Genome Evolution in the Nematode *Pristionchus pacificus*

D i s s e r t a t i o n

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Andreas M. Weller
aus Stuttgart

Tübingen
2013

Tag der mündlichen Qualifikation:	20.12.2013
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Ralf J. Sommer
2. Berichterstatter:	Prof. Dr. Boris Macek

Contents

Summary

Evolution is the ultimate reason why planet earth is full of complex living systems. Evolutionary theory is thus at the heart of biology and necessary to understand our world. A complete picture of evolutionary processes must incorporate multiple levels of organisations and timescales. The nematode *Pristionchus pacificus* is a suitable model organism that allows such interdisciplinary studies with the long-term goal of reconstructing the evolutionary history of the species. In my work, I focused on two aspects of *P. pacificus* evolution: organismal interactions and genome evolution.

Selection pressures in nature often derive from interactions between organisms. *P. pacificus* lives in a necromenic association with scarab beetles. We extracted nematodes from 114 *Geotrupes* beetles and found diverse communties. The only life stage of *P. pacificus* present was the Dauer stage. Beetle-associated bacteria are a source of food as well as a potential pathogen to *P. pacificus*. We isolated 23 bacterial strains from beetle-associated nematodes and studied their interaction in laboratory assays. Bacterial pathogenicity is species-specific and nematodes are able to sense and avoid harmful bacterial strains.

Genome evolution is shaped both by natural selection and a multitude of stochastic processes. First, we studied natural variants in 104 strains of *P. pacificus* and observed strong linkage blocks. These blocks allow deleterious alleles to avoid natural selection and thus increase in frequency stronger than expected. Second, we analyzed the properties of 802 *de novo* mutations derived from 22 Mutation Accumulation Lines. Mutations occur randomly across the genome, but their spectrum is strongly biased towards creation of AT nucleotides. In natural variants, this ATbias is counterbalanced over time and the GC content of the *P. pacificus* genome is stable. Third, we could observe recombination in hybrid lines of *P. pacificus* and show that there is intraspecific variation in recombination rates. In addition, recombination is positively correlated to natural variants, probably because it promotes natural selection by reshuffling alleles. Finally, an analysis of transmission rate distortion revealed strong hybrid incompatibilities even between closely related strains of *P. pacificus*.

Zusammenfassung

Alle lebenden Organismen auf der Erde sind ein Produkt der Evolution. Deshalb ist die Evolutionstheorie von zentraler Bedeutung in der Biologie und essentiell für das Verständnis unserer Welt. Ein vollständiges Bild von Evolutionsprozessen muss verschiedene Größenordungen und Zeitskalen mit einbeziehen. Der Fadenwurm *Pristionchus pacificus* als Modellsystem erlaubt solche interdisziplinären Ansätze, wobei das langfristige Ziel in der kompletten Rekonstruktion der Evolutionsgeschichte der Art liegt. In meiner Arbeit habe ich mich auf zwei Aspekte der Evolution von *P. pacificus* konzentriert: das Verhältnis zu anderen Organismen sowie Genomevolution.

Selektionsdruck in der Nature entsteht oft durch Interaktionen zwischen Organismen. *P. pacificus* lebt in einer nekromenischen Assoziation mit Blatthornkäfern. Wir extrahierten alle Fadenwürmer von 114 Mistkäfern der Gattung *Geotrupes*. Diese Käfer werden von Fadenwürmer verschiedenster Arten befallen, wobei das einzige Befallsstadium von *P. pacificus* die Dauerlarve ist. Auf den Käfern lebende Bakterien stellen für *P. pacificus* sowohl Nahrung als auch ein potentielles Krankheitsrisiko dar. Wir isolierten 23 Bakterienstämme aus Fadenwürmer von Käfern und untersuchten ihre Interaktionen unter Laborbedingungen. Die Schädlichkeit von Bakterien hängt von der Fadenwurmart ab. Die Fadenwürmer sind in der Lage schädliche Bakterienstämme zu erkennen und zu vermeiden.

Genomevolution hängt sowohl von natürlicher Auslese als auch von verschiedenen Zufallsprozessen ab. Zuerst untersuchten wir Allelvarianten in 104 Stämmen von *P. pacificus* und stellten starke Genkopplungsblocks fest. Diese Blocks erlauben nachteiligen Allelen der natürliche Auslese zu entgehen und sich dadurch stärker als erwartet auszubreiten. Als nächstes analysierten wir Eigenschaften von 802 Mutationen aus 22 ingezüchteten Mutationslinien von *P. pacificus*. Mutationen erfolgen zufällig über das Genom verteilt, aber die Art der Mutation zeigt einen starken Hang zur Erzeugung von AT Nukleotiden. In Allelvarianten in der Natur wird diese Ungleichheit allmählich ausgeglichen, so dass der GC Gehalt des *P. pacificus* langfristig stabil bleibt. Drittens konnten wir Rekombinationsprozesse in Hybridlinien direkt beobachten und innerartliche Variation feststellen. Rekombinationsprozesse sind außerdem positiv korreliert mit Allelvarianten, wahrscheinlich weil Rekombination die natürliche Auslese durch Vermischen von Allelen verstärkt. Zum Schluss untersuchten wir die Vererbungsrate von elterlichen Allelen in nahe verwandten Stämmen von *P. pacificus* und stellten starke Fälle von Hybridinkompatibilität fest.

Publications

Teile der Arbeit wurden bereits veröffentlicht oder sind zur Veröffentlichung eingereicht:

1. Christian Rödelsperger, Richard A. Neher, **Andreas M. Weller**, Gabi Eberhardt, Hanh Witte, Werner Mayer, Christoph Dieterich and Ralf J. Sommer:
   Weakly deleterious mutations dominate genetic diversity in *Pristionchus pacificus*
   Genetics, under revision.

2. **Andreas M. Weller**, Christian Rödelsperger, Gabi Eberhardt, Ruxandra I. Molnar and Ralf J. Sommer:
   Mutation accumulation lines and natural variants in the nematode *Pristionchus pacificus* indicate opposing forces of de novo mutations and counterbalancing selection
   Genetics, accepted with minor revision.

3. **Andreas M. Weller**, Werner E. Mayer, Robbie Rae and Ralf J. Sommer:
   Quantitative Assessment of the Nematode Fauna Present on *Geotrupes* Dung Beetles Reveals Species-Rich Communities with a Heterogeneous Distribution
   Journal of Parasitology 2010, 96(3):525-531.

4. Robbie Rae, Metta Riebesell, Iris Dinkelacker, Qiong Wang,
   Andreas M. Weller, Christoph Dieterich and Ralf J. Sommer:
   Isolation of naturally associated bacteria of necromenic *Pristionchus* nematodes and

fitness consequences

Journal of Experimental Biology 2008, 211(12):1927–1936.

Acknowledgments

The **Max Planck Society** funded this work and allowed me access to the Department of Evolutionary Biology at the Max Planck Institute for Developmental Biology in Tübingen, Germany. Working with the international colleagues and first-class equipment of hte Max Planck Society has been a joy and a privilege, but has probably also spoilt all future labs for me. Thank you, german taxpayer!

I am grateful to my supervisor **Ralf J. Sommer** for the constant support over the last 10 years, and especially for believing my outreagous claims to be suited for yet another area of biology three different times over these years. My PhD could have been achieved easier and quicker by sticking to one field, but the path I took broadened my horizon immensely.

Each new field couldn't have been mastered without a personal mentor that invested countless hours of precious one-on-one instructions:

- **Robbie** thaught me about the noisy world of ecology, how to extract funny things from beetles and not get yourself killed by unknown microbes in the S2 lab. I hope my transition to the UK will be eased by his intercultural advice: 'They're all weirdos on the island' and 'Fail to prepare, prepare to fail!'.
- **Ruxandra** was very patient in introducing me to the sometimes painstakingly precise work in a molecular biology lab, even when yet another PCR failed because I picked the wrong temperature or treated the DNA too roughly. She also challenged my typically ignorant Western European attitude of Eastern Europe as a homogenous mass of unfriendly people. Thanks for that!

- **Christian** helped me transition from wielding pipettes to writing code, a world that I ultimately enjoyed much more than the wetlab work. I'm happy about my progress from simple bash commands to writing multi-module Python interfaces, and part of it is due to Christians seemingly unchanging attitude of patience and a Sphinx-like enigmatic smile.

Working with **Fabio** and **Richard** was another enlightening culture shock. The world of mathematics is often scary to a biologist, but also impressively effective in solving problems and actually quite fun once you understand what it's about.

Amit, **Simone** and **Nadine** were both pleasant labmates and part-time mentors in bioinformatics or molecular biology. Thanks!

The TAs are often overlooked in the accomplishments achieved, but the lab probably wouldn't run a day without them. So thanks to **Iris**, **Hanh**, **Gabi**, **Metta** and **Heike**!

I'm also grateful to **Adrian** and **Dan** for their sometimes tough but always friendly and contructive criticism during lab meetings. After presenting my work to this audience for years, a job interview is nothing to be scared of anymore.

A general acknowledgement to **every lab member** whom I didn't mention in person, but who provided a positive working atmosphere and answered many big and small questions over the years.

My friendship to the mexican duo **Jose Arcadio Farias Rico** and **Toledo Saacnicteh** was a highlight of my PhD and provided a much-needed respite from thinking about my work during lunch breaks. I can't remember a single boring conversation over the years, and I learned a lot outside of my discipline about the arcane field of protein design (which feels more like engineering than biology), experimental music or cross-cultural sociology.

I work to live, not live to work, so my energy and good mood in the lab derive from a fullfilling life outside of work. I'm grateful for **Dr. Suska Weller** for being my wife and enduring both my good and bad moods, my absent-mindedness at times and my ability to focus on the good things in life at the expense of mundane duties.

The same is true for **my parents**, all 4 of them. Without your encouragement or at least tolerance for my endless curiosity for weird animals I wouldn't have managed to grow from a small kid breeding spiders to a big kid with a PhD in breeding nematodes.

Aim of the Thesis

This page provides a general overview of the research questions behind my PhD thesis. The goals of the indivual projects are detailed at the end of each introductory chapter.

The articles in this thesis originate from two different fields of biology, namely ecology and genomics, which reflects the development of my own research interest. In a way, this personal change of perspective is a reflection of the changes in biology as a research discipline. The last 5 years since the broad availability of massively parallel next generation sequencing (NGS) were a disruptive time for all fields of biology and even created completely new fields like genomics. When I started working on my PhD in September 2007, the technology used in my later projects on genome evolution was not invented yet.

The overarching goal of my thesis is to study evolutionary change using the model nematode *Pristionchus pacificus*. A thorough understanding of evolution needs to consider different levels of organisation and size: from the macroscopic interactions of predator and prey to the biochemical properties of DNA nucleotides, and from differences between individuals to the level of populations and species. My first project on nematode communities on beetles focused on the macroscopic: which biotic factors rule the dispersal of *Pristionchus* around the world? And what is the environment that the Dauer larvae evolved in? In my second and third projects, I tried to understand the non-adaptive processes that govern genome evolution: according to which rules do mutations introduce novel variants into the gene pool of *P. pacificus*? And how do these variants become rearranged, selected for and eventually fixed over time?

Part I

Background

Chapter 1

Pristionchus pacificus: a well-rounded nematode

1.1 Evo-devo Studies

Pristionchus pacificus is not the first nematode used as a model organism. In the 1960s, Sydney Brenner introduced *Caenorhabditis elegans* to study development, in particular neural development. The fixed cell lineage and translucent body combined with easy culturing and short life cycle soon made *C. elegans* famous as an ideal laboratory animal. Since then, all kinds of processes from developmental biology to behaviour have been studied in this nematode, resulting in three Nobel prizes to date.

Even though the 'model organism approach' was very successful, drawing conclusions from a single species to a whole phylum and even beyond is problematic [Hong and Sommer 2006]. To allow a second point of comparison, *Pristionchus pacificus* was introduced as a satellite model to *C. elegans* in 1996 [Sommer et al. 1996b]. Just like *C. elegans*, *P. pacificus* is a self-fertilizing hermaphrodite with a life cycle of only 3 days. As these animals are omnivores with a body size of only 1mm, culture is cheap and easy on agar plates with Escherichia coli OP50. Since 1996, many techniques available in *C. elegans* have been established in P.pacificus as well, e.g. cell lineaging [Sommer 2001], cell ablations [Jungblut and Sommer 2000] and storage in liquid nitrogen. *P. pacificus* is amenable to forward and reverse genetics through mutagenesis [Gutierrez

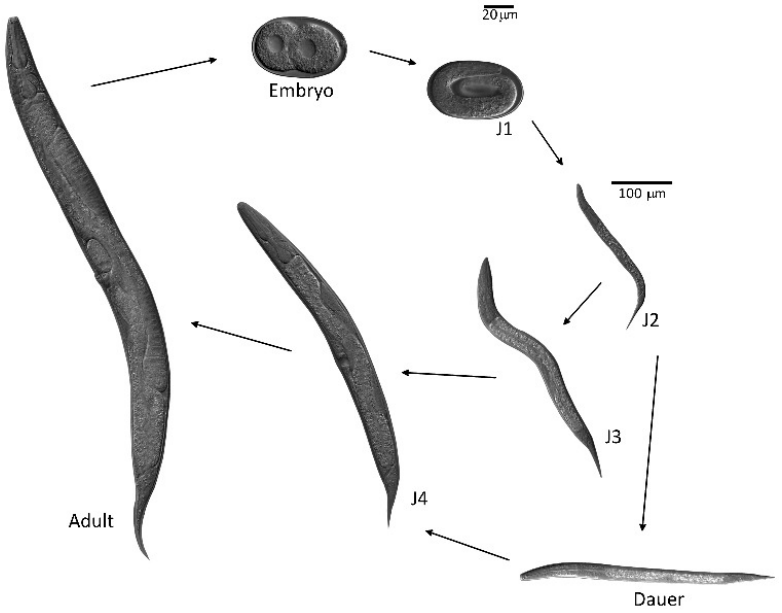


Figure 1.1 – *Pristionchus pacificus* life cycle. Pictures courtesy of Metta Riebesell.

and Sommer 2007], deletion libraries [Tian et al. 2008] and DNA-mediated transgenesis [Schlager et al. 2009].

The first line of research involving *P. pacificus* was a comparative study of vulva development [Sommer and Sternberg 1996; Eizinger and Sommer 1997]. The main conclusion was that a conserved developmental process can be regulated through very different molecular mechanisms. In *P. pacificus*, the somatic gonad together with the posterior body region control vulva development via the Wnt pathway alone [Zheng et al. 2005; Tian et al. 2008; Wang and Sommer 2011], while in *C. elegans* the same goal is achieved by cooperation of the Wnt, EGF and Notch pathways [Eisenmann et al. 1998; Myers and Greenwald 2007; Green et al. 2008]. Further studies have shown that the Wnt pathway in *P. pacificus* has evolved to control other processes such as gonad development [Rudel et al. 2008].

In the last 10 years, a second developmental process has become the focus of attention: development of the Dauer stage [Ogawa et al. 2009, 2011; Sinha et al. 2012]. Like all ecdysozoans, nematode life cycles have discrete stages separated by molting. *P. pacificus* goes through four larval stages called J1-J4 . The first larval stage develops entirely inside the egg shell, molting into J2 upon hatching. Depending on environmental signals, J2 animals will 'decide' between two options: either hatch into the normal J3 stage or become a Dauer larvae, which is arrested in development and resistant

to adverse conditions like starvation or oxygen depletion [Riddle and Albert 1997; Ogawa et al. 2009]. Similar to vulva formation, different molecular mechanisms control entry and exit into the Dauer stage in *P. pacificus* and *C. elegans*.

1.2 Phylogeny and Ecology

Nematodes are arguably one of the most species-rich and diverse animal phyla. They exist in all habitats on earth and comprise different lifestyles from free-living to animal and plant parasites. Blaxter [Blaxter et al. 1998] built a complete nematode phylogeny using the Small Subunit Sequences of the ribosomal RNA gene (SSU) from 53 taxa. According to this phylogeny, *P. pacificus* belongs to the Diplogastridae family in Clade V, while *C. elegans* is in the paraphyletic sister family Rhabditida. The two genera diverged approximately 300 Million years ago.

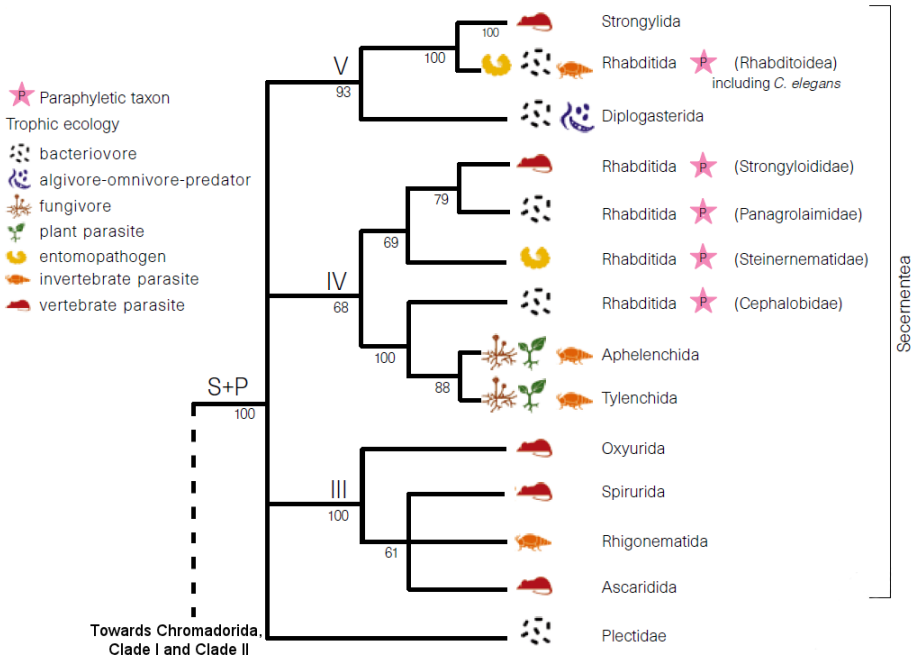


Figure 1.2 – Clade III to V of the phylum Nematoda Dendrogram based on MP and NJ analysis. Numbers at the branch point represent the maximal percentage bootstrap support. Clades I and II are excluded. Paraphyletic taxa are starred. From [Blaxter et al. 1998], modified by [Molnar 2012].

