Title: Disease Outbreaks Select for Mate Choice and Coat Color in Wolves

Authors: Sarah Cubaynes[1*], Ellen E. Brandell[2], Daniel R. Stahler[3], Douglas W. Smith[3], Emily S. Almberg[4], Susanne Schindler[5], Robert K. Wayne[6], Andrew P. Dobson[7], Bridgett M. vonHoldt[7], Daniel R. MacNulty[8], Paul C. Cross[9], Peter J. Hudson[2], Tim Coulson[10]

Affiliations:

[1] CEFE, Univ Montpellier, CNRS, EPHE-PSL University, IRD, Montpellier, France

*Corresponding author. Email: sarah.cubaynes@cefe.cnrs.fr

[2] Center for Infectious Disease Dynamics, Department of Biology, Pennsylvania State University; State College, Pennsylvania, USA. 16802.

[3] Yellowstone Center for Resources; Yellowstone National Park, WY 82190, USA.

[4] Wildlife Division, Montana Fish Wildlife & Parks; 1400 S 19th Avenue, Bozeman, Montana, USA.

[5] School of Biological Sciences, University of Bristol; Bristol, U.K.

[6] Department of Ecology and Evolutionary Biology, UCLA Ecology & Evolutionary Biology; 2312 Life Sciences Building, BOX 951606, Los Angeles, CA, USA.

[7] Department of Ecology and Evolutionary Biology; 106A Guyot Hill, Princeton, New Jersey, USA.

[8] Department of Wildland Resources and Ecology Center, Utah State University; Logan, Utah, USA.

[9] U.S. Geological Survey, Northern Rocky Mountain Science Center, 2327 University way; Suite 2, Bozeman, Montana, USA.

[10] Department of Zoology, University of Oxford; Mansfield Road, Oxford, OX1 3SZ, UK.

Abstract: We know much about pathogen evolution and the emergence of new disease strains but less about host resistance and how it is signaled to other individuals and subsequently maintained. The cline in frequency of black-coated wolves across North America is hypothesized to result from a relationship with canine distemper virus (CDV) outbreaks. We test this hypothesis using cross-sectional data from wolf populations across North America that vary in the prevalence of CDV and the allele that makes coats black, longitudinal data from Yellowstone National Park, and modeling. The frequency of CDV outbreaks generates fluctuating selection that results in heterozygote advantage that in turn impacts the frequency of the black allele, the optimal mating behavior, and the black wolf cline across the continent.

One-Sentence Summary: Disease epidemics drive the evolution of morphology and behavior in wolves.

Variation in color is frequently used to assess the quality of potential mates and their fit to environmental conditions (*1-3*). In many species, color covaries with aspects of the environment such as altitude, weather, the presence of specific parasites, food resources, and predators, and an individual's color can signal its condition or immunological status (*4- 6*). Honest signals need coloration to be correlated with fitness-associated traits, and under these conditions this may select for particular mate choice strategies as individuals choose partners who maximize the expected fitness of their offspring (*7*). When the environment varies spatially, generating a cline in selection pressures, this could lead to landscape-level variation in coloration (*8,9*) and spatial variation in strategies of mate choice behavior (*1,4*).

Although rare at high latitudes, black wolves increase in frequency along a south-west cline towards forested areas (*10*) in North America, with the highest frequencies at each latitude observed along the Rocky Mountains (Fig. 1(A)). The absence of geographical barriers that prevent gene flow coupled with molecular signals of selection points to regional variation in coat color being due to a cline in selection pressures (*11*).

Coat color in wolves is determined by genotype at the *K locus* gene *CBD103* (*12*). The ancestral wild-type *k* allele allows a normal *Agouti* and *Mc1r* gene interaction resulting in gray coat color, whereas a three-nucleotide deletion in the *K locus* gene causes the protein to prevent *Agouti* function, leading to dominant inheritance of a black coat (*12, 13*) . Following a single introgression event into a North American wolf population in the last 7,250 years, the black allele has undergone a selective sweep, revealing one of the most rapid spreads of an adaptive variant known in vertebrates (*14*). The homozygote KK and the heterozygote Kk have indistinguishable black pelage but very different fitnesses (*15*). Therefore, coat color phenotype is not itself under direct selection, but the black allele must have a function that impacts fitness directly or through pleiotropic effects, conferring a strong selective advantage in certain environments (*11,14,16*). As the *K* locus encodes for a β-defensin protein that plays a direct role in innate and adaptive immunity in mammals (*17*), we postulate that the *K* locus is involved in immunity to respiratory infections such as the morbillivirus that causes CDV – a pathogen of carnivores (*18*) that can cause significant mortality among immunologically naive individuals, which are often juveniles (*19*).

An ability to fight disease can generate a fitness cost in the absence of these threats (*20,21*) . We examine whether environment-dependent fitness benefits of certain genotypes could explain the North American cline in wolf coat color frequency. CDV infects most carnivores, and the frequency of outbreaks vary depending on the composition of the carnivore communities (*22, 23*). To test the prediction that coat color varies with CDV occurrence, we analyzed twelve wolf populations to examine whether the probability of a wolf being black was predicted by the presence of CDV antibodies.

