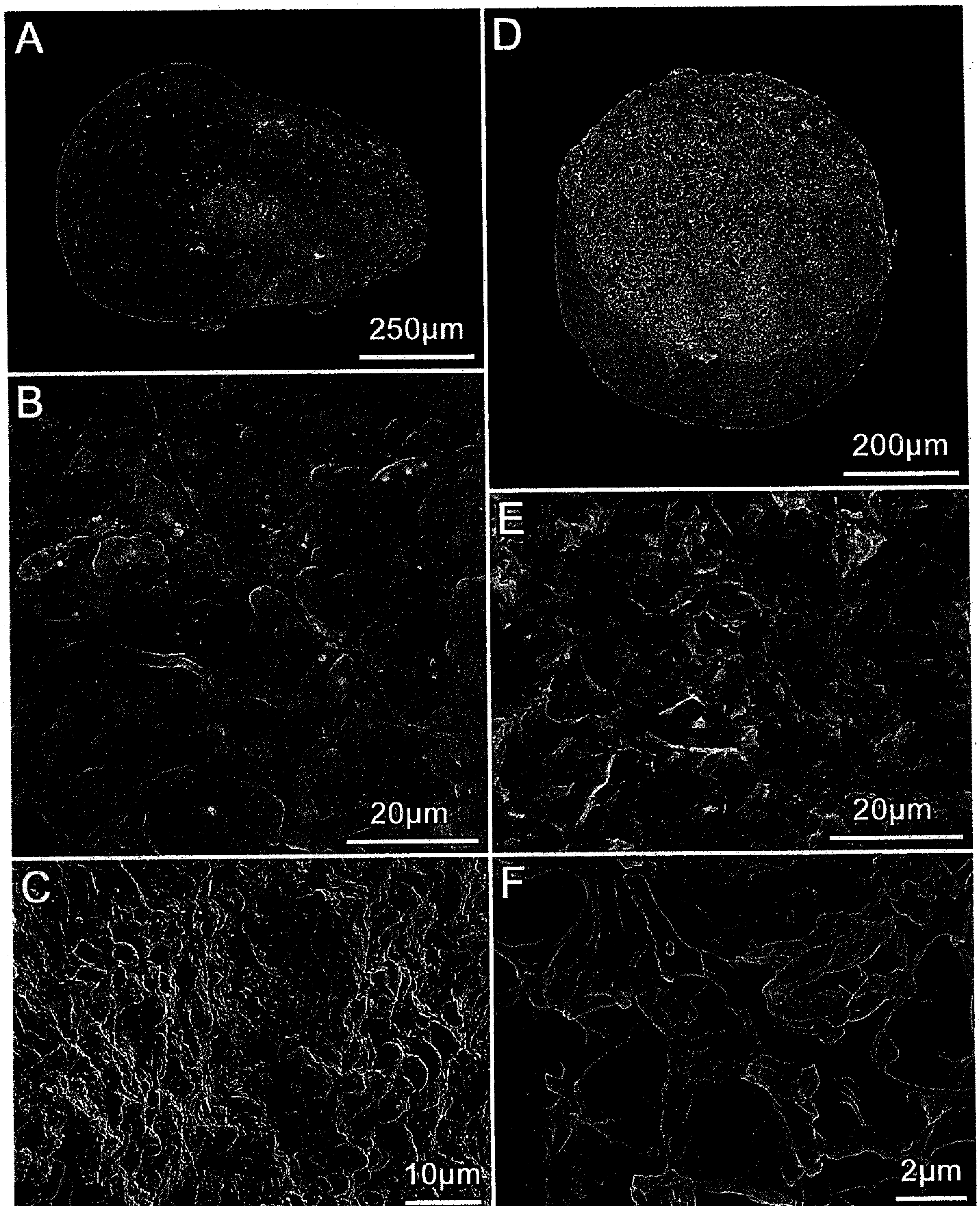


Fig. S5. SEM images of carbon elongates from YDB layer in Arlington Canyon Section, Santa Rosa Island, California (basal black layer). (*A*) Large, subangular carbon elongate showing smooth and glassy surface. (*B*) Surface microstructure at about midpoint (to the left) of elongate in *A* showing relatively smooth relief (botryoidal texture) due to melting. (*C*) Surface microstructure of edge (to the right) of carbon elongate shown in *A* showing roughness (with hollow bean-like structures of unknown origin). (*D*) Bisected, carbon elongate with relatively rounded exterior showing interior structure of complex, nonreticulate walls and voids. (*E*) Irregular, complex, nonreticulate interior of carbon elongate shown in *D* that illustrates well-vitrified and brittle thin walls of amorphous carbon separating voids. (*F*) Higher magnification image of complex, irregular interior of same elongate as in *D* and *E*. Walls are made of massive, highly vitreous, amorphous carbon.

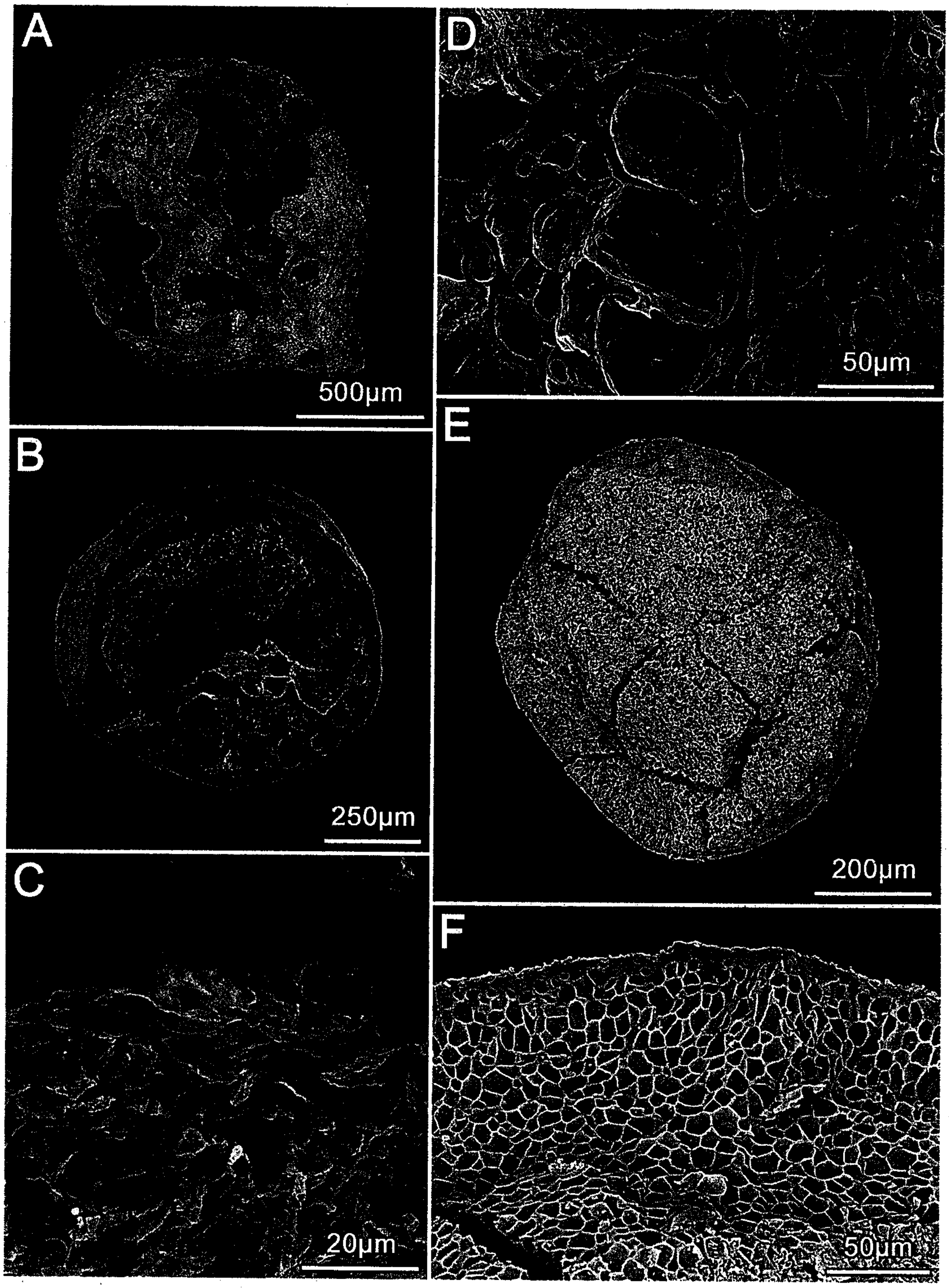


Fig. S6. SEM images of carbon elongates and carbon spherules from YDB layer (basal black layer), Arlington Canyon section, Santa Rosa Island, California. (*A*) Carbon elongate, strongly vitrified walls throughout (surface and interior) containing large voids. (*B*) Bisected carbon elongate, strongly vitrified structure throughout, with thick, massive rind containing voids and hollow center with complex, irregular structure. (*C*) More magnified image of carbon elongate shown in *B* illustrating massive, complex outer crust and irregular, complex interior walls and voids. (*D*) Interior vesicles of strongly vitrified carbon elongate. (*E*) Bisected carbon spherule showing typical internal reticulate (honeycomb) structure and thin, nonreticulate crust. (*F*) Close-up of carbon spherule interior shown in *E* with well-organized reticulate (honeycomb) structure and thin, nonreticulate crust. Carbon spherules differ from carbon elongates by having well-organized, reticulate, rather than irregular, complex interiors.

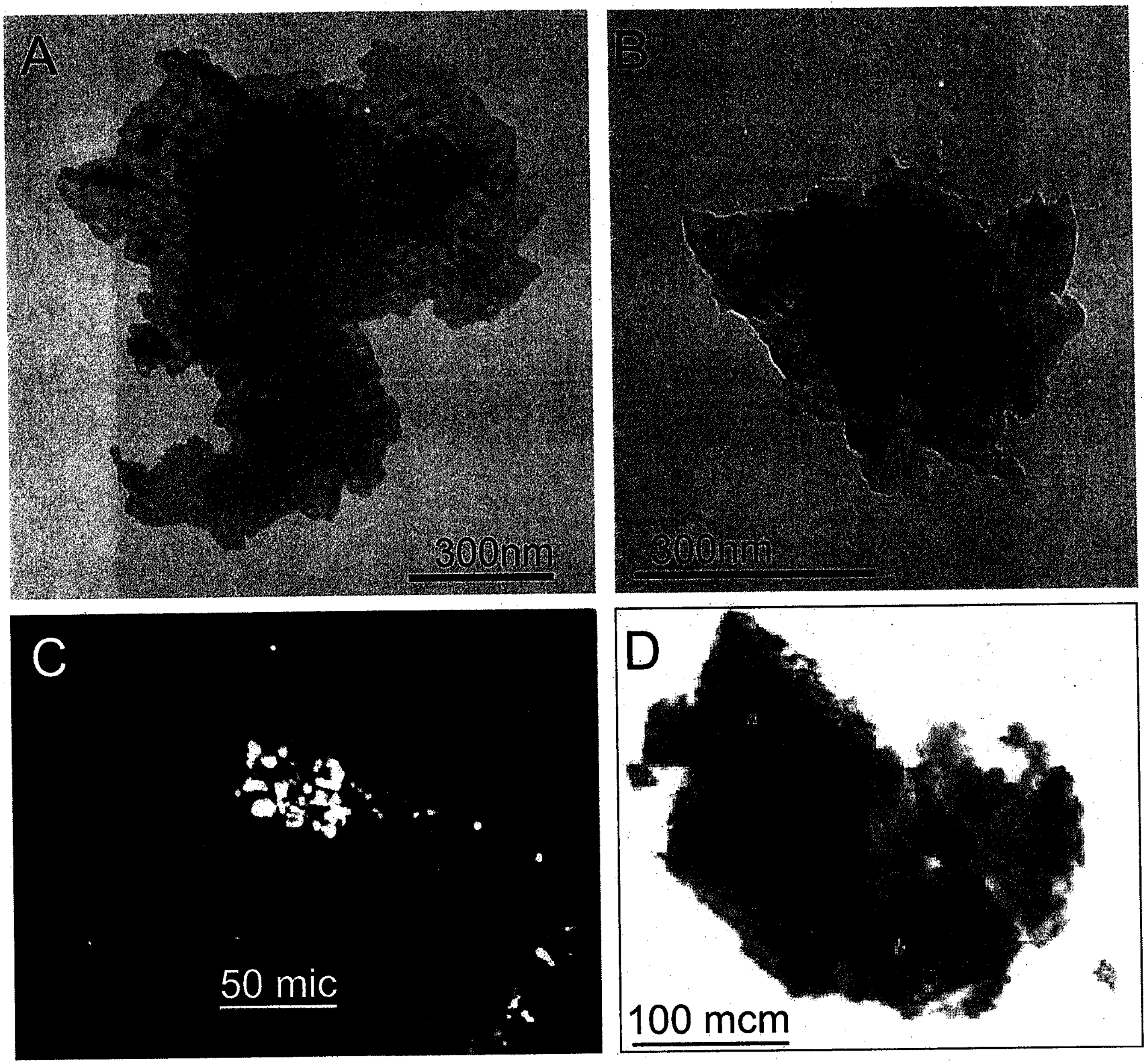


Fig. S7. Images of hexagonal diamond clusters. Diamond clusters were obtained from: (*A*) ≈12.95 ± 0.05 ka deposits in Arlington Canyon (TEM); (*B*) African Ureilite NWA 2971 (TEM); (*C*) African Ureilite NWA 2971 (cathodoluminescence, SEM); (*D*) Belilovka impact crater (5). The SEM image (*C*) shows in situ hexagonal diamonds set in thermally decomposed graphite (dark).

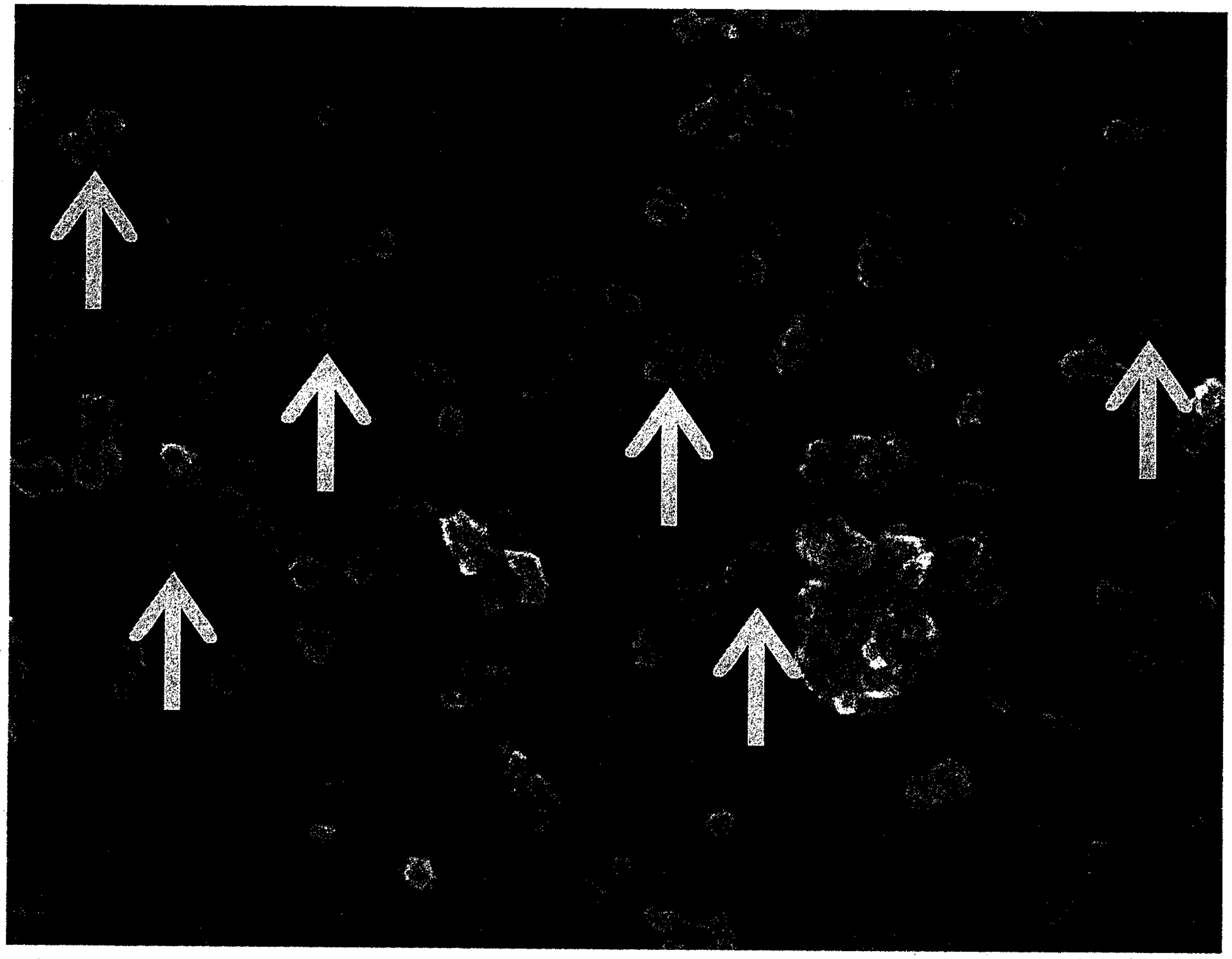


Fig. S8. SEM micrograph of grape-clustered soot (examples marked by yellow arrows) from Arlington Canyon. Highest concentrations were identified in level AC 343 (2,500 ± 250 ppm), with trace amounts found in AC344 (13 ± 1 ppm). High abundances of soot are consistent with impact-triggered fires and sudden burial favoring preservation.

Table S1. Quantitative distribution of organic carbon, other carbon forms (e.g., wood and herbaceous charcoal, various carbon particles, soot), and diamonds in sediment samples from Arlington Canyon (AC-003)

			Charcoal		Carbon			Diamonds					Soot	
								N-Diamonds			Cubic	Hexagonal		
Sample no.	cmbs	% Organic C	Wood no./cm^3	Herbaceous, %	Spherules, no./kg	Elongate, no./kg	Glasslike, g/kg	In CS, ppb*	In Elongate, ppb*	Total, ppb	No. detected	No. detected	ppm	±
AC318	95–99†	1.870	115	7.83	0	192	0.000	0	0	0				
AC319	115–120†	2.080	2	0.00	0	0	0.000	0	0	0				
AC320	122–125	2.044	0	0.00	0	0	0.000	0	0	0				
AC321	145–148	2.601	0	0.00	0	0	0.000	0	0	0				
AC322	166–169	3.870	0	0.00	0	0	0.000	0	0	0				
AC323	179–183	2.175	18	0.00	0	0	0.000	0	0	0				
AC324	195–198	1.991	413	8.72	0	23	0.001	0	0	0				
AC325	215–217	3.459	131	0.00	0	112	0.001	0	0	0				
AC326	226–229	1.920	31	0.00	0	30	0.001	0	0	0				
AC327	238–241	2.437	2	0.00	0	0	0.000	0	0	0				
AC328	245–248	1.416	22	0.00	0	31	0.001	0	0	0				
AC329	267–270	1.688	76	1.32	0	66	0.000	0	0	0				
AC330	297–300	2.838	65	0.00	0	0	0.001	0	0	0				
AC331	340–343	1.769	5	0.00	0	0	0.000	0	0	0				
AC332	383–386	1.648	13	0.00	0	18	0.000	0	18	18	1000			
AC333	392–396	3.583	261	0.00	0	767	0.001	0	1342	1342		1		
AC334	403–406	2.959	430	0.23	0	173	0.001	0	363	363				
AC335	413–416	2.785	0	0.00	0	0	0.000	0	0	0				
AC348	459–464	3.025	51	0.00	8	42	0.001	0	35	35		40		
AC347	464–469	3.577	58	0.00	31	81	0.001	0	26	26		24		
AC346	469–475	3.075	132	0.00	68	193	0.001	0	0	0				
AC345	475–480	3.929	102	0.00	13	113	0.001	0	0	0				
AC344	480–485	3.720	148	0.00	190	264	0.001	53	0	53			13	1
AC343	485–491	4.472	849	0.24	85	249	0.001	12	0	12			2500	250
AC342	491–493	8.719	435	6.90	38	0	0.000	0	0	0			0	0
AC341	493–498	4.069	258	1.55	166	373	0.001	73	0	73			0	0
AC340	498–503	4.310	268	0.37	274	714	0.001	0	855	855			0	0

Diamond polymorphs include n-diamonds in carbon spherules and elongates and hexagonal and cubic found only in association with elongates.
*Embedded in carbonaceous matrix.
†Profile 2.

Experimental and statistical reevaluation provides no evidence for *Drosophila* courtship song rhythms

David L. Stern[a,1], Jan Clemens[b], Philip Coen[c], Adam J. Calhoun[b], John B. Hogenesch[d], Ben J. Arthur[a], and Mala Murthy[b,e]

[a]Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147; [b]Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08544; [c]University College London, Institute of Neurology, London WC1E 6BT, United Kingdom; [d]Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229; and [e]Department of Molecular Biology, Princeton University, Princeton, NJ 08544

Edited by Joseph S. Takahashi, Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, TX, and approved August 9, 2017 (received for review May 6, 2017)

From 1980 to 1992, a series of influential papers reported on the discovery, genetics, and evolution of a periodic cycling of the interval between *Drosophila* male courtship song pulses. The molecular mechanisms underlying this periodicity were never described. To reinitiate investigation of this phenomenon, we previously performed automated segmentation of songs but failed to detect the proposed rhythm [Arthur BJ, et al. (2013) *BMC Biol* 11:11; Stern DL (2014) *BMC Biol* 12:38]. Kyriacou et al. [Kyriacou CP, et al. (2017) *Proc Natl Acad Sci USA* 114:1970–1975] report that we failed to detect song rhythms because (*i*) our flies did not sing enough and (*ii*) our segmenter did not identify many of the song pulses. Kyriacou et al. manually annotated a subset of our recordings and reported that two strains displayed rhythms with genotype-specific periodicity, in agreement with their original reports. We cannot replicate this finding and show that the manually annotated data, the original automatically segmented data, and a new dataset provide no evidence for either the existence of song rhythms or song periodicity differences between genotypes. Furthermore, we have reexamined our methods and analysis and find that our automated segmentation method was not biased to prevent detection of putative song periodicity. We conclude that there is no evidence for the existence of *Drosophila* courtship song rhythms.

Drosophila | courtship song | song rhythms

When a male vinegar fly (*Drosophila melanogaster*) encounters a sexually receptive female, he performs a series of courtship behaviors, including the production of songs containing pulses and hums (or sines) via unilateral wing vibration (Fig. 1*A*). Every parameter of song displays extensive quantitative variation within a bout of singing, including the amplitude and frequency of pulses and sines and the timing of individual pulse and sine events (1–8). Like humans during conversation, *Drosophila* males modulate their song based on sensory feedback from their communication partner (3, 4).

Visual inspection of songs reveals that the mean interpulse interval varies over time (Fig. 1*B*). This observation was first made in 1980 by Kyriacou and Hall (9) and they reported that the mean cycled with a periodicity of about 55 s and was controlled, in part, by the *period* gene, a gene required for circadian rhythms (10). Later papers demonstrated that evolution of a short amino acid sequence within the *period* protein caused species-specific differences in this periodicity (10–13). These reports attracted considerable interest because they implicated the *period* gene in ultradian rhythms, in addition to its well-known role in circadian rhythms (14), and because they illustrated how genetic evolution can cause behavioral evolution.

Despite this progress, the molecular mechanisms causing this periodicity remained unknown. To further advance study of these rhythms, previously we searched for this periodicity using sensitive methods and failed to find evidence for song rhythms (1). We were mindful, however, that Kyriacou and Hall (15) had argued that the presence or detectability of the rhythms was sensitive to assay conditions and methods of analysis. One of us, therefore, replicated the methods of Kyriacou and Hall as closely as possible, but, again, song rhythms could not be detected (2).

Kyriacou et al. (16) have recently questioned our previous conclusions. Here, we focus on three major assertions that they claim call our conclusions into doubt. First, we examine their central claim that manual analysis of songs, but not automated analysis, reveals genotype-specific song rhythms. We find that reanalysis of their manually annotated data provides no statistical support for genotype-specific rhythms. We also find no evidence for song rhythms in the original dataset and a new larger dataset. Second, we examined their claim that the original recordings contained insufficient data to detect rhythms and find that this claim is not supported by simulation studies. Third, we examine their claim that the high false-negative rate of the automated song segmenter decreased the probability of detecting song rhythms and we find no evidence that the missing pulse events biased our analysis of song rhythms. Further, we identify the major sources of false-negative events in automated song analysis and illustrate that minor modifications to initialization parameters substantially improve performance of the song segmenter. Kyriacou et al. (16) also raised a number of minor concerns—such as how to choose an appropriate interpulse-interval cutoff, whether temperature was controlled appropriately in our experiments, and whether songs produced beyond the first few minutes of courtship should be analyzed—that we consider peripheral to the central questions raised and therefore we have addressed these concerns (which are also unsupported by reanalysis) in *SI Appendix*.

Results

Earlier papers that identified song cycles used several unusual methods of data analysis that are useful to review. First, continuous

Significance

Previous studies have reported that male vinegar flies sing courtship songs with a periodic rhythm of approximately 55 s. Several years ago, we showed that we could not replicate this observation. Recently, the original authors have claimed that we failed to find rhythms because (*i*) our flies did not sing enough and (*ii*) our software for detecting song did not detect all song events. They reported that they could detect rhythms in song annotated by hand. We show here that we cannot replicate their observation of rhythms in the hand-annotated data or in other datasets. We also show that our original methods were not biased against detecting rhythms. We conclude that song rhythms cannot be detected.

Author contributions: D.L.S., J.C., P.C., A.J.C., J.B.H., B.J.A., and M.M. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The raw and segmented song data for the new song recordings are available in the Dryad Digital Repository, datadryad.org/ (doi:10.5061/dryad.td223).

[1]To whom correspondence should be addressed. Email: sternd@janelia.hhmi.org.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1707471114/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1707471114

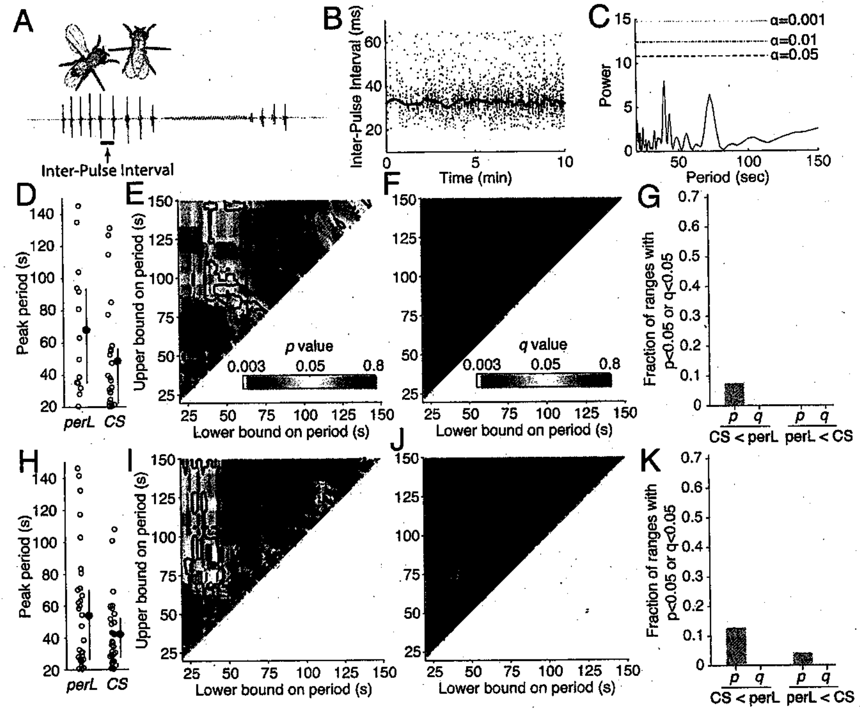


Fig. 1. Genotype-specific periodicity cannot be detected in *Drosophila* courtship song. (*A*) *Drosophila* males produce courtship song, composed of pulses (red) and sines (blue), by extending and vibrating a wing. The interpulse interval is the time between consecutive pulses within a single train of pulses. (*B*) The average interpulse interval varies over time. (Purple line is the running mean with sliding window of 200 samples.) (*C*) Lomb–Scargle periodogram analysis of the interpulse-interval data from *B* plotted for the range of 20–150 s. None of the peaks are significant at $P < 0.05$. (*D*) Comparison of the peak power between 20 and 150 s from the Lomb–Scargle periodograms for the song data for the genotypes periodL (perL) and Canton-S (CS) manually annotated by Kyriacou et al. (16). Red points and lines represent mean ± 1 SD for each genotype. (Right-tailed *t* test $P = 0.06$. Rank sum $P = 0.10$.) (*E*) *P* values for period windows with different lower and upper bounds. (*F*) False discovery rate *q* values for the windows shown in *E*. (*G*) Fraction of ranges with significant comparisons (p or $q < 0.05$) for either the test of Canton-S less than periodL or periodL less than Canton-S. (*H*–*K*) Same as *D*–*G* for newly collected song data from the same genotypes annotated using FlySongSegmenter. (*H*) Right-tailed *t* test $P = 0.06$; rank sum $P = 0.45$.

interpulse-interval data were binned into 10-s intervals. We reported previously that binning the data, together with the analysis of relatively short songs, creates peaks in spectrogram analysis that fall within an artificially narrowed frequency range, corresponding approximately to the frequency range originally reported for the periodicity, and reduces the significance of periodiogram peaks (ref. 2 and see below). Despite the fact that this procedure squeezes periodogram results into a narrow frequency range, few songs contained peaks reaching a significance level of $P < 0.05$ (4 of 149 songs, figure 3*A* of ref. 2), strongly suggesting that these peaks represent signals that cannot be distinguished from noise. All of the previously reported "statistically significant" comparisons of different genotypes are derived from analysis of mainly nonsignificant periodogram peaks. In this reevaluation, we do not discuss binning, but instead focus on other methodological issues.