The members of the 28 genera and more than 300 species in the Diplogastridae family live mostly

as insect parasites or predators [Sudhaus and Fürst von Lieven 2003]. A robust phylogeny of the genus *Pristionchus* is a necessary foundation for all comparative studies. Werner Mayer resolved the tree of the genus *Pristionchus* as well as its phylogenetic position within the Diplogastridae using SSU sequences and other ribosomal protein genes [Mayer et al. 2007, 2009].

Unlike in *C. elegans*, the ecology of *Pristionchus* nematodes is well understood [Herrmann et al. 2006a,b, 2007; Mayer et al. 2007; Herrmann et al. 2010; Morgan et al. 2012]. They are found free living in soil and in a necromenic association with scarab beetles. Nematodes in the Dauer stage rest on the beetle [Weller et al. 2010]. After it's death, they resume development into J4 and reproduce on the rich nutrients of the carcass. The beetle-host relationship is highly species-specific, enabled through an interception of the beetles communication via sex pheromones [Hong et al. 2008].

C. elegans and *P. pacificus* not only live in different habitats, but also differ in their feeding habits. *C. elegans* is specialized in order to eat bacterial food and, like all members of the family Rhabditidae, possesses a grinder in the posterior bulb of the pharynx that is used to crush bacterial cells. *P. pacificus* on the other hand does not have a grinder, but a teeth-like structure at the tip of the pharynx, which allows it to feed on a diversity of sources including bacteria, fungi and other nematodes [Sommer et al. 1996a]. *P. pacificus* is generally more resistant than *C. elegans* to pathogenic bacteria such as *Pseudomonas aeruginosa*, *Staphylococcus aureus* and insecticidal *Bacillus thuringiensis* [Wei et al. 2003; Rae et al. 2008].

More than 700 *Pristionchus* strains have been collected globally. Most *P. pacificus* strains originate from the islands of Japan and La Réunion. A tree based on the nad6-nad4L mtDNA fragments shows a split into four distinct clades roughly based on region of origin. All clades contain strains from La Réunion, but only Clade B is restricted to this island [Herrmann et al. 2010; Morgan et al. 2012].

1.3 Genomics

C. elegans was the first metazoan to have its complete genome sequenced in 1998 [The *C. elegans* Sequencing Consortium 1998]. 10 years later, *P. pacificus* followed suit with a genome based on shotgun sequencing [Dieterich et al. 2008]. Since then, more than 100 natural *Pristionchus* strains have been re-sequenced. Compared to *C. elegans*, its genome is bigger (169Mb), contains more

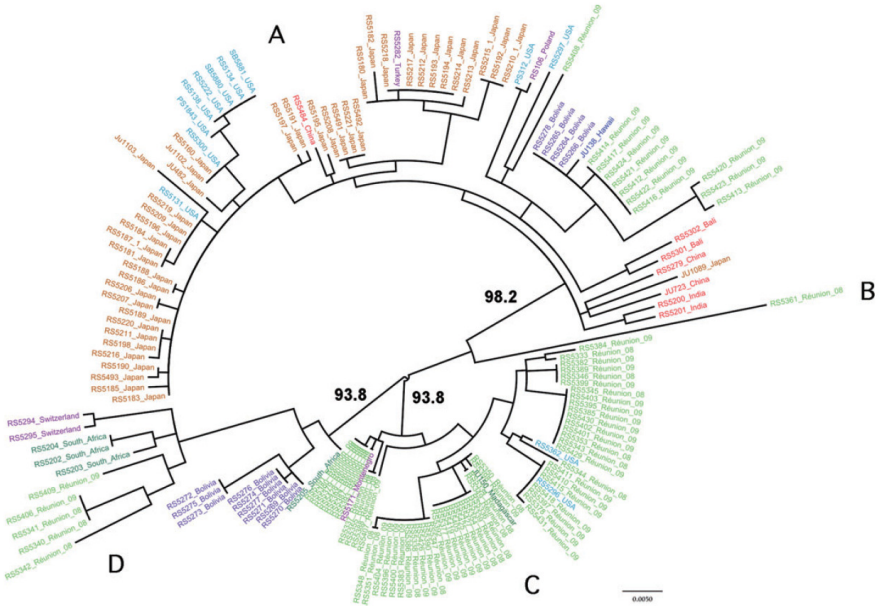


Figure 1.3 – *Pristionchus pacificus* phylogenetic tree. Unrooted ML tree based on mitochondrial *nad6-nad4L* fragment. Taxon labels consits of strain number, sampling site and sampling year. Branch values show NJ bootstrap support vales for the major clades. Figure from [Herrmann et al. 2010], modified by [Molnar 2012].

transcripts and protein coding genes (29.000) and consists of more GC-bases (43%) [Dieterich et al. 2008; Borchert et al. 2010].

A comparison of the *P. pacificus* genome with 5 other published genomes revealed 11.000 predicted genes without a homolog in any other species. These genes are considered to be pioneer genes [Borchert et al. 2010]. Their high number might be due to the recent lifestyle expansion from exclusively free living to a necromenic beetle association. Several cellulase genes were acquired through Horizontal Gene Transfer (HGT) from bacteria [Mayer et al. 2011]. An analysis based on codon-usage patterns further showed that many *P. pacificus* genes are derived from insect genomes [Rödelsperger and Sommer 2011]. Both HGT events seem plausible given the close association of *Pristionchus* to microbes and beetles.

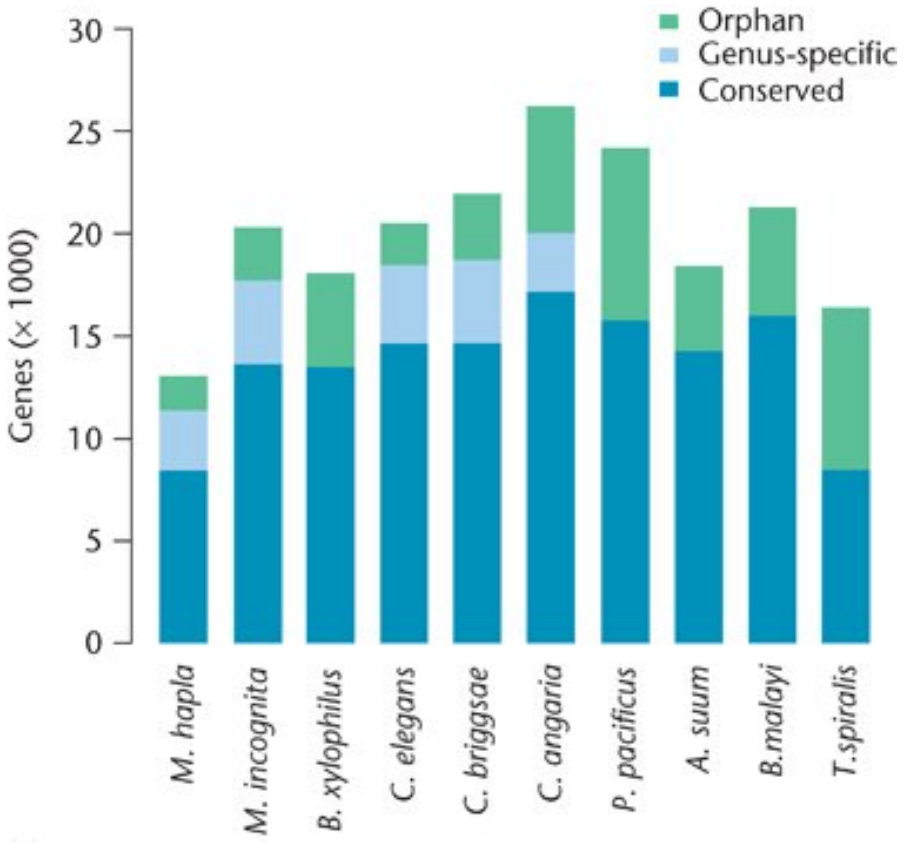


Figure 1.4 – *Pristionchus pacificus* gene orthology. Proteome comparison of 6 genomes, including *Pristionchus pacificus*. Proteomes are divided according to the inferred homology relations. Figure modified from [Rödelsperger et al. 2013].

1.4 Personal research

Bringing the tritrophic interactions between *Pristionchus* nematodes, host beetles and bacteria into the laboratory allows an in-depth study of an important biotic factor in *Pristionchus* evolution. We isolated bacterial strains from the wild and assayed their influence on different nematode species to answer the following questions:

- Do *Pristionchus* and *Caenorhabditis* nematodes associate with different bacteria?
- Is there species-specific pathogenicity?
- Can nematodes recognize and avoid pathogenic bacteria?

The necromenic beetle association is the key difference in ecology between *Pristionchus* nematodes and its relatives such as *C. elegans*. An integrative analysis of the evolutionary change between these taxa therefore depends on a detailed understanding of this novel lifestyle. The usual way of collecting new *Pristionchus* strains by placing dead beetles on agar plates is effective, but it does not provide information on the number or life stage of the nematodes originally present on the beetle host.

I used direct nematode extraction from Geotrupes beetles via Baermann funnels combined with 18S SSU molecular barcoding to elucidate the communities living on dung beetles. I was interested in the following questions:

- Does *Pristionchus* rest on beetles in the Dauer stage or are reproducing individuals present?
- How high and how homogeneous are infection rates?
- Which other nematodes are present on these beetles, and is there an association between *Pristionchus* and other nematode genera?

Chapter 2

Mutations: the source of variation

2.1 Types of mutation

A central mechanistic cause of evolutionary change are mutations [Lynch 2007]. By definition, all changes in DNA sequence are mutations. This includes base-substitution mutations that exchange one nucleotide for another, insertions or deletions (Indels) and even large duplications. In the following chapter, the word mutation will only refer to base-substitution mutations. Mutations are abbreviated as e.g. C:G $\rightarrow$ T:A, meaning a change from cytosine to thymine on the one DNA strand and therefore a change from guanine to adenine on the other strand. DNA nucleotides are classified into either purines (A and G) or pyrimidines (C and T). Mutations involving a change from one type to the other are called transversions, while transitions happen between two nucleotides of the same type.

2.2 Causes of mutation

There are many possible causes for mutations. The reason for a mutation can be within the internal machinery of the cell (e.g. errors introduced during replication or DNA repair) or an external mutagen such as UV-radiation or certain chemicals. In general, 12 different types of mutation are possible, but the mutation spectra are very variable between species.

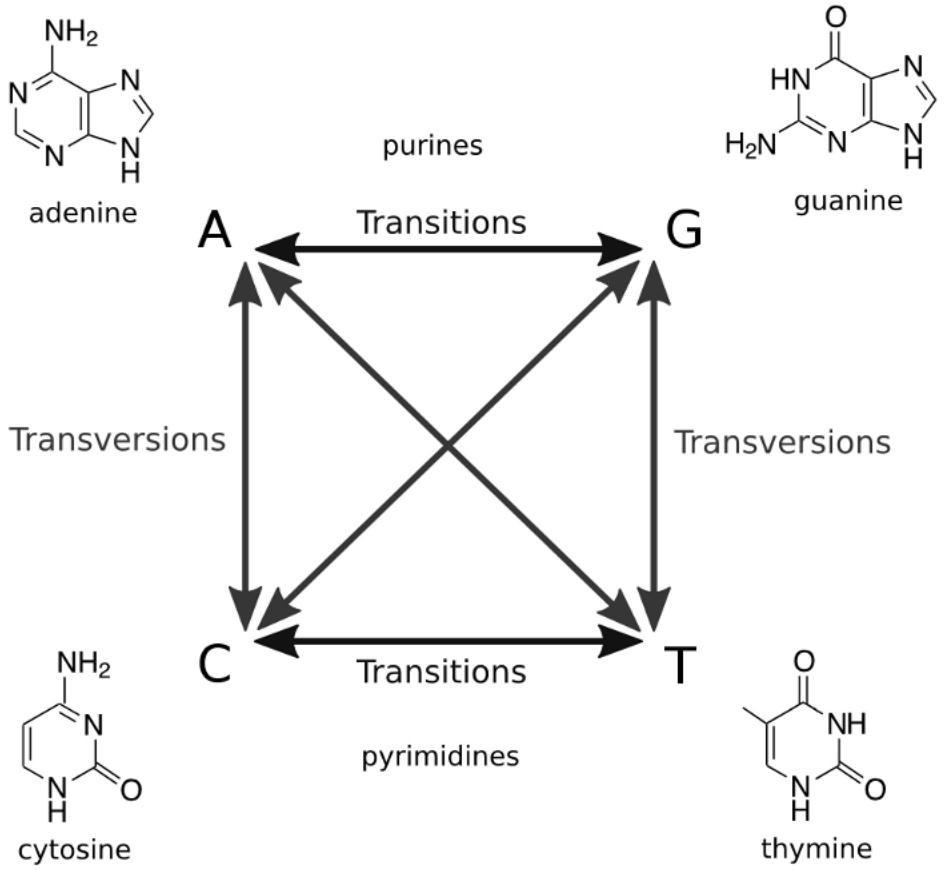


Figure 2.1 – Base-substitution mutations

Due to biochemical reasons, mutations are much more likely to involve cytosine or guanine than adenine or thymine. Spontaneous deamination of cytosine is one of the most common forms of mutation, leading to C:G → T:A transitions. A second common form is the oxidative conversion of guanine to 8-oxo-guanine, which leads to a C:G → A:T transversion if it is not repaired by the cellular machinery.

2.3 Measuring mutation via Mutation Accumulation Lines

Studying mutation in natural variants poses several challenges: first, the rate of spontaneous mutations cannot be estimated [Lynch 2010]. Second, natural selection and drift act on novel variants in nature, making some of them more likely to become fixed in a population than others. Natural variants between species thus do not accurately reflect the spectrum of spontaneous mutations . These major challenges in the study of mutational processes is the rationale behind the creation of laboratory-raised Mutation Accumulation Lines (MA lines) [Mukai 1964].

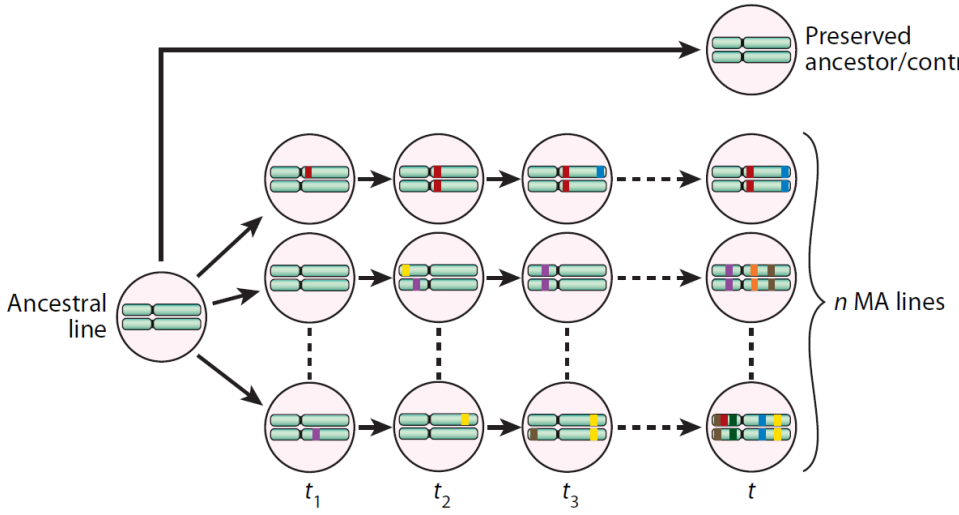


Figure 2.2 – Mutation Accumulation Lines An individual animal gives rise to n MA lines, which are propagated for t generations. Mutations initally happen in a single chromosome and either become homozygous or dissapear. Figure from [Halligan and Keightley 2009].

In MA lines, mutations are allowed to accumulate in parallel lines for many generations [Mukai 1964]. They are propagated under the most benign conditions possible, resulting in random fixation or removal of all but the most deleterious mutations. The first MA lines have been carried out by Muller as early as the 1920s [Muller]. Since then, MA lines were e.g. created in bacteria [Barrick et al. 2009], yeast [Lynch et al. 2008], *Arabidopsis thaliana* [Ossowski et al. 2010], *Daphnia pulex* [Seyfert et al. 2008], *Drosophila melanogaster* [Keightley et al. 2009] and *Caenorhabditis elegans*. [Denver et al. 2009, 2012]. Due to their small size and short generation time, nematodes are ideal candidates to carry out MAline experiments.

2.4 Mutation rate

A genome-wide estimation of the mutation rate is only possible since the second generation of DNA sequencers, i.e. around 2007 [Lynch 2007]. In multicellular organisms, the mutation rate (per base and generation) correlates positively with genome size and negatively with population size, i.e. large organisms tend to have more mutations. According to Lynch [Lynch 2007], the reason for this correlation is that the effective population size (and therefore random genetic drift) provides the lower limit for an improvement of DNA repair fidelity.