Wolves seropositive for distemper are more likely to be black, especially at older ages (Fig. 1(B)). We constructed a model to assess the individual- and population-level effects of distemper on the probability of a wolf being black (24). We predicted the probability of being seropositive for distemper while standardizing for age and confounding factors (Fig. S2). The population effect is the positive correlation between the population-level disease exposure and whether an individual is black or gray (Fig. 1(C)). The individual effect captures whether a previously CDV exposed individual is more likely to be black, perhaps due to being more likely to survive the infection and sampled later. If an individual was seropositive for distemper, its probability of being black increases from 25% to 32% ($p = 0.03$).

These results are consistent with our hypothesis that CDV exposure is positively associated with coat color frequency, but do not provide strong support. Next, we report an analysis of the Yellowstone wolf population using individual life history and coat color data collected since reintroduction in 1995/6 (*25*). The population consists of ~55% gray wolves (genotype *kk*) and 45% black wolves (genotypes *Kk* and *KK*) with only 5% of these being homozygote (Fig. 2). Previous research has revealed that female gray wolves have 25% higher reproductive success in all years compared to black females, and that CDV outbreaks generated a 50% reduction in female reproductive success independent of coat color (*26*). However, because so few genotyped black homozygote individuals have reproduced (n=5) there is insufficient statistical power to examine whether there is any difference in reproductive performance between black genotypes. A survival advantage of black heterozygotes over the other two genotypes across all years has previously been reported (*16*), but we do not know whether coat color and CDV infection interact to influence survival and therefore the relative fitness of the three genotypes.

We used longitudinal data to explore how annual age-specific survival rates varied between 1998 and 2020 between homozygote black, heterozygote black, and homozygote gray wolves with individual exposure to CDV during five CDV outbreaks. We developed a mark-recapture model that included transitions between susceptible, exposed, and immune states (*24*). We also included information on permanent dispersal and non-natural and known natural deaths while simultaneously modeling recapture rates. Pack identity and year were included as random effects.

Mirroring the results of our broad-scale surveys, our analyses reveal that black heterozygote wolves have higher survival compared to gray wolves, but only in CDV-infected individuals (Fig. 3). Turning to reproductive success, gray wolves have 25% higher reproductive success in all years, with average rates of reproduction suppressed for all females in years with CDV outbreaks (*26*). Because inheritance at the K locus is Mendelian, if the survival advantage to the heterozygote exposed to CDV compensates for the reduced fertility of black females it may be advantageous to mate with a partner of the opposite color to maximize the likelihood of producing heterozygote offspring when epizootics are frequent. We consequently hypothesize that fluctuating frequency-dependent selection due to CDV outbreak frequency can alter the relative fitness of the genotypes, resulting in a fitness advantage to heterozygotes when disease is frequent enough, selecting for the disassortative mating strategy observed in Yellowstone (*27*).

We constructed a stochastic, demographic, two-sex model of the dynamics of the three genotypes. We use a mating function that allows us to alter the mating preference from random, through disassortative to assortative and evaluate which mate choice strategy was optimal under various disease outbreak frequencies (*24*).

When the model is parameterized with initial starting conditions equal to the observed coat color frequencies at reintroduction into Yellowstone and observed CDV outbreaks, the simulations capture the observed dynamics of coat color frequency adequately when we assumed random or disassortative mating (Fig. 2). The model did not perform well with assortative mating, in accordance with the excess of black-gray pairs reported in Yellowstone (*27*). Consequently, despite its simplicity, our model captures the dynamics of wolf coat color genotypes in Yellowstone.

Our model predicted that the frequency of black wolves depended upon the frequency of CDV outbreaks and mate choice strategy (Fig. 4(B)). Under all mate choice strategies, the frequency of black wolves increased with the frequency of CDV outbreaks. The rate of increase was steepest when wolves mated assortatively, and shallowest when they mated disassortatively. In Yellowstone wolves mate disassortatively, but is this adaptive?

Disease-induced mortality selects for the evolution of mate choice, but the evolutionarily stable strategy (ESS) changes on either side of a threshold in disease frequency. Below an outbreak frequency of 0.2 (≈ 1 outbreak every 5 years), an assortative mating strategy was the ESS (Fig. 4(B)), while above it, a disassortative mating strategy has greater fitness. Random mating was never the ESS. The black allele is always eliminated in the absence of CDV when the ESS was assortative mating, while disassortative mating results in a stable polymorphism.

Our modelling results are consistent with our hypothesis that the frequency of disease outbreaks is responsible for the observed cline in coat color seen across North America and explains why Yellowstone wolves mate disassortatively. We would expect an assortative mating strategy when CDV outbreaks occur less than once every decade. Although our results are consistent with observation, recent lab experiments challenging wolf cell cultures with a range of pathogens have so far failed to discern genotype-specific responses to CDV (*28*). Although elegant, this work cannot answer the question we have addressed since it fails to capture susceptibility to infection and the complexity of immune responses expected within free-living individuals (*28*). In addition, genetic findings have reported positive selection on coat color genes and MHC and immunity genes along a gradient of temperature and humidity (*11*), a finding that is consistent with our conclusions.