No Evidence That Manual Song Segmentation Reveals Genotype-Specific Song Rhythms. Kyriacou et al.'s (16) core finding is that different genotypes displayed different periodic rhythms of the interpulse interval. This is also the most important discovery reported in earlier papers on this subject (10–12, 17). Kyriacou et al. (16) manually annotated recordings made by Stern (2) from a wild-type strain, *Canton-S*, and a strain carrying a *period* gene mutation, per^L, for flies they categorized as singing "vigorously." We reanalyzed these data and the automatically segmented data (2). Flies homozygous for per^L display circadian rhythms that are longer than normal (14), and earlier papers have reported that per^L confers longer periods on the interpulse-interval rhythm (9–12). Kyriacou et al. (16) report a difference in the mean song period between *Canton*-S and per^L with the manually annotated data, but not with the automatically segmented data, suggesting that song cycles exist and display genotype-specific frequencies and that the automatically segmented data are biased against detecting the song rhythm.

Kyriacou et al. (16) used several methods to measure periodicity in the original time series, which we discuss in more detail in the next paragraph. For ≈85% of these songs, these methods do not yield statistically significant signals in the frequency range of 20–150 s. Because most songs do not yield statistically significant peaks, Kyriacou et al. (16) identified the peak with maximum power in the range of 20–150 s for each song and compared these values between genotypes. This is an unorthodox approach to data analysis. It is equivalent to sampling outliers from a distribution of random noise and then performing further statistics with these data. Nonetheless, Kyriacou et al. (16) detected genotype-specific song rhythms using this method and so, below, we accept this premise and investigate whether there is statistical support for genotype-specific rhythms in the data. We start by examining whether there is evidence for rhythms in individual songs.

The general model proposed for these song rhythms is that the interpulse interval varies, on average, with a regular periodicity (9). Therefore, it should be possible to detect this rhythmicity with appropriate methods of periodogram analysis. We have previously used Lomb–Scargle periodogram analysis (18–20) because this method does not require evenly spaced samples and Kyriacou et al. (16) also adopted this method. For example, the Lomb–Scargle periodogram of the time series in Fig. 1*B* is shown in Fig. 1*C*. In this case, despite the obvious variation in interpulse-interval values observed in Fig. 1*B*, there is no significant periodicity between 20 and

150 s. Kyriacou et al. (16) also used Cosinor (21) and CLEAN (22) for periodogram analysis. CLEAN does not produce a significance value for periodogram peaks, so it is difficult to interpret. We find that Cosinor exhibits a high false-positive rate (*SI Appendix*, Fig. S1) and should be avoided for this type of analysis.

Kyriacou et al. (16) state that wild-type *D. melanogaster* songs exhibit periodicity between 20 and 150 s. Previously, they reported that rhythms occurred with 50–60 s periodicity (9). Increasing the width of the periodicity window from 50–60 s to 20–150 s increases the probability of detecting significant periods, but, even given this wide frequency range, we observed that only 4 of the 25 manually annotated *Canton-S* songs and 3 of the 25 automatically segmented songs contained periodogram peaks that reached a significance level of $P < 0.05$. (When we binned data in 10-s bins, these values declined to 0 of 25 manually annotated and 1 of 25 automatically segmented songs.) These significant peaks are not localized to any particular narrow frequency range (*SI Appendix*, Figs. S1 and S5).

One reason to study nonsignificant peaks would be if periodicity is weak and not detected reliably by periodogram analysis. This seems unlikely, since simulated song rhythms can be detected with high confidence (refs. 1 and 2 and see below). Nonetheless, if periodogram analysis is underpowered, then we expect to observe that the major peak in most songs should display nearly significant periodicity. In fact, we observe that 72% of *P* values are greater than 0.2 (*SI Appendix*, Fig. S2). There is therefore no evidence that songs contain weak periodicity.

An alternative possible reason to include nonsignificant periodogram peaks in downstream analysis is that the signal to noise of the periodicity is extremely low. An analog in neuroscience is that neural signals sometimes cannot be detected with high signal to noise and that only by averaging over many trials of a stimulus presentation can a neural response be detected robustly. We therefore examined the power distribution averaged over all of the results for each genotype. These plots are essentially flat, suggesting that there is no signal hidden in the fluctuations of individual periodograms (*SI Appendix*, Fig. S3).

Given these observations, further analysis of these data seems unwarranted. However, Kyriacou et al. (16) compared the maximum periodogram peaks between 20 and 150 s for the *Canton-S* and per^L recordings and found that the manually annotated data showed a statistically significant difference in the mean period, although the automatically segmented data did not (figure 3*D* of Kyriacou et al.; ref. 16). This is the key result of their paper. We therefore attempted to replicate this observation. For the manually annotated data from each song, we identified the peak in the periodogram of maximum power falling between a period of 20 and 150 s. In contrast to their published results, we found that the average of the periods with maximum power (most of which were not significant) was not significantly different at $P < 0.05$ between the genotypes *Canton-S* and per^L (Fig. 1*D*). We have no explanation for this discrepancy between our statistical analysis and theirs.

Because there is no biological or quantitative justification for the particular frequency ranges examined in any study, we wondered whether the results were sensitive to the frequency range examined. We explored a wide range of possible frequency ranges and found that the test statistic was sensitive to the precise frequency range selected (Fig. 1*E*). Most frequency windows do not generate a statistically significant difference between the genotypes (Fig. 1 *E* and *G*), and false discovery rate correction for multiple testing (23, 24) yields no frequency ranges with significant results (Fig. 1 *F* and *G*).

Thus, there is no support for the specific results reported by Kyriacou et al. (16) and there is no statistical support for defining song interpulse-interval cycle periods as occurring within any particular window. Most importantly, our analysis indicates that genotype-specific analysis of nonsignificant periodogram peaks has no justification. It is difficult to reconstruct precisely what steps in the analysis led previous reports to identify statistically significant genotype-specific differences, but it is possible that previous studies may have serendipitously selected frequency ranges that yielded significant results and/or did not properly control for multiple testing.

Newly Collected Data Provide No Evidence for Genotype Specific Song Periodicities. Although we could not reproduce results reported by Kyriacou et al. (16), we decided to take their observation at face value as a preliminary result and test directly whether genotype specific song rhythms could be detected in an expanded dataset. We recorded song from 33 *Canton-S* males and 34 $period^L$ males. We identified the strongest periodogram peak in the frequency range of 20–150 s for each song and found no significant difference between these genotypes (Fig. 1*H*). We then compared test statistics across a wide set of frequency ranges, as described above. We identified some frequency ranges that yielded significant results in the predicted direction (Fig. 1*I*), with $period^L$ rhythms slower than *Canton-S* rhythms, but for three reasons we believe these results are spurious. First, and most importantly, none of these ranges are significant after false discovery rate correction (Fig. 1*J*). Second, multiple frequency ranges support the opposite conclusion, that *Canton-S* rhythms are slower than $period^L$ rhythms (Fig. 1*K*). Third, the frequency ranges yielding significant comparisons only partially overlap with the ranges found for the original dataset (cf. Fig. 1 *E* and *I*). In conclusion, there is not only no evidence that song rhythms exist, there is also no evidence that reported genotype-specific differences in a song rhythm exist.

Putative song cycles cannot be identified in most automatically segmented song (2) and, as we showed above, in most manually annotated song. In addition, when statistically significant periodicity is detected, the frequencies of this periodicity do not cluster in a specific frequency range, but instead are spread randomly across the entire frequency range examined (*SI Appendix*, Fig. S5; figure 4 of Stern; ref. 2). Finally, no genotype comparisons are significant after correcting for multiple comparisons (Fig. 1). All together, these results imply that the few statistically significant periodicities that can be found do not carry biological significance.

No Evidence That Low-Intensity Courtship Provided Insufficient Data to Detect Song Rhythms. Although we found no statistical evidence for the existence of song rhythms or of genotype-specific rhythms, we feel it is important to rebut several other statements made by Kyriacou et al. (16). They state that rhythms can be detected only in songs produced by vigorously singing males and write: "sporadic songs could not possibly provide any test for song cycles." It is not clear if they mean that rhythms can be detected only in songs with many pulses or that only flies that sing songs with many pulses ("vigorous singers") produce rhythms. Kyriacou et al. (16) manually annotated songs from flies that they categorized as vigorous, and we showed above that significant periodicity can be found in only a minority of these songs and that these significant values are not localized to a particular frequency range (*SI Appendix*, Figs. S1*D* and S5*A*). Therefore, it is unlikely that only flies that sing songs with many pulses produce periodicity. We therefore performed simulations to determine whether rhythms can be detected only in songs with many pulses.

We previously investigated songs from 45-min courtship recordings that contained at least 1,000 interpulse-interval measurements (2). Kyriacou et al. (16) argued that more than 180 interpulse-interval measurements per minute (or ≈5,000 events in a 45-min recording) should be identified to allow identification of song rhythms. To examine this claim, we performed a statistical power analysis using songs with variable numbers of interpulse-interval measurements, where statistical power corresponds to the proportion of times periodicity is detected in songs where periodicity has been artificially imposed on song data (Fig. 2). We started with six 45-min recordings of *Canton-S* from Stern (2) that contained more than 10,000 interpulse-interval measurements. None of these six songs yielded statistically significant power in the frequency range between 50 and 60 s (the range originally defined to contain rhythms; ref. 9) and one song produced a marginally significant peak at 31.7 s ($P = 0.04$), which falls between 20 and 150 s (the range used by Kyriacou et al.; ref. 16). Fig. 2 *D* and *E* illustrate the interpulse-interval data and periodogram for one of these songs. Therefore, these songs do not contain strong periodicity in the predicted range and can serve as a template to examine the power of Lomb–Scargle periodogram analysis to detect simulated rhythms imposed on these data.

The initial reports of periodic cycles in the interpulse-interval reported rhythms with a mean period of 55 s and an amplitude of ≈2 ms (9). Therefore, we imposed a 55-s rhythm with an amplitude of 2 ms on the six songs containing more than 10,000 interpulse-interval measurements (Fig. 2 *A–C*). We detected the simulated 55-s rhythm in all six songs with *P* values <10*e*-74 (example shown in Fig. 2 *F* and *G*). We then randomly removed data points from the songs iteratively and calculated the fraction of times we could detect the simulated rhythm with $P < 0.05$. We removed data randomly from the dataset to simulate the effect of failing to detect individual events in the song, and we also removed chunks of data (in 10-s bins) to simulate large gaps between song bursts, such as might be generated during low-intensity courtship. We found that in both scenarios we could randomly remove at least 90% of the data and still detect simulated rhythms at least 80% of the time (example shown in Fig. 2 *H* and *I*; summary statistics shown in Fig. 2*J* and *SI Appendix*, Fig. S4*A*). That is, as long as songs contained at least 1,000 interpulse-interval measurements, Lomb–Scargle periodogram analysis detected simulated rhythms with power greater than 0.8. Similar results were found when we analyzed only the first 400 s of songs (*SI Appendix*, Fig. S4 *C* and *D*). Furthermore, periodicity could be detected with power greater than 0.8 when the amplitude of simulated periodicity was greater than at least 1 ms (*SI Appendix*, Fig. 4*B*). These results were robust to noise in the original periodicity. Song with a signal-to-noise ratio of as low as 0.25 could be detected with power >0.7 with sample sizes of at least 1,000 interpulse-interval measurements (Fig. 2*K*). Similarly, periodicity could be detected reliably when we simulated a non-sinusoidal rhythm (*SI Appendix*, Fig. S4*E*) and when periodicity was imposed for only a fraction of the total song (*SI Appendix*). Thus, Lomb–Scargle periodogram analysis is a sensitive method for detecting simulated periodicity, even in the presence of noise or discontinuities in the waveform.

Songs containing at least 1,000 interpulse intervals provide sufficient data to identify putative song cycles. In fact, we find that songs can be deeply corrupted by the absence of large segments of song and simulated periodicity can still be detected.

No Evidence That the Automated Fly Song Segmenter Biased the Results. Kyriacou et al. (16) expressed concern that our automated fly song segmenter displayed a low true positive rate (the segmenter failed to detect ≈50% of the pulses identified through manual annotation) and produced some false-positive calls (≈4% of events scored as pulses by the automated segmenter appear to be noise). They suggest that these incorrect pulse event assignments could bias estimation of the mean interpulse interval and, therefore, decrease the signal to noise of the periodic cycle, making it difficult to detect a periodic signal. In principle, a large

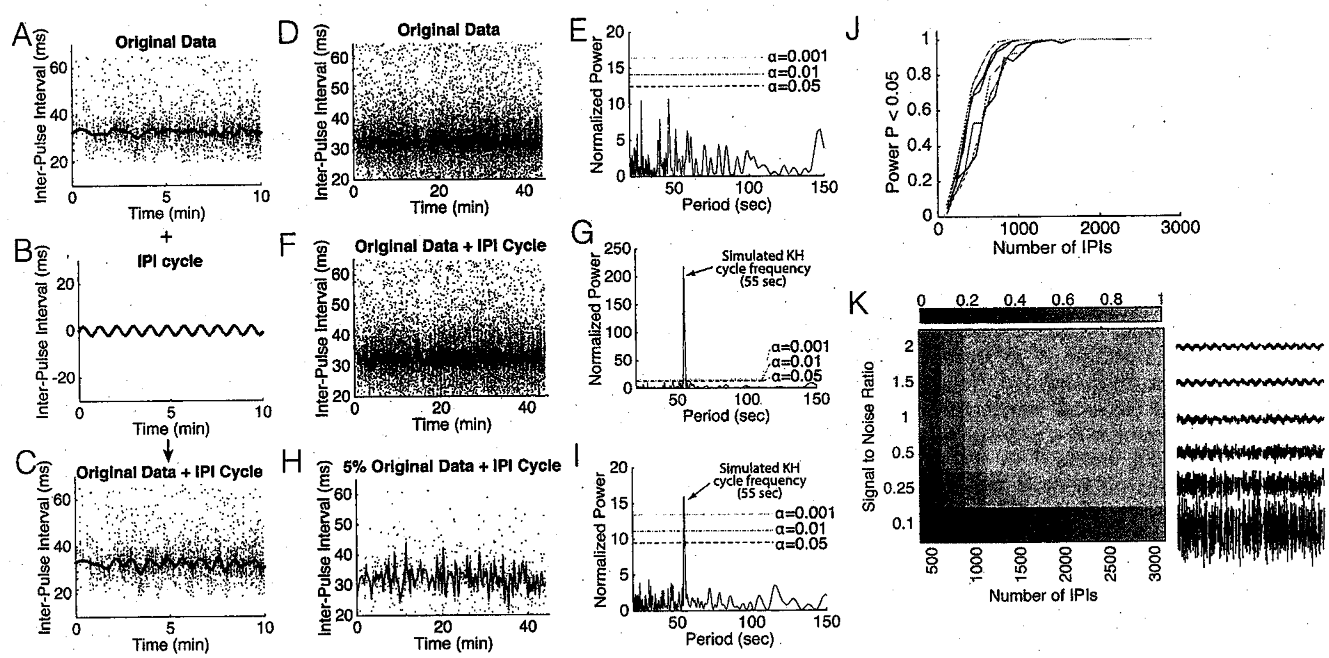


Fig. 2. Simulations to explore power to detect rhythms, should they exist. (*A–C*) Example of how a periodic cycle was added to raw interpulse-interval (IPI) data. Purple line in *A* illustrates the running mean of the raw data. Blue line in *B* shows a periodic rhythm with an amplitude of 2 ms and a period of 55 s. Original data with simulated periodicity is shown in *C*. (*D*) One example of 45 min of interpulse-interval data. Purple line shows running mean. (*E*) Lomb–Scargle periodogram of data in *D* does not detect periodicity. (*F*) Data from *D* with a 55-s periodicity imposed. (*G*) Lomb–Scargle periodogram of data in *F* now reveals a highly significant peak at 55 s, consistent with the simulated Kyriacou–Hall (KH) periodicity. (*H*) Random removal of 95% of the interpulse-interval data from *F*. (*I*) Lomb–Scargle periodogram of the data in *H* detects significant periodicity. (*J*) Power analysis of six songs (each song a different color) containing more than 10,000 interpulse-interval events after 55-s periodicity was added and individual interpulse-interval events were removed randomly. Power equals the fraction of times out of 100 that a song contained a rhythm with significant periodicity between 50 and 60 s at $P < 0.05$. (*K*) Power to detect simulated noisy periodicity versus number of IPIs remaining after random removal of IPIs. Means of simulations for six songs containing more than 10,000 interpulse-interval measurements are shown. Examples of simulated noisy rhythms are shown to the *Right*. Colorbar shows power to detect simulated rhythm.

sample of incorrect calls could bias results, so we investigated whether this was the case for our prior analyses. We used Kyriacou et al.'s (16) manually annotated dataset to investigate the potential for bias and to evaluate performance of the automated segmenter.

When a single pulse event is not detected, the interpulse interval is then calculated as the sum of the two neighboring real intervals. On average, this is approximately double the average interpulse interval. The average interpulse interval for the *Canton-S* recordings reported in Stern (2) is ≈35 ms with a SD of ≈7 ms. Therefore, skipping a single pulse event is expected to result in interpulse-interval measurements of ≈70 ms, but with considerable variance. Following Kyriacou and Hall (15) and Stern (2), we used a heuristic threshold of 65 ms to reduce the number of spurious interpulse-interval values. Therefore, in the specific case when a single pulse in a train is missed, approximately one-third of the incorrectly scored doublet interpulse-interval measurements would be shorter than 65 ms and are expected to contaminate the original dataset.

However, this scenario applies only when one undetected pulse is flanked by two pulses that are detected. Skipping more than one pulse would always result in interpulse-interval measurements that are excluded by the 65-ms threshold. We found, however, that only 9% of the pulses missed by automated segmentation were singletons (*SI Appendix*, Fig. S6*A*). These incorrect interpulse intervals contribute to a slight excess of interpulse intervals with high values (*SI Appendix*, Fig. S6*B*). Lowering the interpulse-interval threshold would, therefore, remove most or all spurious interpulse intervals. Since our power analysis, discussed above, revealed that periodogram analysis was robust to random removal of interpulse-interval events, as long as songs still contained at least 1,000 values, loss of a small number of interpulse intervals is not expected to hamper detection of rhythms. After reducing the interpulse-interval threshold to 55 ms, we still found no compelling evidence for significant periodicity in the original data (*SI Appendix*, Fig. S7). Therefore, we explored the effect of reducing the interpulse-interval cutoff even further. In this case, we used all 68 *Canton-S* songs from Stern (2) and retained for analysis only those songs that contained at least 1,000 interpulse-interval measurements after imposing the new interpulse-interval threshold. We explored a range of cutoff values from 25 to 65 ms. We found that we could detect the simulated rhythm in most songs with at least 1,000 interpulse-interval measurements remaining after thresholding, even when the threshold was as low as 25 ms (Fig. 3). Therefore, we can find no evidence that pulses missed by the automated song segmenter or the specific interpulse-interval threshold used in Stern (2) prevented detection of song rhythms.

Although detection of putative song rhythms is robust to dropped pulses in songs that retain at least ≈1,000 interpulse intervals, it is worth reviewing briefly why the segmenter failed to detect certain pulses in recordings reported in Stern (2). The first step of song segmentation involves detection of pulse-like signals and sine-like signals (1). In subsequent steps, the segmenter filters out many kinds of sounds that were originally classified as song pulses. Both the initial detection of pulses and subsequent filtering steps are sensitive to multiple parameters. These parameters are specified before segmentation and can be modified to enhance performance of the segmenter for different recordings. We identified two primary causes for missed pulses. First, Stern (2) recorded song in larger chambers than those used previously with these microphones (1), to match the chamber size used by Kyriacou and Hall (9). This larger chamber with one microphone had reduced sensitivity compared with the original smaller chamber. The segmenter thus tended to miss pulses of lower amplitude, which are hard to automatically differentiate from noise, and this explains ≈35% of the missed pulses (*SI Appendix*, Fig. S8 *A* and *C*).

The second major cause of missed pulses is that *Drosophila* males produce pulses with a range of carrier frequencies (tones).

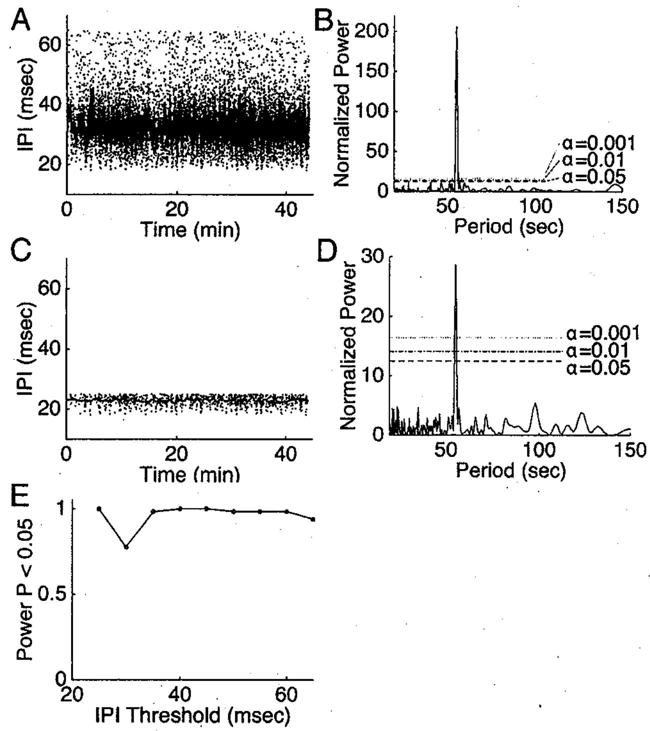


Fig. 3. The specific interpulse-interval threshold does not influence the statistical power to detect putative song rhythms. (*A*) Example of one original song with 55-s periodicity artificially imposed on the original interpulse-interval data. (*B*) Lomb–Scargle periodogram of data in *A*, revealing strong signal at 55 s. (*C*) Same simulated data as in *A* with all interpulse-interval values greater than 25 s removed. (*D*) Lomb–Scargle periodogram reveals strong signal of the simulated periodicity at 55 s, even though the data were thresholded at 25 s. (*E*) Power to detect simulated periodicity versus interpulse-interval threshold for songs retaining at least 1,000 interpulse-interval values after thresholding.

The higher frequency pulses tend to resemble other nonsong noises, like grooming, and a user can set parameters in the segmenter to attempt to exclude these nonsong noises based on the carrier frequency of the event. Stern (2) used parameters to minimize the false-positive rate, including a relatively low carrier frequency cutoff for pulses. The lower pulse frequency threshold used by Stern (2) explains ≈42% of the missed pulses (*SI Appendix*, Fig. S8 *B* and *D*). Using the same software with different parameters (from Coen et al.; ref. 4) recovers many of these high-frequency pulses without substantially increasing the false-positive rate (*SI Appendix*, Fig. S8 *C–F*).

Above, we showed that including more pulse events, by manual annotation, did not increase the probability of detecting song rhythms. Therefore, there is no evidence that the data resulting from the song segmenter parameters used in Stern (2) generated a dataset that was biased against detection of song rhythms. While the song segmenter does not detect all pulse events that can be detected by manual annotation, the segmenter does provide datasets that are several orders of magnitude larger than those that can be generated by manual annotation, which has allowed discovery of multiple new phenomena related to *Drosophila* courtship song (3–7). In addition, the sensitivity of the song segmenter can be improved with optimization of initial parameters, as expected of any segmentation algorithm.

Discussion

We cannot detect a periodic cycling of the interpulse interval in *Drosophila* courtship song even in the songs manually annotated by Kyriacou et al. (16) and used as evidence for periodicity in their paper. Although it is impossible to prove a negative, our results agree with previous analyses that have concluded that there is no statistical evidence that these rhythms exist (1, 2). In particular, by exploring some of the relevant parameter space with statistical tests on the song that was manually annotated by Kyriacou et al. (16), we find that subsets of parameters sometimes produce P values lower than 0.05, but that (*i*) few regions of parameter space generate "significant" results, (*ii*) these significant regions are scattered apparently randomly in parameter space, and (*iii*) none of these significant results survive multiple test correction (Fig. 1).

Previously, we offered one explanation for how apparent song rhythms may have been detected. We found that binning data from short songs confined the periodogram peaks with maximum power close to the range reported as the song cycle (2). While few of these peaks reached statistical significance, previous authors have accepted these peaks as "signal" and performed statistical analyses to compare the peaks between genotypes. All statistically significant results from earlier papers were derived mainly from nonsignificant peaks in periodogram analysis and from relatively small sample sizes (usually fewer than 10 flies of each genotype), so it is questionable whether these derivative statistics are valid. Genotype-specific periodicities reported in earlier papers may have resulted, by chance, from studies of a small number of short songs that fortuitously led to occasional apparent replication of the original observations.

There may be a more prosaic explanation for the initial discovery of song cycles. Every fly produces highly variable interpulse intervals. In addition, a running average of these data reveals that the average interpulse-interval cycles up and down (Fig. 1*B*), similar to the temporally binned data first reported by Kyriacou and Hall (9). There is no debate about this observation. The claim in dispute is that the average interpulse-interval cycles regularly. We can find no evidence for this claim. It is easy to imagine, however, that visual examination of short recordings of song would make it appear as if the mean interpulse-interval cycled regularly.

The extraordinary within-fly variation in the interpulse interval and in the mean interpulse interval may result from multiple causes, including the possibility that male flies respond to ever-changing cues during courtship and modulate their interpulse interval to optimize their chances of mating. Individual *Drosophila* males modulate specific aspects of their courtship song based on their own patterns of locomotion and in response to feedback from females, including the transition between sine and pulse song (4) and the amplitude of pulse song (3). There is additional evidence that males modulate the carrier frequency of sine song (1). We hypothesize that male flies also modulate their interpulse interval in response to specific internal or external cues.

We can find no statistical evidence for periodicity of the interpulse interval in individual courtship songs and no evidence that comparisons of the strongest periodogram peaks from each song identify genotype-specific rhythms. These results hold both for the songs manually annotated by Kyriacou et al. (16) and for two independent large datasets automatically annotated with FlySongSegmenter using optimized parameters. At this time, a conservative assessment of the problem is that *Drosophila* courtship song rhythms and genotype-specific effects on these rhythms cannot be replicated.

Methods

Computer code for all analyses described in this paper is available at https://github.com/murthylab/noIPIcycles. Code for the version of FlySongSegmenter used in Cohen et al. (5) is available at https://github.com/murthylab/songSegmenter. The raw and segmented song data for the new song recordings are available at https://www.janelia.org/lab/stern-lab/tools-reagents-data. Further methods can be found in *SI Appendix*.

ACKNOWLEDGMENTS. We thank Elizabeth Kim for recording the new samples of flies. M.M. is a Howard Hughes Medical Institute Faculty Scholar and was also funded by an NIH New Innovator Award and an NSF Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative Early Concept Grant for Exploratory Research (EAGER) Award.