2.5 Consequences of mutation

Depending on their location relative to the structure of genes, mutations may have dramatically different consequences for the fitness of the organism in which they happen. Mutations in intergenic regions and introns will not affect fitness at all, unless they happen to modify a regulatory element. Changes to the sequence of exons can be either nonsynonymous or synonymous, depending on whether they change the triplet code for respective amino acid or not. The most serious fitness effects are usually caused by STOP mutations, which change the base triplet to a signal to the ribosome to terminate protein synthesis.

Besides altering the function of genes, mutations also influence the nucleotide composition of genomes, i.e. the content of GC nucleotides. Most organisms have a GC-content between 30% and 50%, but there are extremes on both ends of the distribution between 20% in *Plasmodium falciparum* [Gardner et al. 2002] and 72% in *Streptomyces coelicolor* [O'Rourke et al. 2009]. The reason for this large spectrum is still under debate [Lynch 2007]. The predominance of C:G $\rightarrow$ T:A transitions and C:G $\rightarrow$ A:T transversions among the mutational spectrum causes a mutational bias toward thymine and adenine (AT-bias). The AT-bias differs between study organisms, but generally lies between two and three mutation towards AT for each mutation in the opposite direction.

2.6 Personal research

In this chapter, I have explained the impact of mutations on genomes and the challenges associated with studying mutations. *P. pacificus* as a model species provides a unique opportunity for in-depth analysis of mutation processes with its combination of genomic tools, easily generated MA lines and large collection of hundreds of wild strains.

We sequenced the nuclear genome of 104 *P. pacificus* strains and the sister species *P. exspectatus* to investigate the following questions:

- What is the global population structure of *P. pacificus*?
- How diverged are the subpopulations?
- What is the influence of linkage on natural selection in *P. pacificus*?

Knowing the baseline mutation rate has several practical implications for evolutionary biology, because it is a key number for important simulations such as phylogenetic trees. Even though its on the heart of natural selection, the process from the original mutation events to eventual fixation of an allele in the population is not completely understood. This is partly because it has been impossible until very recently to gather the necessary data using high-throughput approaches.

In my personal research, I combined whole genome sequence analysis of MA lines and natural strains to find answers to the following questions:

- What is the baseline mutation rate in *P. pacificus*?
- What is the spectrum and location of de novo mutations?
- Does this spectrum exert pressure on the genomic nucleotide composition?
- What happens to the spectrum of novel alleles during their spread in natural populations?
- What are the consequences for the equilibrium of the genomic nucleotide composition?

Part II

Results and Discussion

Chapter 3

Studies on *Pristionchus* ecology

3.1 Isolation of naturally associated bacteria of necromenic *Pristionchus* nematodes and fitness consequences

Robbie Rae, Metta Riebesell, Iris Dinkelacker, Qiong Wang, **Andreas M. Weller**,
Christoph Dieterich and Ralf J. Sommer:
Journal of Experimental Biology 2008, 211(12):1927–1936.

3.1.1 Synopsis

This study describes the interaction of nematodes with 23 bacterial strains found either in soil or on three common *Pristionchus* host beetles. We show that several of the beetle-associated bacteria are harmful for nematodes in a species-specific manner.

3.1.2 Contributions

Identification of nematode species by SSU sequencing was done by me and Christoph Dieterich. In total, my contribution to this work was about 15%.

3.2 Quantitative Assessment of the Nematode Fauna Present on *Geotrupes* Dung Beetles Reveals Species-Rich Communities with a Heterogeneous Distribution

Andreas M. Weller, Werner E. Mayer, Robbie Rae and Ralf J. Sommer
Journal of Parasitology 2010, 96(3):525-531.

3.2.1 Synopsis

The necromenic association of *Pristionchus* nematodes is a central feature of their ecology and global distribution. The ability to arrest development and endure harsh conditions are a necessary condition for this association and might even serve as a preadaptation for parasitism.

There is a diversity of other animals sharing the *Geotrupes* host with *Pristionchus*. To quantify these communities, we extracted all nematodes from 114 individuals of *G. stercorosus* and determined their species by sequencing of the 18S SSU gene. In total, we found 5002 nematodes distibuted over the beetles in a very heterogenous fashion. Most of nematodes belonged to the genera *Pristionchus*, *Koerneria* and *Pelodera*. All *Pristionchus* individuals were in the Dauer stage.

3.2.2 Contributions

Beetle collection and nematode extraction was done by myself. Nematodes were prepared for sequencing by myself and Robbie Rae. SSU sequencing and species determination was done by Werner Mayer. Experiments planning and manuscript preparation were done by myself and Ralf J. Sommer. In total, my contribution to this work was about 70%.

Chapter 4

Studies on *Pristionchus* genome evolution

4.1 Weakly deleterious mutations dominate genetic diversity in *Pristionchus pacificus*

Christian Rödelsperger, Richard A. Neher, **Andreas M. Weller**, Gabi Eberhardt, Hanh Witte, Werner Mayer, Christoph Dieterich and Ralf J. Sommer
Genetics, In Review.

4.1.1 Synopsis

This study describes the population structure of *P. pacificus* and the effects of selection on natural variants. We sequenced 104 natural strains of *P. pacificus* and found that 40% of non-synonymous variants are weakly deleterious due to linked selection.

4.1.2 Contributions

Bioinformatics analyses were carried out by Christian Rödelsperger, Richard Neher and myself. In total, my contribution to this work was about 10%.

4.2 Mutation accumulation lines and natural variants in the nematode *Pristionchus pacificus* indicate opposing forces of *de novo* mutations and counterbalancing selection

Andreas M. Weller, Christian Rödelsperger, Gabi Eberhardt, Ruxandra I. Molnar and Ralf J. Sommer

Genetics, In Review.

4.2.1 Synopsis

This study describes mutations in *P. pacificus* found by whole-genome sequencing of mutation accumulation lines (MA lines). We sequenced the whole nuclear genome of 22 MA lines that were inbred for 150 generations to directly study mutation patterns unbiased by natural selection. We observe a mutation rate of $2x10^{-9}$ per base and generation. The mutation spectrum is clearly biased towards A/T nucleotides, resulting in an overall ATbias of 5.3. We considered the ATbias in 655,705 natural variants and found a lower mean ATbias of 2.4. In addition, the bias of a variant is negatively correlated with its minor allele frequency (MAF), which serves as a proxy for variant age. These results suggests that the high ATbias of *de novo* mutations is neutralised over time by a counterbalancing force.

4.2.2 Contributions

Preparation of MA lines for sequencing, DNA extraction and Illumina library preparation was done by myself and Gabi Eberhardt. Bioinformatics analyses were carried out by myself and Christian Rödelsperger. Experiments planning and manuscript preparation were done by myself and my supervisor Ralf J. Sommer. In total, my contribution to this work was about 80%.

Part III

Unpublished work

Chapter 5

Intraspecific variance in recombination rates of the nematode *Pristionchus pacificus*

5.1 Introduction

Recombination processes are central to the evolution of animal and plant genomes, but the forces that shape these processes themselves are only little understood. Closely related species or diverged strains of the same species provide unique opportunities to study the control of recombination processes. Such studies will be key to an understanding of these processes both from a mechanistic and evolutionary point of view, leading to insights in diverse topics such as natural selection, speciation and human health.

An important and surprising observation was that recombination rates can change quickly even within species, which suggests that natural selection can target recombination directly (see [Smukowski and Noor 2011] for a comprehensive review). In recent years, whole-genome sequencing allows a fine-grained analysis of crossover events in several model species [Mancera et al. 2008; Comeron et al. 2012; Yang et al. 2012; Bessoltane et al. 2012; Roesti et al. 2013]. Crossover rates per meiosis vary between species, from relatively low in *Arabidopsis thaliana* (9CO [Yang et al. 2012]) to dramatically high in honey bee (40CO [Bessoltane et al. 2012]) and yeast (90CO

[Mancera et al. 2008]). Besides global recombination rates, there is also variation within genomes. While yeast and mammals show extreme local differences known as 'hotspots' and 'coldspots', there is a much softer clustering of events in other organisms. Some studies even suggest that hotspots contribute to their own extinction via meiotic drive [Jeffreys and Neumann 2002].

Several factors have been found so far that are ubiquitously associated with recombination, e.g. gene density, repeats and conserved sequence motifs such as prdm9 in mammals [Baudat et al. 2010; Berg et al. 2010]. The most intriguing finding from an evolutionary point of view is the correlation between recombination and within-species diversity, but not between-species divergence. Most authors agree that this is due to the effects of recombination on natural selection via Hill-Robertson interference [Andolfatto 2001; Begun et al. 2007; McGaugh et al. 2012], but some have argued that a part [Kulathinal et al. 2008] or even all [Spencer et al. 2006] of the connection is due to mutagenicity directly caused by recombination events.

In this study we analyzed recombination events by creating recombinant inbred lines (RILs) between the *P. pacificus* reference strain PS312 and three closely related isolates. We genotyped 69 of these RILs by whole-genome sequencing and analyzed the rate and location of crossovers and their relationship to intra-species nucleotide diversity and gene density. We found that crossover frequencies and locations differ between strains and that crossovers are linked to natural variants and gene-dense regions.

5.2 Results

We created three sets of Recombinant Inbred Lines (RILs) by crossing the reference strain PS312 to three other *P. pacificus* strains (caxwa, caxsa and caxba, see Materials and Methods). After inbreeding until the F5, we sequenced the whole genome of 69 RILs, generating between 11.5 million and 74.6 million reads (mean: 27 million) per RIL. On the 169MB genome of *P. pacificus*, which translates to a coverage between 6.2X and 42.3X, with a mean coverage of 15X. A conservative filtering pipeline was applied to create a Gold Standard Set of reliable SNPs (see Materials and Methods). Between 32.942 (caxwa) and 137.137 (caxba) SNPs could be confidently genotyped in each RIL, resulting in an average marker distance of 4550bp and 1100bp, respectively. Recombination events were detected as switches between the maternal and paternal phase of markers.

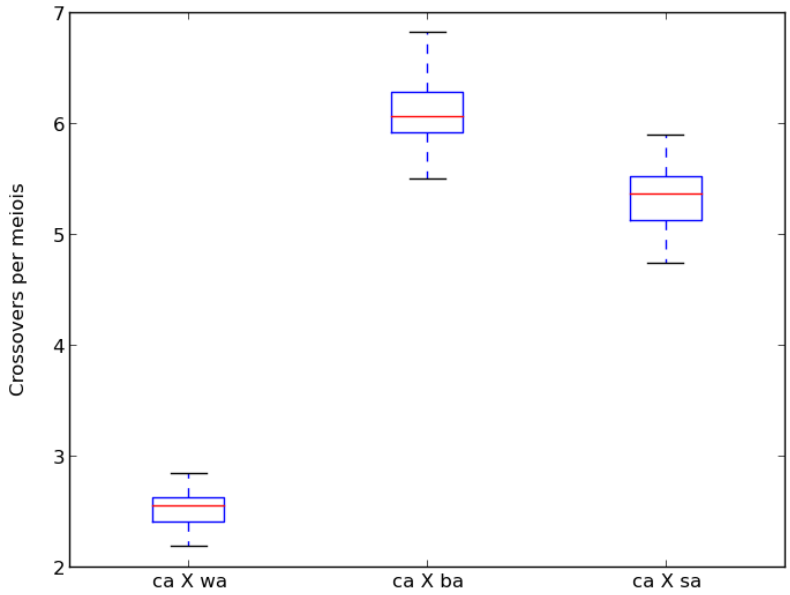


Figure 5.1 – Variance in recombination rates between the three crosses.

First, we focused on crossover events. In total, 1024 crossovers were detected in the 69 RILs. A meaningful comparison of crossover frequencies between the crosses was enabled by repeating the recombination detection in a down-sampled set of 20.000 random markers. After 100 down-sampling iterations, we found on average of 103, 248 and 217 combined crossover events in all RILs of caxwa, caxba and caxsa, respectively (Figure 5.1). According to these results, there is a highly significant (One-way ANOVA, p < 10e-100) variation in crossover rates between the crosses. This variation suggests that at least some of the recombination hot spots vary even between closely related strains of the same species and are therefore highly unstable. A simulation of the inbreeding process allowed us to draw conclusions from the total number of observed events after 5 generations to the events actually happening per meiosis (see Materials and Methods). We plotted crossover events as centiMorgan per Megabase (cM/Mb) across the chromosomes and found strong recombination hotspots (Figure 5.2). These hotspots are not biased for the ends of chromosome arms and are only partially shared between crosses. We investigated the correlation of crossover events with other genomic features in a 2Mb-window around event locations (Figure 5.3). Crossovers occur in regions dense in genes, but not in DNA repeats. We do observe a correlation of crossovers to intra-species diversity. This result provides support for the

hypothesis that recombination increases diversity via Hill-Robertson interference instead of direct mutagenicity.

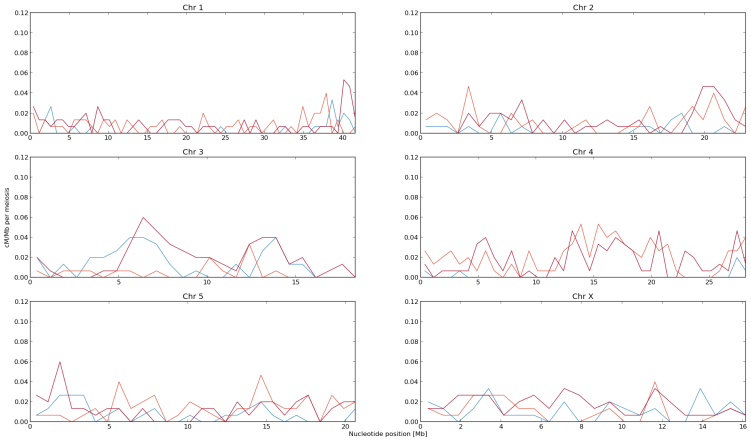


Figure 5.2 – Recombination rates across the chromosomes in the three crosses.

Finally, a linkage map was created by first calculating a separate linkage map for each cross. A consensus of each contig was treated as a single virtual marker to reduce computational load. In a second step, we employed a minimum-spanning tree approach to merge the three linkage maps into a unified consensus map. This map includes millions of SNP markers and thus represents the most comprehensive resource for genetic mapping in *P. pacificus* to date. It is available online on www.pristionchus.org.

5.3 Discussion

In this study we analyze rate and location of crossovers and their relationship to genomic features. After creating three different sets of recombinant inbred lines and genotyping 23 animals of each set, we analyze recombination processes and create a consensus linkage map for *P. pacificus*.

Our crossing scheme maximizes comparability between crosses and excludes sex-specific effects by creating all RILs using a constant maternal animal from our reference strain RS2333. However, it is inherently impossible to distinguish between the natural recombination processes in the respective strains and effects caused by incompatibilities between the parental genomes. The expectation

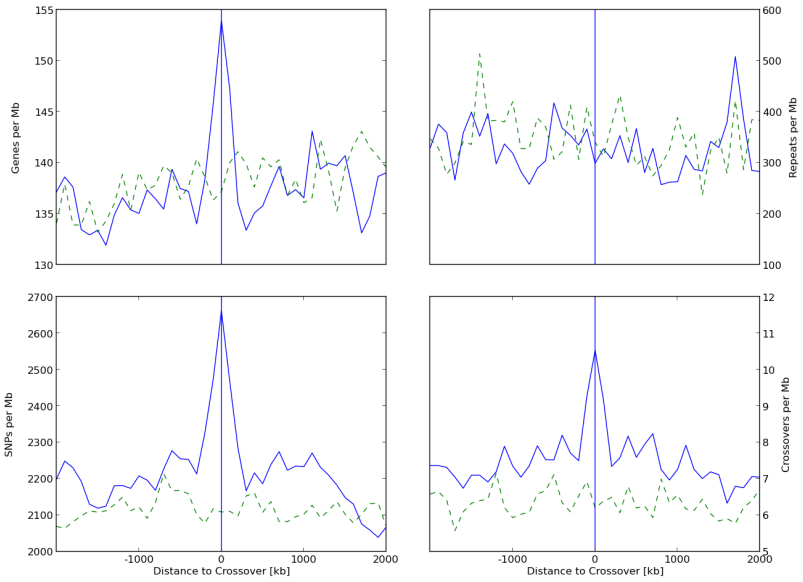


Figure 5.3 – Correlation of crossover locations to genomic features.

of natural recombination processes in the crosses is supported by the fact that crossover rates in a cross are not correlated to the sequence divergence between parental strains. There are strong sequence similarities between our crosses. As a consequence, we would a priori expect a strong conservation of recombination processes, too, but we find highly significant differences in rates between our crosses. Our downsampling approach excludes marker locations or marker densities as the cause for these differences. As we chose closely related strains of *P. pacificus* for our experiment, these results show that recombination processes are highly variable in frequency and location over relatively short time periods.