Our results are unlikely to be specific to wolves. In many insects, amphibians, reptiles, birds and non-human mammals, disease resistance is associated with coloration (*2*,*4*,*16*,*29*), a trait that can act as a signal for pathogen resistance in mate choice (*1*,*3*). Recent findings have identified associations between disease-resistance MHC genes and coloration in mammals (*14*,*30*), amphibians (*31*), reptiles (*32*), and birds (*33*), possibly via pleiotropic effects or the action of 'supergenes'(*2*,*34*). As an example, in some bird species, carotenoid-dependent coloration (*29*) can drive mate choice via associations with disease-resistant MHC genes that influence sensory functions of odor, vision and hearing (*30*).

When coloration is genetically determined and disease resistance is heritable and associated with coloration, a preference for a mate of a specific color will enhance fitness by maximizing the chances of producing resistant offspring in environments with frequent and virulent enough pathogens. When the environment varies spatially, alternative mating strategies could explain the maintenance of color polymorphism (*34*), through negative-frequency dependent selection, as shown here for wolves. Incidental color biproducts of immune response genes may consequently be widespread drivers of mating behaviors observed across a diverse array of animal species. It is possible we have significantly underestimated the role of pathogens in generating the diversify of morphological and behavioral traits observed in nature (*31*,*35*).

CDV requires a high population density to persist, and because wolves live at low densities, CDV cannot be endemic within a population (*22*). Instead, it requires a broad community of carnivores to persist and intermittent spillover transmission back to wolves. The reservoir community, species, or population is not well understood for CDV in North America, but we show that CDV prevalence is positively associated with human density (Fig. S4). CDV evolved from human measles epidemics that decimated indigenous South American populations and

the virus spilled over into the abundant dog population and evolved into CDV before invading North America in the 1760s (*36*). Wolves have genes for black coat color as an indirect consequence of Europeans invading the Americas, and the introgression happened between 1598 and 7248 years ago. It is therefore likely that other pathogens or mechanisms have contributed to the rapid spread of the black allele.

None of our analyses on their own provides conclusive support for the hypothesis that the frequency of black wolves across North America is determined by the frequency of CDV outbreaks but each separate, complementary line of evidence provides support. These results are important because: 1) they reveal how the frequency of disease outbreaks imposes selection on immune function generating heterozygote advantage only under certain environments similar to that seen in sickle cell disease in humans. In the absence of CDV in the environment, the dominant *K* allele is expected to be lost as an assortative mating strategy would be selected, with the frequency of CDV outbreaks determining the frequency of the derived *k* allele. We provide support that variation in CDV outbreak frequency has generated the cline in wolf coat color observed across North America. An incidental biproduct of genetic variation at the *K* locus is coat color variation, a striking phenotypic pattern that has long puzzled researchers. 2) sexual selection has operated on this incidental cue to sculpt wolf behavior, with Yellowstone wolves mating disassortatively to maximize their fitness. This shows how the effects of a pathogen influence not only selection for resistance, but how this is signaled between hosts and thus the mate choice behavior mechanism that results in host diversity. 3) the study shows the true value of coupling geographically restricted intensive long-term individual-based studies of wild populations with continent-wide cross-sectional samples from multiple populations. We are able to link statistical results across these disparate forms of data using evolutionarily explicit population modelling. In doing this, we learn that the maintenance of genetic, morphological, and behavioral variation both within and between populations of a charismatic carnivore is a result of environmentally determined fluctuating frequency-dependent selection.

References and Notes

1. A. Roulin, The evolution, maintenance and adaptive function of genetic colour polymorphism in birds. Biological Reviews 79, 815-848 (2004).
2. A. Orteu, C D. Jiggins. The genomics of coloration provides insights into adaptive evolution. Nature Reviews Genetics 8, 461-75 (2020).
3. M. Andersson, L. W. Simmons, Sexual selection and mate choice. Trends in Ecology & Evolution 21, 296-302 (2006).
4. J. Côte et al. Melanin-based coloration and host–parasite interactions under global change. Proceedings of the Royal Society B: Biological Sciences 285, 20180285(2018)
5. D. M. Hoppe. The influence of color on behavioral thermoregulation and hydroregulation. The behavioral significance of color 35-74 (2018).
6 E. S. Briolat et al. Diversity in warning coloration: selective paradox or the norm?. Biological Reviews 94, 388-414 (2019).
7. M. Milinski, T. C. Bakker, Female sticklebacks use male coloration in mate choice and hence avoid parasitized males. Nature 344, 330-333 (1990).

8. S. Antoniazza et al., Local adaptation maintains clinal variation in melanin-based coloration of European barn owls (Tyto alba). Evolution 64, 1944-1954 (2010).

9. A. D. Potash, et al. Ecological drivers of eastern fox squirrel pelage polymorphism. Frontiers in Ecology and Evolution, 8, 119 (2020).

11. C. Küpper, et al. A supergene determines highly divergent male reproductive morphs in the ruff. Nature Genetics. 48, 79–83 (2015).

12. P. S. Gipson et al., Color patterns among wolves in western North America. Wildlife Society Bulletin 30, 821-830 (2002).

13. T. M. Anderson et al., Molecular and evolutionary history of melanism in North American gray wolves. Science 323, 1339-1343 (2009).

14. M. Musiani et al., Differentiation of tundra/taiga and boreal coniferous forest wolves: genetics, coat colour and association with migratory caribou. Molecular Ecology 16, 4149-4170 (2007).