1. Arthur BJ, Sunayama-Morita T, Coen P, Murthy M, Stern DL (2013) Multi-channel acoustic recording and automated analysis of Drosophila courtship songs. *BMC Biol* 11:11.
2. Stern DL (2014) Reported Drosophila courtship song rhythms are artifacts of data analysis. *BMC Biol* 12:38.
3. Coen P, Xie M, Clemens J, Murthy M (2016) Sensorimotor transformations underlying variability in song intensity during Drosophila courtship. *Neuron* 89:629–644.
4. Coen P, et al. (2014) Dynamic sensory cues shape song structure in Drosophila. *Nature* 507:233–237.
5. Ding Y, Berrocal A, Morita T, Longden KD, Stern DL (2016) Natural courtship song variation caused by an intronic retroelement in an ion channel gene. *Nature* 536:329–332.
6. Shirangi TR, Wong AM, Truman JW, Stern DL (2016) Doublesex regulates the connectivity of a neural circuit controlling Drosophila male courtship song. *Dev Cell* 37:533–544.
7. Shirangi TR, Stern DL, Truman JW (2013) Motor control of Drosophila courtship song. *Cell Rep* 5:678–686.
8. Ewing AW, Bennet-Clark HC (1968) The courtship songs of Drosophila. *Behaviour* 31:288–301.
9. Kyriacou CP, Hall JC (1980) Circadian rhythm mutations in Drosophila melanogaster affect short-term fluctuations in the male's courtship song. *Proc Natl Acad Sci USA* 77:6729–6733.
10. Zehring WA, et al. (1984) P-element transformation with period locus DNA restores rhythmicity to mutant, arrhythmic Drosophila melanogaster. *Cell* 39:369–376.
11. Wheeler DA, et al. (1991) Molecular transfer of a species-specific behavior from Drosophila simulans to Drosophila melanogaster. *Science* 251:1082–1085.
12. Kyriacou CP, Hall JC (1986) Interspecific genetic control of courtship song production and reception in Drosophila. *Science* 232:494–497.
13. Ritchie MG, Halsey EJ, Gleason JM (1999) Drosophila song as a species-specific mating signal and the behavioural importance of Kyriacou & Hall cycles in D. melanogaster song. *Anim Behav* 58:649–657.
14. Konopka RJ, Benzer S (1971) Clock mutants of Drosophila melanogaster. *Proc Natl Acad Sci USA* 68:2112–2116.
15. Kyriacou CP, Hall JC (1989) Spectral analysis of Drosophila courtship song rhythms. *Anim Behav* 37:850–859.
16. Kyriacou CP, Green EW, Piffer A, Dowse HB (2017) Failure to reproduce period-dependent song cycles in Drosophila is due to poor automated pulse-detection and low-intensity courtship. *Proc Natl Acad Sci USA* 114:1970–1975.
17. Kyriacou CP, van den Berg MJ, Hall JC (1990) Drosophila courtship song cycles in normal and period mutant males revisited. *Behav Genet* 20:617–644.
18. Lomb NR (1976) Least-squares frequency analysis of unequally spaced data. *Astrophys Space Sci* 39:447–462.
19. Scargle JD (1982) Studies in astronomical time series analysis. II - Statistical aspects of spectral analysis of unevenly spaced data. *Astrophys J* 263:835–853.
20. Ruf T (1999) The Lomb-Scargle periodogram in biological rhythm research: Analysis of incomplete and unequally spaced time-series. *Biol Rhythm Res* 30:178–201.
21. Refinetti R, Lissen GC, Halberg F (2013) Procedures for numerical analysis of circadian rhythms. *Biol Rhythm Res* 38:275–325.
22. Roberts DH, Lehár J, Dreher JW (1987) Time series analysis with CLEAN. I. Derivation of a spectrum. *Astron J* 93:968–989.
23. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Ser B* 57:289–300.
24. Colquhoun D (2014) An investigation of the false discovery rate and the misinterpretation of p-values. *R Soc Open Sci* 1:140216.

Supplementary Information Appendix

The first section of this supplementary information appendix contains supplementary figures that are cited in the main paper. In addition, at the end of this appendix, we address several issues raised in Kyriacou et al. (1) that we did not have space to address in the main manuscript.

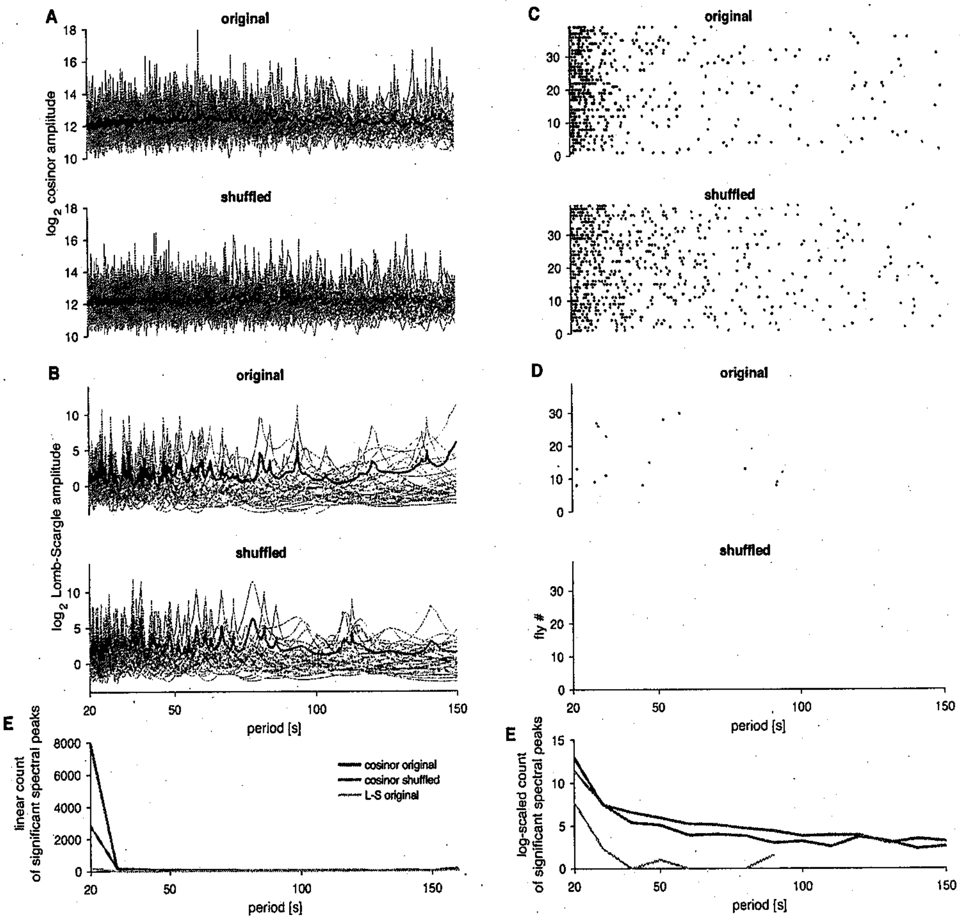


Figure S1. Cosinor analysis of IPI cycles produces many false positives. (A,B) Amplitude of cosinor fits (A) and Lomb-Scargle spectral power (B) for periods in the range 20 to 150 seconds (log2 scale). Spectra for original IPI data (upper panel) and shuffled IPI data (lower panel) are similar. Grey lines show spectra for individual flies, blue/orange lines show population averages for original and shuffled data, respectively. (C,D) Frequencies of significant peaks in the cosinor (C) and Lomb-Scargle (D) spectra for original (upper panel, blue dots) and shuffled data (lower panel, orange dots). One line per fly. (E) Distribution of significant periods over all flies shows an enrichment of short periods (20-30ms). (F) Same distribution as in E but with logarithmic y-scale to highlight counts for high periods. There is no enrichment for longer periods, suggesting that they are false positives.

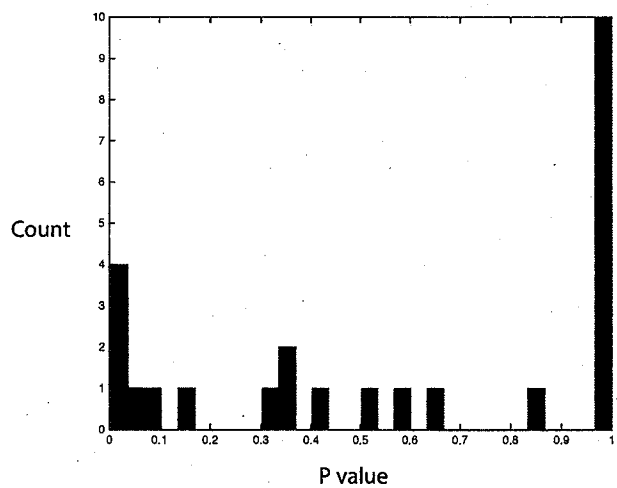


Figure S2. Distribution of p-values for the Lomb-Scargle periodogram peaks with maximum power between 20 and 150 sec for the *Canton-S* song data manually annotated by Kyriacou et al (3). Four of the peaks exhibit p-values < 0.05 and there is not an obvious excess of low p-values.

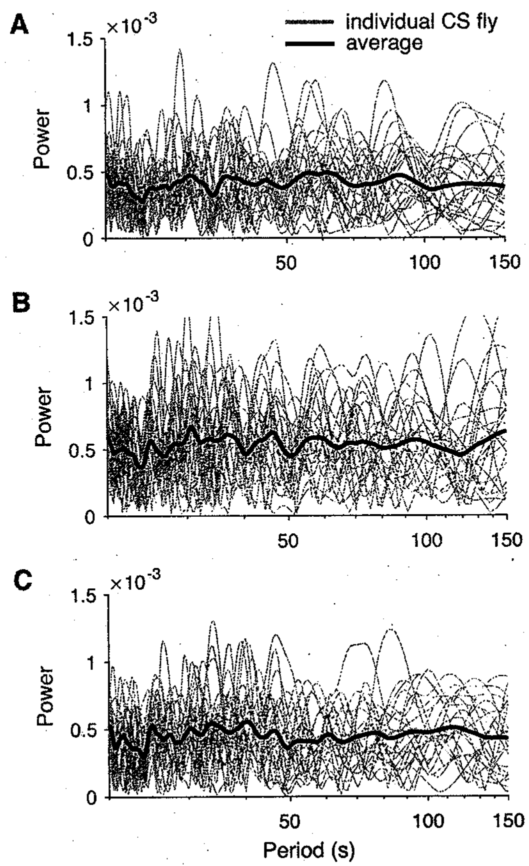


Figure S3. Lomb-Scargle periodograms for *Canton-S* song recordings. (A) From data manually annotated by Kyriacou et al. (3). (B) Automatically segmented data from Stern (3). (C) Automatically segmented data using segmentation parameters from Coen et al. (5). Individual recordings are shown in grey and average over all recordings is shown in black.

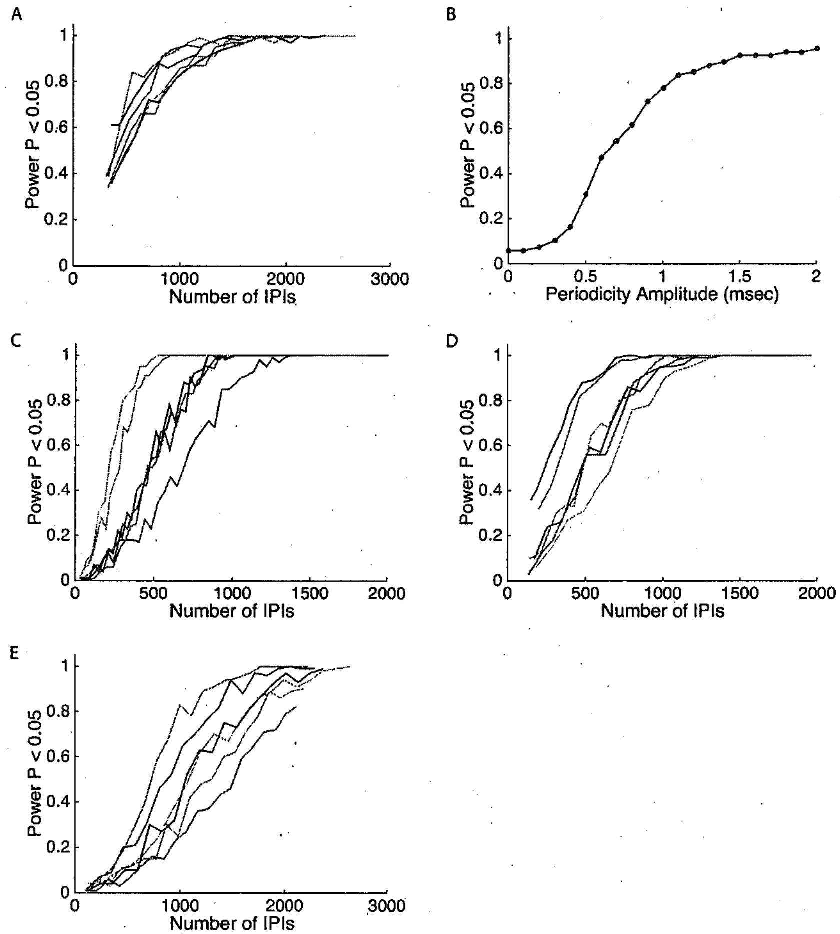


Figure S4. Statistical power analysis under multiple scenarios. (A) Power analysis after ten-second bins of inter-pulse interval data were removed randomly. The plots show the proportion of times out of 100 that periodicity was found between 50-60 sec with $P < 0.05$ for each of six songs containing more than 10,000 inter-pulse interval events. (B) Dependence of power to detect simulated periodicity on periodicity amplitude. Simulated periodicity of 55 sec with amplitude between 0 and 2 msec was imposed on sixty-eight Canton-S songs containing at least 1000 inter-pulse interval measurements. Power equals the fraction of songs that displayed power between 50 and 60 sec at $P < 0.05$. (C, D) Simulated periodicity was added to six songs containing at least 10,000 inter-pulse interval (IPI) events in 45 minutes and then only the first 400 seconds of the song were analyzed. One hundred times, inter-pulse interval data were dropped either randomly (C) or 10 sec bins were dropped randomly (D) and Lomb-Scargle periodogram analysis was performed. (E) Power to detect a sawtooth rhythm. Sawtooth periodicity was added to six songs containing at least 10,000 inter-pulse interval (IPI) events in 45 minutes.

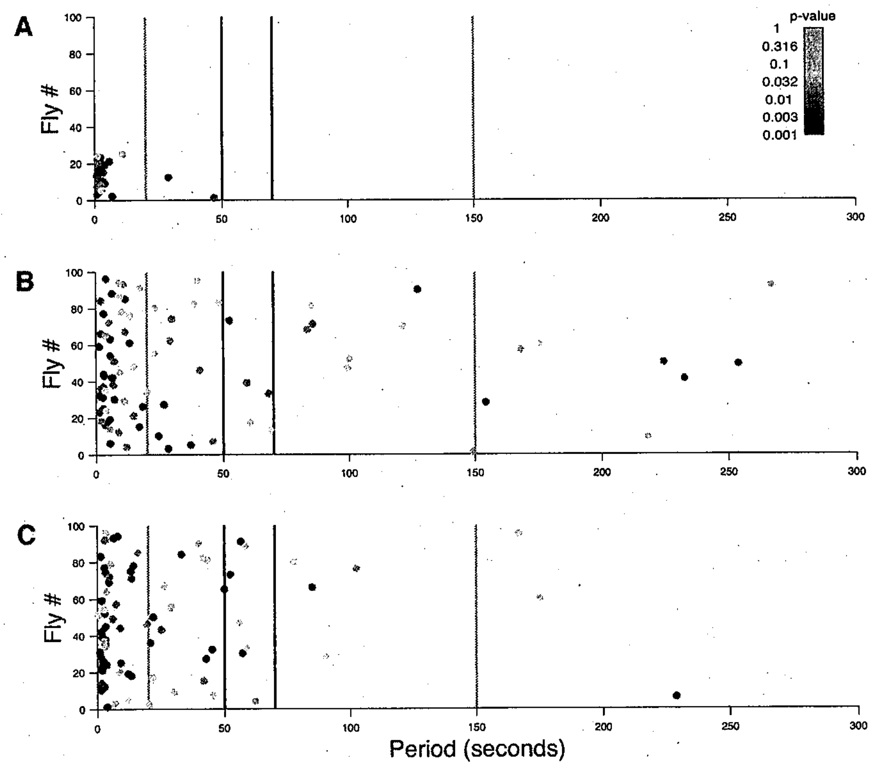


Figure S5. Period of maximum Lomb-Scargle periodogram peaks for inter-pulse interval measurements from multiple individual recordings. (A) Songs manually annotated by Kyriacou et al. (3). (B) Songs automatically segmented in Stern (2). (C) Songs from Stern (2) automatically segmented using parameters defined in Coen et al. (5). In all cases, the vast majority of significant rhythms cluster in the highest frequency range (low period). But both non-significant and significant peaks are distributed widely and apparently at random across the frequency range. In each plot, the 50-70 sec period range is defined by the vertical black lines and the 20-150 sec range defined by Kyriacou et al. (3) is shown with vertical gray lines.

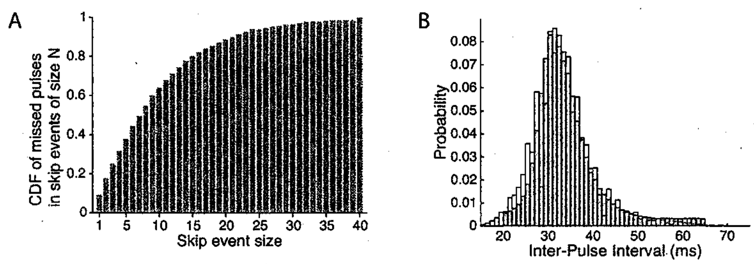


Figure S6. Missed pulse events cause a minor change to the distribution of inter-pulse interval events. (A) Cumulative density function of the number of consecutively missed inter-pulse interval values in the data from Stern (2) illustrates that only 9% of missed pulses were singletons that might alter retained inter-pulse intervals. (B) Histogram of inter-pulse interval data from all *Canton-S* recordings from Stern (2014) in orange and from all manually annotated *Canton-S* recordings from Kyriacou et al. (3) in white. The automatically scored data display a slight excess of inter-pulse interval (IPI) values in the range of approximately 50-65 msec, which are unlikely to significantly alter downstream analysis, as shown by analysis of inter-pulse interval cutoffs in the following panels.

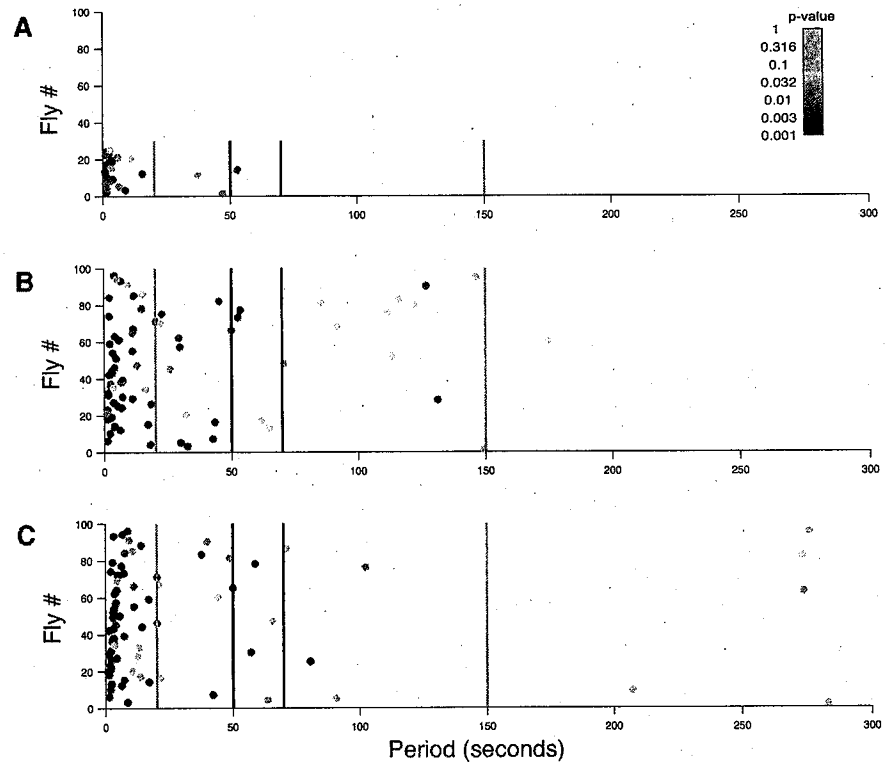


Figure S7. Period of maximum Lomb-Scargle periodogram peaks for inter-pulse interval measurements from multiple individual recordings with an inter-pulse interval cutoff of 55 msec. (A) Songs manually annotated by Kyriacou et al. (3). (B) Songs automatically segmented in Stern (2). (C) Songs from Stern (2) automatically segmented using parameters defined in Coen et al. (5). In all cases, the vast majority of significant rhythms cluster in the highest frequency range (low period). But both non-significant and significant peaks are distributed widely and apparently at random across the frequency range. In each plot, the 50-70 sec period range is defined by the vertical black lines and the 20-150 sec range defined by Kyriacou et al. (3) is shown with vertical gray lines.

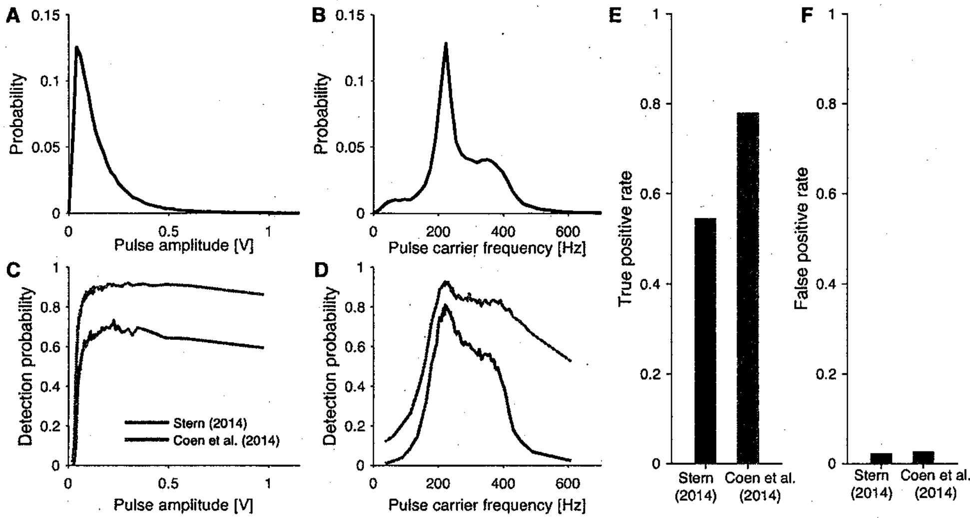


Figure S8. Modification of initialization parameters of FlySongSegmenter influences its performance in detecting pulses. (A, B) Distribution of pulse amplitudes (A) and carrier frequencies (B) for the pulses manually annotated in Kyriacou et al. (3). (C, D) Probability of detecting manually annotated pulses by the automated song segmenter using either the initialization parameters from Stern (2) or Coen et al. (5) versus pulse amplitude (C) or pulse carrier frequency (D). (E, F) True (E) and false (F) positive rate of pulse detection using parameters from Stern (2) and Coen et al. (5) for the pulses manually annotated in Kyriacou et al. (3).

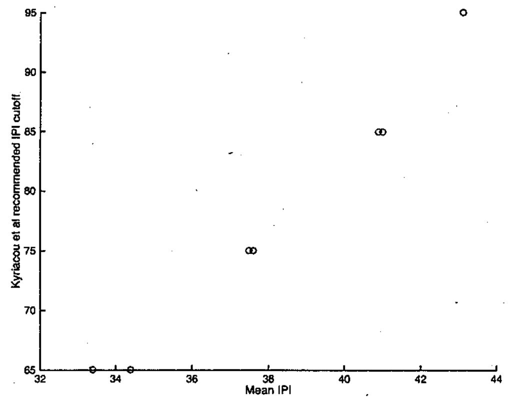
Inter-pulse interval cut-off and temperature control

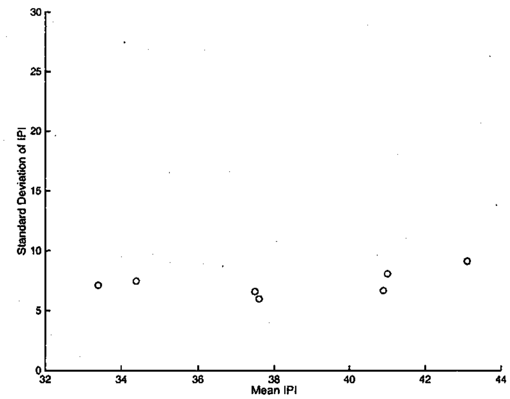
Under the heading "Problem 2: Inappropriate upper IPI cut-offs and poor temperature control," Kyriacou et al. (1) state that Stern (2) used an inappropriate upper inter-pulse interval cutoff for some of the songs and that temperature was not controlled during experiments. We address each concern in turn.

Inter-pulse interval cut-off: Kyriacou et al (1, 3) recommended that the IPI cut-off should scale with the mean inter-pulse interval for a genotype. They did not indicate precisely how the cut-off should scale with the mean. In their table S1, they indicated a "more appropriate cutoff" for each genotype without a quantitative description of how this cutoff should be calculated. The mean inter-pulse intervals and standard deviations calculated from all songs with > 1000 IPIs are shown below along with their recommended upper cut-off.

	per^{01}	*per*^{L}	*per*^{S}	*D. simulans*	CantonS	CantonS Manual	*per*^{L} Manual
Mean IPI	41.0	37.6	40.9	43.1	34.4	33.4	37.5
Recommended IPI cut-off	85	75	85	95	65	65	75
Std Dev IPI	8.09	5.99	6.72	9.13	7.46	7.11	6.59

The mean inter-pulse interval varies by less than 10 msec, but the recommended cut-offs vary by 30 msec. The slope of the regression of mean inter-pulse interval and the recommended cut-off is 3.1 ($y = 3.1x - 40$). In essence, Kyriacou et al. assume that the standard deviation in inter-pulse interval increases considerably faster than the mean inter-pulse interval (plot below left). We find, in contrast, that the standard deviation in inter-pulse interval is relatively constant across genotypes ($y = 0.14x + 1.8$ for automated data) (plot below right). Changing the cutoff by the change in the mean, rather than 3X faster than the mean, is justified by these observations.

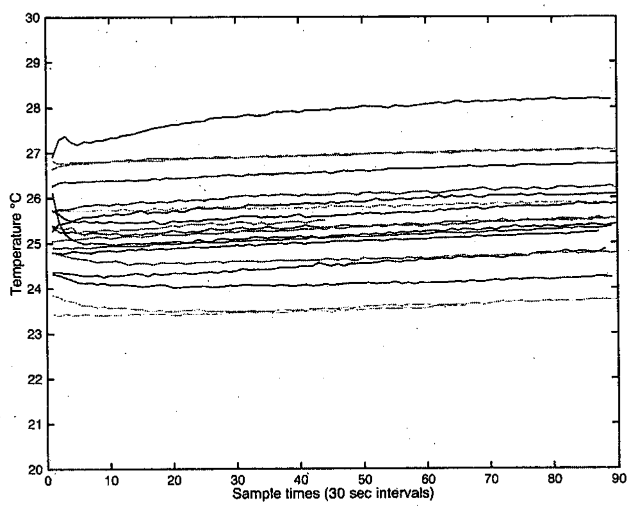




Even more importantly, however, in the main manuscript we report simulations where we progressively reduced the IPI cutoff for song with simulated rhythms. We find that the upper cut-off can be reduced from 65 ms to at least as low as 25 ms and simulated periodicity can still be detected as long as the song retains at least 1000 inter-pulse interval events. It is unlikely, therefore, that any particular IPI cutoff has any influence on the ability to detect song periodicity.

Temperature: Environmental temperature is known to influence the inter-pulse interval of courtship songs. There is no report that temperature can influence the proposed rhythm in the inter-pulse interval, but Kyriacou et al (1) claimed that the experiments reported in Stern (2) had poor temperature control and that this might cause problems with analysis.

We re-examined the data and found that, indeed, average temperature did vary between recording sessions with a range of approximately 4.3°C. However, within each 45-minute recording session, temperature varied on average with a range of 0.52°C. On average, temperatures in the chambers increased slightly over the course of the recording session, likely due to the heat produced by the electronics. In the plot below, we show the temperature for each experiment shown in a different color over each approximately 45-minute recording.



While these slight differences in temperature over the course of each experiment are expected to have a subtle effect on the inter-pulse interval, it is not clear that song periodicity should *disappear* as a result of these small temperature changes. One might imagine that the periodicity might differ at different temperatures, but the essential point of Stern (2), emphasized by results in this paper, is that periodicity itself could not be detected.

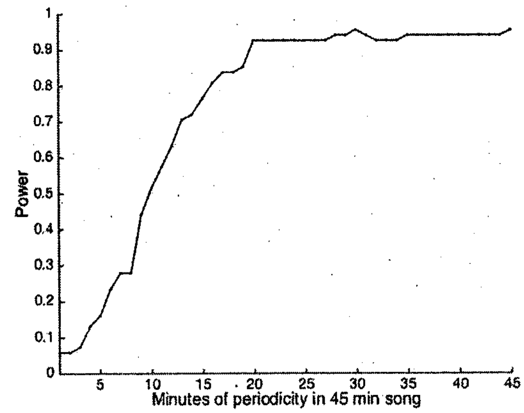
Kyriacou et al also stated that differences in mean IPI should be incorporated into changes in the IPI cutoff. Different experiments were recorded at temperatures that varied by at most ~4°C, although most experiments were recorded at temperatures of between ~25°C and ~26°C. Variation of ~4°C is expected to alter mean IPI by only ~5 msec (4). So, this might justify a change in the IPI cutoff of up to a maximum of 5 msec, which is unlikely to alter any of the statistics substantially. In addition, we showed in the main manuscript that changing the IPI cutoff by up to 40 msec (from 65 msec to 25 msec) has little effect on the ability to detect simulated rhythms, so a small change in the IPI cutoff is unlikely to resolve the question of whether IPI periodicity exists.

Length of courtship

Under the heading "Problem 3: Unrealistic length of courtship," Kyriacou et al. (1) state that "courtship interactions under natural conditions are brief," lasting less than 30 sec and therefore question the use of 45 minute recordings of song. (Of course, if courtship really lasted less than 30 sec, then 50-60 sec periodicity could not be detected.) The key reference the authors cite for natural courtships (5) indeed reported that the majority of courtship interactions lasted less than 30 seconds, however, none of the 153 courtship interactions observed in that study ended in copulation. It is possible that most or all of the females studied were not virgin and were unwilling to participate in courtship. Therefore, these data are not relevant to the question of how long courtship between a male and virgin female persists in nature.