Likely cellular mechanisms behind variable recombination rates are differences in either binding motifs (like the motif that binds to PRDM9 in mammals [Baudat et al. 2010; Berg et al. 2010], or zinc-finger proteins (like trem in *D. melanogaster* [Lake et al. 2011]) that control recombination initiation. However, there is increasing evidence that variation in epigenetic features might reflect on recombination processes as well [Myers et al. 2005; Ptak et al. 2005]. From an evolutionary point of view, these differences between strains might be either a result of random drift or of differential selection in the wild. An intruiging third possibility is that unintentional artifical selection acted on the strains during their laboratory culture to increase recombination rates and thereby adaptability to new conditions [Groenen et al. 2009]. Previous studies have shown

that recombination rate conservation is partially dependent on the scale of analysis [Myers et al. 2005; Coop and Przeworski 2007]. While 'fine scale' studies (i.e. kilobases) tend to find divergence between closely related species, 'broad scale' studies (i.e megabases) do not, a general trend that is supported by our results. A proposed reason for this might be that the need to ensure proper chromosome disjunction during cell division places a tighter control over 'broad scale' than 'fine scale' processes [Hassold et al. 2004; Fledel-Alon et al. 2009].

Our data indicates a correlation between crossover locations and intra-species diversity. Such patterns have been observed in several organisms since the landmark study by Begun and Aquadro (Begun1992Levels) first showed it in D.melanogaster. The consensus explanation for this association is that either positive (i.e. selective sweeps) or negative (i.e. background selection) natural selection reduces nucleotide diversity at sites with reduced recombination. This is further supported by a recent study in which Rödelsperger et al show that genomic diversity in *P. pacificus* is shaped by strong effects of linked selection (see Chapter 4.1).

In summary, we have demonstrated that properties of recombination processes are shaped by natural selection and fluctuate within the species *P. pacificus*.

5.4 Materials and Methods

We crossed females of the *P. pacificus* strain PS312 (California, USA) carrying a dpy-mutation as visual marker with males of the strains PS1843 (Washington, USA), RS5205 (South Africa) and RS5302 (Bali, Indonesia), respectively. In this paper, these crosses are referred to as caxwa, caxsa and caxba. Successful mating was confirmed by reversal to the WT phenotype in the F1. Cross-progeny F1 females were selected to found individual Recombinant Inbred Lines (RILs). The RILs were inbred by self-fertilization until the F5 to obtain enough animals for DNA extraction.

DNA extraction and Illumina library preparation was performed as described before (see Chapter 4.1). Briefly, we used the MasterPure DNA purification kit to extract DNA from a pool of animals, which was then sheared by a Covaris S2 system. We prepared Ilumina libraries using standard reagents and protocol of the Illumina TruSeq v2 sample preparation kit. Selection for an insert size of 300-400bp was performed by agarose gel excision. We determined the final concentration of our libraries on a Bioanalyzer DNA 1000 chip, then normalized them into 10nM. Sequencing of pooled 12 plexes was done in-house on an Illumina HiSeq2000 sequencer.

As described before (see Chapter 4.1), stampy (version 1.0.13, Lunter2011Stampy) was used for inital alignment of the reads to the reference genome (*P. pacificus* Hybrid Assembly). Duplicate reads were removed using samtools ('rmdup' command, version 0.1.17, Li2009Sequence). Previous work on the *P. pacificus* reference genome has uncovered issues arising from duplicated or deleted regions in some strains. In order to avoid artefacts in recombination detection, we created a Gold Standard set by filtering known polymorphisms between our strains through a conservative pipeline. First, we called SNPs against the RS2333 reference genome by a custom Python script using Â the raw read pileup of samtools version (0.1.17, Li2009Sequence) as input. SNPs were accepted to the Gold Standard if they conformed to the following criteria: A) minimum coverage of 7, B) at least 90% of reads supporting the alternative base, C) no indels called within a 300bp window around the location and D) an RS2333 reference base covered by at least 7 reads and supported by 90% of reads. Gold Standard SNPs were then genotyped in a RIL according to similar criteria: coverage >= 7 and at least 90% of reads confirming to either RS2333 or the other maternal strain, respectively. Crossover events were detected by a custom Python script. As these events are usually supported by many markers, they are easy to see and robust to misplaced markers. For the detection of Gene Conversions however, a secondary filtering step was crucial to avoid an overwhelming number of false-positives, as each artefact marker (i.e. of random genotype) would be regarded as a GC in 50% of rils. First, the segregation distortion rate was calculated for each position. Markers with less than 10% of genotypes from one parent were dismissed as artefact snps. GC events were then found via a custom Python script with a cutoff for maximum tract length of 15kb (see [Comeron et al. 2012]).

We were interested in recombination events per meiosis, but were constrained by our organism to sequence after the F5. Simply dividing the numbers observed in the F5 by five would underestimate the crossover frequency, as crossovers in homozygous regions cannot be observed, so we used a simulation to empirically determine the ratio of recombination events that would still be observable in the F5. A simulated diploid organism with a single chromosome consisting of 50.000 randomly placed markers was subjected to a crossover event per generation until the F5. After 10.000 iterations, the mean number of observed crossovers was used to calculate the correction factor c ($c = 0.54$).

For each interval, the number of observed crossover events after the F5 was multiplied by c to obtain the number of crossovers per meiosis. Chromosome-wide plots of centiMorgan per Megabase (cM/Mb) were drawn based on the consensus linkage map described below. These and

all other plots in this paper were drawn using the Matplotlib plotting library [Hunter 2007].

The reason for varying crossover frequencies in the crosses might be easier event detection in crosses with more markers. To allow a meaningful comparison of crosses, we modified the original Python script for event detection to only consider a random subset of 20.000 markers per cross. Numbers of crossovers detected using the down-sampled marker set were plotted after 100 iterations.

We investigated correlations between crossover locations and natural variants, de novo mutations and repeats. Natural variants between all 104 sequenced *P. pacificus* strains were described previously (see Chapter 4.1). Locations of de novo mutations were used from a Mutation Accumulation Line experiment in *P. pacificus* PS312 (see Chapter 4.2). Repeats in the reference genome were detected using repeatmasker (version 4.0.1). We fetched a 2MB region around each crossover location, then divided each region into 20 windows of 100kb each. We then plotted the mean response variable for each set of windows in the same distance to their respective crossover. Control locations were placed randomly across the genome and analyzed equally.

For creation of the linkage maps, we treated each contig as an individual marker to reduce computational load. First, we calculated the consensus genotype for each RIL and contig. We then used MSTmap [Wu et al. 2008] to assemble a linkage map per cross. Finally, we merged all 3 cross-specific maps into a consensus map using MergeMap [Wu et al. 2008].

Chapter 6

Strong hybrid incompatibilities between strains of *Pristionchus pacificus*

6.1 Introduction

In general, each allele of a gene has an equal probability of being present in a gamete. Deviations from this 50/50 ratio are called *segregation distortion.* A transmission ratio distortion might be pre-zygotic if a gene increases it's chance of ending up in a gamete beyond 50% (true *meiotic drive*). Such meitoic drive has been observed in many organisms, including *D. melanogaster* [Kusano et al. 2001], *M. musculus* [Bauer et al. 2005] and humans [Liu et al. 2013]. The cause for a transmission ratio distortion might also be post-zygotic if both gametes are equally likely to be produced, but a certain combination of alleles is less likely to grow into an adult organism (*hybrid incompatibility*). An intriguing study [Seidel et al. 2008] in *C. elegans* showed how the linked gene pair *zeel-1* and *peel-1* induces lethality in embryos not inheriting it.

In this study, I systematically analyzed genomic data from sequenced *P. pacificus* hybrid lines to investigate pattern of hybrid incompatibilities in this species.

6.2 Results and Discussion

I visualized the ratio of parental alleles in 6 types of *P. pacificus* hybrid lines as the percentage of PS312 California alleles (Figure 6.1) . There is a strong segregation rate distortion in a chromosome and cross specific manner. Most notable is a consistent distortion towards higher percentage of PS312 California alleles on Chromosome III. This suggests a common causal locus in PS312 that manages to either prevent alleles of other strains from ending up in the zygote or killing off hybrids not carrying it like in the *zeel-1*/*peel-1* case. The causal loci must have been lost during laboratory culture between 1997 and 2004. This provides an ideal starting point to study the molecular basis of this transmission distortion as we know that only a few mutations between these two strains are present on Chromosome III (unpublished results).

6.3 Materials and Methods

I collected genomic data on 6 *P. pacificus* hybrid lines from 3 different previous experiments, each involving a hybrid cross between the reference strain PS312 California and another strain. The first data source was from an own experiment in which I crossed PS312 California (1997) with RS2333 (the 2004 derivative of the same strain) and sequenced a pool of several thousand individuals on a single Illumina GAII lane. The second data source were the 66 recombinant inbred lines described in Chapter 5 plus 22 sequenced RILs from a cross between PS312 California and RS106 Poland that were not analyzed in Chapter 5. Finally, Drs. Christian Rödelsperger and Eduardo Moreno shared the genomes of 102 RILs from a cross between PS312 California and RSB001 LaReunion which were sequenced on Illumina HiSeq machines using the RAD-Seq method. Parental ratios were visualized using the linkage map described in Chapter 5 and the the Matplotlib plotting library [Hunter 2007].

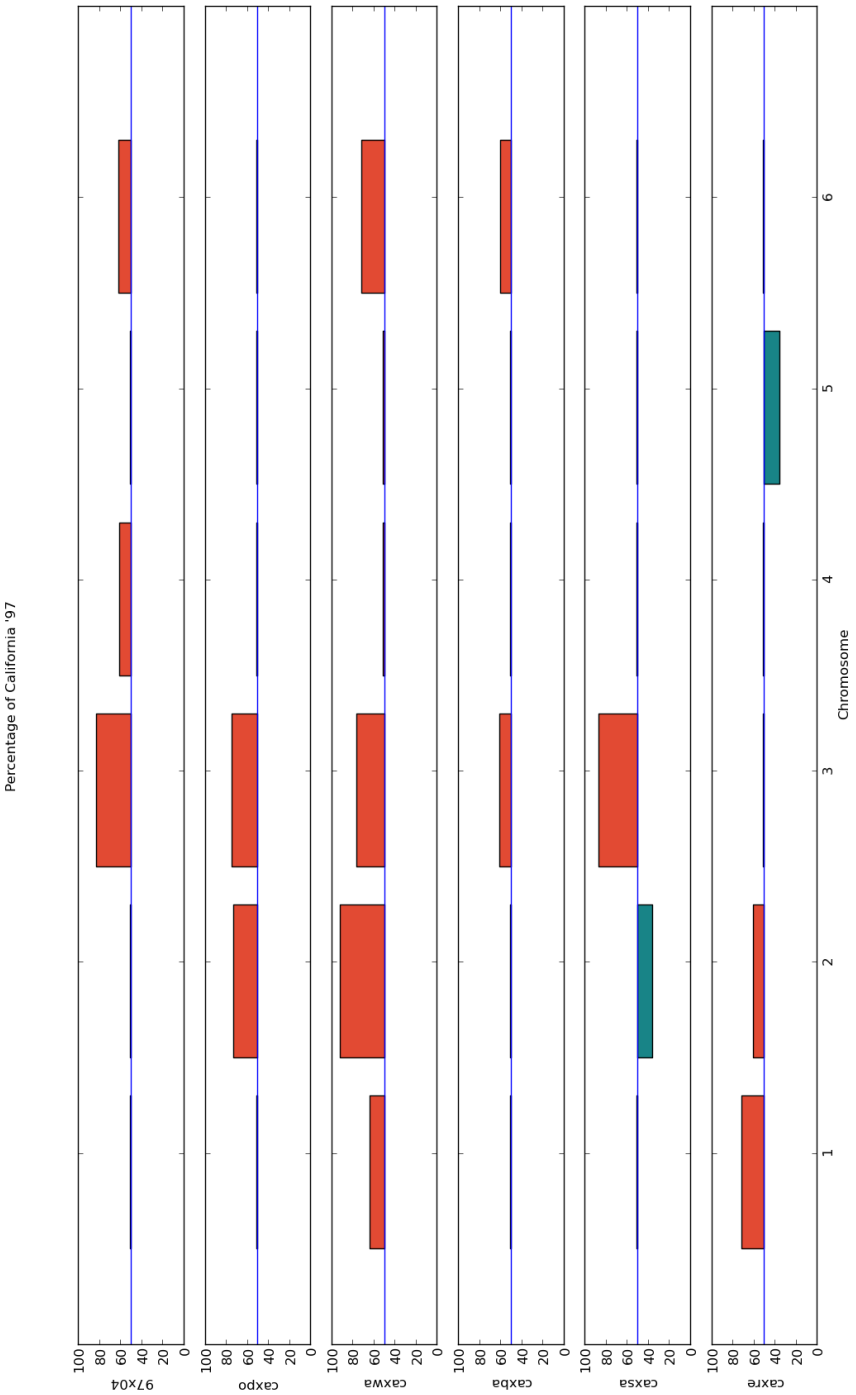


Figure 6.1 – Deviation from expected 50% transmission ratio in 6 hybrid strains of *P. pacificus*. Bars indicate percentage of PS312 California alleles in the offspring for each chromosome.

List of Abbreviations

C,T,G,A	Cytosine, Thymine, Guanine, Adenine
DNA	Deoxyribonucleic Acid
HGT	Horizontal Gene Transfer
Indel	Insertion and Deletion Mutation
J1-J4	*P. pacificus* larval stages
MA line	Mutation Accumulation Line
MP	maximum parsimony
NGS	next generation sequencing
NJ	neighbour-joining
rDNA	ribosomal DNA
SFS	site frequency spectra
SSU	small subunit rDNA
UV	Ultraviolet Light

List of Figures

Curriculum Vitae

Andreas M. Weller

Experience

2013–today **Bioinformatics Scientist III**, *Life Technologies*, Oxford, UK.

2013–2013 **Project Manager**, *Computomics*, Tübingen.

Education

2007–2013 **PhD**, *Max-Planck-Institute for Developmental Biology*, Tübingen.
PhD project title: 'Genome Evolution in the Nematode *Pristionchus pacificus*'

2003–2007 **Diplom in Biology**, *University of Tübingen*.
Thesis title: "Ecology and Biogeography of the Nematode *Pristionchus pacificus*"

2002–2003 **Civil Service**, *Krankenhaus Herrenberg*.

1993–2002 **Abitur**, *Andreae Gymnasium Herrenberg*.

Ursrainer Ring 103 – 72076 Tubingen – Germany

Bibliography

Andolfatto Peter. Adaptive hitchhiking effects on genome variability. *Current Opinion in Genetics & Development*, 11(6):635–641, December 2001.

Barrick Jeffrey E., Yu Dong S., Yoon Sung H., Jeong Haeyoung, Oh Tae K., Schneider Dominique, Lenski Richard E., and Kim Jihyun F. Genome evolution and adaptation in a long-term experiment with Escherichia coli. *Nature*, 461(7268):1243–1247, October 2009.

Baudat F., Buard J., Grey C., Fledel-Alon A., Ober C., Przeworski M., Coop G., and de Massy B. PRDM9 Is a Major Determinant of Meiotic Recombination Hotspots in Humans and Mice. *Science*, 327(5967): 836–840, February 2010.

Bauer Hermann, Willert Jurgen, Koschorz Birgit, and Herrmann Bernhard G. The t complex-encoded GTPase-activating protein Tagap1 acts as a transmission ratio distorter in mice. *Nat Genet*, 37(9):969–973, September 2005.

Begun David J., Holloway Alisha K., Stevens Kristian, Hillier LaDeana W., Poh Yu-Ping, Hahn Matthew W., Nista Phillip M., Jones Corbin D., Kern Andrew D., Dewey Colin N., Pachter Lior, Myers Eugene, and Langley Charles H. Population Genomics: Whole-Genome Analysis of Polymorphism and Divergence in Drosophila simulans. *PLoS Biol*, 5(11):e310+, November 2007.

Berg Ingrid L., Neumann Rita, Lam Kwan-Wood G., Sarbajna Shriparna, Odenthal-Hesse Linda, May Celia A., and Jeffreys Alec J. PRDM9 variation strongly influences recombination hot-spot activity and meiotic instability in humans. *Nat Genet*, 42(10):859–863, October 2010.

Bessoltane Nadia, Toffano-Nioche Claire, Solignac Michel, and Mougel Florence. Fine Scale Analysis of Crossover and Non-Crossover and Detection of Recombination Sequence Motifs in the Honeybee (Apis mellifera). *PLoS ONE*, 7(5):e36229+, May 2012.

Blaxter ML, De Ley P, Garey JR, Liu LX, Scheldeman P, Vierstraete A, Vanfleteren JR, Mackey LY, Dorris M, Frisse LM, Vida JT, and Thomas WK. A molecular evolutionary framework for the phylum Nematoda.

Nature, 392:71–75, 1998.

Borchert N, Dieterich C, Krug K, Schütz W, Jung S, Nordheim A, Sommer RJ, and Macek B. Proteogenomics of *Pristionchus pacificus* reveals distinct proteome structure of nematode models. *Genome Research*, 20: 837–846, 2010.

Comeron Josep M., Ratnappan Ramesh, and Bailin Samuel. The Many Landscapes of Recombination in Drosophila melanogaster. *PLoS Genet*, 8(10):e1002905+, October 2012.

Coop Graham and Przeworski Molly. An evolutionary view of human recombination. *Nat Rev Genet*, 8(1): 23–34, January 2007.

Denver Dee R., Dolan Peter C., Wilhelm Larry J., Sung Way, Lucas-Lledó J. Ignacio, Howe Dana K., Lewis Samantha C., Okamoto Kazu, Thomas W. Kelley, Lynch Michael, and Baer Charles F. A genome-wide view of Caenorhabditis elegans base-substitution mutation processes. *Proceedings of the National Academy of Sciences*, 106(38):16310–16314, September 2009.