15. Z. Fan, D. Ortega-Del Vecchyo, R. K. Wayne, EPAS1 variants in high altitude Tibetan wolves were selectively introgressed into highland dogs. PeerJ 5, e3522 (2017).

16. S. I. Candille et al., A β-defensin mutation causes black coat color in domestic dogs. Science 318, 1418-1423 (2007).

17. R. M. Schweizer et al., Natural selection and origin of a melanistic allele in North American gray wolves. Molecular Biology and Evolution 35, 1190-1209 (2018).

18. T. Coulson et al., Modeling effects of environmental change on wolf population dynamics, trait evolution, and life history. Science 334, 1275-1278 (2011).

19. A. L. Ducrest et al. Pleiotropy in the melanocortin system, coloration and behavioural syndromes. Trends in ecology & evolution 23, 502-510 (2008).

20. D. Yang et al., β-defensins: linking innate and adaptive immunity through dendritic and T cell CCR6. Science 286, 525-528 (1999).

21. S. L. Deem, L. H. Spelman, R. A. Yates, R. J. Montali, Canine distemper in terrestrial carnivores: a review. Journal of Zoo and Wildlife Medicine 31, 441-451 (2000).

22. E. S. Almberg, L. D. Mech, D. W. Smith, J. W. Sheldon, R. L. Crabtree, A serological survey of infectious disease in Yellowstone National Park's canid community. PloS One 4, e7042 (2009).

25. M. Aidoo et al., Protective effects of the sickle cell gene against malaria morbidity and mortality. The Lancet 359, 1311-1312 (2002).

26. M. Fu et al. Major histocompatibility complex variation and the evolution of resistance to amphibian chytridiomycosis. Immunogenetics 69, 529-536 (2017).

27. E. S. Almberg, P. C. Cross, D. W. Smith, Persistence of canine distemper virus in the Greater Yellowstone Ecosystem's carnivore community. Ecological Applications 20, 2058-2074 (2010).

28. H. Hillebrand, On the generality of the latitudinal diversity gradient. The American Naturalist 163, 192-211 (2004).

29. D. W. Smith, D. R. Stahler, D. R. MacNulty, Yellowstone Wolves: Science & Discovery in the World's First National Park. (University of Chicago Press, Chicago, 2021).

30. D. R. Stahler et al. The adaptive value of morphological, behavioural and life-history traits in reproductive female wolves. Journal of Animal Ecology 82, 222-234 (2013).

31. P. W. Hedrick et al. Negative-assortative mating for color in wolves. Evolution, 70, 757-466 (2016).

32. E. E. Brandell et al. A metapopulation model of social group dynamics and disease applied to Yellowstone wolves. . Proceedings of the National Academy of Science 118, e2020023118 (2021).

33. T. K. Fuller, L. D. Mech, J. F. Cochrane, in Wolves: Behavior, Ecology, and Conservation, L. D. Mech, L. Boitani, Eds. (Chicago, University of Chicago Press, 2003), pp. 161–191.

34. S. Tuljapurkar, S. H. Orzack, Population dynamics in variable environments I. Long-run growth rates and extinction. Theoretical Population Biology 18, 314-342 (1980).

35. R. A. Johnston et al., K Locus effects in Gray wolves: experimental assessment of TLR3 Signaling and the gene expression response to canine distemper virus. Journal of Heredity, 112, 458-468 (2021).

36. R. M. Schweizer et al. Genetic subdivision and candidate genes under selection in North American grey wolves. Molecular ecology 25, 380-402 (2016).

37. A. Roulin, A.-L. Ducrest, Association between melanism, physiology and behaviour: a role for the melanocortin system. European Journal of Pharmacology 660, 226-233 (2011).

38. M. J. P. Simons et al. What does carotenoid-dependent coloration tell? Plasma carotenoid level signals immunocompetence and oxidative stress state in birds–a meta-analysis. PlosOne e43088 (2012).

39. P.S. Santos et al. The best smellers make the best choosers: mate choice is affected by female chemosensory receptor gene diversity in a mammal. Proceedings of the Royal Society B, 285, 2018-2426 (2018).

40. A. L. Trujillo et al. Spatiotemporal adaptive evolution of an MHC immune gene in a frog-fungus disease system. Heredity, 126, 640-655 (2021).

41. J.D. Hacking et al. Specific MHC class I supertype associated with parasite infection and color morph in a wild lizard population. Ecology and Evolution, 8, 9920-9933 (2018).

42. H. Y. Liu et al. Cape Feather Coloration Signals Different Genotypes of the Most Polymorphic MHC Locus in Male Golden Pheasants (Chrysolophus pictus). Animals, 11, 276 (2021).

43. L. M. San-Jose, A. Roulin. Toward understanding the repeated occurrence of associations between melanin-based coloration and multiple phenotypes. The American Naturalist, 192, 111-130 (2018).

44. C. Küpper et al. A supergene determines highly divergent male reproductive morphs in the ruff. Nature genetics, 48, 79-83 (2016).

45. M. E. LaCava et al. Spatio-temporal analyses reveal infectious disease-driven selection in a free-ranging ungulate. Royal Society Open Science, 8, 210802 (2021).

46. E. W. Uhl et al., New world origin of canine distemper: Interdisciplinary insights. International Journal of Paleopathology 24, 266-278 (2019).