Kyriacou et al. (1) further question the use of 45 minute recordings because circadian rhythms can dampen quickly, citing (6). Reference (6) reports on dampening of circadian rhythms during real-time luminescence recording from cultured explanted rat superchiasmatic loci over the course of approximately 10 days. One can imagine multiple reasons why cultured cells would display a dampened rhythm over 10 days. It is not clear how this is relevant to a presumptive song rhythm over a roughly 45-minute time span.

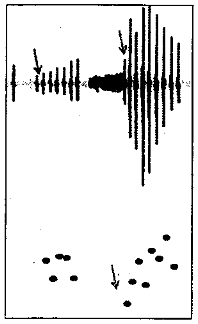
Nonetheless, we decided to investigate this issue more closely. First, we examined the power to detect periodicity in songs if the periodicity was present for only the first N minutes of the song. Periodicity was imposed on the first N minutes of 45 minute recordings for 68 Canton-S songs that contained more than 1000 inter-pulse interval measurements and the average probability of detecting this periodicity with LS periodogram analysis is reported as power in the plot below. We retained power greater than 0.8 as long as periodicity persisted for at least the first 16 minutes. In addition, the probability of detecting periodicity rose above random (P = 0.05) with as little as three minutes of periodicity. Thus, it is extremely unlikely that we would have failed to have detected periodicity in the song recordings as long as periodicity persisted for more than a few minutes.

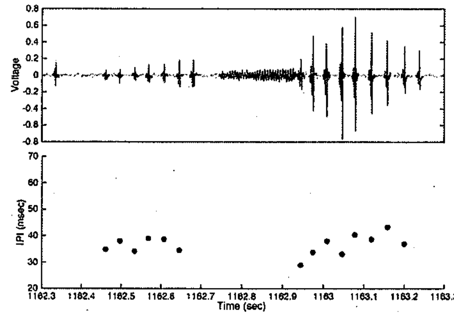


Furthermore, if we perform power analysis only on songs 400 sec long, then we retained power of > 0.8 as long as these short songs contained at least 1000 inter-pulse interval events (Fig. S4a), even when pulses were dropped in 10-sec bins (Fig. S4b). Thus, there is no evidence that the length of courtship recordings generated data that are biased against detecting courtship rhythms.

Reanalysis of Stern's primary matlab song records

Kyriacou et al. (1) observed an apparent error (blue arrow below) in the calling of an inter-pulse interval in Figure 1b of Stern (2) and report this in Fig S2 of their paper. Figure 1b in Stern (2), reproduced below on left, was derived from experiment PS_20130625111709_ch3, sample points approximately 1162.3 sec to 1163.3 sec. We have re-examined the original data and find that the apparently missing inter-pulse interval is in fact found in the csv file that was provided with the original manuscript, but was inadvertently deleted during construction of the figure. We have replotted the data below on the right.





References

1. Kyriacou CP, Green EW, Piffer A, Dowse HB, Takahashi JS (2017) Failure to reproduce period-dependent song cycles in Drosophila is due to poor automated pulse-detection and low-intensity courtship. *PNAS*. doi:10.1073/pnas.1615198114.
2. Stern DL (2014) Reported Drosophila courtship song rhythms are artifacts of data analysis. *BMC Biol.*
3. Kyriacou CP, van den Berg MJ, Hall JC (1990) Drosophila courtship song cycles in normal and period mutant males revisited. *Behav Genet* 20(5):617–644.
4. Peixoto AA, Hall JC (1998) Analysis of temperature-sensitive mutants reveals new genes involved in the courtship song of Drosophila. *Genetics* 148(2):827–38.
5. Gromko M, Markow T (1993) Courtship and remating in field populations of Drosophila. *Anim Behav* 45:253–262.
6. Yamazaki S, Takahasi JS, Takahashi JS (2005) Real-Time Luminescence Reporting of Circadian Gene Expression in Mammals. *Methods Enzymol* 393:288–301.

Failure to reproduce *period*-dependent song cycles in *Drosophila* is due to poor automated pulse-detection and low-intensity courtship

Charalambos P. Kyriacou[a,1], Edward W. Green[b], Arianna Piffer[c], and Harold B. Dowse[d]

[a]Department of Genetics, University of Leicester, Leicester LE1 7RH, United Kingdom; [b]Division of Translational Immunology, Deutsches Krebsforschungszentrum, 69120 Heidelberg, Germany; [c]Medical School, San Raffaele University, 20132 Milan, Italy; and [d]School of Biology and Ecology, University of Maine, Orono ME 04469-5751

Edited by Joseph S. Takahashi, Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, TX, and approved January 13, 2017 (received for review September 13, 2016)

Stern has criticized a body of work from several groups that have independently studied the so-called "Kyriacou and Hall" courtship song rhythms of male *Drosophila melanogaster*, claiming that these ultradian ~60-s cycles in the interpulse interval (IPI) are statistical artifacts that are not modulated by mutations at the *period* (*per*) locus [Stern DL (2014) *BMC Biol* 12:38]. We have scrutinized Stern's raw data and observe that his automated song pulse-detection method identifies only ~50% of the IPIs found by manual (visual and acoustic) monitoring. This critical error is further compounded by Stern's use of recordings with very little song, the large majority of which do not meet the minimal song intensity criteria which Kyriacou and Hall used in their studies. Consequently most of Stern's recordings only contribute noise to the analyses. Of the data presented by Stern, only per^L and a small fraction of wild-type males sing vigorously, so we limited our reanalyses to these genotypes. We manually reexamined Stern's raw song recordings and analyzed IPI rhythms using several independent time-series analyses. We observe that per^L songs show significantly longer song periods than wild-type songs, with values for both genotypes close to those found in previous studies. These *per*-dependent differences disappear when the song data are randomized. We conclude that Stern's negative findings are artifacts of his inadequate pulse-detection methodology coupled to his use of low-intensity courtship song records.

Drosophila | courtship song | cycles | *period* gene | interpulse interval

During courtship, the *Drosophila melanogaster* male vibrates his wing toward the female and produces a series of pulses and hums (1). The pulses have a variable interpulse interval (IPI), which ranges from 15 to 80 ms but usually averages between 30 and 40 ms, whereas sympatric *Drosophila simulans* mean IPIs vary from 45 to 80 ms depending on the strain (2, 3). In 1980, in these pages, the first of a series of studies by Kyriacou, Hall, and their collaborators revealed that superimposed on these IPIs was a low-amplitude oscillation of about 60 s in *D. melanogaster* and 40 s in *D. simulans* (4–12). Furthermore, in *D. melanogaster*, these cycles were modulated in a predictable fashion by the circadian rhythm *period* (*per*) mutations (13): per^L males with long 29-h circadian cycles also sang with long ~80-s song cycles, whereas circadian arrhythmic per^{01} males showed a corresponding song phenotype (2, 4, 9, 10). These cycles were shown to have functional significance in playback experiments in which females were shown to be most responsive to both their species-specific IPI and cycle (2, 14–16).

The work of Kyriacou, Hall, and collaborators was performed in the late 1970s and 1980s using extremely laborious analog technology, so to attain any kind of throughput, IPIs were binned into consecutive 10-s intervals and a mean IPI calculated on a minimum of 10 IPIs (2, 4–6, 8). Initially, sine/cosine functions were fitted to the mean IPIs (2, 4–6, 8), and this was later complemented by spectral analyses, with both types of statistical methods having associated significance tests (9–12, 17). Males were raised in small single-sex groups and were usually recorded at 3 d of age with an unreceptive <24-h virgin female, thereby prolonging the courtship. The vast majority of courtships were robust for 5–7 min, generating many hundreds of IPIs per song. Those few courtships in which one-third or more 10-s bins were empty (<10 IPIs) were not analyzed. The results from several studies (including several performed blind) revealed repeatedly and consistently that wild-type flies sang with ~50- to 65-s cycles, and that these periods were altered by *per* mutations (2, 6–8, 12).

Using different methods, Alt et al. (18) obtained similar results to the original song-cycle work, whereas Ritchie et al. (19) replicated and extended the playback experiments. It also became apparent that song cycles could be masked under conditions where the male and female were confined in small cells, when inappropriate upper IPI cut-off limits were used, or when gaps in the song record were artificially added (9, 20, 21). Against this background, Stern has used fully automated pulse detection to conclude that IPI cycles are nonexistent and consequently there are no *per*-specific differences among genotypes or species (22). We have therefore reanalyzed Stern's primary song recordings that are provided on his website (22). Among other serious errors, we observe that Stern's automated method only detects ~50% of song pulses. Unambiguous manual logging of pulses from Stern's primary song records using his software platform FlySongSegmenter (23) reveals *per*-dependent song rhythms with mean period values close to those published in previous studies.

Significance

The study of ~60-s courtship song rhythms in *Drosophila* and their modulation by *period* clock mutations plays an important historical role in developing the molecular basis of the circadian oscillator. Carried out mostly in the 1980s using extremely laborious analogue methods, key features of the work were replicated by independent groups in the following decade. Recently, a study by Stern, using automated methods for song detection, has failed to reproduce these findings. By manually logging the same songs and comparing the results to the corresponding automated analyses we observe that Stern's method detects only ~50% of the song, but in addition, incorporates spurious errors. Unambiguous manual reanalysis confirms the *period*-dependent nature of fly song rhythms as originally reported.

Author contributions: C.P.K. and H.B.D. designed research; C.P.K., E.W.G., A.P., and H.B.D. performed research; C.P.K., E.W.G., A.P., and H.B.D. analyzed data; and C.P.K., E.W.G., and H.B.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The raw data has been deposited in Dryad (dx.doi.org/10.5061/dryad.80c1f).

[1]To whom correspondence should be addressed. Email: cpk@leicester.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1615198114/-/DCSupplemental.

Results

A Critical Assessment of Stern's Methods. Based on Stern's claims that his automated pulse-picking algorithms are highly accurate (22, 23), we initially examined the processed IPI results that Stern deposited on his website as .csv files. Each file has a series of time coordinates stretching to 45 min, with the corresponding IPI (22). We immediately noticed several problems.

Problem 1: Poor intensity courtship song. Stern analyzed all courtships exceeding 1,000 IPIs over a 45-min period, and many courtships only just exceed this threshold, giving an average of ~four IPIs per 10 s time bin (Fig. 1*A*). Such sporadic songs could not possibly provide any test for song cycles. The songs analyzed by Kyriacou and Hall were vigorous with few gaps, with most songs containing an average of 30–40 IPIs per 10-s bin, and if extrapolated to a 45-min recording would translate to 8,100–10,800 IPIs or 5,400–7,100 IPIs even if one-third of the bins were empty (<10 IPIs). We fortuitously found song transcripts for two genotypes processed by hand in the mid-1980s: $per^{01}/w^{+}Y$ and $per^{+}/w^{+}Y$. The lengthier songs that did not result in copulations provided an estimate of song period (7). The average number of IPIs per 10-s bin is 33.3, predicting 8,979 IPIs in 45 min (Dataset S1). Furthermore the average proportion of gaps (<10 IPIs per 10 s) is 9.1% with 12 of 27 songs having no gaps at all (Fig. 1*B* and Dataset S1). These are unremarkable courtships in that most songs analyzed by Kyriacou and Hall (2, 4–12) showed this kind of vigor.

Fig. 1*A* shows examples of the Canton-S courtships that met Stern's criterion for analysis (>1,000 IPIs). Fig. 1*A*, *Top* shows 45 min of a Canton-S song with just over 1,000 IPIs. The IPIs have been placed in 10-s bins as means. Only 23 points of a possible 270 are present, making any attempt to fit a cycle to these mean IPI values by spectral analysis (or any method) completely spurious, whether analyzing the complete dataset, or successive 5-min clips as performed by Stern (22), as each clip includes an average of only two to three data points. Of 312 courtships, 148 recorded by Stern had >1,000 IPIs (Fig. 1*B*) and were included in his analysis, yet one-third of these contained only 1,000–2,000 IPIs. Progressively more vigorous Canton-S songs are illustrated in Fig. 1*A*, but until songs contain >5,000 IPIs there are too many gaps for effective determination of any periodicity in the first 5–7 min, strikingly contrasting with Kyriacou and Hall's robust song production in the first few minutes of courtship (4–12) (Dataset S1).

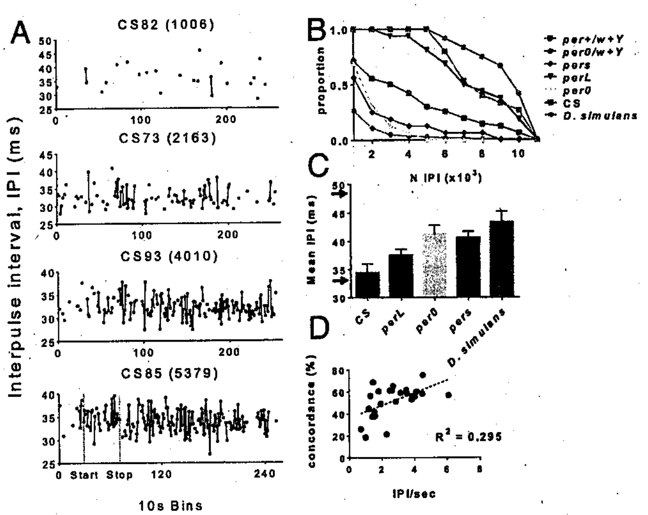


Fig. 1. Stern's song data. (*A*) Each panel shows songs analyzed by Stern (22) for 45 min, with the mean IPI calculated for every 10-s time bin. Numbers in brackets refer to the number of IPIs. Lines connect consecutive means based on 10 or more IPIs. Note that these are very sparse until the courtship has >5,000 IPIs, and even so the courtship is intermittent at the beginning of the recording period. "Start" and "Stop" on bottom panel show the points at which the reanalysis of 400 s of CS85 song analysis was initiated and ended (Dataset S2). (*B*) Song vigor is very low in all Stern genotypes except per^{L}. *x* axis: Song vigor as number of IPIs in 45 min. *y* axis: Proportion of songs. The vast majority of all genotypes except per^{L} generate considerably fewer than 5,000 IPIs, which was the threshold at which we analyzed songs in this study. All $per^{01}/w{+}Y$ and $per^{+}/w^{+}Y$ songs analyzed in the mid-1980s (7) would produce >5,000 IPIs in 45 min (see also Dataset S1). (*C*) Mean IPIs (±SEM) for each genotype. The arrows show the species typical IPIs for *D. melanogaster* (black) and *D. simulans* (gray). (*D*) The concordance between Stern's automated pulse calls compared with manual detection is poor and correlates with song vigor. The songs of the first 25 Canton-S males (including CS71; Dataset S3) were manually processed and the concordance between those IPIs correctly called by Stern and the manual method plotted against the vigor of each song (number of manually scored IPIs per second of courtship).

From Stern's processed IPI records in .csv files, we were able to calculate the number of IPIs generated in the 45-min recording period for all five genotypes (22). Of Stern's 96 wild-type songs, 28 meet a criterion of >5,000 IPIs in 45 min; only 4 of 167 *D. simulans* songs reach this criterion, the vast majority have less than 1,000 (Fig. 1*B*); in sharp contrast, per^{L} males sing robustly with 13 of 16 having >5,000 IPIs, but only 1 of 16 of the per^{01} and 2 of 16 per^{s} produce >5,000 IPIs (Fig. 1*B*). Consequently, to reanalyze at least a suitable fraction of Stern's data, we maintained our criterion for analysis as >5,000 IPIs in 45 min, giving us initially 28 Canton-S and 13 per^{L} songs. Two further per^{L} songs had >4,000 IPIs, so we added these to increase the sample size (Fig. 1*B*).

Problem 2: Inappropriate upper IPI cut-offs and poor temperature control. To distinguish an IPI from an interburst interval, Kyriacou and Hall routinely took 2× the mean IPI as the upper IPI limit, and this was applied on a song-by-song basis (24). On inspecting Stern's *D. simulans* IPI records (22), it was clear that instead of raising the maximum IPI cut-offs to match the higher mean IPIs of this species (usually reported as ~48 ms *D. simulans* (1), Stern maintained the *D. melanogaster* cut-off of 65 ms instead of a more realistic 95 ms (22). This approach may explain why this species' mean IPI is uncharacteristically low (43.36 ± 0.38 ms, SEM) (Fig. 1*C*). From Stern's processed IPI files, we calculated the Canton-S mean IPI at ~35 ms, so his cut-off of 65 ms was not unreasonable, and the per^{L} mean at 37.5 ms (Fig. 1*C*). Kyriacou and Hall would have applied an average cut-off of 75 ms for these per^{L} songs on an individual song-by-song basis. In addition, nearly all of the per^{01} and per^{s} songs gave mean IPIs > 40 ms (Fig. 1*C*), so it would have been more appropriate to increase the IPI cut-off to ~85 ms.

We subsequently interrogated Stern's primary song Matlab files to obtain the distribution of his individual IPI calls from the songs of all five of his genotypes using a revised upper IPI cut-off of 75 ms for per^{L}, 85 ms for per^{01} and per^{s}, and 95 ms for *D. simulans* and calculated the number of IPIs that would have been included using these more realistic values. We then compared these values to those obtained by Stern (22) using his exclusive 65-ms IPI cut-off (Table S1). Stern missed a tolerable 3% of per^{L} IPIs, which rose to 6% and 7% for per^{s} and per^{01}, respectively. However, for *D. simulans*, Stern missed ~20% of the IPIs by using the inappropriate *D. melanogaster* cut-off (Table S1). Within Stern's Matlab files, we also found his temperature recordings for each song. The large numbers of Canton-S and *D. simulans* songs had been recorded between 23.6 and 27.9 °C, and the *per* mutants between 23.5 and 24.6 °C. Weighted mean recording temperatures for Canton-S were 25.5 °C, for *D simulans* 25.8 °C, for per^{L} 24.6 °C, and for per^{01} and per^{s} 23.8 °C. The inverse correlation between temperature and IPI (25) therefore accounts for why the latter two mutants have generally longer mean IPIs compared with per^{L} and wild type. We were surprised that this basic environmental variable, which has important implications for setting upper IPI cut-offs, was so poorly controlled.

Problem 3: Unrealistic length of courtship. Courtship interactions under natural conditions are brief, and observations in nature reveal that most interactions last less than 30 s, although a few courtships last minutes (26). Consequently, we question Stern's (22) approach of analyzing 45-min records, given that circadian cycles can damp quickly, particularly when they are generated in the periphery (27) as song rhythms are likely to be (28). This may be reflected in the observation that for the 25 most vigorous Canton-S songs we selected (see below), 19 of 25 songs had higher mean IPIs in the 1st compared with the 45th minute (34.98 vs. 33.92 ms, paired t, $P = 0.0039$, distribution χ^2, $P = 0.009$).

Analysis of Stern's Processed IPI Files Do Not Reveal *period*-Dependent Song Cycles. We took Stern's (22) processed IPI files for the most vigorous Canton-S and *per*^{L} songs and selected the first 400 s of song to match the 5- to 7-min records analyzed by Kyriacou and Hall in their studies (2, 4–12). When we initiated our analysis at the same point as Stern, most songs had fewer than 10 IPIs in more than one-third of the 10-s time bins (Fig. 1*A*) and would therefore have been discarded by Kyriacou and Hall. We therefore shifted up from the beginning of Stern's song analysis, until we found an ~400-s section that was vigorous. For example, the *Bottom* panel of Fig. 1*A* shows the start and stop points of our analysis for song CS85. Even then, if there were still many gaps, we extended the record for another 50–80 s so that we had at least 30 mean IPI data points and no more than one-third of time bins had <10 IPIs. If a song had not become sufficiently vigorous after shifting forward for 300 s from the beginning of Stern's analysis, we did not proceed further. Dataset S2 shows the exact points from Stern's records that each analysis was initiated and for how long it was extended. In this way we were able to study 25 of the 98 Canton-S songs and 14 of the 16 *per*^{L} songs. We used cosinor analysis, which is particularly suitable for short time series (29), Lomb-Scargle (L-S) periodogram as used by Stern (22), and CLEAN spectral analysis (30) (Fig. 2). These three methods do not require equidistant time points. We also analyzed the binned data by autocorrelation (after using interpolation to fill gaps; see *SI Methods*), where we took regularly repeating peaks in the correlogram as evidence for rhythmicity (31, 32). We did not use the autocorrelation-derived song period because it is only resolved to the nearest 10 s. Finally we took the unbinned IPIs and used the L-S method to investigate periodicity in the raw IPI datasets.

Because the mean IPIs are calculated from 10-s bins, the minimum period (Nyquist) is twice the bin size, so we took the highest peak in the spectral analyses >20–150 s as the period, irrespective of whether it was significant or not (arrows in Fig. 2). Very few of the peaks in the L-S periodiogram were significant for any song. This result is not unexpected given the small number (30–45) of time bins (33), whereas most of the cosinor plots were significant (Dataset S2). CLEAN has an associated Monte Carlo procedure that gives the 95% confidence limits for the period, not whether there is significant rhythmicity (34), so we used the highest period observed in CLEAN as an independent spectral analysis to support the L-S method (Fig. 2 and Dataset S2).

Fig. 2 illustrates the analyses for binned and raw (unbinned) IPIs for the first vigorous songs from Stern's list for each genotype, Canton-S (CS2) and *per*^{L1} (perL1). The corresponding panels in Fig. 2 show the IPI means/time plots for the two songs, respectively, for Stern's automated analysis. The adjacent columns show the cosinor, L-S, and CLEAN analyses of the binned data. The right hand column illustrates the corresponding L-S periodogram for the raw, unbinned IPIs. From the CS2 cosinor analysis, the most significant period is 22 s (arrowed in Fig. 2) ($P = 0.011$), whereas for both L-S and CLEAN it is 48.8 s (*Top* row in Fig. 2). The raw IPIs give the highest L-S peak at 0.8 s; but if we inspect the 20- to 150-s domain to compare with the binned results, it is 48.5 s. For perL1 in row 3 of Fig. 2, we observe two prominent cosinor periods: the first at 27 s ($P = 0.077$) and the second at 77 s. L-S and CLEAN give periods of 26.2 s and 33.5 s, respectively. The raw L-S analysis for perL1 generates the highest peak at 1.2 s, with a much lower peak at 33.9 s that represents the best period in the 20- to 150-s range (see Fig. S1 for autocorrelations).

The song periods observed for all of the Canton-S and *per*^{L} songs are illustrated in Fig. 3*A*, and although there are slight average increases in period length for *per*^{L} compared with Canton-S

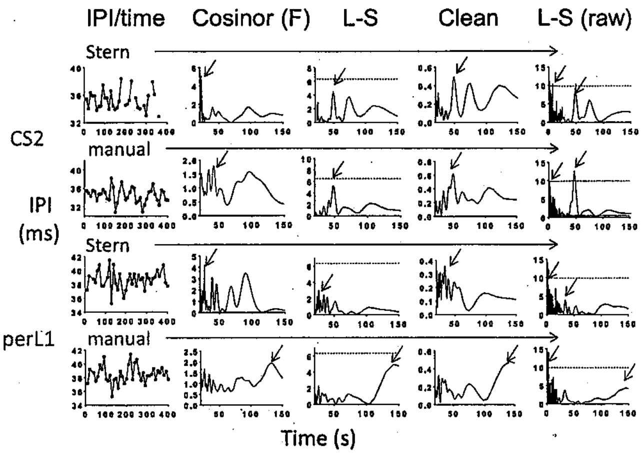


Fig. 2. Reanalyzed Canton-S and *per*^{L} songs. *Upper* two rows: top row shows the mean IPI/10 s time plot for song CS2 for 400 s from the point we initiated our reanalyses from Stern's processed .csv files (Dataset S2), the corresponding cosinor *F*-plot, the L-S periodogram and CLEAN spectrogram followed by the L-S plot on the unbinned, raw IPIs. The second row shows the manually processed results for this song. *Lower* two rows: corresponding results for song perL2 (blue). Arrows show the peaks in the different time series analyses. The two arrows on the L-S raw IPI plots show the peaks in the high frequency and 20- to 150-s range. Dotted lines show 95% confidence limits for the L-S periodogram.

when using the two spectral analyses (66 vs. 51 s and 54 vs. 43 s for L-S and CLEAN, respectively), these are not significantly different, and are not evident in the mean periods determined by cosinor (ANOVA genotypes $F_{1,\ 111} = 1.63$, methods $F_{2,\ 111} = 1.03$). We also compared the two genotypes within each statistical method separately using the nonparametric Mann–Whitney U test and, again, none were significant. Fig. 3 *D* and *E*, *Left*, show the mean song periods based on the highest peaks in the 20- to 150-s and 0- to 150-s domains for the raw IPI logged automatically by Stern (22). Again, there are no significant differences between the genotypes either by parametric or nonparametric analysis (see also Dataset S2).

We conclude that even when we select the few vigorous songs, Stern's analysis (22) based on automated pulse detection provides little evidence for any *per* allele-specific differences in song periods. We further agree with Stern that when raw IPI data are examined, the highest peaks found by the L-S method are usually in the very high frequencies.

Reanalysis of Stern's Primary Matlab Song Records. The figures Stern presents suggest that his algorithms are exceptionally efficient at picking pulses (22, 23); however, we did notice an error in figure 1B of ref. 22, suggesting a problem with his automated pulse detection method (Fig. S2). Consequently, we examined Stern's primary song files manually using Stern's FlySongSegmenter platform (23), which provides a visual trace and a corresponding frequency plot for each pulse with an additional acoustic monitoring option. Indeed FlySongSegmenter provides an excellent platform for manual pulse-song analysis that generates very little ambiguity.

Stern's Automated Method for Detecting Pulses and IPIs Is Highly Unreliable. We selected at random song CS71, which according to Stern produced 2,189 IPIs, and we manually scored the song during the first 300 s of courtship (Fig. 4). Pulses detected by Stern's automated method are shown in blue on the song trace in Fig. 4, with manual analysis plotted in red and green, the latter color identifying "pulse-skipping" where Stern's analysis has failed to detect a pulse that is flanked by two others, thereby generating a very long IPI as an artifact. Given that the original analyses of Kyriacou and Hall (2, 4–12) were based on mean IPIs, the implications of even a single spurious long IPI contributing to the

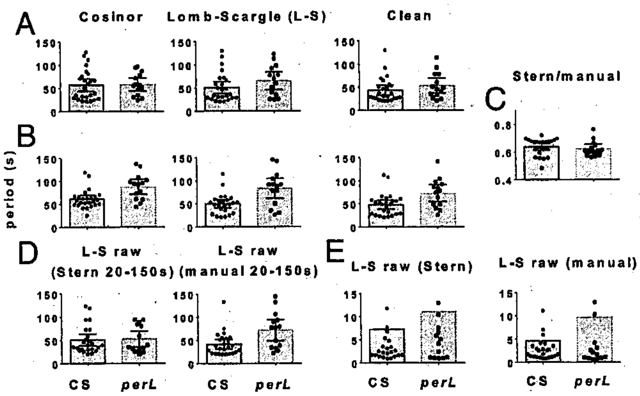


Fig. 3. Song rhythms of Canton-S and per^L songs. (*A*) Mean periods for 10-s binned Canton-S (pink) and perL (blue) songs (+95% confidence limits) from cosinor, L-S, and CLEAN calculated from Stern's processed IPI files. (*B*) Corresponding results from manually corrected method. (*C*) Proportion of IPIs observed by automated compared with manual method (Stern/Manual). (*D*) L-S derived periods from unbinned raw IPIs from Stern's automated method compared with manually processed song pulses (20- to 150-s range). (*E*) Same as in *D*, but L-S periods from 0 to 150 s (single values of 94.7 s for CS2 and 94.5 s for perL10 from Stern's automated analysis, and 47.1 s for CS2 and 93.8 s for perL14 from the manual analyses are omitted from this figure, but not from the calculations, in order not to compress the *y* axis; see Dataset S2 for all results).

mean are obvious. Above the song trace in Fig. 4 is shown the frequency plot for each pulse, which is extremely helpful in determining whether the signal represents a real pulses or noise. For example, one IPI detected by Stern is generated by noise (third panel, top row in Fig. 4). This pattern of errors is reflected in the whole 300-s segment where 184 IPIs are manually logged with Stern correctly calling 48, but he also adds an additional 4 very long IPIs (but <65 ms) by pulse-skipping plus the spurious IPI produced by wing noise, giving an unacceptably low concordance rate between manual and automated analyses of 26% (48 of 184) (Fig. 4).

To pursue this further, we manually analyzed the IPIs for the first 25 wild-type songs that sang for at least 60–90 s, including our analysis of CS71. We were astonished to see that Stern's concordance rate compared with the manual analysis varied between 18% and 75%, with an average of 51.9%, but with an additional 3–4% of spurious pulses (mostly pulse skipping) (Dataset S3). We also observed a significant correlation between song vigor, measured as manually assessed IPIs generated per second of courtship and concordance ($R^2 = 0.295$, $P = 0.005$) (Fig. 1*D* and Dataset S3), supporting our earlier decision to select only the most vigorous songs from Stern's processed IPI files for reanalysis.