Denver Dee R., Wilhelm Larry J., Howe Dana K., Gafner Kristin, Dolan Peter C., and Baer Charles F. Variation in Base-Substitution Mutation in Experimental and Natural Lineages of Caenorhabditis Nematodes. *Genome Biology and Evolution*, 4(4):513–522, January 2012.

Dieterich C, Clifton SW, Schuster LN, Chinwalla A, Delehaunty K, Dinkelacker I, Fulton L, Fulton R, Godfrey J, Minx P, Mitreva M, Roeseler W, Tian H, Witte H, Yang SP, Wilson RK, and Sommer RJ. The *Pristionchus pacificus* genome provides a unique perspective on nematode lifestyle and parasitism. *Nature Genetics*, 40:1193–1198, 2008.

Eisenmann DM, Maloof JN, Simske JS, Kenyon C, and Kim SK. The β-catenin homolog BAR-1 and LET-60 Ras coordinately regulate the Hox gene *lin-39* during *Caenorhabditis elegans* vulval development. *Development*, 3680:3667–3680, 1998.

Eizinger A and Sommer RJ. Nematode Epidermal Cell Fates The Homeotic Gene *lin-39* and the Evolution of Nematode Epidermal Cell Fates. *Science*, 452:452–455, 1997.

Fledel-Alon Adi, Wilson Daniel J., Broman Karl, Wen Xiaoquan, Ober Carole, Coop Graham, and Przeworski Molly. Broad-Scale Recombination Patterns Underlying Proper Disjunction in Humans. *PLoS Genet*, 5 (9):e1000658+, September 2009.

Gardner Malcolm J, Hall Neil, Fung Eula, White Owen, Berriman Matthew, Hyman Richard W, Carlton Jane M, Pain Arnab, Nelson Karen E, Bowman Sharen, Paulsen Ian T, James Keith, Eisen Jonathan A, Rutherford Kim, Salzberg Steven L, Craig Alister, Kyes Sue, Chan Man-Suen, Nene Vishvanath, Shallom Shamira J, Suh Bernard, Peterson Jeremy, Angiuoli Sam, Pertea Mihaela, Allen Jonathan, Selengut Jeremy, Haft Daniel, Mather Michael W, Vaidya Akhil B, Martin David M A, Fairlamb Alan H, Fraunholz Martin J, Roos David S, Ralph Stuart A, McFadden Geoffrey I, Cummings Leda M, Subramanian G Mani, Mungall Chris, Venter J Craig, Carucci Daniel J, Hoffman Stephen L, Newbold Chris, Davis Ronald W,

Fraser Claire M, and Barrell Bart. Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature*, 419:498–511, 2002.

Green JL, Inoue T, and Sternberg PW. Opposing Wnt pathways orient cell polarity during organogenesis. *Cell*, 134(4):646–656, 2008.

Groenen Martien A. M., Wahlberg Per, Foglio Mario, Cheng Hans H., Megens Hendrik-Jan, Crooijmans Richard P. M. A., Besnier Francois, Lathrop Mark, Muir William M., Wong Gane K., Gut Ivo, and Andersson Leif. A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. *Genome Research*, 19(3):510–519, March 2009.

Gutierrez A and Sommer RJ. Functional diversification of the nematode mbd2/3 gene between *Pristionchus pacificus* and *Caenorhabditis elegans*. *BMC Genetics*, 8:57, 2007.

Halligan Daniel L. and Keightley PD. Spontaneous Mutation Accumulation Studies in Evolutionary Genetics. *Annual Review of Ecology, Evolution, and Systematics*, 40:151–172, 2009.

Hassold T., Judis L., Chan E. R., Schwartz S., Seftel A., and Lynn A. Cytological studies of meiotic recombination in human males. *Cytogenetic and Genome Research*, 107(3-4):249–255, 2004.

Herrmann M, Mayer WE, and Sommer RJ. Nematodes of the genus *Pristionchus* are closely associated with scarab beetles and the Colorado potato beetle in Western Europe. *Zoology*, 109:96–108, 2006a.

Herrmann M, Mayer WE, and Sommer RJ. Sex, bugs and Haldane's rule: the nematode genus *Pristionchus* in the United States. *Frontiers in Zoology*, 3:14, 2006b.

Herrmann M, Mayer WE, Hong RL, Kienle S, Minasaki R, and Sommer RJ. The nematode *Pristionchus pacificus* (Nematoda: Diplogastridae) is associated with the oriental beetle *Exomala orientalis* (Coleoptera: Scarabaeidae) in Japan. *Zoological Science*, 24:883–9, 2007.

Herrmann M, Kienle S, Rochat J, Mayer WE, and Sommer RJ. Haplotype diversity of the nematode *Pristionchus pacificus* on Réunion in the Indian Ocean suggests multiple independent invasions. *Biological Journal of the Linnean Society*, 100:170–179, 2010.

Hong RL and Sommer RJ. *Pristionchus pacificus*: a well-rounded nematode. *BioEssays*, 28:651–659, 2006.

Hong RL, Svatos A, Herrmann M, and Sommer RJ. Species-specific recognition of beetle cues by the nematode *Pristionchus maupasi*. *Evolution and Development*, 279:273–279, 2008.

Hunter J. D. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, 9(3):90–95, May 2007.

Jeffreys Alec J. and Neumann Rita. Reciprocal crossover asymmetry and meiotic drive in a human recombination hot spot. *Nat Genet*, 31(3):267–271, July 2002.

Jungblut B and Sommer RJ. Novel cell-cell interactions during vulva development in *Pristionchus pacificus*. *Development*, 127:3295–3303, 2000.

Keightley Peter D., Trivedi Urmi, Thomson Marian, Oliver Fiona, Kumar Sujai, and Blaxter Mark L. Analysis of the genome sequences of three Drosophila melanogaster spontaneous mutation accumulation lines. *Genome Research*, 19(7):1195–1201, July 2009.

Kulathinal Rob J., Bennett Sarah M., Fitzpatrick Courtney L., and Noor Mohamed A. F. Fine-scale mapping of recombination rate in Drosophila refines its correlation to diversity and divergence. *Proceedings of the National Academy of Sciences*, 105(29):10051–10056, July 2008.

Kusano A., Staber C., and Ganetzky B. Nuclear Mislocalization of Enzymatically Active RanGAP Causes Segregation Distortion in Drosophila. *Developmental Cell*, 1(3):351–361, September 2001.

Lake Cathleen M., Nielsen Rachel J., and Hawley R. Scott. The Drosophila Zinc Finger Protein Trade Embargo Is Required for Double Strand Break Formation in Meiosis. *PLoS Genet*, 7(2):e1002005+, February 2011.

Liu Yang, Zhang Liangliang, Xu Shuhua, Hu Landian, Hurst Laurence D., and Kong Xiangyin. Identification of Two Maternal Transmission Ratio Distortion Loci in Pedigrees of the Framingham Heart Study. *Scientific Reports*, 3, July 2013.

Lynch Michael. *The Origins of Genome Architecture*. Sinauer Associates Inc, 1 edition, March 2007.

Lynch Michael. Rate, molecular spectrum, and consequences of human mutation. *Proceedings of the National Academy of Sciences*, 107(3):961–968, January 2010.

Lynch Michael, Sung Way, Morris Krystalynne, Coffey Nicole, Landry Christian R., Dopman Erik B., Dickinson W. Joseph, Okamoto Kazufusa, Kulkarni Shilpa, Hartl Daniel L., and Thomas W. Kelley. A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proceedings of the National Academy of Sciences*, 105(27):9272–9277, July 2008.

Mancera Eugenio, Bourgon Richard, Brozzi Alessandro, Huber Wolfgang, and Steinmetz Lars M. High-resolution mapping of meiotic crossovers and non-crossovers in yeast. *Nature*, 454(7203):479–485, July 2008.

Mayer WE, Herrmann M, and Sommer RJ. Phylogeny of the nematode genus *Pristionchus* and implications for biodiversity, biogeography and the evolution of hermaphroditism. *BMC Evolutionary Biology*, 7:104, 2007.

Mayer WE, Herrmann M, and Sommer RJ. Molecular phylogeny of beetle associated diplogastrid nematodes suggests host switching rather than nematode-beetle coevolution. *BMC Evolutionary Biology*, 9:212, 2009.

Mayer WE, Schuster LN, Bartelmes G, Dieterich C, and Sommer RJ. Horizontal gene transfer of microbial cellulases into nematode genomes is associated with functional assimilation and gene turnover. *BMC Evolutionary Biology*, 11:13, 2011.

McGaugh Suzanne E., Heil Caiti S. S., Manzano-Winkler Brenda, Loewe Laurence, Goldstein Steve, Himmel

Tiffany L., and Noor Mohamed A. F. Recombination Modulates How Selection Affects Linked Sites in Drosophila. *PLoS Biol*, 10(11):e1001422+, November 2012.

Molnar RI. *PhD Thesis*. PhD thesis, 2012.

Morgan K, McGaughran A, Villate L, Herrmann M, Witte H, Bartelmes G, Rochat J, and Sommer RJ. Multi locus analysis of *Pristionchus pacificus* on La Réunion Island reveals an evolutionary history shaped by multiple introductions, constrained dispersal events and rare out-crossing. *Molecular Ecology*, 21: 250–266, 2012.

Mukai T. The genetic structure of natural populations of Drosophila melanogaster. *Genetics*, 50:1–19, July 1964.

Muller H J. Further Studiss of the nature and causes of gene mutations. In *Proceedings of the Sixth International Congress of Genetics*, pages 213–255.

Myers Simon, Bottolo Leonardo, Freeman Colin, McVean Gil, and Donnelly Peter. A Fine-Scale Map of Recombination Rates and Hotspots Across the Human Genome. *Science*, 310(5746):321–324, October 2005.

Myers TR and Greenwald I. Wnt signal from multiple tissues and lin-3 / EGF signal from the gonad maintain vulval precursor cell competence in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences*, 104:20368–20373, 2007.

Ogawa A, Streit A, Antebi A, and Sommer RJ. A Conserved Endocrine Mechanism Controls the Formation of Dauer and Infective Larvae in Nematodes. *Current Biology*, 19:67–71, 2009.

Ogawa A, Bento G, Bartelmes G, Dieterich C, and Sommer RJ. *Pristionchus pacificus daf-16* is essential for dauer formation but dispensable for mouth form dimorphism. *Development*, 138:1281–4, 2011.

O'Rourke Sean, Wietzorrek Andreas, Fowler Kay, Corre Christophe, Challis Greg L., and Chater Keith F. Extracellular signalling, translational control, two repressors and an activator all contribute to the regulation of methylenomycin production in Streptomyces coelicolor. *Molecular Microbiology*, 71(3): 763–778, February 2009.

Ossowski Stephan, Schneeberger Korbinian, Lucas-Lledó José I., Warthmann Norman, Clark Richard M., Shaw Ruth G., Weigel Detlef, and Lynch Michael. The Rate and Molecular Spectrum of Spontaneous Mutations in Arabidopsis thaliana. *Science*, 327(5961):92–94, January 2010.

Ptak Susan E., Hinds David A., Koehler Kathrin, Nickel Birgit, Patil Nila, Ballinger Dennis G., Przeworski Molly, Frazer Kelly A., and Paabo Svante. Fine-scale recombination patterns differ between chimpanzees and humans. *Nat Genet*, 37(4):429–434, April 2005.

Rae Robbie, Riebesell Metta, Dinkelacker Iris, Wang Qiong, Herrmann Matthias, Weller Andreas M., Dieterich Christoph, and Sommer Ralf J. Isolation of naturally associated bacteria of necromenic Pristionchus nematodes and fitness consequences. *Journal of Experimental Biology*, 211(12):1927–1936,

June 2008.

Riddle DL and Albert P. Genetic and environmental regulation of dauer larva development. In Riddle DL, Blumenthal T, Meyer BJ, and Preiss JR, editors, *C. elegans II*, pages 739–768. Cold Spring Harbor Laboratory Press, New York, 1997.

Rödelsperger C and Sommer RJ. Computational archaeology of the *Pristionchus pacificus* genome reveals evidence of horizontal gene transfers from insects. *BMC Evolutionary Biology*, 11:239, 2011.

Rödelsperger Christian, Streit Adrian, and Sommer Ralf J. Structure, Function and Evolution of The Nematode Genome. 2013.

Roesti Marius, Moser Dario, and Berner Daniel. Recombination in the threespine stickleback genomeâpatterns and consequences. *Mol Ecol*, 22(11):3014–3027, June 2013.

Rudel D, Tian H, and Sommer RJ. Wnt signaling in *Pristionchus pacificus* gonadal arm extension and the evolution of organ shape. *Proceedings of the National Academy of Sciences*, 105:10826–10831, 2008.

Schlager B, Wang X, Braach G, and Sommer RJ. Molecular cloning of a dominant roller mutant and establishment of DNA-mediated transformation in the nematode *Pristionchus pacificus*. *Genesis*, 47: 300–304, 2009.

Seidel HS, Rockman MV, and Kruglyak L. Widespread genetic incompatibility in *C. elegans* maintained by balancing selection. *Science*, 319:589–94, 2008.

Seyfert Amanda L., Cristescu Melania E. A., Frisse Linda, Schaack Sarah, Thomas W. Kelley, and Lynch Michael. The Rate and Spectrum of Microsatellite Mutation in Caenorhabditis elegans and Daphnia pulex. *Genetics*, 178(4):2113–2121, April 2008.

Sinha A, Sommer RJ, and Dieterich C. Divergent gene expression in the conserved dauer stage of the nematode *Pristionchus pacificus* and *Caenorhabditis elegans*. *BMC Genomics*, submitted, 2012.

Sommer RJ. As good as they get: cells in nematode vulva development. *Current Opinion in Cell Biology*, 13: 715–720, 2001.

Sommer RJ and Sternberg PW. Evolution of Nematode Vulval Fate Patterning. *Developmental Biology*, 173: 396–407, 1996.

Sommer RJ, Carta K, Kim SY, and Sternberg PW. Morphological, genetic and molecular description of *Pristionchus pacificus* sp n (Nematoda: Neodiplogastridae). 19:511–521, 1996a.

Sommer RJ, Carta LK, Kim S-Y, and Sternberg PW. Morphological, genetic and molecular description of *Pristionchus pacificus* sp. n. (Nematoda : Neodiplogastridae). *Fundamental and Applied Nematology*, 19: 511–521, 1996b.

Spencer Chris C. A., Deloukas Panos, Hunt Sarah, Mullikin Jim, Myers Simon, Silverman Bernard, Donnelly Peter, Bentley David, and McVean Gil. The Influence of Recombination on Human Genetic Diversity.

PLoS Genet, 2(9):e148+, September 2006.

Sudhaus W and Fürst von Lieven A. A phylogenetic classification and catalogue of the Diplogastridae (Secernentea, Nematoda). *Journal of Nematode Morphology and Systematics*, 6:43–89, 2003.

The *C. elegans* Sequencing Consortium . Genome Sequence of the Nematode *C. elegans*: A Platform for Investigating Biology. *Science*, 282:2012–2018, 1998.

Tian H, Schlager B, Xiao H, and Sommer RJ. Wnt signaling induces vulva development in the nematode *Pristionchus pacificus*. *Current Biology*, 18:142–146, 2008.

Wang X and Sommer RJ. Antagonism of LIN-17/Frizzled and LIN-18/Ryk in Nematode Vulva Induction Reveals Evolutionary Alterations in Core Developmental Pathways. *PLoS Biology*, 9:e1001110, 2011.

Wei Jun-Zhi, Hale Kristina, Carta Lynn, Platzer Edward, Wong Cynthie, Fang Su-Chiung, and Aroian Raffi V. Bacillus thuringiensis crystal proteins that target nematodes. *Proceedings of the National Academy of Sciences*, 100(5):2760–2765, March 2003.

Weller AM, Mayer WE, Rae R, and Sommer RJ. Quantitative assessment of the nematode fauna present on Geotrupes dung beetles reveals species-rich communities with a heterogeneous distribution. *The Journal of Parasitology*, 96:525–531, 2010.

Wu Yonghui, Bhat Prasanna R., Close Timothy J., and Lonardi Stefano. Efficient and Accurate Construction of Genetic Linkage Maps from the Minimum Spanning Tree of a Graph. *PLoS Genet*, 4(10):e1000212+, October 2008.

Yang Sihai, Yuan Yang, Wang Long, Li Jing, Wang Wen, Liu Haoxuan, Chen Jian-Qun, Hurst Laurence D., and Tian Dacheng. Great majority of recombination events in Arabidopsis are gene conversion events. *Proceedings of the National Academy of Sciences*, 109(51):20992–20997, December 2012.

Zheng M, Messerschmidt D, Jungblut B, and Sommer RJ. Conservation and diversification of Wnt signaling function during the evolution of nematode vulva development. *Nature Genetics*, 37:300–304, 2005.