47. E.E. Brandell, et al. Patterns and processes of pathogen exposure in gray wolves across North America. Scientific reports, 11,1-14 (2021).

48. J. A. Kerns et al., Linkage and segregation analysis of black and brindle coat color in domestic dogs. Genetics 176, 1679-1689 (2007).

49. K. Erles, J. Brownlie, Expression of β-defensins in the canine respiratory tract and antimicrobial activity against Bordetella bronchiseptica. Veterinary Immunology and Immunopathology 135, 12-19 (2010).

50. B. C. Leonard et al., Activity, expression and genetic variation of canine β-defensin 103: a multifunctional antimicrobial peptide in the skin of domestic dogs. Journal of Innate Immunity 4, 248-259 (2012).

51. M. Afshar, R. L. Gallo, Innate immune defense system of the skin. Veterinary Dermatology 24, 32-e39 (2013).

52. D. Santoro, C. W. Maddox, Canine antimicrobial peptides are effective against resistant bacteria and yeasts. Veterinary dermatology 25, 35-e12 (2014).

53. C. M. van Damme, T. Willemse, A. van Dijk, H. P. Haagsman, E. J. Veldhuizen, Altered cutaneous expression of β-defensins in dogs with atopic dermatitis. Molecular Immunology 46, 2449-2455 (2009).

54. A.-L. Ducrest, L. Keller, A. Roulin, Pleiotropy in the melanocortin system, coloration and behavioural syndromes. Trends in Ecology & Evolution 23, 502-510 (2008).

55. G. Emaresi et al., Pleiotropy in the melanocortin system: expression levels of this system are associated with melanogenesis and pigmentation in the tawny owl (Strix aluco). Molecular Ecology 22, 4915-4930 (2013).

56. A. Bennett, V. Hayssen, Measuring cortisol in hair and saliva from dogs: coat color and pigment differences. Domestic Animal Endocrinology 39, 171-180 (2010).

57. B. Almasi, L. Jenni, S. Jenni-Eiermann, A. Roulin, Regulation of stress response is heritable and functionally linked to melanin-based coloration. Journal of Evolutionary Biology 23, 987-996 (2010).

58. A. Roulin, B. Almasi, K. Meichtry-Stier, L. Jenni, Eumelanin-and pheomelanin-based colour advertise resistance to oxidative stress in opposite ways. Journal of Evolutionary Biology 24, 2241-2247 (2011).

59. P. Galeotti, R. Sacchi, Differential parasitaemia in the tawny owl (Strix aluco): effects of colour morph and habitat. Journal of Zoology 261, 91-99 (2003).

60. J. Gasparini et al., Strength and cost of an induced immune response are associated with a heritable melanin-based colour trait in female tawny owls. Journal of Animal Ecology 78, 608-616 (2009).

61. L. Gangoso et al., MC1R-dependent, melanin-based colour polymorphism is associated with cell-mediated response in the Eleonora's falcon. Journal of Evolutionary Biology 24, 2055-2063 (2011).

62. A. H. Harcourt, White wolves can stand the heat. Science 324, 1141-1143 (2009).

63. M. D. Jimenez et al., Wolf dispersal in the rocky mountains, Western United States: 1993–2008. The Journal of Wildlife Management 81, 581-592 (2017).

64. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 6, 1-48 (2015).

65. M. Plummer, Penalized loss functions for Bayesian model comparison. Biostatistics 9, 523-539 (2008).

66. A. Gelman, D. B. Rubin, Inference from iterative simulation using multiple sequences. Statistical science 7, 457-472 (1992).

67. D. J. Spiegelhalter, N. G. Best, B. P. Carlin, A. Van Der Linde, Bayesian measures of model complexity and fit. Journal of the royal statistical society: Series b (statistical methodology) 64, 583-639 (2002).

68. B. M. Vonholdt et al., The genealogy and genetic viability of reintroduced Yellowstone grey wolves. Molecular ecology 17, 252-274 (2008).

69. E. Afgan et al., The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic acids research 44, W3-W10 (2016).

70. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997, (2013).

71. Broad Institute. (2019) GitHub Repository: http://broadinstitute.github.io/picard/.

72. H. Li et al., The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078-2079 (2009).

73. A. McKenna et al., The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research 20, 1297-1303 (2010).

74. R. Poplin et al., Scaling accurate genetic variant discovery to tens of thousands of samples. BioRxiv, 201178 (2018).

75. P. C. Cross et al., Estimating distemper virus dynamics among wolves and grizzly bears using serology and Bayesian state-space models. Ecology and evolution 8, 8726-8735 (2018).

76. E. S. Almberg, P. C. Cross, A. P. Dobson, D. W. Smith, P. J. Hudson, Parasite invasion following host reintroduction: a case study of Yellowstone's wolves. Philosophical Transactions of the Royal Society B: Biological Sciences 367, 2840-2851 (2012).

77. R. Schultz, B. Thiel, E. Mukhtar, P. Sharp, L. Larson, Age and long-term protective immunity in dogs and cats. Journal of comparative pathology 142, S102-S108 (2010).

78. R. King, A review of Bayesian state-space modelling of capture–recapture–recovery data. Interface Focus 2, 190-204 (2012).