Manual Reanalysis of Stern's Vigorous Courtships Reveals *period*-Dependent Song Cycles in Binned and Unbinned Data. The unsatisfactory performance of Stern's automated method obliged us to manually log ~73,000 IPIs for the corresponding 25 Canton-S and 14 per^L song segments we had earlier studied using Stern's IPI calls. From our manual analysis, which included visual (using the song trace and pulse-frequency plots) (Fig. 4) and acoustic monitoring of each pulse, we observed that Stern's automated method detected ~63% of the IPIs detected manually for both Canton-S and per^L (Fig. 3*C* and Dataset S2), but this is an overestimate because we did not exclude the few percent of spurious IPIs. The observer repeated the manual analysis for each song and obtained a concordance rate of 98–99% with a few ambiguous IPIs (10–15 per song), usually at the beginning or end of a pulse train. Consequently, the intra-observer reliability of the manual method is extremely high. A second inexperienced observer reanalyzed the ~300 s of courtship song CS71 shown in Fig. 4 from the primary Matlab song recording, and detected 170 of the 184 IPIs with 10 additional putative IPIs, giving a concordance of 92.4% with the first observer. Subsequently, the first 300 s of two additional song records (CS6 and CS10) were also compared. The IPI concordance for song CS6 between the two observers improved to 97.5% (with an additional putative 1.5% IPIs detected by the naïve observer) and 99.1% for CS10 (with an additional 1.1% IPIs detected by the naïve observer). Consequently, inter- and intra-observer reliability is extremely high, and is enhanced with practice.

Fig. 2 also illustrates the corresponding manual analyses for songs CS2 and perL1. The arrows on the panels in Fig. 2 show that the peaks for cosinor, L-S and CLEAN are consistent for the binned data for CS2 at ~48 s, whereas for perL1 all peaks lie at ~145 s, supported by the correlograms (Fig. S1). In the unbinned raw data, the L-S profile for CS2 shows the peak at 47 s, whereas for perL1 the highest peak is in the high frequencies, yet in the 20- to 150-s range it falls at 145 s, just as in the binned data. Fig. 3*B*

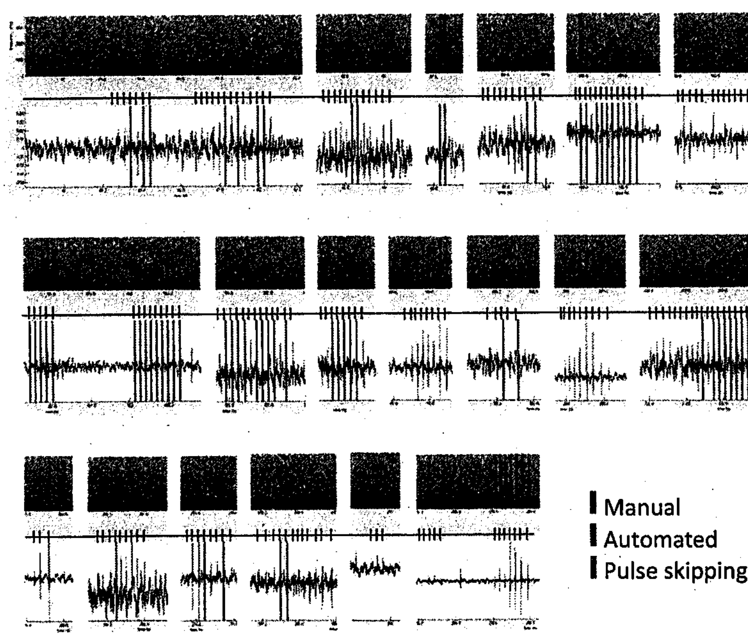


Fig. 4. Stern's automated pulse detection is highly insensitive compared with manual analysis and also generates IPI artifacts. The first 300 s of recording for song CS71 are shown in FlySongSegmenter. Stern's pulse calls in blue, manual pulse calls in red. Seven examples of pulse skipping (green line) by the automated method are shown that create very long IPIs, four of which (<65 ms) were included in Stern's analysis, as was one spurious IPI (third song panel, top row). The *Upper* panel in each row shows the pulse frequency plot.

summarizes the cosinor, L-S, and CLEAN results of binned data from the corresponding manually corrected song files. Statistical analysis revealed a highly significant genotype effect ($F_{1, 111} = 31.2$, $P \sim 0$), with no significant effects for the different statistical methods nor any interaction. The mean period from cosinor of Canton-S songs taken from the binned data were 60.8 s and that of per^L was 87.6 s (Fig. 3*B*). With the L-S and CLEAN algorithms, Canton-S songs vs. per^L gave mean periods of 50.1 s vs. 83.1 s and 48.1 vs. 72.9 s, respectively (Fig. 3*B*). We confirmed all these results with the Mann–Whitney *U* test (cosinor $P = 0.005$, L-S $P = 0.007$, CLEAN $P = 0.009$; all are significant with Bonferroni corrections, even when two-tailed). Furthermore, when we initiated our analysis 5 s through the first bin to generate a maximal displacement of IPIs from their original 10-s bins, cosinor analysis still revealed a significant difference in the periods between the two genotypes (49.6 vs. 74.5 s, $F_{1, 37} = 6.8$, $P = 0.013$ or $P = 0.032$ by Mann–Whitney *U* test, two tailed). Consequently, cosinor analysis does not appear to be particularly sensitive to the IPI position within the bin at which the analysis is initiated. We also examined the raw unbinned IPI data using L-S, and again the high frequencies of 0.8–5 s dominated, with no significant differences between the two genotypes (Fig. 3*E*, *Right*) (Mann–Whitney *U* test, $P = 0.59$). When we took the highest peak in the 20- to 150-s range, we observed a mean period of 46.1 s for Canton-S and 72.2 s for per^L (Fig. 3*D*, *Right*) ($F_{1, 37} = 8.59$ $P = 0.006$, Mann–Whitney *U* test, $P = 0.009$, two tailed).

***period*-Dependent Differences in Song Cycles Disappear in Randomized IPI Datasets.** We then randomized both the binned and corresponding raw IPI datasets, maintaining Stern's interburst intervals of >65 ms for each song and analyzed with L-S. We compared the distributions of the observed to the corresponding randomized binned data by placing the periods into four categories reflecting the per^0 (20–30 s), per^s (30–45 s), wild-type, (45–70 s), and per^L (70–150 s) domains (9), and performed a goodness-of-fit test. For both the binned wild-type and per^L data, the distributions of the observed and randomized binned periods were highly significantly different ($\chi^2 = 20.9$, $P = 0.0001$ for Canton-S with infinite χ^2 for per^L as none of the randomized data period fell into the per^L domain, in contrast to 10 of 14 of the observed per^L periods). A test of independence comparing the observed Canton-S with per^L period distributions was also highly significant, as expected ($\chi^2 = 19.6$, $P = 0.0002$).

For the unbinned IPIs, the peak periods, nearly always in the high frequencies, were placed into four categories corresponding to periods of <1.0, 1–2, 2–5, and >5 s. The goodness-of-fit tests generated significant deviations of the observed from the randomized data (Canton-S $\chi^2 = 26.4$, $P \sim 0$, per^L $\chi^2 = 21.3$, $P < 0.0001$). A test of independence of the observed Canton-S and per^L distributions gave a marginal difference ($\chi^2 = 8.7$, $P = 0.034$). When we repeated the goodness-of-fit test for the peak periods in the 20- to 150-s domain, there was no significant difference between the observed and randomized Canton-S distributions ($\chi^2 = 3.0$ $P = 0.39$) in contrast to per^L ($\chi^2 = 11.2$, $P = 0.011$). A test of independence of the observed 20- to 150-s periods from unbinned data between the two genotypes was also significant ($\chi^2 = 9.91$ $P = 0.019$), further confirming the significance of the period difference between them in this temporal domain.

Stern's Automated Method Performs Poorly Even When the Signal-to-Noise Ratio Is Improved. Finally, we wondered whether the poor performance of Stern's automated pulse detection was also observed in his initial paper, in which he describes the FlySongSegmenter methodology (23). In that study, Stern used much smaller mating cells that focused the flies closer to the sensitive part of the microphone and, consequently, the song pulses would be expected to be louder and generate a higher signal-to-noise ratio, even though under such confined conditions song rhythms are masked (9). We took the 16 Canton-S Tully song records from that work (available as *.wav files) (23) and analyzed up to a few hundred pulses from each song manually, then compared these results to his automated method. We observed an average concordance of 75%, with the best song having an impressive 96% but the worst 11% (Table S2). Furthermore, each song had a number of IPIs automatically detected that were spurious. For example, the song that had the 11% concordance rate had only detected 65 of 329 IPIs but of these 65, 29 were spurious, so only 36 of the IPIs had been successfully detected, generating the concordance value of 11% (36 of 329). The song with the 96% concordance was the only one in which Stern's automated method selected more IPIs than the manual method (111 vs. 93). However, the IPI analysis is severely contaminated because 22 of the 111 pulses were spurious.

Stern's own comparison of his automated and manual methods using 1 min of song from each of nine wild-type songs shows a mean predictive value, *F*, of ~0.9 (23). Our more pragmatic analysis of nearly 66 min of song from 16 wild-type flies from this study suggests a far less-impressive performance for the automated method. Consequently, when using smaller mating cells, the mean concordance rate is enhanced to 75%, but so is the spurious IPI rate to 7% (Table S2), so any improvement in sensitivity using smaller mating cells is largely offset by the higher rate of incorrect pulse detection.

Discussion

Serious methodological flaws have been uncovered in Stern's song analyses (22). We are at a loss to explain why Stern's per^L songs are as vigorous as those studied by Kyriacou and Hall (2, 4–12), whereas all of the other genotypes sing so poorly. Even if we assume Stern missed one-third to one-half of the IPIs of per^{01} and per^s, these males would not sing nearly enough to be included in any song rhythm analysis using Kyriacou and Hall criteria (Fig. 1*C*). The use of courtships with very sparse song production adds a major noise component to Stern's results. The lack of temperature control in Stern's study is startling, given the well-known dependence of IPI on temperature (25). Equally puzzling is the adoption of *D. melanogaster* upper IPI limits when analyzing *D. simulans* songs. However, Stern's apparent reluctance to manually monitor extended samples of his songs in his latter study is astonishing, as he would have immediately observed a concordance rate of only ~50%. On the basis of his critically flawed analysis, Stern has sought to discredit a decade's worth of results on song cycles from Kyriacou, Hall, and collaborators (2, 4–12), as well as confirmatory work from other independent groups (18, 19). From our exhaustive manual reanalyses of nearly 6 h of Stern's courtship song records, we seldom observed consistently reliable automated scoring for any 10-s segment of song from his study (22) (Fig. 4).

We have documented that when the most vigorous of Stern's song records are analyzed manually and IPIs placed into 10-s bins, per^L songs have significantly longer periods than Canton-S using three different time-series analyses. The distribution of periods for each genotype is also significantly different from the corresponding randomized data. For unbinned data, we concur with Stern that the highest peaks in the L-S periodogram are nearly always in the very short period range of 0.8–2 s, and these show a very marginal difference between the genotypes. However, the period distributions of these high-frequency cycles are not random. This may be caused by the general increase in IPI reported to occur during a burst (35), a pattern that would generate high frequencies in the L-S periodogram and be further modulated by interburst intervals. When we examined the periods of the unbinned raw IPIs within the 20- to 150-s domain, we again see significantly longer periods in per^L compared with wild type, which is also supported by comparison of the corresponding period distributions of the two genotypes. Given that each male sings such a broad range of IPI lengths ranging from 15 to 65 ms, it seems unlikely that an individual species-specific IPI value acts as a trigger to release female mating, as originally envisioned (1). It seems more likely that females average IPIs over short periods of time. If each female has a slightly different preferred mean IPI, then the cycling of IPI could act to allow the male to scan the IPI preferences of different potential partners as they court them successively (36). Consequently, mean IPIs taken over short periods of time could represent more biologically relevant stimuli than individual IPIs.

In terms of mechanism, the effects of *per* on ~60-s song rhythms cannot be explained by the canonical circadian transcription–translation loop. However, ultradian (~30 min) hormone rhythms are also known to be altered by circadian clock mutations in hamsters (37), as is the timing of critical period plasticity in cortical neurons during mouse postnatal maturation (38). The fly song cycle may be a pleiotropic effect of *per*, which acts on thoracic ganglia during development, possibly through *per*-expressing glia (28).

Our reinspection of Stern's data and his methods suggests that considerable caution needs to be exercised in accepting conclusions from song studies using the automated pulse detection within FlySongSegmenter (23). Although it might be acceptable to calculate an average IPI value for a complete courtship based on, for example, 2,000 rather than 3,000 IPIs (where it is unlikely to make much overall difference), omitting significant proportions of the data and adding spurious IPIs can generate major problems on 10-s or finer time scales. This problem particularly applies to studies where much more sensitive and dynamic song features are extracted using FlySongSegmenter. For example, changes reported by Stern to occur over the length of a song burst may need to be revised, given that we have documented how discrete and complete song bursts cannot be reliably defined because of undetected pulses (39). Furthermore, quantifying pulse-song intensity is also unlikely to be reliable, because the automated software ignores so much low-intensity pulse song. Indeed any measure that includes pulse song as a variable needs to be reexamined manually (40–42).

In summary, we have demonstrated that Stern's automated pulse detection method is critically flawed (22). Manual reanalysis of his primary songs records reveal IPI periods of ~80 s in per^L and 50–60 s in wild type. We acknowledge that despite our attempts to be as accurate as possible, some human error will inevitably creep into such a comprehensive manual analysis. However, these errors pale into insignificance when compared with the systematic omissions and spurious IPIs generated by Stern's automation. We appreciate that automated methods can never be perfect, but Stern appears to have severely overstretched the trade-off between convenience and accuracy.

Methods

Further details can be found in *SI Methods*.

ACKNOWLEDGMENTS. We thank Dr. D. Heslop for kindly providing us with a working version of CLEAN.

1. Ewing AW, Bennett-Clark HC (1968) The courtship songs of *Drosophila*. *Behaviour* 31:288–301.
2. Kyriacou CP, Hall JC (1986) Interspecific genetic control of courtship song production and reception in *Drosophila*. *Science* 232(4749):494–497.
3. Kawanishi M, Watanabe TK (1980) Genetic variations of courtship song of *Drosophila melanogaster* and *Drosophila simulans*. *Jpn J Genet* 55:235–240.
4. Kyriacou CP, Hall JC (1980) Circadian rhythm mutations in *Drosophila melanogaster* affect short-term fluctuations in the male's courtship song. *Proc Natl Acad Sci USA* 77(11):6729–6733.
5. Kyriacou CP, Hall JC (1985) Action potential mutations stop a biological clock in *Drosophila*. *Nature* 314(6007):171–173.
6. Zehring WA, et al. (1984) P-element transformation with *period* locus DNA restores rhythmicity to mutant, arrhythmic *Drosophila melanogaster*. *Cell* 39(2 Pt 1):369–376.
7. Hamblen M, et al. (1986) Germ-line transformation involving DNA from the *period* locus in *Drosophila melanogaster*: Overlapping genomic fragments that restore circadian and ultradian rhythmicity to *per0* and *per-* mutants. *J Neurogenet* 3(5):249–291.
8. Yu Q, Colot HV, Kyriacou CP, Hall JC, Rosbash M (1987) Behaviour modification by *in vitro* mutagenesis of a variable region within the period gene of *Drosophila*. *Nature* 326(6115):765–769.
9. Kyriacou CP, Hall JC (1989) Spectral analysis of *Drosophila* courtship song rhythms. *Anim Behav* 37:850–859.
10. Kyriacou CP, van den Berg MJ, Hall JC (1990) *Drosophila* courtship song cycles in normal and *period* mutant males revisited. *Behav Genet* 20(5):617–644.
11. Dushay MS, et al. (1990) Phenotypic and genetic analysis of *Clock*, a new circadian rhythm mutant in *Drosophila melanogaster*. *Genetics* 125(3):557–578.
12. Wheeler DA, et al. (1991) Molecular transfer of a species-specific behavior from *Drosophila simulans* to *Drosophila melanogaster*. *Science* 251(4997):1082–1085.
13. Konopka RJ, Benzer S (1971) Clock mutants of *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 68(9):2112–2116.
14. Kyriacou CP, Hall JC (1982) The function of courtship song rhythms in *Drosophila*. *Anim Behav* 30:794–801.
15. Griffith LC, et al. (1993) Inhibition of calcium/calmodulin-dependent protein kinase in *Drosophila* disrupts behavioral plasticity. *Neuron* 10(3):501–509.
16. Greenacre ML, Ritchie MG, Byrne BC, Kyriacou CP (1993) Female song preference and the *period* gene in *Drosophila*. *Behav Genet* 23(1):85–90.
17. Demetriades MC, Thackeray JR, Kyriacou CP (1999) Courtship song rhythms in *Drosophila yakuba*. *Anim Behav* 57(2):379–386.
18. Alt S, Ringo J, Talyn B, Bray W, Dowse H (1998) The period gene controls courtship song cycles in *Drosophila melanogaster*. *Anim Behav* 56(1):87–97.
19. Ritchie MG, Halsey EJ, Gleason JM (1999) *Drosophila* song as a species-specific mating signal and the behavioural importance of Kyriacou & Hall cycles in *D. melanogaster* song. *Anim Behav* 58(3):649–657.
20. Ewing AW (1988) Cycles in the courtship song of male *Drosophila melanogaster* have not been detected. *Anim Behav* 26:1091–1097.
21. Crossley SA (1988) Failure to confirm rhythms in *Drosophila* courtship song. *Anim Behav* 36:1098–1109.
22. Stern DL (2014) Reported *Drosophila* courtship song rhythms are artifacts of data analysis. *BMC Biol* 12(1):38.
23. Arthur BJ, Sunayama-Morita T, Coen P, Murthy M, Stern DL (2013) Multi-channel acoustic recording and automated analysis of *Drosophila* courtship songs. *BMC Biol* 11(1):11.
24. Hall JC, Kyriacou CP (1990) Genetics of biological rhythms in *Drosophila*. *Adv Insect Physiol* 22:221–298.
25. Shorey HH (1962) Nature of the sound produced by *Drosophila melanogaster* during courtship. *Science* 137(3531):677–678.
26. Gromko MH, Markow TA (1993) Courtship and remating in field populations of *Drosophila*. *Anim Behav* 45:253–262.
27. Yamazaki S, Takahashi JS (2005) Real-time luminescence reporting of circadian gene expression in mammals. *Methods Enzymol* 393:288–301.
28. Konopka RJ, Kyriacou CP, Hall JC (1996) Mosaic analysis in the *Drosophila* CNS of circadian and courtship-song rhythms affected by a *period* clock mutation. *J Neurogenet* 11(1-2):117–139.
29. Refinetti R, Lissen GC, Halberg F (2007) Procedures for numerical analysis of circadian rhythms. *Biol Rhythm Res* 38(4):275–325.
30. Roberts DH, Lehar J, Dreher JW (1987) Time series with CLEAN. I. Derivation of a spectrum. *Astron J* 93:968–989.
31. Chatfield C (1996) *The Analysis of Time Series* (Chapman and Hall, London), 5th Ed.
32. Dowse HB (2009) Analyses for physiological and behavioral rhythmicity. *Methods Enzymol* 454:141–174.
33. Logan IG, Rosenberg J (1989) A referee's comments on the identification of cycles in the courtship songs of *Drosophila melanogaster*. *Anim Behav* 37:860.
34. Heslop D, Dekkers MJ (2002) Spectral analysis of unevenly spaced climatic time series using CLEAN: Signal recovery and derivation of significance levels using a Monte Carlo simulation. *Phys Earth Planet Inter* 130:103–116.
35. Ewing AW (1983) Functional aspects of *Drosophila* courtship. *Biol Bull* 58:275–291.
36. Kyriacou CP, Ritchie M, Greenacre ML (1992) Genetics of sexual selection in *Drosophila*. *Am Zool* 32:31–39.
37. Loudon AS, et al. (1994) Ultradian endocrine rhythms are altered by a circadian mutation in the Syrian hamster. *Endocrinology* 135(2):712–718.
38. Kobayashi Y, Ye Z, Hensch TK (2015) Clock genes control cortical critical period timing. *Neuron* 86(1):264–275.
39. Coen P, et al. (2014) Dynamic sensory cues shape song structure in *Drosophila*. *Nature* 507(7491):233–237.
40. Clemens J, et al. (2015) Connecting neural codes with behavior in the auditory system of *Drosophila*. *Neuron* 87(6):1332–1343.
41. Coen P, Xie M, Clemens J, Murthy M (2016) Sensorimotor transformations underlying variability in song intensity during *Drosophila* courtship. *Neuron* 89(3):629–644.
42. Shirangi TR, Wong AM, Truman JW, Stern DL (2016) Doublesex regulates the connectivity of a neural circuit controlling *Drosophila* male courtship song. *Dev Cell* 37(6):533–544.
43. Dowse HB, Ringo JM (1987) Further evidence that the circadian clock in *Drosophila* is a population of coupled ultradian oscillators. *J Biol Rhythms* 2(1):65–76.
44. Kay SM, Marple SGJ (1981) Spectrum analysis, a modern perspective. *IEEE Proc* 69:1380–1419.
45. Bretthorst GL (2001) Nonuniform sampling: Bandwidth and aliasing. *Baysian Interferance and Maximum Entropy Methods in Science and Engineering*, eds Rychert J, Erickson G, Smith CR (Kluwer, Dordrecht, The Netherlands), pp 1–28.

Supporting Information

Kyriacou et al. 10.1073/pnas.1615198114

SI Methods

Our initial analysis used Stern's processed IPI data that is deposited on his website as .csv files. We calculated the number of IPIs in each 10-s segment of real time and selected the most vigorous 400-s clip of courtship that we could find with the constraint that this did not initiate past the first 5 min of the courtship. For most song recordings we could not begin the analysis at the same point as Stern because of large numbers of gaps at the beginning of the song record. If the 400-s clip still contained many gaps we extended it for 50–80 s so that we included at least 35 mean IPI points. We were able to generate vigorous song clips for 25 Canton-S and 14 per^L songs in this way.

After the time series analyses were performed on Stern's IPI calls, we then used each of the corresponding song clips for manual analysis by obtaining the primary Matlab song files from Stern's website. If within that song clip, Stern had initiated his IPI analysis in the middle of a song burst (because he had missed the preceding IPIs), we initiated at the beginning of the song burst. We took the mean IPI for each song and applied an upper cut-off that was twice the mean (to the nearest 5 ms). We randomized Stern's binned and raw data using the MATLAB routine RANDPERM before applying the L-S method. For the unbinned data we maintained the interburst intervals in Stern's data. We also used the more conservative cosinor method to analyze the IPIs when we initiated 5 s into the first 10-s bin, so we were 180° out of phase with our initial manual analysis. We included an extra 5 s of manually derived IPIs at the end of the song clip to complete the final bin in the series.

Manual analysis relied on using the FlySongSegmenter pulse-song plots as well as the pulse frequency plots. Song pulses have intrapulse frequencies of 150–1,000 Hz and the frequency plots, as well as the facility to listen to every pulse made scoring of IPIs largely unambiguous. A second naïve observer was also asked to score three of the songs (see main text) and interobserver reliability was initially 92.4% but improved with practice to 99% and was far superior to the ~50% concordance between Stern's automated method with manual analysis. The interobserver reliability raw data files are deposited in Dryad.

Time Series Analysis

We used the CLEAN (30) and L-S spectral methods (see Stern, 22), as well as cosinor analysis (29). Cosinor is widely-used in circadian research and is a similar curve-fitting procedure as used by Kyriacou and Hall in their studies (4–12). It is particularly useful for short time series. We also analyzed the binned data by autocorrelation, a well-known method for ascertaining whether there is periodicity in data which requires data sampled at regular intervals and is commonly used in circadian research (31). To do this required interpolation to create uniformly sampled datasets with no missing values and was performed for the data assembled in 10-s bins by the Maximum Entropy method using a variation on an algorithm developed previously to allow computation of signal-to-noise ratios in circadian rhythms (43). Zeroes (gaps, <10 IPIs) were first replaced by linear interpolation to make further processing possible. Next, a five-coefficient autoregressive model was fitted to the data using MESA (43). The original dataset, containing the zero bins, was then processed using the model to replace any zeroes encountered by this rule: $X(t) = a1X(t-1) + a2X(t-2) + \ldots + a5X(t-5)$, where $X(t)$ is the value to be predicted and the series are the coefficients from the model. This works even if there are several consecutive missing values, as the estimated values are used in the subsequent iterations and the method is identical to the way in which autoregressive modeling allows extension of the autocorrelation function in MESA (32, 44).

In the binned data, the effective sampling rate is once per 10 s. The minimum frequency that can be calculated is twice this value, 1/20 s, or the Nyquist frequency (31), so we accepted periods that were >20 s to 150 s. The situation for the Nyquist changes for raw data that are sampled at unequal intervals. Here, Nyquist is determined by the "effective dwell time," which is the interval one would use to obtain the existing time series if the data were to be sampled at regular intervals, here less than or equal to the shortest recorded IPI, ΔT. The Nyquist period then is simply $2\Delta T'$ (45). However, as the minimum IPI was 15 ms, the peak periods were never in that high-frequency range with almost all lying between 0.8 and 5 s.

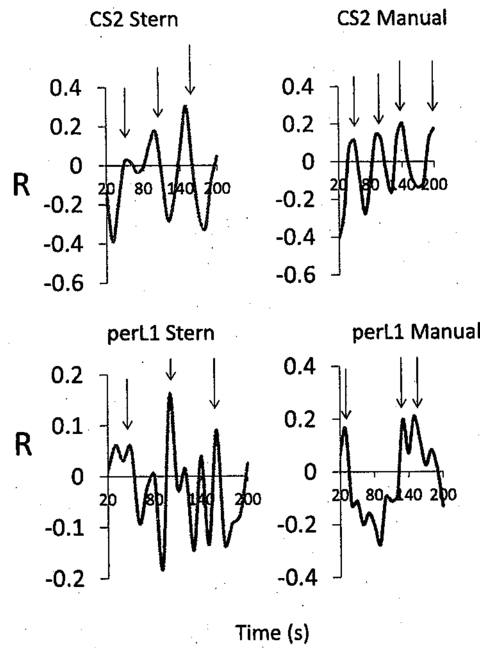


Fig. S1. Correlograms for songs CS2 and perL1. The mean IPIs per 10 s generated by Stern's automated pulse detection for songs CS2 and perL2 (see Fig. 2) generate the correlograms in blue, and the same songs analyzed manually, in red. Arrows show the repeated pattern, which is 50 s for CS2 (both Stern and manual analyses) and 50 s for Stern's perL1 analysis, but 130–150 s for the same song when the IPIs are computed manually. Note how the manual analyses generate correlograms with more distinct repeating patterns.

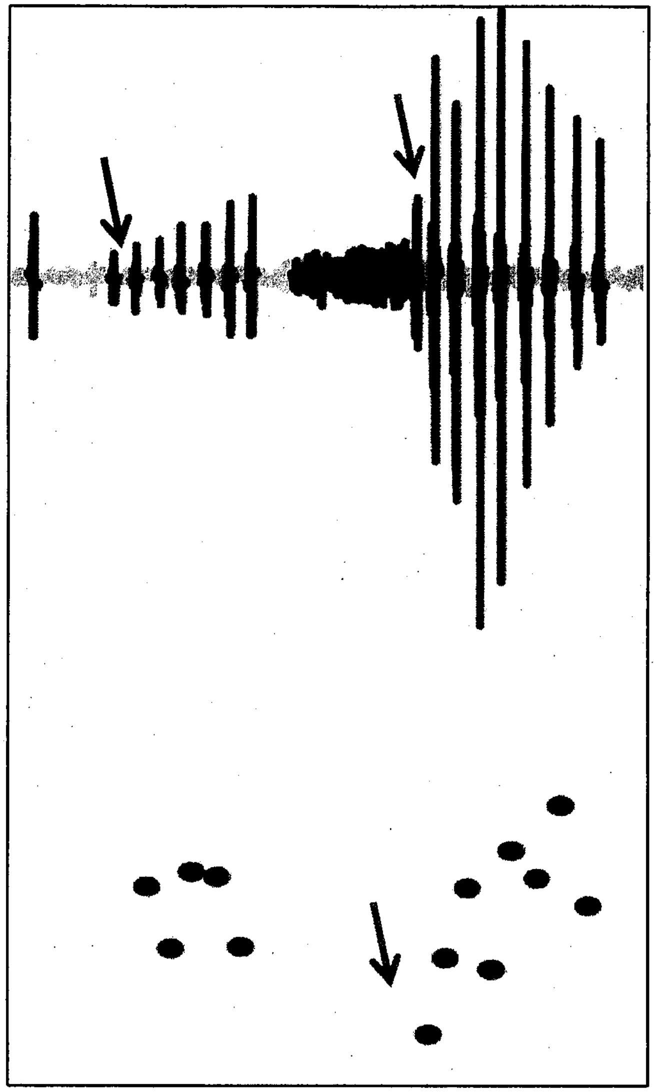


Fig. S2. Stern's automated method of pulse detection appears to generate errors [reproduced with permission from figure 1B of Stern (22)]. In the first pulse train, eight pulses (possibly nine) can be seen and seven IPIs, but the computer appears to detect seven pulses but only five IPIs are logged (black dots below the song trace). In the following song burst, the first two pulses and corresponding IPI are detected and generate a very short IPI, which is logged (green arrows). This IPI is shorter than the IPI that should have been logged in the previous song burst (marked with blue arrow) so cannot be less than the lower threshold value of 15 ms.