Articles

J. Parasitol., 96(3), 2010, pp. 525–531

QUANTITATIVE ASSESSMENT OF THE NEMATODE FAUNA PRESENT ON *GEOTRUPES* DUNG BEETLES REVEALS SPECIES-RICH COMMUNITIES WITH A HETEROGENEOUS DISTRIBUTION

Andreas M. Weller, Werner E. Mayer, Robbie Rae, and Ralf J. Sommer*

Max-Planck Institute for Developmental Biology, Department for Evolutionary Biology, Spemannstrasse 37, D-72076 Tübingen, Germany. *e-mail: ralf.sommer@tuebingen.mpg.de*

ABSTRACT: *Pristionchus* spp. nematodes exhibit several traits that might serve as pre-adaptations to parasitism. Under harsh environmental conditions, these nematodes can arrest development and form dauer larvae. In addition, they have been shown to live in necromenic association with a range of beetles, including dung beetles (*Geotrupes stercorosus*) on which, for example, *Pristionchus entomophagus* is commonly found. It has been argued that the formation of dauer larvae and the association with invertebrates represent intermediate steps towards parasitism. To better understand necromenic associations, and to gain information on *Pristionchus* spp. abundance and the general species composition on dung beetles, we extracted all the nematode fauna present on 114 individuals of *G. stercorosus*. By direct sequencing using the 18S *SSU*, we provide a barcode for all nematodes isolated from the beetle samples. In total, 5,002 dauer-stage nematodes were sequenced, which included *Pristionchus* spp., *Koerneria* spp. (Diplogastridae), *Pelodera* spp. (Rhabditidae), and Strongyloidea as well as Spirurida. Intensities of infection varied from over 1,000 nematodes isolated from a single *G. stercorosus* to none, with *Pelodera* spp. being the most abundant group isolated. This study presents the first quantitative data on the *Pristionchus* spp. infection of beetles.

Studies of nematodes from the diplogasterid genus *Pristionchus* have revealed both a wealth of genetic and genomic techniques and a good understanding of ecology. Physical and genetic linkage maps and a fully sequenced genome, as well as transgenic techniques, are available (Dieterich et al., 2008; Schlager et al., 2009). There are more than 150 *P. pacificus* isolates collected from around the world and 25 *Pristionchus* species are in culture. With these attributes, *Pristionchus pacificus* serves as a model system in evolutionary developmental biology (evo-devo) and for comparison with *Caenorhabditis elegans* (Zheng et al., 2005; Tian et al., 2008).

More recently, *P. pacificus* has been used for evolutionary ecology and population genetic studies (Herrmann et al., 2007; Zauner et al., 2007). In the natural environment, most *Pristionchus* species are found on beetles in a species-specific manner. For example, *Pristionchus maupasi* and *Pristionchus entomophagus* are found on cockchafers (*Melolontha* sp.) and dung beetles (*Geotrupes* sp.), respectively (Herrmann et al., 2006a), while *P. pacificus* and *Pristionchus uniformis* are predominantly found on the oriental beetle (*Exomala orientalis*) and the Colorado potato beetle (*Leptinotarsa decemlineata*), respectively (Herrmann et al., 2006a, 2007). These nematodes are thought to have a necromenic relationship with their hosts (Kiontke and Sudhaus, 2006), whereby they infest the beetle and, upon beetle death, they feed on a variety of bacteria, fungi, and other nematodes that proliferate on the carcass (Rae et al., 2008). In addition, *Pristionchus* species display unique chemoattraction profiles when exposed to insect and plant semiochemicals (Hong and Sommer, 2006; Herrmann et al., 2007; Hong et al., 2008).

These studies also add to our understanding of the evolution of parasitism, as *Pristionchus* spp. display several traits that may function as pre-adaptations to a parasitic lifestyle, e.g., the formation of dauer larvae, high toxicity tolerance, and low oxygen tolerance (Dieterich and Sommer, 2009). Pre-adaptations are defined as adaptations of an organism to its current environment that might be co-opted for new functions in the process of acquiring a different niche during the course of evolution (Osche, 1956). The importance of pre-adaptations in the evolution of parasitism has been discussed for many years by multiple authors, largely on hypothetical grounds (Osche 1956; Anderson, 1984; Poulin, 2007). Still, there are controversies on the term pre-adaptation and in other scientific disciplines, such as evolutionary psychology, the word "co-opted adaptation" has been proposed instead of "pre-adaptation" (Buss et al., 1998). As experimental data have remained scarce, *Pristionchus* spp. nematodes could emerge as a model system, in this context, due to their combination of genetic tools and a necromenic host association (Dieterich and Sommer, 2009; Ogawa et al., 2009).

The development of molecular tags, e.g., sequencing of 18S small subunit ribosomal RNA gene (*SSU*), provides a means for rapid analysis of the molecular operational taxonomic units (MOTU) present in nematode communities (Blaxter et al., 1998; Floyd et al., 2002; Griffiths et al., 2006). Using a similar molecular approach, we profiled the natural, resident nematode community present on dung beetles *Geotrupes* (*Anoplotrupes*) *stercorosus* L. from the Schönbuch Forest in Tübingen, Germany. These beetles are highly abundant in the forest habitat and are hosts for a large number of nematode species including *Pristionchus* spp., *Koerneria* spp., and *Pelodera* spp. (Kühne, 1996; Herrmann et al., 2006a). The dung beetle family has 3 subfamilies (Aphodiinae, Geotrupinae, and Scarabaeinae) (Janssens, 1960) and can be grouped into dwellers, tunnellers, or rollers (Cambefort and Hanski, 1991) according to dung life-style. The Geotrupinae are tunnellers and dig a vertical tunnel beneath the dung heap, then move dung to the shaft base (D'hondt et al., 2008). Larvae can remain in pupal cells for an average of 5–6 mo for *G. spiniger* and 9–10 mo for *G. vernalis* (Kühne, 1996), giving ample time for nematode colonization.

The main aim of the present study was to gather information on the number and life stage of *Pristionchus* spp. individuals on *G. stercorosus* to provide a qualitative assessment of the necromenic beetle association. Additionally, we wanted to identify and quantify the rest of the nematode community.

Received 2 September 2009; revised 11 November 2009, 7 December 2009; accepted 21 January 2010.
*To whom correspondence should be addressed.
DOI: 10.1645/GE-2319.1

TABLE 1. Number of nematodes isolated from 114 *Geotrupes stercorosus* individuals.

Beetle no.	Sex	*Pristionchus*	*Koerneria*	Diplogasterida	*Pelodera*	Spirurida	Other	Total worms
1	f	0	0	0	3	1		4
2	m	0	0	0	3	1		4
3	f	0	0	0	0	1		1
4	m	1	0	0	19	0		20
5	m	0	0	0	3	0	*Rhabditophanes* (4)	7
6	f	0	0	0	9	0		9
7	m	0	0	0	5	0		5
8	m	0	0	0	15	0	*Rhabditophanes* (1)	16
9	m	0	0	0	1	0		1
10	f	0	0	0	0	5	*Rhabditophanes* (2)	7
11	f	0	0	0	0	0		0
12	f	1	0	0	0	0		1
13	m	0	0	0	4	10		14
14	f	0	0	0	0	0		0
15	f	0	0	0	6	2		8
16	f	0	0	0	1	0		1
17	m	0	0	0	4	2	*Fictor* (1)	7
18	f	0	0	0	0	2		2
19	m	0	0	0	1	1		2
20	f	0	0	0	1	0		1
21	m	0	0	0	0	0		0
22	m	0	0	0	3	0		3
23	f	1	0	0	0	0		1
24	f	0	0	0	0	1		1
25	m	0	0	0	0	1		1
26	f	0	1	0	3	0		4
27	m	0	0	0	0	0		0
28	f	0	0	0	0	0		0
29	m	0	1	0	0	0		1
30	m	0	0	0	0	0		0
31	f	0	0	0	0	0		0
32	f	0	0	0	2	0		2
33	f	0	0	0	0	1		1
34	m	0	0	0	0	0		0
35	m	0	0	0	0	0		0
36	f	0	0	0	0	0		0
37	f	0	0	0	0	0		0
38	f	1	0	0	0	0		1
39	m	0	0	0	1	2		3
40	m	0	0	0	7	0		7
41	m	0	0	0	0	0		0
42	m	0	0	0	1	0		1
43	m	0	0	0	30	0		30
44	m	0	0	1	110	0		111
45	f	0	0	0	0	0		0
46	f	0	0	1	0	3		4
47	f	0	0	1	1	3		5
48	f	0	0	1	2	3		6
49	f	0	0	1	11	29		41
50	f	0	0	1	85	8		94
51	f	0	0	16	282	0	*Pellioditis* (1), *Oscheius* (1)	300
52	f	1	0	12	515	0	*Heterorhabditis* (1)	529
53	f	0	0	0	0	0		0
54	f	0	0	0	0	0		0
55	f	0	0	1	2	0		3
56	f	0	0	6	0	0		6
57	f	0	0	0	15	0		15
58	m	0	0	0	0	0		0
59	m	0	0	0	0	20		20
60	m	0	1	2	0	34		37
61	m	0	0	0	0	0		0

Table 1. Continued.

Beetle no.	Sex	*Pristionchus*	*Koerneria*	Diplogasterida	*Pelodera*	Spirurida	Other	Total worms
62	f	0	0	0	0	0		0
63	f	0	0	1	0	0		1
64	f	0	0	1	9	0		10
65	f	0	0	0	50	0		50
66	f	0	0	0	0	0		0
67	f	0	0	0	0	14		14
68	m	0	0	0	0	0		0
69	f	0	0	0	0	0		0
70	f	0	0	0	0	12		12
71	f	7	0	0	2	0		9
72	f	2	0	0	5	0		7
73	f	0	0	0	0	0		0
74	f	0	0	0	16	0		16
75	f	0	0	380	0	0		380
76	f	0	0	0	9	0		9
77	f	0	2	0	34	0		36
78	m	0	2	1	0	0		3
79	f	0	0	0	0	7		7
80	m	0	0	0	6	0		6
81	f	0	2	0	0	43		45
82	m	0	1	1	480	0		482
83	f	0	2	0	7	0		9
84	f	0	0	0	0	40		40
85	f	0	0	0	0	10		10
86	f	1	0	0	41	0		42
87	f	0	0	0	4	3		7
88	f	0	0	0	0	0		0
89	m	0	0	0	2	0		2
90	f	0	0	0	5	11		16
91	f	0	0	0	0	0		0
92	f	0	4	0	0	19		23
93	f	0	4	0	0	0		4
94	f	1	0	0	103	0		104
95	f	0	3	32	1522	0		1557
96	f	0	4	0	207	14		225
97	m	0	5	1	0	0		6
98	m	0	2	2	1	12		17
99	f	0	1	0	0	0		1
100	m	0	3	0	1	40		44
101	f	0	1	0	1	0	Strongyloida (1)	3
102	f	0	1	0	0	46		47
103	f	0	4	0	0	0	Strongyloida (1)	5
104	f	0	3	0	3	8		14
105	f	0	3	1	8	17		29
106	f	1	0	0	191	0		192
107	f	0	7	0	32	0		39
108	m	0	2	0	2	7		11
109	m	0	3	0	0	63		66
110	f	0	2	1	0	0		3
111	f	0	0	0	36	11		47
112	m	0	2	1	1	0		4
113	m	0	4	0	2	1		7
114	f	0	1	1	2	0		4
Total		17	71	466	3,927	508		5,002
Mean		0.15	0.63	4.12	34.73	4.49		44.23
SD		1.74	6.75	56.33	399.64	48.60		495.13

MATERIALS AND METHODS

Geotrupes stercorosus sampling regime

In our first experiment, we collected 100 *G. stercorosus* beetles from 3- to 4-day-old horse dung from the Schönbuch Forest, Tübingen, Germany (GPS: N 48°32′ 29.91″, O 9° 1′ 15. 3″). These beetles were used to test the efficiency of the nematode extraction methods employed. For more detailed analysis, a total of 114 *G. stercorosus* individuals were collected on 28 May, 2 June, 11 September, and 9 October 2008. In all experiments, *G. stercorosus* beetles were stored in non-airtight plastic boxes, taken back to the laboratory and immediately killed by cutting in 2 halves with sterile scissors before nematode extraction. All nematodes present were extracted and identified using direct sequencing.

Nematode extraction via Baermann funnels

To isolate *Pristionchus* spp. nematodes, beetles are usually cut transversally and placed on nematode growth media (NGM) agar plates (for details see, Herrmann et al., 2006a). This allows simple detection of *Pristionchus* spp. nematodes within 2 to 10 days, but does not record number, species, or life stage of the nematode community originally present. Therefore, we used Baermann funnels to extract nematodes from the beetles.

Genitalia of all *G. stercorosus* were removed prior to nematode extraction, as they typically contain thousands of *Koerneria* sp. dauer larvae (Kühne, 1995) and would make any sequencing effort impossible. *Geotrupes stercorosus* beetles were cut into pieces of approximately 2 × 2 mm, using scissors, and placed in modified Baermann funnels (Hooper, 1986) filled with M9/Triton X-100 (0.2% v/v) for 12–15 hr. All nematodes were then individually placed into single worm lysis buffer for PCR and *SSU* sequencing (Floyd et al., 2002; Herrmann et al., 2006a).

Efficacy of nematode extraction using Baermann funnels or agar plates

To compare the efficacy of the modified Baermann technique with the standard *Pristionchus* spp. extraction method, 100 *G. stercorosus* beetles were either killed and placed on 6-cm NGM agar plates or homogenized and processed via Baermann funnels. After 12 hr, the Baermann extract was placed on separate agar plates and all plates were monitored for 14 days. Nematodes that morphologically resembled species in the diplogasterid family, e.g., they lacked a pharyngeal grinder and striations on cuticle, were picked, and species were determined via *SSU* sequencing.

Nematode species identification using 18S *SSU*

Genomic DNA from individual nematodes was isolated using the NaOH digestion method by Floyd et al. (2002). Briefly, single worms were added to 20 µl of 0.25 M NaOH and incubated at 22 C overnight. The worm mixture was then heated to 99 C for 3 min before the addition of 4 µl of 1 M HCl, 10 µl of 0.5 M Tris-HCl (pH 8.0), and 5 µl of 2% Triton X-100. The mixture was then heated to 99 C for 3 min, frozen to −20 C, and then heated for a further 3 min at 99 C. Two µl of the extract were then used for PCR. DNA was amplified using the primers SSU18A (5′-AAGATTAAGCCATGCATG-3′) and SSU26R (5′-CATTCTTGG-CAAATGCTTTCG-3′). PCR was carried out in 25-µl reactions containing 2.5 mM of $MgCl_2$, 0.16 mM of each deoxynucleoside triphosphate, 0.5 µM of each primer, 2 µl of the lysate, and 2 units of *Taq* DNA polymerase (Amersham Biosciences, Piscataway, New Jersey). The mixture was then subjected to the following PCR conditions: 2 min at 95 C, 35 cycles including 15 sec at 95 C, 50 C for 15 sec, 72 C for 2 min, followed by 7 min at 72 C. PCR products were then diluted 10-fold and added to the BigDye® terminator sequencing mix (Applied Biosciences, Darmstadt, Germany) which contained the sequencing primer SSU9R (5′-AGCTGGAATTACCGCGGCTG-3′). Gene sequences of nematodes were aligned using SeqMan (DNASTAR, Inc., Madison, Wisconsin) and compared with GenBank database sequences using Blast searches using sequence similarity matches higher than 90% (Altschul et al., 1997).

Data analysis

Data were analyzed by using JMP 5.1, SAS, and Microsoft Excel. Efficiency of the 2 nematode extraction techniques was compared using Pearson's chi-square test. Differences between numbers of nematodes collected from the 114 *G. stercorosus* individuals were analyzed using each paired Student's *t*-test. Species composition on female and male *G. stercorosus* beetles was analyzed using a Wilcoxon rank-sum test. Statistical significance was defined as having a *P*-value less than 0.05.

RESULTS

Efficiency of *Pristionchus* spp. extraction using Baermann funnels or agar plates

From 100 *G. stercorosus* individuals collected on the same date, and randomly assigned for nematode extraction via agar plate or Baermann funnel, 24 of 50 were positive for *Pristionchus* using agar plates compared to 16 of 50 *G. stercorosus* undergoing Baermann extraction, which is not significantly different ($P = 0.1025$).

Identification and analysis of nematodes present on *G. stercorosus*

In total, 5,002 nematodes from 114 *G. stercorosus* individuals were obtained using the Baermann funnel procedure. Seventy-eight percent of the collected beetles contained nematodes (Table I). All nematodes that emerged were in the dauer stage, with the exception of an obligate parasitic species (Spirurida).

The *SSU* of all individuals was sequenced. The nematodes were grouped into 5 categories: *Koerneria* spp. (Diplogastridae), *Pristionchus* spp. (Diplogastridae), diplogastrid nematodes from other genera, *Pelodera* spp. (Rhabditida), and obligate parasitic nematodes (Spirurida). For percentages of infested beetles per group, see Figure 1.

Pelodera spp. were the most abundant nematode found on *G. stercorosus* and differed significantly in number from the other groups ($P < 0.01$). There were no significant differences between numbers of *Koerneria* spp., *Pristionchus* spp., and Spirurida ($P > 0.5$). Infestation by different nematodes varied greatly among single beetles, ranging from no infections to more than 1,500 individuals per beetle. For each group, we also determined whether there was a preference in the numbers of nematodes found on either male or female *G. stercorosus*. However, we found no significant differences between the numbers of diplogastrid spp. ($\chi^2 = 0.2290$, df = 1, $P = 0.632$), *Koerneria* spp. ($\chi^2 = 0.3428$, df = 1, $P = 0.558$), *Pelodera* spp. ($\chi^2 = 0.0130$, df = 1, $P = 0.909$), Spirurida spp. ($\chi^2 = 0.0527$, df = 1, $P = 0,818$) or *Pristionchus* spp. ($\chi^2 = 2.8669$, df = 1, $P = 0.0904$).

DISCUSSION

This study presents the first quantitative assessment of the nematode fauna present on *Geotrupes* spp. dung beetles and reveals species-rich communities with a heterogeneous distribution. In particular, *Pristionchus* spp. nematodes and other phoretic and necromenic nematodes were observed on *Geotrupes* spp. While this study cannot directly address the evolution of parasitism, it quantifies necromenic associations of nematodes. Seventy-eight percent of the collected beetles contained nematodes, showing that *G. stercorosus* is an important and common host for a diverse range of nematode species. We found high variability in the numbers of nematodes infesting dung beetles, possibly due to the amount of time individual beetles spent buried in horse dung prior to being collected.