79. S. Cubaynes et al., Density-dependent intraspecific aggression regulates survival in northern Yellowstone wolves (Canis lupus). Journal of Animal Ecology 83, 1344-1356 (2014).

80. J.-D. Lebreton, K. P. Burnham, J. Clobert, D. R. Anderson, Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecological monographs 62, 67-118 (1992).

81. O. Gimenez, J.-D. Lebreton, R. Choquet, R. Pradel, R2ucare: An R package to perform goodness-of-fit tests for capture-recapture models. bioRxiv, 192468 (2017).

82. O. Devineau et al., Evaluating the Canada lynx reintroduction programme in Colorado: patterns in mortality. Journal of applied Ecology 47, 524-531 (2010).

83. R. Pradel, J. E. Hines, J.-D. Lebreton, J. D. Nichols, Capture-recapture survival models taking account of transients. Biometrics, 60-72 (1997).

84. M. Genovart, R. Pradel, D. Oro, Exploiting uncertain ecological fieldwork data with multi-event capture–recapture modelling: an example with bird sex assignment. Journal of animal ecology 81, 970-977 (2012).

85. Y.S., Su, Yajima, M., Su, M.Y.S. and SystemRequirements, J.A.G.S., 2015. Package 'R2jags'. R Package Version 0.03-08.

86. D. R. MacNulty, D. W. Smith, L. D. Mech, L. E. Eberly, Body size and predatory performance in wolves: is bigger better? Journal of Animal Ecology 78, 532-539 (2009).

87. K. A. Cassidy, D. R. MacNulty, D. R. Stahler, D. W. Smith, L. D. Mech, Group composition effects on aggressive interpack interactions of gray wolves in Yellowstone National Park. Behavioral Ecology 26, 1352-1360 (2015).

88. L. D. Mech, S. M. Barber-Meyer, J. Erb, Wolf (Canis lupus) generation time and proportion of current breeding females by age. PloS one 11, e0156682 (2016).

Acknowledgments: SC is supported by a grant from the Agence Nationale de la Recherche (ANR-18-CE02-0011, MathKinD). TC and SC acknowledge the Natural Environment Research Council for funding. D.R.M., R.K.W., D.R.S. and D.W.S. thank NSF for support (separate grants to R.K.W. and D.W.S., D.R.S., and D.R.M.). EEB and PJH are grateful to the Verne Willaman Professorship that supported aspects of this work. We thank Wei Wang, Lance Parsons, Alexandra DeCandia, and Christopher Lawrence for assistance with developing the K locus amplicon-sequencing assay and genotyping pipeline, and Vincenzo Gervasi and Romain Villoutreix for comments on a previous version of the manuscript.

Author contributions: DWS and DRS run the Yellowstone wolf project and along with DRM and EEB they conducted or oversaw all data collection and developed the ideas with SC, TC, ESA, EEB, PC, PH, APD, and RKW. BV conducted genotyping. ESA, PC and EEB collected and analysed the serological data and provided expertise on disease dynamics along with APD and PH. SC, EEB and PC conducted the statistical analyses and modelling, with expertise from TC and SS. TC and SC wrote the paper with input from all coauthors. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Competing interests: Authors declare that they have no competing interests.

Data and materials availability: Data used in this paper can be downloaded from Dryad (add DOI) and code used for the statistical analysis and population modelling are available at Zenodo (add DOI). [To be added following peer-review]

Figures

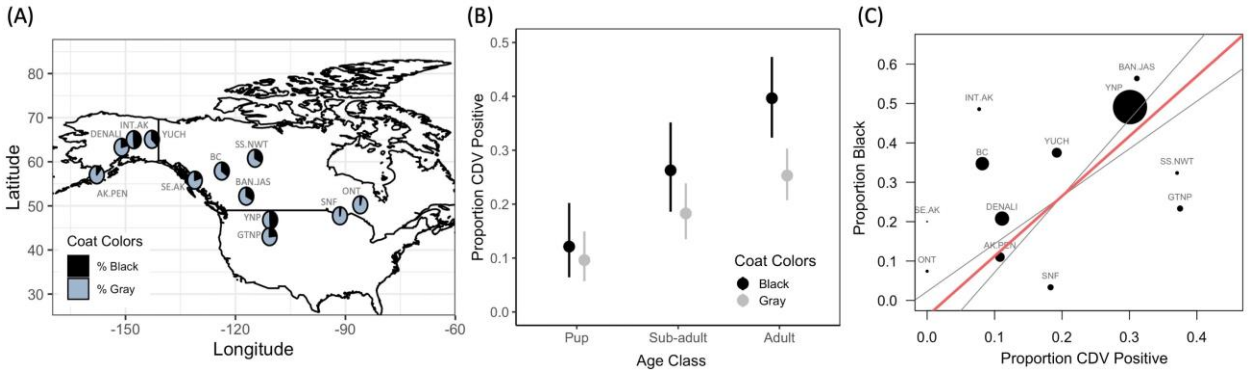

Fig. 1. CDV and coat color occurrence across North America. (A) Wolf sampling locations and proportion of each coat color phenotype (see 24 for location codes and methods). YNP and GTNP are offset for visual purposes. (B) The proportion of pup, sub-adult and adult wolves seropositive for distemper among n = 1134 from the wolf populations sampled in (A) with 95% confidence intervals. (C) Relationship between CDV prevalence and proportion black coat color. Red line is a restricted major axis regression weighted by sample size with grey lines showing 95% bounds of the regression estimate. Circles scaled to sample size (n = 1166).