Table S1. The effects of different upper IPI cut-offs

IPI	per^{01}	per^{L}	per^{S}	*D.simulans*	CantonS
Stern	27,388	120,764	31,772	131,890	360,478
Kyriacou and Hall	29,207	124,361	33,674	157,925	360,478
Missing (%)	6.64	2.98	5.99	19.74	0
IPI cut-off (ms)	85	75	85	95	65

The number of IPIs detected by Stern from his processed .csv files for all genotypes with 65-ms upper cut-off are shown. The number of IPIs that are detected using the more appropriate cutoff of 85 ms for per^{01} and per^{s}, 75 ms for per^{L}, and 95 ms for *Drosophila simulans* were calculated from his raw song records retrieved from his Matlab files. The percentages of IPIs undetected by Stern using these longer cut-offs are illustrated.

Table S2. **Automated versus manual IPI detection in study by Arthur et al. (23)**

Song	Starts	Ends	Duration (s)	Stern (*n*)	Manual (*n*)	S/M	Spurious IPI (*n*)	Concord	Spurious/ manual
1	118.87	347.7	228.83	82	96	0.85	2	0.83	0.02
3	69.44	174.72	105.28	296	438	0.68	8	0.66	0.03
4	220.92	400.64	179.72	733	757	0.97	18	0.94	0.02
17	64.08	257.39	193.31	351	461	0.76	9	0.74	0.03
18	120.11	279.23	159.12	65	329	0.20	29	0.11	0.45
19	41.41	277.88	236.47	361	426	0.85	8	0.83	0.02
20	28.65	377.19	348.54	111	93	1.19	22	0.96	0.20
21	271.84	501.04	229.2	536	662	0.81	8	0.80	0.01
22	7.49	170.03	162.54	443	520	0.85	8	0.84	0.02
23	113.53	410.22	296.69	508	662	0.77	45	0.70	0.09
24	0.87	260.87	260	1047	1210	0.87	12	0.86	0.01
25	414.3	800.52	386.22	664	714	0.93	23	0.90	0.03
26	191.15	341.12	149.97	59	100	0.59	8	0.51	0.14
28	271.73	797.78	526.05	164	206	0.80	3	0.78	0.02
29	83.99	300	216.01	826	1004	0.82	5	0.82	0.01
30	89.01	360	270.99	271	373	0.73	5	0.71	0.02
		Total	3948.94	6517	8051		213		
		Mean				0.79		0.75	0.07

We downloaded 16 songs from Arthur et al. (23) in which the songs were recorded in small mating cells. We compared Stern's automated IPI detection with manual detection. We show the number of pulses detected by both methods, and the number of spurious pulses detected by Arthur et al. (23) (caused by pulse skipping and noise). When we removed these from Stern's analysis, the concordance rates between the two methods vary from 11 to 96%, with an average of 75%. However, this concordance rate is contaminated by a mean spurious IPI rate of 7%. Consequently, smaller mating cells, which generate better signal-to-noise ratios produce better Stern/manual proportions compared with the songs in Stern (22), but these are offset by a higher spurious pulse detection rate.

Dataset S1. Spreadsheet showing courtship vigor from songs analyzed by Hamblen et al. (7)

Dataset S1

The spreadsheet shows the number of IPIs calculated from each 10-s bin for two control genotypes. The mean number of IPIs per bin is shown at the bottom and the number that would be predicted in a 45-min record. The number of bins with 10 or more IPIs, gaps (<10 IPIs) and proportion of gaps are also shown.

Dataset S2. Spreadsheet of all of the results of the song reanalyses

Dataset S2

The results for each song are illustrated, including the start and end time for each reanalysis (from Stern's .csv files) and the results obtained with cosinor, L-S and CLEAN methods using Stern's automated analysis (labeled "Stern") and the manual analysis (labeled "Manual"). Actual significance levels are shown for the cosinor analysis and asterisks for the L-S analyses (*$P < 0.1$,**$P < 0.05$, ***$P < 0.01$, ****$P < 0.005$, *****$P < 0.001$). The unbinned data (RAW) were analyzed with L-S only for both the 0- to 150-s and 20- 150-s domains. "??" under autocorrelation means that there was no obvious cycle.

Dataset S3. Spreadsheet showing initial automated versus manual pulse detection study for 25 Canton-S songs from Stern (22)

Dataset S3

Canton-S songs analyzed by Stern and manual method for first 24 wild-type songs plus CS71. All were analyzed from the beginning of singing for at least ~60–90 s, except song CS7 which stopped after ~30 s and did not resume. The numbers of IPIs detected by the two methods are shown as are the proportions of IPIs detected by Stern. Also shown is the vigor of the song as IPI/s of courtship time, the numbers of pulse skippings that gave rise to long IPs and any noise treated as an IPI by Stern. These spurious IPIs are shows as a proportion of the number of IPIs detected by Stern and the concordance rate (number of correct IPI calls by Stern as a proportion of the total number of manual calls) is also given.

EXHIBIT C

Salsbury, Daniel

From:	Mark Z. Jacobson <jacobson@stanford.edu>
Sent:	Wednesday, October 11, 2017 4:27 PM
To:	Salsbury, Daniel
Cc:	Mark Anthony Delucchi; PNASNews
Subject:	Clarification to 2015 PNAS article
Attachments:	Correction-PNAS15.pdf

Dear Daniel,

We would like to submit for publication to PNAS the attached clarification to our previously published 2015 PNAS article. This clarification has no impact on any plot or energy data result in our original paper but is necessary to clarify our hydropower assumption because the original text describing this assumption was not clear, resulting in several people requesting clarification of the assumption or drawing their own incorrect interpretation of the assumption. Thank you for considering this request.

Sincerely,
Mark Jacobson

P.S., I tried to submit this online, but there was no submission option for errata.

--

Mark Z. Jacobson
Professor of Civil and Environmental Engineering
Director, Atmosphere/Energy Program
Stanford University
Yang & Yamazaki Environ. and Energy Bldg
473 Via Ortega, Room 397
Stanford, CA 94305-4020
Phone: 650-723-6836
Fax: 650-723-7058
jacobson@stanford.edu
Twitter: @mzjacobson
www.stanford.edu/group/efmh/jacobson/

Errata

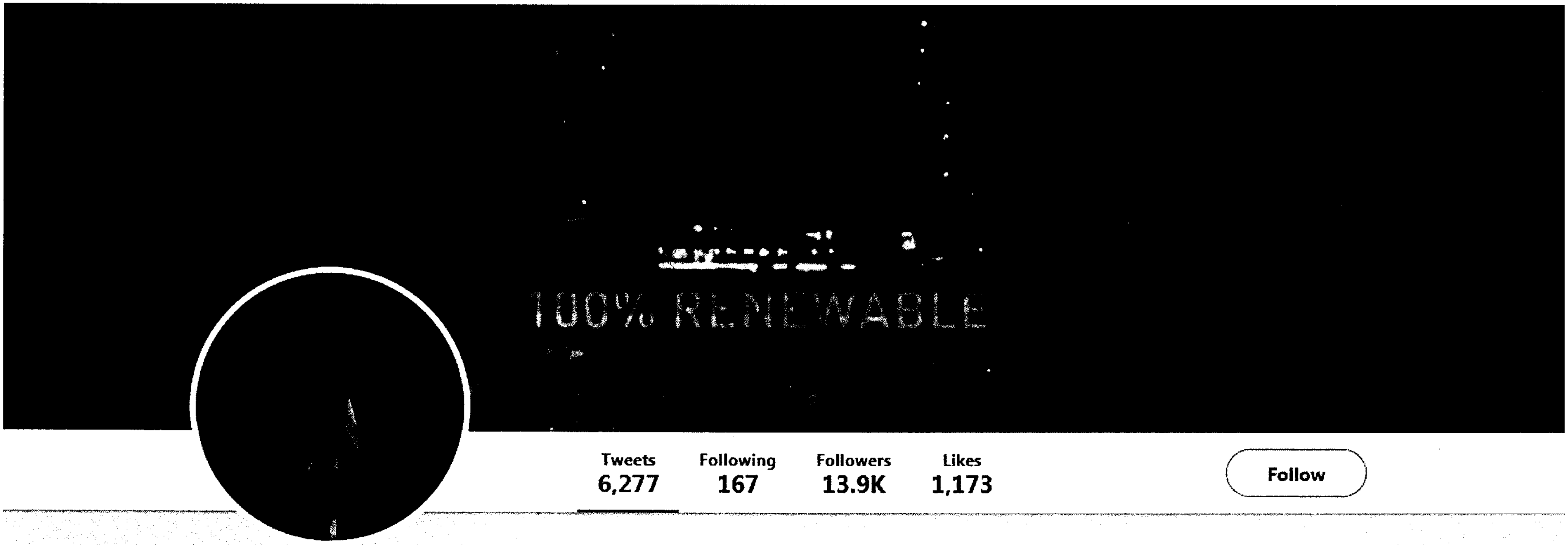
Clarification to "A low-cost solution to the grid reliability problem with 100% penetration of intermittent wind, water, and solar for all purposes, " by Mark Z. Jacobson, Mark A. Delucchi, Mary A. Cameron, and Bethany A. Frew, first published December 8, 2015; 10.1073/pnas.1510028112 (Proc Natl Acad Sci USA 112:15060-15065).

The authors clarify Footnote 4 of Table S2 (Supplementary Information) to state, "As stated in Section 5.4 of [1] but reiterated here, 9.036 GW of the 87.48 GW of previously-installed hydropower in this table are Canadian installations providing pre-existing imported hydropower. (The difference between the 87.48 GW here and the 87.86 GW in [1] is that the former is for the 48 contiguous United States and the latter is for all 50 states). The 87.48 GW in this table is not only the contemporary installed hydropower capacity, it is also the maximum *potential* annually averaged discharge rate of hydropower both today and in 2050 in this study. Thus, this maximum *potential* annually averaged rate is held constant over time here. The *actual* annually averaged discharge rate of hydropower in this study for 2050 is 45.92 GW (Table 2), which is much less than the 87.48 GW maximum *potential* annually averaged value. However, as indicated in Figures 2b, S4b, and S5b, it is assumed here that 1,282.5 GW of turbines are added to existing hydropower dams to increase the maximum *instantaneous* discharge rate of hydropower to a total 1,370 GW without changing the reservoir size or maximum *potential* annually averaged discharge rate of hydropower of 87.48 GW. Thus, while the peak discharge rate may increase significantly for some hours, it decreases significantly for others to ensure the *actual* annually averaged discharge rate of hydropower is not much different from today and much less than maximum annual value, 87.48 GW. This can be accomplished by modifying powerhouses to increase either the number or capacity of turbines and the instantaneous flow rate of water to them, by either adding pipes around or above dams or widening penstocks through dams. The cost of electrical equipment (turbines, generators, and transformers) in a hydropower plant ranges from ~$560/kW for 500 MW plants to ~$200-$300/kW for 1000 MW plants (Figs. 4.5 and 4.7 of [2]). We start with the cost for a large 1000-MW plant and add costs for pipes or widening penstocks and for equipment housing and contingencies due to possible supply shortages to arrive at an estimated total cost of the additional hydropower turbines of roughly $385 (325-450) per kW. This amounts to ~$494 billion for all of the additional turbines proposed here, which would increase the total all-sector capital cost in Table 2 by a mean of just over 3%. We believe this cost increase has no impact on the main conclusions of this study. Even if costs were much higher, there are multiple other low-cost solutions with zero added hydropower turbines but more CSP and batteries instead, not only for North America, but also for 20 world regions, so the increase in hydropower peak instantaneous discharge is just one of several options."

1. M.Z. Jacobson, M.A. Delucchi, G. Bazouin, Z.A.F. Bauer, C.C. Heavey, E. Fisher, S.B. Morris, D.J.Y. Piekutowski, T.A. Vencill, T.W. Yeskoo, 100% clean and renewable wind, water, sunlight (WWS) all-sector energy roadmaps for the 50 United States, Energy and Environmental Sciences, 8 (2015) 2093-2117.
2. IRENA (International Renewable Energy Agency), Renewable Energy Technologies: Cost analysis series. Hydropower, Vol. 1(3), IRENA, Abu Dhabi, 2012.

EXHIBIT D

Home Moments Search Twitter Have an account? **Log in**



Tweets	Following	Followers	Likes
6,277	167	13.9K	1,173

Follow

Mark Z. Jacobson

@mzjacobson

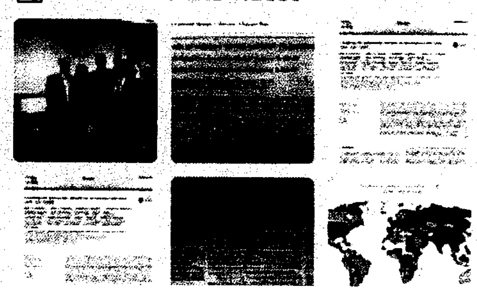
Professor of Civil and Environmental Engineering; Director, Atmosphere/Energy Program; Research on climate, air pollution, clean and renewable energy systems

Stanford University

stanford.edu/group/efmh/jac...

Joined February 2010

166 Photos and videos



Tweets **Tweets & replies** **Media**

Pinned Tweet

Mark Z. Jacobson @mzjacobson · Aug 23

Time to act: @NBCnews: 139 countries can go #100% clean, reenwable nbcnews.to/2xt3uQM @MichaelEMann @NaomiOreskes @NaomiAKlein @BillNye

139 countries could move to 100% clean energy under researchers' p...

A research group based at Stanford University has drawn a plan for nations to reduce global warming by relying on solar and wind power.

Mark Z. Jacobson @mzjacobson · 13 Jul 2016
Misinfo from zealots @JesseJenkins & @theBTI who deny obvious&never published on 100% WWS hardly useful @bradplumer

> **brad plumer** @bradplumer
> Very good tweetstorm on why 100% renewables might not be the best path to zero carbon emissions. twitter.com/JesseJenkins/s...

1 · 2 ·

Mark Z. Jacobson @mzjacobson · 13 Jul 2016
They rely on decarb studies that ignore most storage, load reduc w/electrification, true nuc costs, etc. web.stanford.edu/group/efmh/jac... @bradplumer

1 · 1 · 1

brad plumer @bradplumer · 13 Jul 2016
jesse really doesn't strike me as a zealot in the slightest, but i will read this—thanks

7 · 1 · 5

JesseJenkins
@JesseJenkins

Follow

Replying to @bradplumer @mzjacobson

idea that anyone who disagrees w/you is a zealot is really poor form for a scholar. We can just disagree w/out that.

12:12 PM - 13 Jul 2016

2 Retweets **6** Likes

1 · 2 · 6

Janne M. Korhonen @jmkorhonen · 13 Jul 2016
Replying to @JesseJenkins @bradplumer
agree. Never seen behavior like this in academia. @mzjacobson studies are outliers & he shd tolerate critique.

2 · 4

3 · 5 · 5

Mark Z. Jacobson
@mzjacobson

Follow

Replying to @Atomicrod

@Atomicrod fakes data by showing video of Wigley falsely stating our plans add hydro web.stanford.edu/group/efmh/jac ...

8:02 AM - 14 Dec 2015

3

Rod Adams @Atomicrod · 14 Dec 2015
Replying to @mzjacobson
@mzjacobson Other responses, including one from your Stanford colleague. "Unrealistic assumptions give unrealistic results." Any response?

1

Mark Z. Jacobson @mzjacobson · 14 Dec 2015
@Atomicrod There is no specific comment to respond to by @kenCaldeira. I'm sure he hadn't read study just as Wigley hadn't. It just came out

1

Rod Adams @Atomicrod · 14 Dec 2015
@mzjacobson Disingenuous to say your study "just came out." You've been making 100% renewable case for years. @KenCaldeira

1

Mark Z. Jacobson @mzjacobson · 14 Dec 2015
@Atomicrod The 139 country paper just came out. 48 contiguous states also 0 new hydro. @KenCaldeira did not read new paper; only smeared

1 · 1

Ken Caldeira @KenCaldeira · 14 Dec 2015
@mzjacobson What smear? Quotation?@Atomicrod

1

Mark Z. Jacobson @mzjacobson · 14 Dec 2015
@KenCaldeira @Atomicrod See video. You make blanket criticism w/ any specifics=smear. If you read paper, should have corrected Wigley too

1

Mark Z. Jacobson
@mzjacobson

Follow

Amazing how @saeverley @EnergyInDepth flat out lie about paper at web.stanford.edu/group/efmh/jac ... Table 9 clearly shows 2 mil net 40-yr jobs created

10:05 AM - 5 Jan 2016

17 Retweets **12** Likes

1 17 12

ConvivialPeter @PeterDBuckland · 12 Jan 2016

Replying to @mzjacobson

A #renewableenergy economy will provide millions of jobs. #sustainability dailykos.com/story/2016/1/1... @mzjacobson @saeverley @EnergyInDepth

We Don't Need No Education: Oil Group Misleads ...

Sometimes, oil and gas spokespeople are sneaky about the source of their funding, hiding behind multiple layers of shell groups in order to appear neutral. Energy in D...

dailykos.com

1 2

Steve Everley @saeverley · 5 Jan 2016
@mzjacobson I think we're done here. Post is based on data tables you provided, and terms in your data/study. I'm not gonna keep saying that

1 2 1

Mark Z. Jacobson
@mzjacobson

Follow

Replying to @saeverley

@saeverley I am going to be very clear publicly that you have intentionally falsified data, Mr. Everly.

5:07 PM - 5 Jan 2016

2

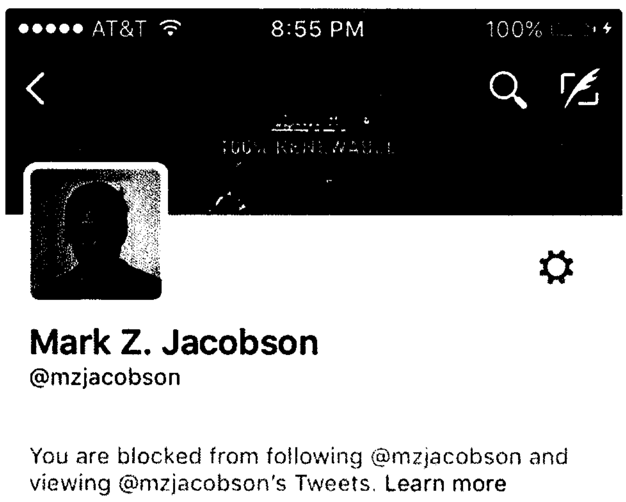
Steve Everley @saeverley · 5 Jan 2016
Replying to @mzjacobson
@mzjacobson Threats now? I haven't intentionally falsified anything. We disagree about what constitutes "permanent."

1 1 1

Steve Everley @saeverley · 5 Jan 2016
@mzjacobson Or, rather, you appear to disagree about what your own data and study show. Either way, it's all sourced and clear.

2 1

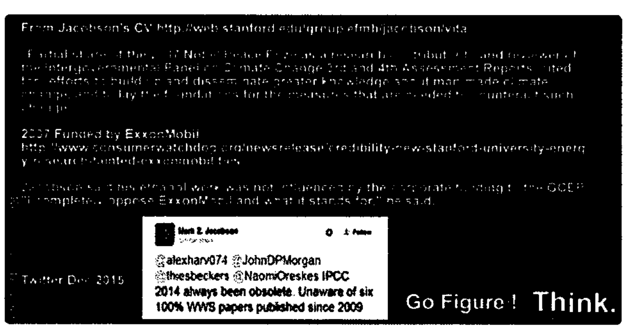
Steve Everley @saeverley · 5 Jan 2016
Stanford professor pushing 100% renewables calls me a liar for exposing unsavory detail in his work, then blocks me



2 5 1

1 more reply

Thor 4 @MSR_Future · 6 Jan 2016
Replying to @mzjacobson
@mzjacobson @saeverley Will be in very good company.



@PhelimMcAleer Data faker @saeverley lies that numbers he used were part of study mediamatters.org/blog/2016/01/1...

We Don't Need No Education: Oil Group Misleads ...
Sometimes, oil and gas spokespeople are sneaky about the source of their funding, hiding behind multiple layers of shell groups in order to appear neutral. Energy in D...
dailykos.com

1

Phelim McAleer @PhelimMcAleer · 13 Jan 2016
.@mzjacobson @saeverley but why did you delete the data?

1

Mark Z. Jacobson @mzjacobson · 13 Jan 2016
@PhelimMcAleer I informed @saeverley Jan 5 that #s he was using were test #s, not real twitter.com/mzjacobson/sta... yet he still used

> **Mark Z. Jacobson** @mzjacobson
> @saeverley Columns M,N,Q,R = dead test columns not used for anything. 'Tables for 50-state paper' tab B551-E601 give all numbers+references

2

Steve Everley @saeverley · 13 Jan 2016
.@mzjacobson @PhelimMcAleer They are (or, well, were) listed as supporting data for your E&ES 2015 study.

2 1 1

Phelim McAleer @PhelimMcAleer · 13 Jan 2016
.@saeverley @mzjacobson is this correct?

2

Mark Z. Jacobson @mzjacobson · 13 Jan 2016
@PhelimMcAleer All real data used in papers still there.Dead non-real #s removed only because @saeverley abuse by falsely claiming were real

1

Mark Z. Jacobson
@mzjacobson

Follow

@PhelimMcAleer Not only did @saeverley intentionally use fictitious #s, he then cries wolf when the fictitious #s removed due to his abuse

11:17 AM - 13 Jan 2016

2

Thor 4 @MSR_Future · 13 Jan 2016
Replying to @mzjacobson
.@mzjacobson @PhelimMcAleer @saeverley

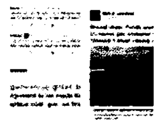
Jacobson admits using fictitious numbers in his research. & cries like a child when highlighted.

1

Dave Quast @davequast · 13 Jan 2016
Replying to @mzjacobson
@mzjacobson @PhelimMcAleer @saeverley intentionally published "fictitious #s" - ok.

1

Mark Z. Jacobson @mzjacobson · 21 Jun 2016
Don't mislead.Refusing expensive 2025 relicense that costs more than WWS to replace it not same as shutting down now



Nicholas Thompson @thompn4
Looks like @mzjacobson now supports closing low carbon sources. twitter.com/mzjacobson/sta... twitter.com/mzjacobson/sta...

3 1 2

Nicholas Thompson @thompn4 · 21 Jun 2016
Not misleading. All that money spent on new WWS doesn't help at all to reduce emissions. brings us back to baseline. @mzjacobson

1 4

Nicholas Thompson @thompn4 · 21 Jun 2016
We shouldn't cheer for the closure of low carbon sources. We should keep all low carbon running, work to replace coal/gas. @mzjacobson

1 2 7

Mark Z. Jacobson
@mzjacobson

Follow

Replying to @thompn4

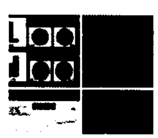
You don't know the first thing about solving the climate, air pollution, and energy security problem. Stop pretending you do.

1:49 PM - 21 Jun 2016

2 Likes

3 2

Thor 4 @MSR_Future · 21 Jun 2016
Replying to @mzjacobson
.@mzjacobson You deny history for your own self-gratification. WWS is a joke. twitter.com/MSR_Future/sta...
#Climatechange



Thor 4 @MSR_Future
French #Energiewende should be model for #EnergyUnion not the German. Results are stuning & self evident

Mark Z. Jacobson
@mzjacobson

Follow

Misinfo from zealots @JesseJenkins & @theBTI who deny obvious&never published on 100% WWS hardly useful @bradplumer

> **brad plumer** @bradplumer
> Very good tweetstorm on why 100% renewables might not be the best path to zero carbon emissions. twitter.com/JesseJenkins/s...

10:50 AM - 13 Jul 2016

2 Retweets

1 2

Mark Z. Jacobson @mzjacobson · 13 Jul 2016
Replying to @mzjacobson
They rely on decarb studies that ignore most storage, load reduc w/electrification, true nuc costs, etc. web.stanford.edu/group/efmh/jac... @bradplumer

1 1 1

brad plumer @bradplumer · 13 Jul 2016
jesse really doesn't strike me as a zealot in the slightest, but i will read this—thanks

7 1 5

Mark Z. Jacobson @mzjacobson · 13 Jul 2016
Someone who finds any excuse to support nuclear, despite facts on weapons, meltdown, waste risk; timelag; C-emis; cost is zealot

2 1 2

Nathan Macher @NathanNuclear · 13 Jul 2016
what are you referring to when you write timelag?

1 1

Mark Z. Jacobson @mzjacobson · 13 Jul 2016
Timelag between planning, siting, permitting, financing, constructing plant 10-19 yrs avg U.S. versus 2-5 yrs for wind/solar

1

Nathan Macher @NathanNuclear · 13 Jul 2016
Nuclear plants can and are built quicker. It does take time and planning but France shows that nuclear scales up quickly.

Mark Z. Jacobson
@mzjacobson

Follow

I don't pay attention to non-experts, especially if they've worked at a nuclear advocacy org, BTI, where they must criticize other solutions

> **Duncan S. Campbell** @duncan__c
> @mzjacobson what are your thoughts on this podcast? Particularly critique of UTES for all end-users in America? twitter.com/TheEnergyGang/...

12:13 PM - 14 Apr 2017

7 Retweets **9** Likes

7 7 9

Duncan S. Campbell @duncan__c · Apr 14
Replying to @mzjacobson
@JesseJenkins and @greentechmedia are non-experts?

2 8

JesseJenkins @JesseJenkins · Apr 14
Yeah if u listen to the podcast, it's pretty apparent that @Stphn_Lacey @shaylekann & I are just making this up off the top of our heads ;-)

2 1 17

Christopher Clack @clacky007 · Apr 14
What constitutes an expert? PhD in atmospheric science for energy modeling? Bit much. Expertise is part education, part experience!

3 6

Rauli Partanen @Kaikenhuippu · Apr 15
An expert to me is a person who is able to read, does read widely on subject(s) and changes his mind if proper evidence is presented.

3

Kevin Leahy @KevinSLeahy · Apr 14
Replying to @mzjacobson
You have the Stanford Energy Modeling Forum folks not far away-might be good to talk with them sometimes.

1 8

Kennedy Maize @kennedymaize · Apr 14
Replying to @mzjacobson

Mark Z. Jacobson
@mzjacobson

Follow

Paper by @clacky007 @kencaldeira is intentionally scientifically fraudulent with falsified data twitter.com/mzjacobson/sta ...

Carnegie Science @carnegiescience
"Full toolbox" is needed to solve the #ClimateChange problem say @KenCaldeira @clacky007 and other experts carnegiescience.edu/news/%E2%80%9C...

1:19 PM - 19 Jun 2017

4 Retweets **5** Likes

1 4 5

Claudia Kemfert @CKemfert · Jun 19
Replying to @mzjacobson @clacky007 @KenCaldeira
100 % renewable energy economically+ technically feasible,nuclear+CCS too costly. Smart energy transition needs flexible demand+load+storage

1 1

Christopher Clack @clacky007 · Jun 19
Read our paper. See if you agree with that assessment afterwards.

1 1 3

Claudia Kemfert @CKemfert · Jun 19
I agree with @mzjacobson and his assessment, also after reading your paper.

2 2 8

Christopher Clack @clacky007 · Jun 19
So no issue with dispatching hydro to 1,300 GW when only 145 GW exists? No transmission modeling, othr big assumptions. Yet, thats all fine?

3 1 6

Mark Z. Jacobson @mzjacobson · Jun 19
Specific example of data falsification by @clacky007 @KenCaldeira on hydro issue in response 3 at

Mark Z. Jacobson
@mzjacobson

Follow

Replying to @clacky007 @elisadquin

More fraud by @clacky007 - Inc hydro discharge rate intentional assumption, not error-More realistic than his covering US w/transmission

4:15 PM - 19 Jun 2017

2

Christopher Clack @clacky007 · Jun 19

Replying to @mzjacobson @elisadquin

No fraud. Just the truth. I know you just assume all that transmission with no modeling, we studied it explicitly as a pathway.

The model assumes a short- and long-distance [redacted] (T&D) system that carries power from distributed and centralized WWS generators to storage and load centers. Costs of and power losses during T&D are accounted for (Table 2, footnote), but power flows through individual lines or substations are not explicitly modeled. The model also accounts for storage costs and

1 1

Mark Z. Jacobson @mzjacobson · Jun 19

Let's see you get approval for all that transmission and to cover 1/3 of Maine with wind turbines. Your modeling ignores reality

1

Christopher Clack @clacky007 · Jun 19

If folks read our paper, they will see the details for placement of wind. If you have issue with it send it through peer review.

1

Mark Z. Jacobson @mzjacobson · Jun 19

Why bother. Nobody pays attention to your paper, which is why you need to attack others and tweet "look at my paper"

1

Elisa @elisadquin · Jun 19

[redacted]

EXHIBIT E

Allegedly False And Misleading Statements
Identified In Exhibit 12 To Plaintiff's Complaint

In addition to the three allegedly "egregious" statements identified in his complaint, plaintiff alleges that Exhibit 12 to his complaint identifies additional "falsehoods and misstatements" that appeared in Dr. Clack's article. Complaint ¶ 84. The left column of the table below sets forth each of the statements that plaintiff's Exhibit 12 identifies as being incorrect or misleading. The right column summarizes the reasons the statement cannot provide the basis for a defamation claim.

It should be noted that some of the alleged falsehoods and misleading statements identified in Plaintiff's Exhibit 12 (and repeated below) are not accurate quotations from Dr. Clack's published article.