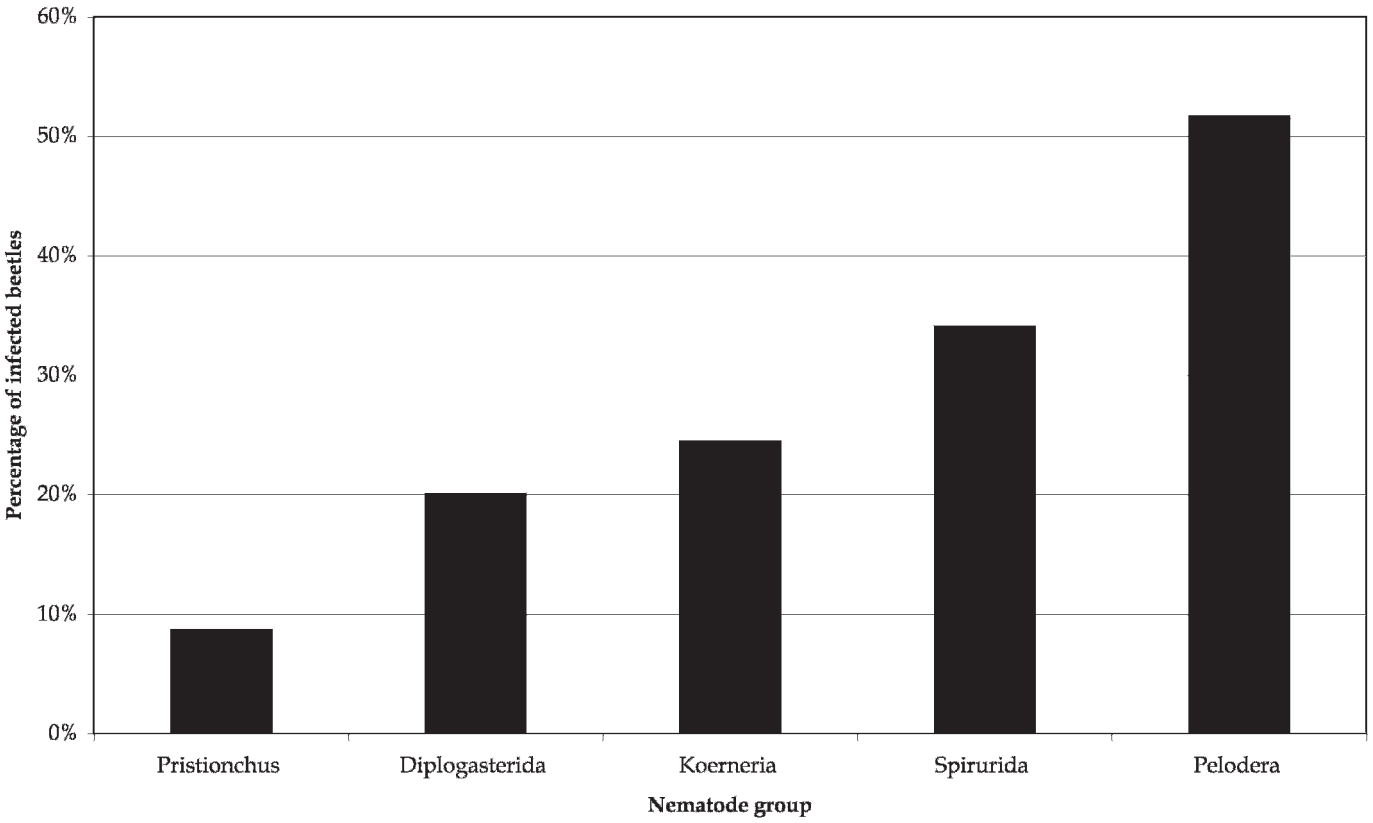


FIGURE 1. Percentage of beetles infected by each nematode group.

Pristionchus spp. were isolated in low abundance, with a maximum of 7 individuals present on a single beetle. Species of nematodes found included *Pristionchus* sp. 6, *P. entomophagus*, and *Pristionchus lheritieri*, which is in accordance with Herrmann et al. (2006), who found the same 3 nematode species on *G. stercorosus* from the same location. By using the agar plate assay, it was previously difficult to determine what life stage of *Pristionchus* spp. existed in the natural environment. This was primarily because this extraction process relies on the discovery of adult nematodes. Using the Baermann funnel technique, we extracted *Pristionchus* spp. from *G. stercorosus* in the dauer stage only and not in any other life stage. This finding stresses the importance of the dauer stage as a prerequisite for an extended association with host insects, which is regarded as a necessary condition for the evolution of parasitism (Osche, 1956; Dieterich and Sommer, 2009).

We sampled *G. stercorosus* from horse dung, which is a patchily distributed, rapidly changing habitat that exists for only a limited time (Kiontke, 1996). It is, therefore, important for nematodes to move from these areas to other areas more suitable for feeding and reproduction. The majority of the nematodes isolated from our study are known to utilize insects as transport to, and from, dung heaps. *Pelodera coarctata* is commonly found in dung, where it reproduces until the food source is depleted and then uses dung beetles (*Aphodius* spp.) as a phoretic carrier (Matthews, 1998). *Pelodera strongyloides* is also found in animal manure, but is thought to have a phoretic relationship with mammals. Kiontke (1996) studied the nematode *Diplogaster coprophila*, which is a phoretic nematode that uses sepsid flies for transport from exhausted dung heaps and is one of the most abundant nematode species in partially decomposed cow dung. In an in-depth study of *Geotrupes* spp. beetles and brood balls from laboratory and field, Kühne (1996) found *Diplogaster* spp., *Diplogasteroides* spp., *Tylopharynx* spp., *Diploscapter* spp., *Rhabditis* spp., *Acrobeloides* spp., and *Aphelenchoides* spp. In our study, we also encountered strongyloid and spirurid species, which are obligate parasitic nematodes. Upon dissection of beetles, these obligate parasites were found in the gut of *G. stercorosus* (data not shown). Most spirurid nematodes have at least 1 paratenic invertebrate host, with reproduction taking place in a final, vertebrate host. The members of the Spirurida found in this study, however, were all in the adult stage. Therefore, we assume that, in this case, *G. stercorosus* is the final host. When *G. stercorosus* is killed and placed on agar plates, their nematodes frequently go unnoticed because they do not exit the beetle carcass and die within a few days, which is a disadvantage of using agar plates as a nematode extraction technique.

In addition, we recorded species of *Rhabditophanes* (Rhabditida), *Oscheius* (Rhabditida), and *Pellioditis* (Rhabditida), but in very low numbers. These nematodes are thought to be associated with insects and have been isolated from pseudoscorpions, hemipterans, and dipterans (Stock and Camino, 1991; Curcic et al., 2004; Stock et al., 2005). We also found a species of *Heterorhabditis* (Heterorhabditida), which is an entomopathogenic nematode that causes mortality to insects by use of its symbiotic bacteria (*Photorhabdus* spp.) (Forst et al., 1997). Although these latter nematodes are lethal parasites for many insect species, they are known to use invertebrates such as

earthworms, isopods, and pine weevils as phoretic carriers (Eng, 2005; Campos-Herrera et al., 2006; Kruitbos et al., 2009). The detection of these genera and species shows the advantage of the molecular identification method used, as they might have easily been overlooked by morphological approaches.

The numbers of each nematode species isolated from *G. stercorosus* varied considerably, especially among species of *Pelodera*, with numbers ranging from zero up to 1,522 individuals on a single beetle. We also found that the *Koerneria* sp. is highly abundant in the genital chamber of dung beetles, with populations of more than 1,000 individuals present (data not shown). The genital region of *Geotrupes* spp. seems to be an area where nematodes commonly accumulate. Similarly, Kühne (1996) found a maximum of 2,000 *Diplogaster henrichae* and *Diplogaster hirschmannae* dauer larvae in the genital region of *G. spiniger*. In much smaller numbers, we found up to 63 spirurids on a single *G. stercorarius*, presumably an important host for these nematodes.

Pristionchus spp. are present on dung beetles, but only in very low numbers, alongside a diverse community of hundreds of parasitic, phoretic, and necromenic nematodes. All *Pristionchus* spp. were found in their dauer stage, the life stage best suited to life on beetles. Often, the emergence of *Pristionchus* spp. nematodes from beetles killed in the laboratory can only be seen after several days (Herrmann et al., 2006a). This observation, together with the small number of *Pristionchus* spp. dauer larvae on dung beetles, suggests that these nematodes emerge relatively late in the decomposition of the beetle. Future studies, using a combination of laboratory-based genetic manipulation and field studies as described here, may provide the opportunity to examine the *Pristionchus* spp.–host relationship in more detail and to explore the extent of possible pre-adaptations to parasitism.

ACKNOWLEDGMENTS

We thank Iris Dinkelacker and Gabi Bartelmes for help in sequencing nematodes, as well as Drs. Dan Bumbarger and Adrian Streit for critically reading the manuscript. This study was funded by the Max Planck Society.

LITERATURE CITED

Altschul S. F, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Research **25:** 3389–3402.

Anderson, R. C. 1984. The origin of zooparasitic nematodes. Canadian Journal of Zoology **62:** 317–328.

Blaxter, M. L., P. De Ley, J. R. Garey, L. X. Liu, P. Scheldeman, A. Vierstraete, J. R. Vanfleteren, L. Y. Mackey, M. Dorris, L. M. Frisse et al. 1998. A molecular evolutionary framework for the phylum Nematoda. Nature **392:** 71–75.

Buss, D. M, M. G. Haselton, T. K. Shackelford, A. L. Bleske, and J. C. Wakefield. 1998. Adaptations, exaptations and spandrels. American Psychologist **5:** 533–548.

Cambefort, Y., and I. Hanski. 1991. Dung beetle population biology. *In* Dung beetle ecology, I. Hanski, and Y. Cambefort (eds.). Princeton University Press, Princeton, New Jersey, p. 36–50.

Campos-Herrera, R., D. Trigob, and C. Gutiérrez. 2006. Phoresy of the entomopathogenic nematode *Steinernema feltiae* by the earthworm *Eisenia fetida*. Journal of Invertebrate Pathology **92:** 50–54.

Croll, N. A., and B. E. Matthews. 1977. Biology of nematodes. Glasgow, Scotland, xxx p.

Curcic, B. P. M., W. Sudhaus, R. N. Dimitrijevic, V. T. Tomic, and S. B. Curcic. 2004. Phoresy of *Rhabditophanes schneideri* (Butschli) (Rhabditida: Alloionematidae) on pseudoscorpions (Arachnida: Pseudoscorpiones). Nematology **6:** 313–317.

D'hondt, B., B. Bossuyt, M. Hoffmann, and D. Bonte. 2008. Dung beetles as secondary seed dispersers in temperate grassland. Basic and Applied Ecology **9:** 542–549.

Dieterich, C., S. W. Clifton, L. Schuster, A. Chinwalla, K. Delehaunty, I. Dinkelacker, L. Fulton, R. Fulton, J. Godfrey, P. Minx et al. 2008. The *Pristionchus pacificus* genome provides a unique perspective on nematode lifestyle and parasitism. Nature Genetics **40:** 1193–1198.

———, and R. J. Sommer. 2009. How to become a parasite—Lessons from the genome of nematodes. Trends in Genetics **25:** 203–209.

Eng, M. S., E. L. Preisser, and D. R. Strong. 2005. Phoresy of the entomopathogenic nematode *Heterorhabditis marelatus* by a non-host organism, the isopod *Porcellio scaber*. Journal of Invertebrate Pathology **88:** 173–176.

Floyd, R., E. Abebe, A. Papert, and M. Blaxter. 2002. Molecular barcodes for soil nematode identification. Molecular Ecology **11:** 839–850.

Forst, S., B. Dowds, N. Boemare, and E. Stackebrandt. 1997. *Xenorhabdus* and *Photorhabdus* spp.: Bugs that kill bugs. Annual Review of Microbiology **51:** 47–72.

Griffiths, B. S., S. Donn, R. Neilson, and T. J. Daniell. 2006. Molecular sequencing and morphological analysis of a nematode community. Applied Soil Ecology **32:** 325–337.

Herrmann, M., W. E. Mayer, R. L. Hong, S. Kienle, R. Minasaki, and R. J. Sommer. 2007. The nematode *Pristionchus pacificus* (Nematoda: Diplogastridae) is associated with the oriental beetle *Exomala orientalis* (Coleoptera: Scarabaeidae) in Japan. Zoological Science **24:** 883–889.

———, ———, and R. J. Sommer. 2006a. Nematodes of the genus *Pristionchus* are closely associated with scarab beetles and the Colorado potato beetle in Western Europe. Zoology **109:** 96–108.

———, ———, and ———. 2006b. Sex, bugs and Haldane's rule: The nematode genus *Pristionchus* in the United States. Frontiers in Zoology **3:** 14.

Hong, R. L., and R. J. Sommer. 2006. Chemoattraction in *Pristionchus* nematodes and implications for insect recognition. Current Biology **16:** 2359–2365.

———, A. Svatos, M. Herrmann, and R. J. Sommer. 2008. Species-specific recognition of beetle cues by *Pristionchus maupasi*. Evolution & Development **10:** 273–279.

Hooper, D. J. 1986. Extraction of free-living stages from soil. *In* Laboratory methods for work with plant and soil nematodes. Her Majesty's Stationery Office, London, U.K., p. 5–30.

Janssens, A. 1960. Insectes Coléoptères Lamellicornes, Institut Royal des Sciences Naturelles de Belgique, Brussels, Belgium, 411 p.

JMP, Version 5.1, SAS Institute Inc., Cary, North Carolina, 1989–2007.

Kiontke, K. 1996. The phoretic association of *Diplogaster coprophila* Sudhaus and Rehfeld, 1990 (Diplogastridae) from cow dung with its carriers, in particular flies of the family Sepsidae. Nematologica **42:** 354–366.

———, and W. Sudhaus. Ecology of *Caenorhabditis* species (January 09, 2006). WormBook, the online review of *C. elegans* biology. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.37.1. Available at: http://wormbook.org.

Kruitbos, L. M., S. E. Heritage, and M. J. Wilson. 2009. Phoretic dispersal of entomopathogenic nematodes by *Hylobius abietis*. Nematology **11:** 419–427.

Kühne, R. 1996. Relations between free-living nematodes and dung-burying *Geotrupes* spp. (Coleoptera: Geotrupini). Fundamental and Applied Nematology **19:** 263–271.

Matthews, B. E. 1998. An introduction to parasitology. Cambridge University Press, Cambridge, U.K., 192 p.

Microsoft Excel. 2004. Version 11.5.2, Washington, D.C.

Ogawa, A., A. Streit, A. Antebi, and R. J. Sommer. 2009. A conserved endocrine mechanism controls the formation of dauer and infective larvae in nematodes. Current Biology **19:** 67–71.

Osche, G. 1956. Die Präadaption freilebender Nematoden an den Parasitismus. Zoologischer Anzeiger **19:** 391–396.

Poulin, R. 2007. Evolutionary ecology of parasites. Princeton University Press, Princeton, New Jersey, 342 p.

Rae, R., M. Riebesell, I. Dinkelacker, Q. Wang, M. Herrmann, A. Weller, C. Dieterich, and R. J. Sommer. 2008. Isolation of naturally associated bacteria of necromenic *Pristionchus* nematodes

and fitness consequences. Journal of Experimental Biology **211:** 1927–1936.

Schlager, B., X. Wang, G. Braach, and R. J. Sommer. 2009. Molecular cloning of a dominant Roller mutant and establishment of DNA-mediated transformation in the nematode model *Pristionchus pacificus*. Genesis **47:** 300–304.

Stock, S. P., A. M. Caiceda, and P. A. Calatayud. 2005. *Rhabditis* (*Oscheius*) *colombiana* n. sp. (Nematoda: Rhabditidae), a necromenic associate of the subterranean burrower bug *Cyrtomenus bergi* (Hemiptera: Cydnidae) from the Cauca Valley, Colombia. Nematology **7:** 363–373.

———, and N. B. Camino. 1991. *Pellioditis pellio* (Schneider) (Nematoda, Rhabditidae) parasitizing *Scaptia* (*Scaptia*) *Lata* (Guerin-Meneville) (Diptera, Tabanidae). Memorias Do Instituto Oswaldo Cruz **86:** 219–222.

Tian, H., B. Schlager, H. Xiao, and R. J. Sommer. 2008. Wnt signaling by differentially expressed Wnt ligands induces vulva development in *Pristionchus pacificus*. Current Biology **18:** 142–146.

Zauner, H., W. E. Mayer, M. Herrmann, A. Weller, M. Erwig, and R. J. Sommer. 2007. Distinct patterns of genetic variation in *Pristionchus pacificus* and *Caenorhabditis elegans*, two partially selfing nematodes with cosmopolitan distribution. Molecular Ecology **207:** 1267–1280.

Zheng, M., D. Messerschmidt, B. Jungblut, and R. J. Sommer. 2005. Conservation and diversification of Wnt signaling function during the evolution of nematode vulva development. Nature Genetics **37:** 300–304.