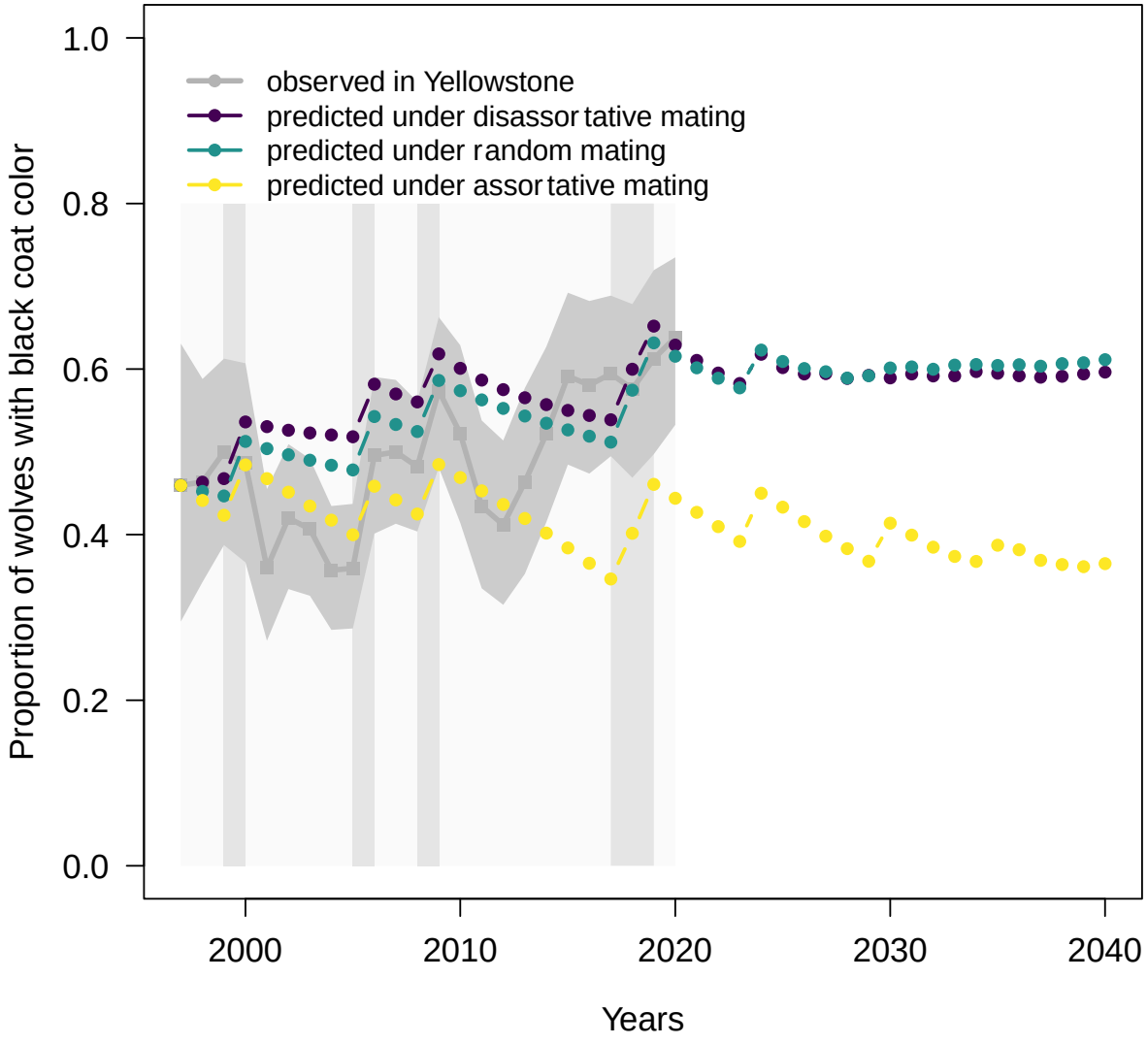

Fig. 2. Observations and predictions of the frequency of black wolves in Yellowstone. Dark gray line are point estimates with the shaded area representing 95% CIs. Vertical gray lines represent past CDV outbreaks. Up to 2020, colored lines and dots represent a single model projection accounting for past CDV outbreaks for random (green), disassortative (purple) and assortative (yellow) mating. After 2020, colored shaded areas represent 95% CI estimated from 500 model runs assuming the same frequency of CDV outbreaks (annual probability of 0.2).