Alleged Falsehood Or Misleading Statement In Dr. Clack's Article	**Why Statement Cannot Form The Basis For A Defamation Claim**
"A number of analyses, meta-analyses and assessments, including those performed by the Intergovernmental Panel on Climate Change, the National Oceanic and Atmospheric Administration, the National Renewable Energy Laboratory, and the International Energy Agency, have concluded that deployment of a diverse portfolio of clean energy technologies makes a transition to a low-carbon-emission energy system both more feasible and less costly than other pathways." Complaint, Ex. 12 at 1.	This statement does not say or imply anything about plaintiff, his honesty, his integrity or his character. The statement, and plaintiff's explanation of why he believes the statement is wrong (*see* Complaint, Ex. 12, Rebuttal Points 1-3), merely reflect a difference of opinion on how studies by other scientists should be interpreted.
"In contrast, Jacobson et al. [Jacobson MZ, Delucchi MA, Cameron MA, Frew BA (2015) Proc Natl Acad Sci USA 112(49):15060-15065] argue that it is feasible to provide 'low-cost solutions to the grid reliability problem with 100% penetration of WWS [wind, water [cont'd] and solar power] across all energy sectors in the continental United States between 2050	This statement does not say or imply anything about plaintiff's honesty, integrity or character. Moreover, Plaintiff does not assert that the statement is false. *See* Complaint Ex. 12, Rebuttal Point 4. He merely asserts that there are also other studies that examined "100% or close to 100% clean, renewable energy." *Id.*

and 2055', with only electricity and hydrogen as energy carriers." Complaint, Ex. 12 at 2.	
"In this paper, we evaluate that study and find significant shortcomings in the analysis. In particular, we point out that this work used invalid modeling tools, contained modeling errors, and made implausible, and inadequately supported assumptions." Complaint Ex. 12 at 2.	This statement does not say or imply anything about plaintiff's honesty, integrity or character. It merely reflects an opinion on the adequacy of plaintiff's methods and analysis. Moreover, plaintiff contends that this statement is incorrect or misleading only because the paper's subsequent discussion includes the three allegedly "egregious" statements addressed in the Academy's Memorandum. *See* Complaint Ex. 12, Rebuttal Point 5. As demonstrated in the Academy's Memorandum, there is nothing false or defamatory about any of those allegedly "egregious" statements.
"Policy makers should treat with caution any visions of a rapid, reliable, and low-cost transition to entire energy systems that relies almost exclusively on wind, solar and hydroelectric power." Complaint Ex. 12 at 3.	This statement does not say or imply anything about plaintiff, his honesty, his integrity or his character. It just expresses an opinion on what policy makers should do. Moreover, Plaintiff does not assert that the statement is false. *See* Complaint Ex. 12, Rebuttal Point 6. He merely asserts that his own study did not rely almost exclusively on wind, solar and hydroelectric power. *Id.*
"A number of studies, including a study by one of us, have concluded that an 80% decarbonization of the U.S. electric grid could be achieved at reasonable cost [1, 2]. The high level of decarbonization . . . ." Complaint, Ex. 12 at 3.	This statement does not say or imply anything about plaintiff, his honesty, his integrity or his character. Moreover, Plaintiff does not assert that the statement is false. *See* Complaint Ex. 12, Rebuttal Point 7. He merely asserts that "80% decarbonization of the electricity sector alone" is not a "high level of decarbonization," and that the referenced studies are different "in aim and scope" than a study that eliminates 100% of GHG and air-pollution emissions from all energy sectors. *Id.*
". . . is facilitated by an optimally configured continental high voltage transmission network. There appears to be some consensus that substantial amounts of greenhouse gas emissions could be avoided with widespread deployment of solar and wind electric generation technologies along with supporting [cont'd] infrastructure. Further, it is not in question that it would be	These statements do not say or imply anything about plaintiff, his honesty, his integrity or his character. Moreover, Plaintiff does not assert that the statements are false. *See* Complaint Ex. 12, Rebuttal Point 11. He merely asserts that none of Dr. Clack's references performed "a 'deep carbonization' grid integration study" and that there are studies which "obtain stable grids with 100% or near 100% clean, renewable energy . . . ." *Id.*

theoretically possible to build a reliable energy system while excluding all bioenergy, nuclear energy, and fossil-fuel sources. Given unlimited resources to build variable energy production facilities, while expanding the transmission grid and accompanying energy storage capacity enormously, one would eventually be able to meet any conceivable load. Yet in developing a strategy to effectively mitigate global energy-related CO_2 emissions, it is critical that the scope of the challenge to achieve this in the real world is accurately defined and clearly communicated. Wind and solar are variable energy sources, and some way must be found to address the issue of how to provide energy if their immediate output cannot continuously meet instantaneous demand. The main options are to: (1) curtail load (i.e., modify or fail to satisfy demand) at times when energy is not available, (2) deploy very large amounts of energy storage, or (3) provide supplemental energy sources that can be dispatched when needed. It is not yet clear how much it is possible to curtail loads, especially over long durations, without incurring large economic costs. There are no electric storage system available today that can affordably and dependably store the vast amounts of energy needed over weeks to reliably satisfy demand using expanded wind and solar power generation alone. These facts have led many U.S. and global energy system analyses [1-10] to recognize the importance of a broad portfolio of electricity generation technologies including sources that can be dispatched when needed." Complaint, Ex. 12 at 3-5.	
"Previously analyses have found that the most feasible route to a low-carbon energy future is one that adopts a diverse portfolio of technologies." Complaint, Ex. 12 at 4.	This statement does not say or imply anything about plaintiff, his honesty, his integrity or his character. The statement, and plaintiff's explanation of why he believes the statement is wrong (Complaint, Ex. 12, Rebuttal Points 1 & 8), merely involve a difference of opinion on how studies by other scientists should be interpreted.

"In contrast, Jacobson et al. (2015) consider whether the future primary energy sources for the United States could be narrowed to almost exclusively wind, solar and hydroelectric power . . . ." Complaint, Ex. 12 at 4.	This statement does not say or imply anything about plaintiff's honesty, integrity or character. Moreover, plaintiff's reason for disagreeing with the statement (that his paper "did not rely 'almost exclusively on wind, solar and hydroelectric power'") merely reflects a difference of opinion over nomenclature – *i.e.*, the meaning of the term "almost exclusively." *See* Complaint, Ex. 12, Rebuttal Points 6 & 9.
". . . and suggest that this can be done at 'low-cost' in a way that supplies all power with a probability of loss of load 'that exceeds electric-utility-industry standards for reliability.' We find that their analysis involves errors . . . ." Complaint, Ex. 12 at 4.	These statements do not say or imply anything about plaintiff's honesty, integrity or character. The statements merely reflect an opinion on the adequacy of plaintiff's analysis. Moreover, plaintiff contends that these statement are incorrect or misleading only because the paper's subsequent evaluation includes the three allegedly "egregious" statements addressed in the Academy's Memorandum. *See* Complaint Ex. 12, Rebuttal Points 5 & 10. As demonstrated in the Academy's Memorandum, there is nothing false or defamatory about any of those allegedly "egregious" statements.
"Jacobson et al. [11], along with additional colleagues in a companion article [12], attempt to demonstrate the feasibility of supplying all energy end uses [in the continental United States] with almost exclusively Wind, Water and Solar (WWS) power (no coal, natural gas, bioenergy, or nuclear power), while meeting all loads, at reasonable cost. Reference [11] does include 1.5% generation from geothermal, tidal and wave energy. Throughout the remainder of the paper, we denote the scenarios in ref. [11] as 100% wind, solar and hydroelectric power for simplicity. Such a scenario may be a useful way to explore the hypothesis that it is possible to meet the challenges associated with reliably supplying energy across all sectors almost exclusively with large quantities of a narrow range of variable energy resources. However, there is a difference between presenting such [cont'd] visions as thought experiments and asserting, as the authors do, that rapid and complete conversion to an almost 100% wind, solar and	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, Plaintiff does not assert that anything in these statements is false. *See* Complaint Ex. 12, Rebuttal Point 12. His response to these statements merely reflects his views on the scope and objectives of various studies. *Id.*

hydroelectric power system is feasible with little downside [12]. It is important to understand the distinction between physical possibility and feasibility in the real world. To be clear, the specific aim of Jacobson et al. [11] is to provide "low-cost solutions to the grid reliability problem with 100% penetration of WWS [wind, water and solar power] across all energy sectors in the continental United States between 2050 and 2055". Relying on 100% wind, solar and hydroelectric power could make climate mitigation more difficult and more expensive than it needs to be. For example, the analysis by Jacobson et al. (11, 12] exclude from consideration several commercially available technologies such as nuclear and bioenergy that could potentially contribute to decarbonization of the global energy system, while also helping assure high levels of reliability in the power grid. Further, Jacobson et at. [11, 12] exclude carbon capture and storage technologies for fossil fuel generation. An additional option not considered in the 100% wind, solar and hydroelectric studies is bioenergy coupled with carbon capture and storage (CCS) to create negative emissions within the system, which could help with emissions targets. With all available technologies at our disposal, achieving an 80% reduction in greenhouse gas emissions from the electricity sector at reasonable costs is extremely challenging, even using a new continental scale high voltage transmission grid. Decarbonizing the last 20% of the electricity sector, as well as decarbonizing the rest of the economy that is difficult to electrify (e.g., cement manufacture, aviation), is even more challenging. These challenges are deepened by placing constraints on technological options." Complaint, Ex. 12 at 5.	
"In our view, to demonstrate that a proposed energy system is technically and economically feasible, a study must, at a minimum, demonstrate through transparent inputs,	This statement does not say or imply anything about plaintiff, his honesty, his integrity or his character. Plaintiff's assertion that the statement "baselessly implies that [his]

outputs, analysis . . . ." Complaint, Ex. 12 at 5.	analysis was not transparent" (Complaint Ex. 12, Rebuttal Point 13) would, at most, reflect a difference of opinion over the sufficiency of a scientific analysis.
". . . and validated modeling [13] that the required technologies have been commercially demonstrated at scale at a cost comparable to alternatives; that the technologies can, at scale, provide adequate and reliable energy; that the deployment rate required of such technologies and their associated infrastructure is plausible and commensurate with other historic examples in the energy sector; and the deployment and operation of the technologies do not violate environmental regulations. We demonstrate that ref. [11] and [12] do not meet these criteria and, accordingly, do not show the technical, practical or economic feasibility of a 100% wind, solar and hydroelectric energy vision. As we detail below and in the Supporting Information, ref. [11] contains modeling errors, ..." Complaint, Ex. 12 at 6.	This statement does not say or imply anything about plaintiff's honesty, integrity or character. The statement, and plaintiff's comments on it (Complaint Ex. 12, Rebuttal Point 14), merely reflect a difference of scientific opinion concerning the adequacy of plaintiff's analysis.
". . . incorrect, implausible and/or inadequately supported assumptions, and the application of methods inappropriate to the task. In short, the analysis performed in ref. [11] does not support the claim that such a system would perform at reasonable cost and provide reliable power. The vision proposed by the studies in ref. [11, 12] narrows generation options . . . ." Complaint, Ex. 12 at 6.	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, Plaintiff does not assert that anything in these statements is false. *See* Complaint Ex. 12, Rebuttal Points 12 &15. His response to these statements merely reflects his views on the scope and objectives of various studies. *Id.*
". . . yet includes a wide range of currently uncosted innovations that would have to be deployed at large scale (e.g., replacement of our current aviation system with yet-to-be-developed hydrogen-powered planes). The system in ref. [11] assumes the availability of multi-week energy storage systems that are not yet demonstrated at scale and deploys them at a capacity twice that of the entire U.S. generating and storage capacity today." Complaint, Ex. 12 at 6.	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, plaintiff's reasons for disagreeing with the statements merely reflect a difference of opinion over whether his proposed storage capacity should be characterized as "twice that of the entire U.S. generating and storage capacity today" or "2/5 of the energy modeled." *See* Complaint Ex. 12, Rebuttal Point 16.
"There would be underground thermal energy storage systems deployed in nearly every [cont'd]	This statement does not say or imply anything about plaintiff, his honesty, his integrity or his character. Moreover, plaintiff's disagreement

community to provide services for every home, business, office building, hospital, school, and factory in the United States." Complaint, Ex. 12 at 6.	with the statement merely reflects a difference of opinion over the feasibility of using underground thermal storage systems on a large scale. *See* Complaint Ex. 12, Rebuttal Point 17.
"Yet the analysis does not include an accounting of the costs of the physical infrastructure (pipes, distribution lines) to support these systems. An analysis of district heating [14] showed that having existing infrastructure is key to effective deployment because of the high upfront costs of with [sic] the infrastructure is prohibitive. It is not difficult to match instantaneous energy demands for all purposes with variable electricity generation sources in real time as needed to assure reliable power supply if one assumes, as the authors of the ref. [11] do, that there exists a nationally integrated grid, that most loads can be flexibly shifted in time, that large amounts of multi-week and seasonable energy storage will be readily available at low cost, and that the entire economy can easily be electrified or made to use hydrogen. But adequate support for the validity of these assumptions is lacking. Furthermore, the conclusions in ref. [11] rely heavily on free, non-modeled hydroelectric capacity expansion (adding turbines that are unlikely to be feasible without major reconstruction of existing facilities) at current reservoirs, without consideration of hydrological constraints or the need to for [sic] additional supporting infrastructure (penstocks, tunnels, space);" Complaint, Ex. 12 at 7.	These statements do not say or imply anything about plaintiff, his honesty, his integrity or his character. Moreover, plaintiff does not assert that the statements are false. *See* Complaint Ex. 12, Rebuttal Point 18. He merely asserts that the addition of turbines accounts for "only a small portion of grid balancing" and provides cost data not included in his published paper. *Id.*
". . . massive scale-up of hydrogen production and use . . . ." Complaint, Ex. 12 at 7.	This statement does not say or imply anything about plaintiff, his honesty, his integrity or his character. Moreover, plaintiff does not assert that the statement is false. *See* Complaint Ex. 12, Rebuttal Point 19. He merely asserts that his proposed scale-up of hydrogen is less than scale-ups of transmission and nuclear power in other studies. *Id.*
". . . unconstrained, non-modeled transmission expansion with only rough cost estimates . . ."	This statement does not say or imply anything about plaintiff, his honesty, his integrity or his

Complaint, Ex. 12 at 7.	character. Moreover, plaintiff's disagreement with the statement merely reflects a difference of opinion over nomenclature – plaintiff contends that his cost estimates should be characterized as "uncertain," rather than as "rough." *See* Complaint Ex. 12, Rebuttal Point 20.
". . . and free time-shifting of loads at large-scale in response to variable energy provision. None of these are going to be achieved without cost. Some assumed expansions, such as the hydroelectric power output, imply operating facilities way beyond existing constraints that have been established for important environmental reasons. Without these elements, the costs of the energy system in ref. [11] would be substantially higher than claimed." Complaint, Ex. 12 at 7.	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, plaintiff's disagreement with the statements merely reflects a difference of opinion over the cost of implementing plaintiff's proposal. *See* Complaint Ex. 12, Rebuttal Point 21.
"In evaluating the 100% wind, solar and hydroelectric power system [11], we focus on four major issues that are explored in more detail below and in the Supporting Information. (i) We note several modeling errors presented in ref. [11] that invalidate the results in the studies . . . ." Complaint, Ex. 12 at 8.	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, plaintiff's disagreement with the statements merely reflects a difference of opinion on whether his study involved any modeling errors. *See* Complaint Ex. 12, Rebuttal Point 22.
". . . particularly with respect to the amount of hydropower available and the demand response of flexible loads (SI Appendix section S1)." Complaint, Ex. 12 at 9.	This statement does not say or imply anything about plaintiff's honesty, integrity or character. Moreover, plaintiff's disagreement with this statement is based on the allegedly "egregious" statements addressed in the Academy's Memorandum. *See* Complaint Ex. 12, Rebuttal Points 23, 25 & 26. As demonstrated in the Academy's Memorandum, there is nothing defamatory about any of those allegedly "egregious" statements.
"(ii) We examine poorly documented and implausible assumptions, including: the cost and scalability of storage technologies; the use of hydrogen fuels; lifecycle assessments of technologies; cost of capital and capacity factors of existing technologies; and land use (SI Appendix section S2). (iii) We discuss the studies' lack of electric power system [cont'd]	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, plaintiff's disagreement with the statements either involve a difference of opinion over whether a particular climate/ weather model has been sufficiently validated by scientists (*see* Complaint Ex. 12, Rebuttal Point 24) or the allegedly "egregious" statements addressed in the Academy's

modeling of transmission, reserve margins and frequency response; despite claims of system reliability (SI Appendix section S3). (iv) Lastly, we argue that the climate/weather model used for estimates of wind and solar energy production has not demonstrated the ability to accurately simulate wind speeds or solar insolation at the scales needed to assure the technical reliability of an energy system relying so heavily on intermittent energy sources (SI Appendix section S4)." Complaint, Ex. 12 at 9.	Memorandum. *See* Complaint Ex. 12, Rebuttal Points 24-26. As demonstrated in the Academy's Memorandum, there is nothing defamatory about any of the allegedly "egregious" statements.
"As we detail in Supporting Information section S1 of this paper, ref. [11] includes several modeling mistakes that call into question the conclusions of the study. For example, the numbers given in the Supporting Information of ref. [11] imply that maximum output from hydroelectric facilities cannot exceed 145.26 GW (see our Section S1.1), about 50% more than exists in the U.S. today [15], yet Figure 4(b) of ref. [11] (our Fig. 1) shows hydroelectric output exceeding 1,300 GW." Complaint, Ex. 12 at 9.	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, plaintiff's disagreement with these statements is based on the allegedly "egregious" statements addressed in the Academy's Memorandum. *See* Complaint Ex. 12, Rebuttal Point 25. As demonstrated in the Academy's Memorandum, there is nothing defamatory about any of those allegedly "egregious" statements.
"Similarly, as detailed in our Section S1.2, the total amount of load labeled as flexible in the figures of ref. [11] is much greater than the amount of flexible load represented in their supporting tabular data. In fact, the flexible load used by LOADMATCH is over double the maximum possible value from their Table 1. The maximum possible from Table 1 from ref. [11] is given as 1,064.16 GW, while Fig. 3 of [11] shows flexible load (in green) used up to 1,944 GW (on day 912.6). Indeed, in all the figures in [11] that show flexible load, the restrictions enumerated in their Table 1 are not satisfied." Complaint, Ex. 12 at 11.	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, plaintiff's disagreement with these statements is based on the allegedly "egregious" statements addressed in the Academy's Memorandum. *See* Complaint Ex. 12, Rebuttal Point 26. As demonstrated in the Academy's Memorandum, there is nothing defamatory about any of those allegedly "egregious" statements.
"In the analysis in ref. [11], the flexible loads can be accumulated in eight-hour blocks; which raises a serious issue of extreme excess industrial/commercial/residential capacity to utilize the high power for short periods of time. Under these assumptions, there would need to be oversized facilities on both the demand and [cont'd]	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, plaintiff's disagreement with these statements merely reflects a difference of opinion about "how load shifting works" and whether load shifting affects "demand-side" or "generation-side" capacity. *See* Complaint Ex. 12, Rebuttal Point 27.

generation sides to compensate for their respective variabilities. These errors are critical, as the conclusions reached by ref. [11] depend on the availability of large amounts of dispatchable energy and a large degree of flexibility in demand. Reference [11] also includes a scenario where zero demand response is allowed, and it shows that there is almost no cost changes and the grid is still stable." Complaint, Ex. 12 at 11.	
"Implausible assumptions The conclusions contained in ref. [11] rely on a number of unproven technologies and poorly substantiated assumptions, as detailed in section S2 of our SI Appendix. In summary, the reliability of the 100% wind, solar and hydroelectric power system proposed scheme depends centrally on a large installed capacity of several different energy storage systems [11], which collectively allow their model to flexibly reshape energy demand to match the output of variable electricity generation technologies. The study [11] assumes a total of 2,604 GW of storage charging capacity, more than double the entire current capacity of all power plants in the United States [16]." Complaint, Ex. 12 at 12.	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, plaintiff's reasons for disagreeing with the statements merely reflect a difference of opinion over whether his proposed storage capacity should be characterized as "twice that of the entire U.S. generating and storage capacity today" or "2/5 of the energy modeled." *See* Complaint Ex. 12, Rebuttal Points 16 & 28.
"The energy storage capacity consists almost entirely of two technologies that remain unproven at any scale: 514.6 TWh of underground thermal energy storage (UTES; the largest UTES facility today is 0.0041 TWh; further discussed in Section S2.1 of our SI Appendix), and 13.26 terawatt-hours (TWh) of phase-change materials (PCM; effectively, in research and demonstration phase; further discussed in Section S2.2 of our SI Appendix) coupled to concentrating solar thermal power (CSP). To give an idea of scale, the 100% wind, solar and hydroelectric power system proposed in ref. [11] envisions underground thermal energy storage (UTES) systems deployed in nearly every community for nearly every home, business, office building, hospital, [cont'd]	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, plaintiff's disagreement with the statements merely reflect a difference of opinion about whether underground thermal storage and phase-change materials have been shown to be viable technologies. *See* Complaint Ex. 12, Rebuttal Point 29.

school and factory in the United States, while only a handful exist today." Complaint, Ex. 12 at 12.	
"Although both PCM and UTES are promising resources, neither technology has reached the level of technological maturity to be confidently employed as the main underpinning technology in a study aiming to demonstrate the technical reliability and feasibility of an energy system. The relative immaturity of these technologies cannot be reconciled with the authors' assertion that the solutions proposed in ref. [11] and companion papers are ready to be implemented today at scale, at low cost, and that there are no technological or economical hurdles to the proposed system. "The 100% wind, solar and hydroelectric power system study [11] also makes unsupported assumptions about widespread adoption of hydrogen as an energy carrier, including the conversion of the aviation and steel industries to hydrogen, and the ability to store in hydrogen an amount of energy equivalent to more than a month of current U.S. electricity consumption. Further, in Figure S6 of ref. [11], hydrogen is being produced at a peak rate consuming nearly 2,000 GW of electricity, nearly twice the current U.S. electricity generating capacity." Complaint, Ex. 12 at 13.	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, plaintiff's disagreement with the statements merely reflect a difference of opinion about the reasonableness of plaintiff's assumptions, which plaintiff contends "are aggressive but reasonable." *See* Complaint Ex. 12, Rebuttal Point 30.
"As detailed in Section S2.3 of our SI Appendix, the costs and feasibility of this transition to a hydrogen economy are not appropriately accounted for by ref. [11]. To demonstrate the scale of the additional capacities that are demanded in ref. [11, 12] we plot them along with the electricity generation capacity in 2015 in Fig. 2. The data used for Fig. 2 can be found in two spreadsheets (and references therein) accompanying the manuscript. References [11] and [12] cite each other [cont'd]	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, plaintiff does not assert that any of these statements are false. *See* Complaint Ex. 12, Rebuttal Point 31. His comments merely address differences between the data used in two different papers. *Id.*

about the values of capacity. For example, ref [12], which supposedly includes information for all 50 states, reports ref. [11] [Tables S2] as the source of the numbers. "Then ref. [11], which only includes information for the capacity in the 48 contiguous states, cites ref. [12] [Table 2] as the source of the values. The values in the two papers do not agree, presumably because of the difference in number of states included, so it is unclear how each reference can be the source of the values for the other one." Complaint, Ex. 12 at 14.	
"Additionally, ref. [11] assumes that 63% of all energy-intensive industrial demand is flexible, able to reschedule all energy inputs within an 8-hour window. As discussed in Section S2.4 of our SI Appendix and in the National Research Council "Real Prospects for Energy Efficiency", it is infeasible for many industrial energy demands to be rapidly curtailed." Complaint, Ex. 12 at 14.	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, plaintiff does not assert that the statement mischaracterizes his paper. *See* Complaint Ex. 12, Rebuttal Point 32. He merely asserts that Dr. Clack mischaracterized the National Research Council's study. *Id.*
"Similarly, ref. [11] assumes that the capacity factor (i.e., actual electricity generation divided by the theoretically maximum potential generation obtained by operating continuously at full nameplate capacity) for existing energy technologies will increase dramatically in the future. As described in Section S2.5 of our SI Appendix, the authors of ref. [11] anticipate that individual hydropower facilities are assumed to increase generation by over 30%. They explain this by saying, 'Increasing the capacity factor is feasible because existing dams currently provide much less than their maximum capacity, primarily due to an oversupply of energy available from fossil fuel sources, resulting in less demand for hydroelectricity' [12]. From [12] it is stated that hydroelectric and geothermal capacity factors increase because 'For geothermal and hydropower, which are less variable on short time scales than wind and solar, the capacity- [cont'd]	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, plaintiff does not assert that the statement mischaracterizes his paper. *See* Complaint Ex. 12, First Rebuttal Point numbered 34. He merely asserts that there was no inconsistency between two papers because his paper's statement about using hydropower "only as a last report" was a "new finding." *Id.*

factor multipliers in our analysis are slightly greater than 100% on account of these being used more steadily in a 100% WWS system than in the base year'. In addition to being inconsistent with their statement that hydropower is 'used only as a last resort' [11] . . . ." Complaint, Ex. 12 at 15.	
". . . this explanation demonstrates a fundamental misunderstanding of the operation of electricity markets and the factors determining hydroelectric supply. With near-zero marginal costs (free 'fuel'), hydroelectric generators will essentially run whenever they are available; in those instances where they participate in merchant markets they underbid fossil generators that must at least recover their coal or natural gas costs. The primary factor limiting hydroelectric capacity factor is water supply and environmental constraints, not lack of demand. Further, there appears to be a mistake with the hydroelectric capacity factor adjustment: from EIA it should only go up to 42% not 52.5%." Complaint, Ex. 12 at 15.	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, plaintiff does not assert that Dr. Clack mischaracterized the operation of hydroelectric markets. *See* Complaint Ex. 12, Second Rebuttal Point numbered 34. He merely disagrees with 52.5% number in the quoted language, claiming that his paper used "0.925 x 52.5% = 48.6%." *Id.*
"To illustrate the implausibility of the assumed increase in hydroelectric net generation (dispatched from the plants to the electricity grid) in the face of limited water supply, we plot in Fig. 3 the last 25 years of generation from hydropower in the U.S. along with the average for the studies in ref. [11, 12]. The data used for Fig. 3 can be found in two spreadsheets (and references therein) accompanying the manuscript. Average future generation assumed by ref. [11,12] is 13% higher than the highest peak year in the last 25 and 85% higher than the minimum year in the last 25. So in addition to needing 1,300 GW of peak power from 150 GW of capacity, there also needs to be an extra 120 TWh of hydroelectric generation on top of the 280 TWh available. Further difficulties in raising hydropower capacity factors are described in our SI Appendix, Section 2.5." Complaint, Ex. 12 at 15.	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Plaintiff merely asserts that the "comparison should never have been made" because the figures presented in his paper included both U.S. and Canadian electric power. *See* Complaint Ex. 12, Rebuttal Point 35. But, in an errata sheet submitted to the Academy after the complaint was filed, plaintiff acknowledged that his paper did not clearly disclose he had included Canadian power in his figures.
"Most of the technologies considered in ref. [cont'd]	These statements do not say or imply anything about plaintiff's honesty, integrity or character.