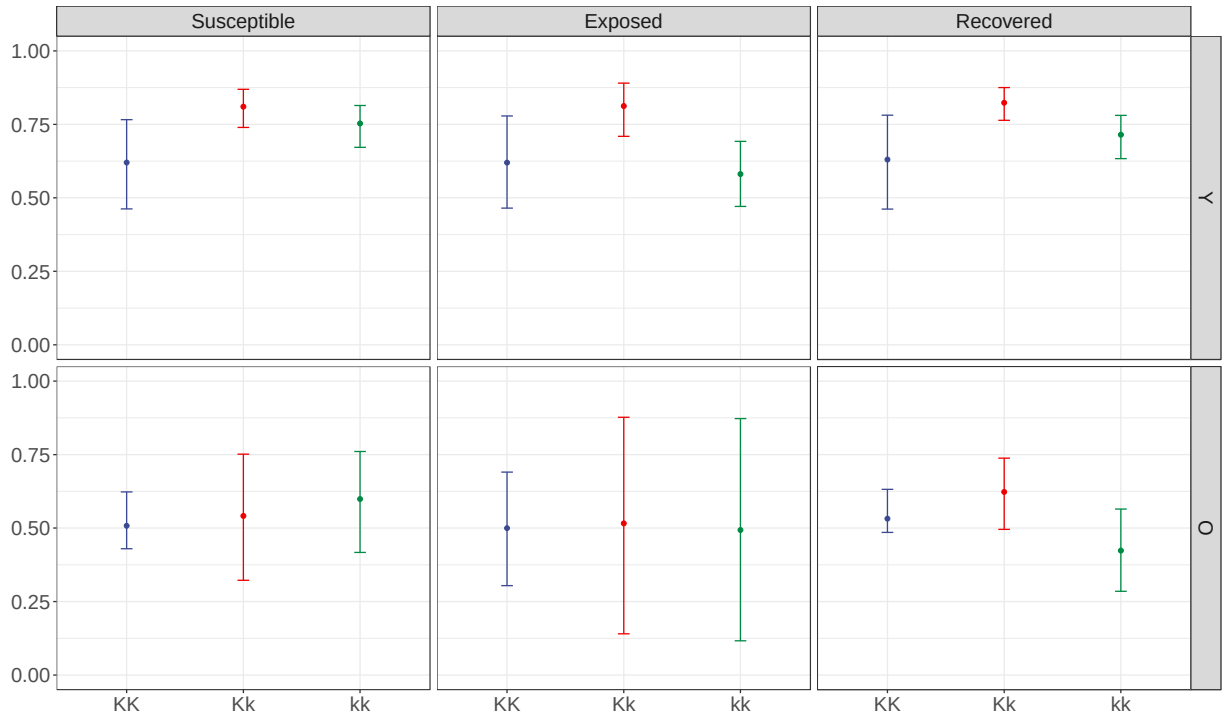

Fig. 3. Survival analysis results. (**A**) Effects of age (top Young, bottom Old), disease status (columns) and K-locus genotype (blue for homozygote black, red for heterozygote black and green for homozygote grey) on survival rates. Medians with 80% credible intervals are displayed.

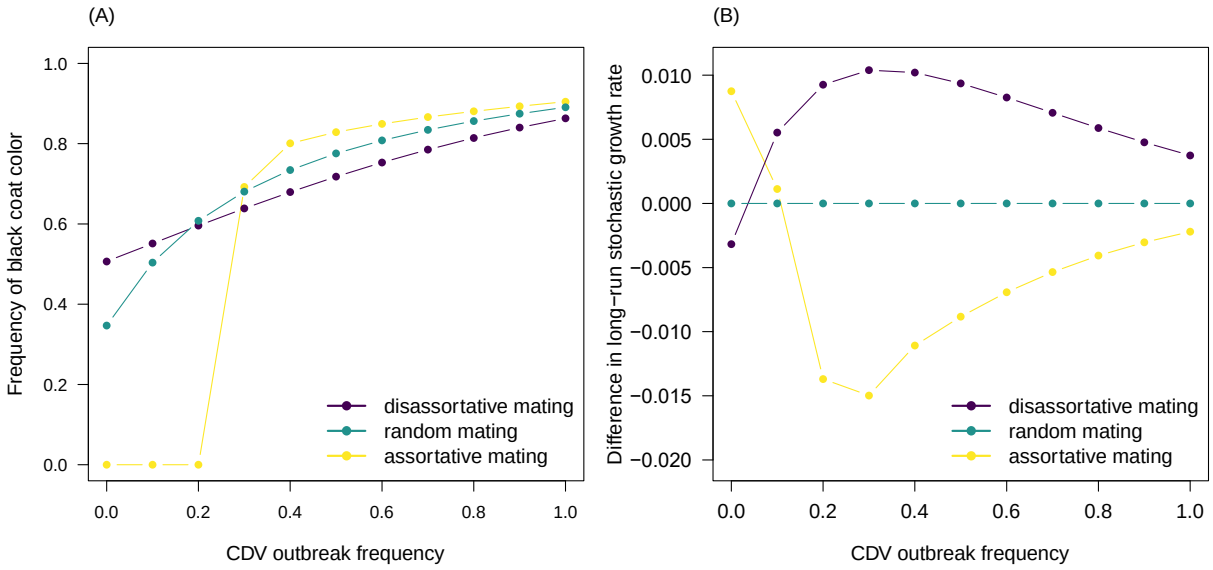

Fig. 4. Model predictions as CDV outbreak frequency varies. (**A**) Effect of CDV outbreak frequency and mating system on the frequency of black wolves, (**B**) difference in strategy fitness relative to random mating (green) for assortative (yellow) and disassortative (purple) mating strategies as a function of CDV outbreak frequency. Lines represent point estimates, shaded polygons represent 95% CIs from 500 simulations.