[11] have high capital costs but relatively low operating costs. As a result, the cost of capital is a primary cost driver in the vision contained in ref. [11]. As discussed in Section S2.7 of our SI Appendix, the baseline value for cost of capital in ref [11] is one-half to one-third of that used by most other studies." Complaint, Ex. 12 at 15.	Moreover, these statements and plaintiff's comments on them merely reflect a difference of opinion on the appropriate discount rates to use in projecting costs. *See* Complaint Ex. 12, Rebuttal Points 36 & 37.
"The 100% wind, solar and hydroelectric energy system studies [11, 12] provide little evidence that the low cost of capital assumed in their study could be obtained by real investors in the capital markets. Using more realistic discount rates of 6 – 9% per year instead of the 3 – 4.5% in ref [11] could double the estimate of an 11 cents/kWh, even before adding in the unaccounted-for capital costs described above. One possible explanation of the lower discount rates used could be that they forecast lower growth (or negative) gross domestic product. In the case of lower growth, there would likely be lower interest rates; however that lower growth may also lead to lower energy demand and investment." Complaint, Ex. 12 at 15.	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, these statements and plaintiff's comments on them merely reflect a difference of opinion on the appropriate discount rates to use in projecting costs. *See* Complaint Ex. 12, First Rebuttal Point numbered 37.
"One of the global leaders of solar PV and wind energy installation in recent years is Germany, which through its 'Energiewende' is attempting to shift toward an 80% renewables energy system. Germany therefore, presents a suitable example against which to benchmark the feasibility of the plan set out in ref. [11] for the United States. In Section S2.8 of our SI Appendix, we describe how ref. [11] assumes that the U.S. will build out new solar, wind and hydro facilities at a sustained rate that, on a per unit GDP basis, is 16 times greater than the average deployment rate in Germany's Energiewende initiative during the years 2007 to 2014, and over 6 times greater than Germany achieved in the peak year of 2011 (see our Fig. S4)." Complaint, Ex. 12 at 16-17.	These statements do not say or imply anything about plaintiff, his honesty, his integrity or his character. Moreover, plaintiff does not say that the statements are false. *See* Complaint Ex. 12, Rebuttal Point 38. He simply asserts that they are irrelevant. *Id.*
"In Fig. 4, we display another metric on the scale of expansion. It shows the rate of [cont'd] installation as Watts per Year per Capita.	These statements do not say or imply anything about plaintiff, his honesty, his integrity or his character. Moreover, plaintiff does not say that the statements are false. *See* Complaint Ex. 12,

Using this metric, we can compare the scale of capacity expansion in ref. [11] with historic data. Figure 4 shows that the plans proposed in references [11] and [12] would require a sustained installation rate that is over 14 times the U.S. average over the last 55 years; and over 6 times the peak rate. For the sake of comparison, Fig. 4 includes the estimated rate for a solution that decarbonizes the U.S. electric grid by 78% by 2030 [1], historical German data and historical Chinese data. We note that ref. [1] considered large-scale storage, but excluded it based upon preliminary results showing that it was not cost effective compared to a national transmission system. The data used for Fig. ?? can be found in two spreadsheets (and references therein) accompanying the manuscript. Sustaining public support for this scale of investment (and this scale of deployment of new wind turbines, power lines, etc.) could prove challenging. One of the reasons this buildout may prove difficult, is that the 100% wind, solar and hydroelectric system relies on energy sources with relatively low areal power density (see Section S2.9 of our SI Appendix for further details). According to NREL, average power densities achieved in land-based wind farms is about 3 W/m^2 with a range of 1-11.2 W/m^2 (although at larger deployment scales, power densities would likely be lower) [17]." Complaint, Ex. 12 at 17.	Rebuttal Point 39. Rather, his comments demonstrate there is a difference of opinion on the relevance and adequacy of the data Dr. Clack cited. *Id.*
"At the average power densities, the scale of wind power envisioned in ref. … would require nearly 500,000 km^2 (134,000-1,500,000 km^2), which is roughly … of the continental U.S. and >1,500 m^2 of land for wind turbines for each Americ . . . . " Complaint, Ex. 12 at 17.	These statements do not say or imply anything about plaintiff, his honesty, his integrity or his character. Moreover, plaintiff's response to the statements merely asserts that Dr. Clack used an incorrect density figure and thereby miscalculated the area needed for wind power. *See* Complaint Ex. 12, Second Rebuttal Point numbered 37.
"Much of this land could be dual use, but the challenges associated with this lev… scale up should not be underestimated. The propose transition in ref. [11] req… unprecedented [cont'd] rates of technology deployment. For example,	These statements do not say or imply anything about plaintiff, his honesty, his integrity or his character. Moreover, plaintiff's disagreement with these statements merely reflects a difference of opinion on whether his paper is consistent with literature published by others.

increased pressu… materials, elevated commodity prices and high demand for wind power installa… produced elevated prices for wind power deployment between 2002 and 2008 [9]. The rejection of many potential sources of low-carbon-emission energy is … on an analysis presented by Jacobson et al. in ref. [20]. A full discussion of paper is beyond the scope of our current evaluation. However, one flaw is its fa… to use other numbers already published detailed studies on life-cycle greenhouse … (GHG) emissions, land-use requirements and human mortality of e… production technologies. Rather than using the results of the many detailed st… available from large international bodies such as those surveyed by … Intergovernmental Panel on Climate Change, ref. [20] presents assessments th… many cases differ in method and granularity to produce results that offer mar… from those generally accepted in scientific and technical communities." Complaint, Ex. 12 at 17.	*See* Complaint Ex. 12, Rebuttal Point 40.
"Selective assessments of life-cycle emissions can be used to favor or disfavor specific technologies. As an example, the lifecycle GHG-emissions for nuclear power generation in ref. [20] include the emissions of the background fossil-based power system during an assumed planning and construction period for up to 19 years per nuclear plant." Complaint, Ex. 12 at 18.	These statements do not say or imply anything about plaintiff, his honesty, his integrity or his character. Moreover, plaintiff's reason for disagreeing with this statement is unclear. He appears to assert only that the statement "up to 19 years" should be "10-19 years," and that he believes Dr. Clack's article should have discussed the costs associated with "nuclear weapons proliferation" and "meltdown risk." *See* Complaint Ex. 12, Rebuttal Point 41.
"Added to this, the effects of a nuclear war, which is assumed to periodically reoccur on a 30-year cycle, is included in the analysis of emissions and mortality of civilian nuclear power. In contrast, those same authors do not consider emissions for the fossil-based power system associated with construction and permitting delays for off-shore wind farms (or the transmission infrastructure needed to connect these farms), which has already been a [cont'd] challenge in the development of U.S. offshore	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, plaintiff's comment on the statements merely assert that he assumed 2-5 years between planning and operation of wind farms, without stating whether he took into account the emissions and transmission infrastructure issues mentioned by Dr. Clack. *See* Complaint Ex. 12, Rebuttal Point 42. In any event, this involves nothing more than a disagreement over the adequacy of a scientific analysis.

wind resources." Complaint, Ex. 12 at 18.	
"While there is extensive experience outside of the U.S. with developing offshore wind resources, very few offshore wind facilities have been permitted in the U.S. territorial waters. The 100% wind, solar and hydroelectric power system [11] envisions more than 150,000 5-MW turbines permitted and built offshore, without delays. **Insufficient power system modeling** The study of a 100% wind, solar and hydroelectric power system [11] purports to report the results of a 'grid integration model'. It is important to understand the limitations of the study with regard to what is usually meant by grid integration. Reliable operation of the grid involves a myriad of challenges beyond just matching total generation to total load. Its role in cascading failures and blackouts illustrates the important role of the transmission system [21]. Reliable grid operation is further complicated by its AC nature, with real and reactive power flows and the need to closely maintain a constant frequency [22]. Margins for generator failures must be provided through operational and planning reserves [23]. The solution proposed by ref. [11, 12] involves fundamental shifts in aspects of grid architecture that are critical to reliable operation. Wind generation, largely located far from load centers, will require new transmission. Solar generation and on-site storage connected to the distribution grid replace capability currently connected to the more-centralized transmission grid. Rotating machines whose substantial inertia is critical for frequency stability are supplanted by synchronous wind and solar generators. While a grid integration study is detailed and complex, the grid model of ref. [11] is spatially zero-dimensional; all loads, generation (sited before the LOADMATCH runs and placed [cont'd] precisely where existing generation resides)	These statements do not say or imply anything about plaintiff's honesty, integrity or character. Moreover, plaintiff's comments on the statements merely assert that he did not ignore transmission capacity expansion, and provide his opinions on the relative merits of his analysis and analyses conducted by other scientists. *See* Complaint Ex. 12, Rebuttal Point 43.

and storage are summed in a single place. Therefore, those authors do not perform any modeling or analysis of transmission. As a result, their analysis ignores transmission capacity expansion, power flow, and the logistics of transmission constraints (see our Section S2.6). Similarly, those authors do not account for operating reserves, a fundamental constraint necessary for the electric grid. Indeed, LOADMATCH used in ref. [11] is a simplified representation of electric power system operations that does not capture requirements for frequency regulation to ensure operating reliability (see further details in Section S3 of our S1 Appendix)." Complaint, Ex. 12 at 18-20.	
"Further, the model is fully deterministic, implying perfect foresight about electricity demand and the variability of wind and solar energy resou… neglecting the effect of forecast errors on reserve requirements [24]. In a sy… where variable renewable resources make up over 95% of U.S. energy su… renewable energy forecast errors would be a significant source of uncertain …the daily operation of power systems. The LOADMATCH model does… demonstrate the technical ability of the proposed system from ref. [11] to op… reliably given the magnitude of the architectural changes to the grid and… degree of uncertainty imposed by renewable resources." Complaint, Ex. 12 at 20.	These statements do not say or imply anything about plaintiff, his honesty, his integrity or his character. Moreover, plaintiff's disagreement with these statements merely reflects a difference of opinion over whether LOADMATCH should be characterized as a "deterministic model" or as being more similar to a "pure stochastic model." *See* Complaint Ex. 12, Rebuttal Point 44.
"Inadequate scrutiny of input climate model The climate model used to generate weather data used by ref. [11] has never been adequately evaluated. For example, results from this model have not been made available to the Climate Model Intercomparison Project [25] or been opened to public inspection in ways similar to the results for major reanalysis projects [26]. As detailed in section S4 of our SI Appendix, the fragmentary results that have been made available show poor correlation with reality in terms of resolution and [cont'd] accuracy. Since the conclusions from ref. [11]	These statements do not say or imply anything about plaintiff, his honesty, his integrity or his character. Moreover, plaintiff's disagreement with the statements merely reflect a difference of opinion over whether a particular climate/weather model has been sufficiently validated by scientists. *See* Complaint Ex. 12, Rebuttal Points 24 & 44.

depend on the weather data used, their conclusions cannot be considered to be adequate without an appropriate evaluation of the weather data used." Complaint, Ex. 12 at 21.	
"Conclusions Many previous studies of deep decarbonization of electric power illustrate that much can be done with wind and solar power, but that it is extremely difficult to achieve complete decarbonization of the energy system even when employing every current technology and tool available, including energy efficiency and wind, hydro, and solar energy, but also carbon capture and storage, bioenergy, and nuclear energy [1-, 8-10]. In contrast, ref. [11] asserts that it is cost-effective to fully decarbonize the U.S. energy system primarily using just three inherently variable generating technologies: solar PV, solar CSP, and wind, to supply more than 95% of total energy in the proposal presented in ref. [11]. Such an extraordinarily constrained conclusion demands a standard of proof that ref. [11] does not meet. The scenarios of ref. [11] can at best be described as a poorly executed exploration of an interesting hypothesis. The study's numerous shortcomings and errors render it unreliable as a guide about the likely cost, technical reliability, or feasibility of a 100% wind solar and hydroelectric power system. It is one thing to explore the potential use of technologies in a clearly caveated hypothetical analysis; it is quite another to claim that a model employing these technologies at an unprecedented scale conclusively demonstrates the feasibility and reliability of the modeled energy system implemented by mid-century. From the information given by ref. [11], it is clear that both hydroelectric power and flexible load have been modeled in erroneous ways, and these errors alone invalidate the study and [cont'd] its result. The study of 100% wind, solar and	These statements do not say or imply anything about plaintiff's honesty, integrity or character. The statements and plaintiffs comments on the statements merely reflect a difference of opinion over plaintiff's methodologies, assumptions and conclusions. *See* Complaint Ex. 12, Rebuttal Point 45.

hydroelectric power systems [11] extrapolates from a few small-scale installations of relatively immature energy storage technologies to assume ubiquitous adoption of high-temperature phase-change materials for storage at concentrating solar power plants, underground thermal energy storage for heating, cooling, and refrigeration for almost every building in the United States, and widespread use of hydrogen to fuel airplanes, rail, shipping, and most energy-intensive industrial processes. For the critical variable characteristics of wind and solar resources, they rely on a climate model that has not been independently scrutinized. The authors of ref. [11] claim to have demonstrated that their proposed system would be low-cost and that there are no economic barriers to the implementation of their vision [12]. However, the modeling errors described above, the speculative nature of the TW-scale storage technologies envisioned, the theoretical nature of the solutions proposed to handle critical stability aspects of the system and a number of unsupported assumptions, including a cost of capital that is a third to a half lower than is used in practice in the real world, undermine that claim. Their LOADMATCH model does not consider aspects of transmission power flow, operating reserves or of frequency regulation that would typically be represented in a grid model aimed at assessing reliability. Further, as detailed above and in the SI Appendix, a large number of costs and barriers have not been considered in ref. [11]. Many researchers have been examining energy system transitions for a long time. Previous detailed studies have generally found that energy system transitions are extremely difficult, and that a broad portfolio of technological options eases that transition. If one reaches a new conclusion by not addressing factors considered by others, by making a large set of unsupported assumptions,	

by using simpler models that do not consider important features, and then performing an analysis that contains critical mistakes, the anomalous conclusion cannot be heralded as a new discovery. The conclusions reached by the study contained within ref. [11] about the performance and cost of a system of "100% penetration of intermittent wind, water and solar for all purposes" are not supported by adequate and realistic analysis and do not provide a reliable guide to whether, and at what cost, such a transition might be achieved. In contrast, the weight of the evidence suggests that a broad portfolio of energy options will help facilitate an affordable transition to a near zero-emission energy system." Complaint, Ex. 12 at 21.	

EXHIBIT F

MANUSCRIPT SUBMISSION

ABOUT // AUTHORS // REVIEWERS // LOGOUT MY HOMEPAGE

Submission Summary

2015-10028R

Mark Z. Jacobson

Direct Submission

A Low-Cost Solution to the Grid Reliability Problem With 100% Penetration of Intermittent Wind, Water, and Solar for all Purposes

Take a guided tour

Upload Files

Files

Manuscript Info

✔ Manuscript Title

✔ Authors

✔ Manuscript Classification

✖ **Author COI and Comments**

✔ PNAS License to Publish

Review Material

Length Estimate

Review Files

✖ Review Data

Save and Exit

Author COI and Comments

Please enter the information requested below. Required fields are marked with an asterisk (*). If a field is not completed, that section will be shaded in red.

Author Conflict of Interest

According to journal policy, "Authors must disclose, at submission, any association that poses a Conflict of Interest in connection with the manuscript and acknowledge all funding sources supporting the work." Failure to disclose a Conflict of Interest, including patent filings, ***at submission*** *may result in additional review, publication delays, or author sanctions. For more information, see the full* Conflict of Interest Policy. *Recent collaborators, defined as people who have coauthored a paper or were a principal investigator on a grant with the author or member within in the past 48 months, must be excluded as editors and reviewers. If conflict of interest is indicated, please provide details in the text box below.*

◉ No, the authors declare no conflict of interest (such as defined by PNAS policy).

○ Yes, the authors declare a conflict of interest (such as defined by PNAS policy). Please specify in the box below.

*** Open Access Option**

If your article is accepted for publication, do you elect to pay an additional fee of $1,450 if your institution does not have a site license or $1,100 if your institution has a current site license to make your paper free online immediately after publication through the PNAS open access option? If you select the open access option, PNAS will confirm whether or not your institution has a current site license upon acceptance of your manuscript and will bill you accordingly. Note: All PNAS articles are free online 6 months after publication.

○ Yes, I select the open access option and agree to pay an additional $1,450 (or $1,100 if my institution has a current site license).

◉ No, I do not select the open access option.

*** Data Sharing**

PNAS requires authors to share their data, materials, and code. Authors should include a statement detailing their data sharing plans (including all data, documentation, and code used in analysis) in the methods section. Any restrictions on access to the data, or any parts of it, must be disclosed at submission and be included in the paper. If any of the data cannot be made available, please explai in detail.

◉ Yes, all data discussed in the paper will be made available to readers.

○ No, there are restrictions on accessing data discussed in the paper.

If you selected No, please use the textbox below to explain in detail any restrictions to accessing the data in your submission. The PNAS Materials and Data Availability policy is found here: http://www.pnas.org/site/authors/journal.xhtml

MANUSCRIPT SUBMISSION

Funding Body
Please indicate the funder(s) of the research described in this manuscript and the associated grant reference numbers.

ABOUT // AUTHORS // REVIEWERS // LOGOUT MY HOMEPAGE

Begin typing the name(s) of the funder(s) into the Funder(s) textbox to select from the standardized list of names for the funders you listed in the acknowledgments section of your manuscript. If no list appears or the funder you are looking for does not appear in the list, please skip this step.

All authors included in your submission will be pre-populated in the "Author" field. Please ASSIGN each author TO THEIR SPECIFIC funder. (To select more than one author hold down the CTRL or Apple key)

Funder(s)	Grant Reference Number *	Authors *	
		Mark Z. Jacobson Mark A Delucchi Mary A Cameron Bethany A Frew	Clear
		Mark Z. Jacobson Mark A Delucchi Mary A Cameron Bethany A Frew	Clear

More Funders

Funding comments:

None

Please enter Funding Body

Database Deposition
Before publication, authors must deposit large data sets in an approved database and provide an accession number for inclusion in the published article. Please indicate below whether you have a crystallographic, sequence, or other type of deposition, or select "None" if database deposition is not applicable to your manuscript. When no public repository exists, authors must provide the data as supporting information or—in special circumstances when this is not possible—on the author's institutional website, provided that a copy of the data is submitted to PNAS.

None
Crystallography, atomic coordinates, and structure factors
NMR, atomic coordinates, chemical shifts, and restraints
Sequence deposition

If you have selected a deposition type, please use the box below to supply all database accession numbers and/or codes (this information will be provided as a footnote to help readers find your data if your manuscript is accepted for publication). If "Other" type of deposition is selected, provide a description below along with any special instructions regarding database deposition you wish to communicate to the editorial staff.

Previously Published Material
If your manuscript includes material (e.g., figures or tables) that were published previously, whether or not you are an author of the earlier publication, you must provide the previous publisher's permission to republish the material in print and online. PNAS will acknowledge that copyright is held by the original publisher. Please indicate whether your manuscript includes previously published material.

○ No, the manuscript does not include previously published material(s).
○ Yes, the manuscript includes previously published material(s). Please describe below.

MANUSCRIPT SUBMISSION

Sample text: Figure 1 published in XYZ journal on DATE.

ABOUT // AUTHORS // REVIEWERS // LOGOUT MY HOMEPAGE

Digital Image Manipulation Statement

According to journal policy, "No specific feature within an image may be enhanced, obscured, moved, removed, or introduced." For more information, see the full Figure Preparation section under the PNAS Journal Policies. If the terms of the figure preparation guidelines are not met for any and all figures being submitted, please use the text box below to explain.

- (●) Yes, the figure(s) for this submission conform to the terms of PNAS policy.
- ( ) There are no figures associated with this submission.
- ( ) No, the figure(s) for this submission do not conform to the terms of PNAS policy. Please specify in the box below.

Preprint Server

Please let us know whether you have posted a version of your submission to a preprint server

- ( ) Yes
- ( ) No

If you selected Yes, please provide the URL or DOI for the preprint version of your article:

URL or DOI

Comments for the Editorial Board and Editor

Please describe why you believe this article is appropriate for publication in PNAS. These comments will be made available to the Editorial Board and assigned Editor.

If published, this paper may serve as the scientific basis by which countries of the world can justify the rapid conversion from fossil fuels to 100% clean, renewable wind, water, and solar (WWS) energy for all sectors (electricity, transportation, heating/cooling, and industry). Currently, such a conversion is largely limited by the fear of grid instability and resulting high cost due to the intermittency of

Comments for Editorial Staff

Please add any additional comments you wish to convey with your submission.

Terms and Conditions

By clicking this button you affirm you have read and agreed to abide by the editorial policies noted above and in the Information for Authors.

By clicking this button you affirm you have read and agreed to abide by the editorial policies noted above and in the Information for Authors.

☑ I agree with the terms and conditions stated above.

MANUSCRIPT SUBMISSION

*** Consent to Use Human Images in Manuscript Material**

Please confirm whether your manuscript contains recognizable human faces, and whether or not you have received consent to publish these images.

ABOUT | AUTHORS | REVIEWERS | LOGOUT MY HOMEPAGE

Does this manuscript contain photographs or videos that include recognizable human faces?

○ Yes
○ No

Please indicate.

If you answered yes to the previous question, have you received consent to use these images from the persons pictured?

○ Yes
○ No

Guarantor of the Work

PNAS requires the corresponding author to check the box below as a way of certifying that at least one author, usually a senior author, has seen all primary data, has approved the presentation of the data, and will abide by the PNAS policy on Materials and Data Availability *to make materials, data, and associated protocols available to readers.*

I confirm a senior author has reviewed the data being presented in this submission.

☐ I agree

Previous Next

EXHIBIT G

PNAS License to Publish

Please complete both sections of this form and e-mail it to PNASauthorlicense@nas.edu, or fax it to PNAS (202-334-2739). Failure to return the form will delay or prevent publication.

License to Publish

1. In consideration of publication of the WORK (including but not limited to figures, tables, artwork, abstracts, cover images, or summaries submitted with the WORK) in the Proceedings of the National Academy of Sciences USA (PNAS), the author(s) hereby grants to the National Academy of Sciences (USA) for the full term of copyright and any extensions thereto the sole and exclusive, irrevocable license to publish, reproduce, distribute, transmit, display, store, translate, create derivative works from and otherwise use the Work in any language or in any form, manner, format, or medium now known or hereafter developed without limitation throughout the world, and to permit and/or license others to do any or all of the above. In the event that PNAS decides not to publish the WORK, this license shall be terminated and all rights revert to the author(s).
2. In consideration of publication of the WORK in PNAS, the author(s) also grants to the National Academy of Sciences for the full term of copyright and any extensions thereto all of the same rights provided for in clause 1 above, for the Supplemental Information, but on a nonexclusive basis.
3. Ownership of copyright in the WORK remains with the author(s), and provided that, when reproducing the WORK or extracts from it, the author(s) credit first publication of the WORK in PNAS, the author(s) retains the following nonexclusive rights:
   a. The author(s) reserves the right after publication of the WORK by PNAS, to use all or part of the WORK in compilations or other publications of the author's own works, to use figures and tables created by them and contained in the WORK, and to make copies of all or part of the WORK for the author's use for lectures, classroom instruction, or similar uses.
   b. The author(s) reserves the right to post the WORK (including the PNAS-formatted PDF) after publication in PNAS on the author's web page provided that a link to the article in PNAS Online is included.
   c. The author(s) reserves the right to post and update the WORK on pre-print services such as arXiv provided that PNAS-formatted files (HTML and PDF) are not used and that a link to the WORK in PNAS Online is included.
   d. The author(s) reserves the right to deposit his or her final manuscript in the author(s)'s funding body's archive or designated repository on acceptance for publication by PNAS, provided that a link to the WORK in PNAS Online is included; and, the right to request public access 6 months after print publication or immediately upon publication by PNAS if the author(s) has paid the PNAS open access fee. PNAS automatically deposits the PNAS-formatted version of all papers in PubMed Central.
4. The exclusive license granted to PNAS permits the following: after publication of the WORK by PNAS, third parties are automatically granted permission to use figures and tables created by the author(s) and contained in the WORK for noncommercial use (e.g., not for sale or for commercial advantage).
5. The author(s) warrants that the WORK is original, that it has not been published, that all the facts it contains are true and accurate, that it does not infringe upon any copyright, proprietary, or personal right of any third party, and that permissions have been obtained and are attached for any portions of the WORK owned or controlled by a third party. Each author identified on the manuscript must sign this form. If any author's signature does not appear below, the signing author(s) represents and warrants that he or she signs this license as authorized agent(s) for and on behalf of all the authors and that this license is made on behalf of all the authors. The author(s) shall indemnify the Academy and/or its successors and assigns for any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of this warranty or other representations contained herein.
6. The Author(s) authorize(s) PNAS to take such steps as it considers necessary at its own expense in the Author(s)'s name and on their behalf if it believes that a third party is infringing or is likely to infringe the copyright in the WORK.
7. If the WORK has been prepared by an employee within the scope of his or her employment, the employer reserves the right to make copies of the WORK for its own internal use.

Title of WORK

Author's Signature Date

Print Name

Author's Signature Date

Print Name

Author's Signature Date

Print Name

Please use additional form(s) for more authors.

Manuscript No. ________________

Documentation Report:

If any of the following items are pertinent to your paper, we must receive the appropriate documentation or statements before we can publish your paper.

Please check the appropriate boxes and sign below:

- ☐ **None of the four items below applies to this paper.**
- ☐ **Structural Coordinate Deposition and Release.** Authors of papers reporting new or revised structures must deposit their coordinates in the Protein Data Bank or an equivalent public archive, and the accession numbers must be supplied for publication. You agree, by signing below, that the coordinates will be released when the article is published. A footnote reporting the accession numbers will be added to your paper. You must provide the accession numbers on the page proof.
- ☐ **Sequence Database Deposition.** Authors of papers reporting new sequences must submit these data to an appropriate database, and the accession numbers must be supplied for publication. A footnote reporting the accession numbers will be added to your paper. You must provide the accession numbers on the page proof.
- ☐ **Open Access Mandate Waiver.** Authors are responsible for obtaining waivers of any nongovernment institutional open access mandates before publication with PNAS. Many academic institutions require that their authors transfer a nonexclusive author license to the institution and deposit the final author manuscript, with edits from peer review incorporated, into institutional repositories. These policies conflict with the PNAS exclusive license to publish. Authors employed by an institution with such a mandate should obtain a waiver for the nonexclusive license and forward it to PNASAuthorLicense@nas.edu.
- ☐ **Permission to Print Previously Published Material.** If your paper includes material (e.g., figures or tables) that were published previously or modified from a previous publication, whether or not you are an author of the earlier publication, you must provide the previous publisher's permission to republish this material in print and online. We acknowledge that copyright is held by the original publisher. Please state which items were published previously and where:

Author's Signature Date

If WORK is a Work Made for Hire and the employer of the author(s) owns the copyright, this license to publish must be signed by an author and an authorized representative of the employer.

Name of Employer

Authorized Signature Date

Title

Name(s) of Employee(s)

Authorized Signature Date

Title

If all authors are US Government employees: I certify that the WORK was written as part of the official duties of the authors as employees of the US Government. On behalf of all authors, I authorize publication of the WORK in PNAS under terms equivalent to those provided for in the license and agree to all of the warranties specified in the license.

Author's Signature Date

Print Name

Author's Signature Date

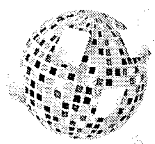
Print Name

EXHIBIT H

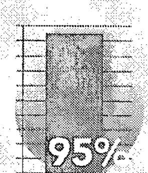
PNAS Submission & Editorial Review

HOW IT WORKS . . .

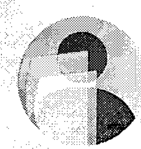
1. OVERVIEW



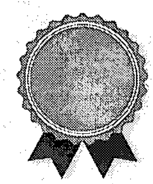
Submissions are welcomed from researchers all over the world. Authors do not need to have a connection to an NAS member to publish in PNAS.



of submitted papers are Direct Submissions—not NAS member Contributed papers.



PNAS is committed to transparency in its editorial review process. FAQs for authors explaining the review process are available at www.pnas.org/site/authors/authorfaq.xhtml.



All articles are evaluated solely on their scientific merit by peers—not by staff editors. Accepted papers must be of exceptional scientific importance and intelligible to a broad scientific audience.

2. SUBMISSION AND REVIEW

TWO CATEGORIES

1. **Direct Submission**—the primary and preferred route for publication in PNAS
   - A member of the NAS does not need to sponsor the paper.
   - Authors receive a first decision within about 3 weeks and have their papers published online as soon as 4–5 weeks after acceptance.
2. **Contributed Papers**—articles authored and submitted by NAS members
   - Members select their own reviewers.
   - Less than 25% of the papers published in PNAS are Contributed by NAS members.

THREE-TIER REVIEW PROCESS FOR ALL RESEARCH ARTICLES

An NAS member makes the final decision for every paper.

1. **Editorial Board Member**—Each paper is assigned to an Editorial Board member from one of the 31 NAS disciplines. For a list of Editorial Board members and their disciplines see www.pnas.org/misc/masthead.xhtml. All articles are subject to evaluation by the Editorial Board.
2. **NAS Member Editor**—If the Editorial Board member decides the paper should proceed further, it is assigned to an NAS Member Editor, a professional scientist and active researcher, who oversees the review process and determines suitability of the work for PNAS.
3. **Independent Peer Reviewers**—Recognized subject experts are selected by the Member Editor to review the work. The reviewers' recommendations are evaluated by the Member Editor, and the final decision to accept an article is made by an Editorial Board member.

3. PUBLISHING OPTIONS: OPEN ACCESS and PNAS PLUS

- Published papers can be made immediately available online at no cost to the reader when the authors choose the **open access** option. All PNAS papers are freely available 6 months after publication.
- **PNAS Plus** allows authors the option to publish an expanded research article with a 10-page limit. PNAS Plus articles follow the same peer review process as regular papers.

IN THE SUPERIOR COURT
FOR THE DISTRICT OF COLUMBIA
Civil Division

MARK Z. JACOBSON, Ph.D.,)	
Plaintiff,)	
v.)	**Civil Action No. 2017 CA 006685 B**
)	**Hon. Elizabeth Carroll Wingo**
CHRISTOPHER T. M. CLACK, Ph.D.)	**Next Court Date: December 29, 2017**
***et al.*,**)	**Event: Initial Conference**
Defendants.)	

[PROPOSED] ORDER

Upon consideration of Defendant National Academy of Sciences' Special Motion to Dismiss Pursuant to the D.C. Anti-SLAPP Act or, In the Alternative, Motion to Dismiss Pursuant to Rule 12(b)(6), it is this ___ day of __________________, 2018, hereby

ORDERED, that the Special Motion to Dismiss is GRANTED in all respects and Plaintiff's claims are dismissed in their entirety with prejudice; and it is further

ORDERED, that Defendant National Academy of Sciences may file a request for attorneys' fees and costs.

SO ORDERED.

____
Judge Elizabeth Carroll Wingo
District of Columbia Superior Court

cc: Paul S. Thaler (D.C. Bar No. 416614)
Counsel for Plaintiff
Mark Z. Jacobson, Ph.D.
(By eService)

Drew W. Marrocco (D.C. Bar No. 453205)
Counsel for Defendant
Christopher T. M. Clack, Ph.D.
(By eService)

Joseph P. Esposito (D.C. Bar No. 375507)
William E. Potts, Jr. (D.C. Bar No. 945824)
Counsel for Defendant National Academy of Sciences
(By eService)