The Journal of Experimental Biology 211, 1927-1936
Published by The Company of Biologists 2008
doi:10.1242/jeb.014944

Isolation of naturally associated bacteria of necromenic *Pristionchus* nematodes and fitness consequences

Robbie Rae, Metta Riebesell, Iris Dinkelacker, Qiong Wang, Matthias Herrmann, Andreas M. Weller, Christoph Dieterich and Ralf J. Sommer*

Max-Planck Institute for Developmental Biology, Department for Evolutionary Biology, Spemannstrasse 37, D-72076 Tübingen, Germany

*Author for correspondence (e-mail: ralf.sommer@tuebingen.mpg.de)

Accepted 31 March 2008

SUMMARY

Nematodes and bacteria are major components of the soil ecosystem. Many nematodes use bacteria for food, whereas others evolved specialized bacterial interactions ranging from mutualism to parasitism. Little is known about the biological mechanisms by which nematode–bacterial interactions are achieved, largely because in the laboratory nematodes are often cultured under artificial conditions. We investigated the bacterial interactions of nematodes from the genus *Pristionchus* that have a strong association with scarab beetles. *Pristionchus* has a different feeding strategy than *Caenorhabditis* and meta-genomic 16S sequence analysis of *Pristionchus* individuals showed a diversity of living bacteria within the nematode gut and on the nematode cuticle. Twenty-three different bacterial strains were isolated from three *Pristionchus*–beetle associations and were used to study nematode–bacterial interactions under controlled laboratory conditions. We show a continuum of bacterial interactions from dissemination, to reduction in brood size and nematode mortality caused by bacteria derived from insect hosts. Olfactory discrimination experiments show distinct chemoattraction and fitness profiles of *Pristionchus* nematodes when exposed to different bacteria. For example, *Pristionchus pacificus* avoids *Serratia marcescens* possibly because of pathogenicity. Also, *P. pacificus* avoids *Bacillus thuringiensis* and insect pathogenic bacteria but is resistant to the human pathogens *Staphylococcus aureus* and *Pseudomonas aeruginosa*, unlike *Caenorhabditis elegans*. *Pristionchus* specifically recognize and respond to bacteria that cause ill health. Bringing the nematode–bacterial interaction into the laboratory allows detailed functional studies, including the genetic manipulation of the interaction in both nematodes and bacteria.

Supplementary material available online at http://jeb.biologists.org/cgi/content/full/211/12/1927/DC1

Key words: *Pristionchus pacificus*, *Caenorhabditis elegans*, nematode–bacterial interactions, *Bacillus thuringiensis*, entomopathogenic bacteria.

INTRODUCTION

Nematodes and bacteria are among the most numerous organisms on Earth with numbers of nematodes thought to exceed 1 million per m^2 (Floyd et al., 2002) and numbers of bacterial cells in 1 g of soil thought to be approximately 10^{10} (Faegri et al., 1977). Many nematodes use bacteria only for food but some species have more specialized interactions, ranging from mutualism to parasitism. For example, entomopathogenic nematodes of the genera *Steinernema* and *Heterorhabditis* exhibit a close symbiotic relationship with the bacteria *Xenorhabdus* and *Photorhabdus*, respectively (Forst et al., 1997) and rely on these bacteria to cause mortality to insect hosts. The slug parasitic nematode *Phasmarhabditis hermaphrodita* is thought to depend on the bacterium *Moraxella osloensis* to kill slugs and snails (Wilson et al., 1995a; Wilson et al., 1995b; Tan and Grewal, 2001; Tan and Grewal, 2002). Filarial nematodes such as *Brugia malayi* require the endosymbiotic bacterium *Wolbachia* for development, fertility and survival (Taylor et al., 2005). The combination of the pine wood nematode, *Bursaphelenchus xylophilus*, and strains of *Pseudomonas fluorescens* are thought to be responsible for causing an increase in mortality of pine trees (Han et al., 2003). Marine nematodes from the family Stilbonematinae have a mutualistic relationship with thiotrophic ectosymbiotic bacteria (Nussbaumer et al., 2004).

Although the interaction of the aforementioned nematodes with specific bacteria has been well documented, little is known about laboratory nematode model organisms, such as *Caenorhabditis elegans* and *Pristionchus pacificus*. *C. elegans* is a model organism for many areas of biology (see The *C. elegans* Research Community, 2005), whereas *P. pacificus* has been established as a satellite organism in evolutionary developmental biology (Hong and Sommer, 2006a). Both species have – at least in part – been selected as model organisms because they can be cultured in the laboratory using artificial *Escherichia coli* OP50 as food (Brenner, 1974; Sommer et al., 1996).

Recent studies started to investigate the environment in which *Pristionchus* can be found in nature. Several field studies revealed that *Pristionchus* nematodes have close associations with scarab beetles and the Colorado potato beetle (*Leptinotarsa decemlineata*) (Herrmann et al., 2006a; Herrmann et al., 2006b). For example, *P. pacificus* was isolated from the oriental beetle (*Exomala orientalis*) in Japan and the United States (Herrmann et al., 2007). Biological surveys of beetle-associated *Pristionchus* species have concentrated on Europe, North-America, Japan and South Africa. In total, more than 1200 *Pristionchus* isolates have been obtained from more than 15 000 surveyed beetles. These isolates fall into 18 distinct species with a specific biogeographic pattern (Mayer et al., 2007).

P. pacificus currently represents the only cosmopolitan species (Zauner et al., 2007).

Pristionchus nematodes show a high species specificity with certain beetles. For example, the two European species *Pristionchus maupasi* and *Pristionchus entomophagus* are found on cockchafers (*Melolontha* sp.) and dung beetles (*Geotrupes* sp.), respectively (Herrmann et al., 2006a). Similarly, the Colorado potato beetle, which lives in Europe and North America, is highly infested with *Pristionchus uniformis* (Herrmann et al., 2006b). Chemoattraction studies have shown that different *Pristionchus* species display unique chemoattraction profiles towards insect pheromones and plant volatiles (Hong and Sommer, 2006b) demonstrating the utility of such assays for probing the nematodes' environment under laboratory conditions. *Pristionchus* chemoattraction is highly diverse and is presumably involved in shaping the specific interaction with host beetles.

In general, nematode–insect associations can be categorized as phoretic, necromenic or parasitic (Kiontke and Sudhaus, 2006). *Pristionchus* has a necromenic association with beetles whereby the infective juvenile nematodes enter an insect, wait for the death of the host and then feed on bacteria and fungi that proliferate on the insect carcass. Necromenic associations are typically more specific than phoretic associations, in which nematodes use insects or other invertebrates for transport but not as food. It has been suggested that necromeny represents a pre-adaptation for the evolution of true parasitism because the nematode is exposed to low oxygen levels, high temperatures and toxic host enzymes (Weischer and Brown, 2000).

In the context of the different life-style of *Pristionchus* nematodes, it is important to note that these nematodes show major morphological and physiological adaptations with respect to feeding when compared with *C. elegans* and other rhabditids. *C. elegans* has a grinder in the terminal bulb of the pharynx, which disrupts food bacteria (such as *E. coli* OP50), and under laboratory conditions bacteria are completely lysed (Fig. 1A,B). By contrast, *Pristionchus* nematodes have a pharynx with a metacorpus and a terminal bulb typical for rhabditid nematodes but do not have a grinder (Fig. 1C) (Fürst von Lieven and Sudhaus, 2000; Chiang et al., 2006). *Pristionchus* worms do not completely lyse bacteria and intact cells can be found in the intestine as revealed by transmission electron microscopy (TEM; Fig. 1D). Interestingly, it has been suggested that *Pristionchus* nematodes might be actively involved in bacterial dissemination in the wild (Chantanao and Jensen, 1968; Poinar, 1983).

Here, we begin an analysis of the tritrophic interactions of *Pristionchus* nematodes with bacteria associated with beetles and soil. To gain insight into this relationship we chose the following aims: (1) to identify (using metagenomic and microbiological techniques) bacteria that are associated with *Pristionchus* nematodes isolated from soil and beetle hosts, i.e. *P. maupasi* (from cockchafers), *P. entomophagus* (from dung beetles) and *P. pacificus* (from the oriental beetle); (2) to investigate the effect these bacteria have on a range of *Pristionchus* species using chemotaxis, survival and fecundity assays; (3) to assess survival of *P. pacificus* to human (*Pseudomonas aeruginosa* and *Staphylococcus aureus*) and insect (*Xenorhabdus nematophila*, *Xenorhabdus* sp. and *Photorhabdus luminescens*) pathogens; (4) to examine whether physiological functions are variable between nematode species and are affected by bacteria.

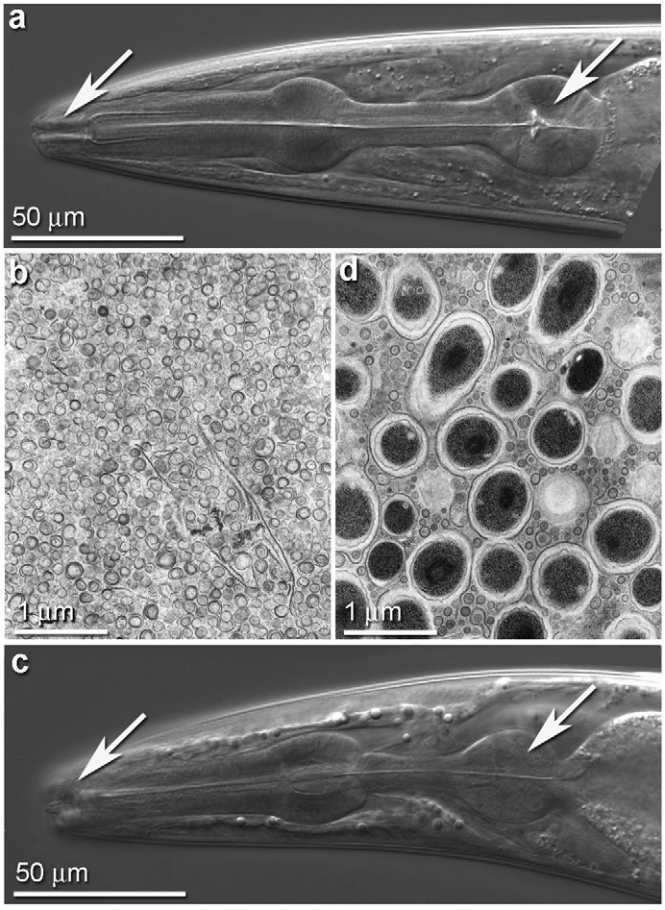


Fig. 1. Pharynx morphology of *Pristionchus pacificus* and *Caenorhabditis elegans*. (A) *C. elegans* pharynx with grinder and long, narrow mouth-like suction pump. (B) *Escherichia coli* OP50 crushed with the *C. elegans* grinder. (C) *P. pacificus* pharynx with no grinder and shorter, broader mouthparts. (D) *E. coli* OP50 is not completely disrupted after passage through the pharynx of *P. pacificus*.

MATERIALS AND METHODS

Nematodes, bacteria strains and culture conditions

Nematode strains used were: *Pristionchus pacificus* (PS0312), *P. entomophagus* (RS0144), *P. maupasi* (RS5015) and *Caenorhabditis elegans* (Bristol N2). Nematodes were grown at 20°C on nematode growth medium (NGM) seeded with *E. coli* OP50 before use in experiments. Bacteria were isolated from beetle nematodes and soil nematodes. Bacteria were maintained on LB agar plates at 30°C. In all experiments bacteria were grown overnight in LB broth at 30°C. Bacteria used in this study were: *Pseudomonas aeruginosa* PA14, *Staphylococcus aureus* Newman, *Photorhabdus luminescens* TT01Tn7GFP (associated with the nematode *Heterorhabditis bacteriophora*), *Xenorhabdus* sp. (associated with *Steinernema scapterisci*), *Xenorhabdus nematophila* (associated with *Steinernema carpocapsae*). Phase variation in *X. nematophila*, *Xenorhabdus* sp. and *P. luminescens* was monitored using NBTA agar (nutrient agar supplemented with Bromothymol Blue and triphenyltetrazolium chloride) and only phase 1 bacteria were used, which are the most virulent. All equipment was sterilized before use and bacteria were only isolated under a laminar flow. To rule out any airborne contaminants 10 LB plates were placed in the laminar flow for 5–8 h and then sealed and stored at 30°C overnight and monitored for colonies the next day. No bacterial colonies or contaminants were recorded, proving that all bacteria isolated were solely from the nematodes.

Isolation of bacteria and nematodes from beetles

Cockchafers (*Melolontha hippocastani*) were collected from Karlsruhe, Germany, Oriental beetles (*Exomala orientalis*) were

collected from Carver, Massachusetts and dung beetles (*Geotrupes* sp.) were collected from the Schönbuch forest, near Tübingen, Germany. Beetles were cut in half transversely with scissors and then placed on 6 cm NGM plates and stored at room temperature. Beetles were inspected daily for 7–14 days for nematodes moving or reproducing on the cadavers. Any nematodes that resembled *Pristionchus* were removed, washed in M9 buffer for 1–2 min and placed on separate LB plates and incubated at room temperature. After 48 h the individual nematodes were removed, lysed and sequenced (see below for methods) to confirm whether the nematodes were the correct species of *Pristionchus* (according to Mayer et al., 2007). Bacteria growing on LB plates were isolated, subcultured and then prepared for sequencing for species identification using PCR amplification of 16S ribosomal RNA genes (Lane, 1991). The bacteria isolated from nematodes from beetles will be referred to as 'beetle-derived' bacteria, hereafter.

We also extracted one adult *P. entomophagus* from a soil sample from the Schönbuch forest, Tübingen and isolated the bacteria that were subsequently excreted from the nematode gut and removed from the nematode cuticle and grew them on an LB plate. The bacteria were then identified (as described below) and will be referred to as 'soil-derived' bacteria.

Bacteria and nematode DNA extraction and PCR amplification

Bacteria of each species were grown overnight in LB broth and DNA was extracted using Aqua Pure genomic DNA kit (Bio-Rad, Hercules, CA, USA). Polymerase chain reaction (PCR) amplification of bacterial 16S rRNA genes was carried out in 20 μl reactions using primer set 27f (5′-AGAGTTTGATCMTGGCTCAG-3′) and 1492r (5′-TACGGYTACCTTGTTACGACTT-3′) (Lane, 1991). Thermal cycling conditions were as follows: 3 min at 95°C followed by 35 cycles of 15 s at 95°C, 30 s at 55°C, 1.5 min at 72°C, and a final step of 8 min at 72°C. A typical reaction contained 2 μl 10× PCR buffer, 2 μl 2 mmol l^{-1} dNTPs, 1 μl 10 μmol l^{-1} 27f, 1 μl 10 μmol l^{-1} 1492r, one unit of Taq DNA polymerase, 12.8 μl H_2O and 1 μl of bacterial DNA. PCR amplicons were visualized by standard agarose gel electrophoresis (Sambrook et al., 1989) and bands were excised using a clean scalpel. DNA was extracted from bands using QIAquick gel extraction kit (Qiagen, Valencia, CA, USA).

After isolation of bacteria, nematodes were removed from the LB plate and identified using the small subunit rRNA gene. Genomic DNA from single nematodes was isolated using the NaOH digestion method of Floyd et al. (Floyd et al., 2002). Briefly, single worms were added to 20 μl of 0.25 mol l^{-1} NaOH and incubated at 25°C overnight. The worm mixture was then heated to 99°C for 3 min before the addition of 4 μl of 1 mol l^{-1} HCl, 10 μl of 0.5 mol l^{-1} Tris-HCl (pH 8.0) and 5 μl of 2% Triton X-100. The mixture was then heated to 99°C for 3 min, frozen to –20°C and then heated for a further 3 min at 99°C. Two microlitres of the extract were then used for PCR. DNA was amplified using the primers SSU18A (5′-AAAGATTAAGCCATGCATG-3′) and SSU26R (5′-CATTCTTGGCAAATGCTTTCG-3′). PCR was carried out in 25 μl reactions containing 2.5 mmol l^{-1} $MgCl_2$, 0.16 mmol l^{-1} each deoxynucleoside triphosphate, 0.5 μmol l^{-1} each primer, 2 μl lysate, 2 units Taq DNA polymerase (Amersham Biosciences, Piscataway, NJ, USA). The mixture was then subjected to the following PCR conditions: 2 min at 95°C, 35 cycles including, 15 s at 95°C, 15 s at 50°C, 2 min at 72°C, followed by 7 min at 72°C. PCR products were then diluted 10–20-fold and added to the Big Dye terminator sequencing mix (Applied Biosciences, Foster City, CA, USA), which contained the sequencing primer SSU9R (5′-AGCTGGAATTACCGCGGCTG-3′). For bacteria we required a minimum length of 200 base pairs for the query sequence (16S rRNA). Gene sequences of nematodes and bacteria were aligned using Seqman (DNA Star, Madison, WI, USA), compared with GenBank database sequences using Blastn searches using sequence similarity matches at 90%.

Metagenomic analysis of bacteria in the *Pristionchus* gut and cuticle

Soil samples were taken from the Schönbuch forest, Tübingen and were added to 9 cm NGM plates and stored at room temperature. The plates were then checked for presence of nematodes every day for the next 7 days. In total, four *P. entomophagus* and four *P. lheritieri* individuals were isolated, washed in M9 buffer and placed in single worm lysis buffer. The resultant suspension was then used for nematode identification (as described above) and bacterial cloning. Bacterial DNA was amplified using conditions and reagents as described above and was then cloned using Topo cloning kit (Invitrogen, Carlsbad, CA, USA) following the manufacturer's guidelines. The full-length gene was then ligated into pCR4-TOPO (Invitrogen) and transformed into Top 10 chemically competent bacterial cells for sequence analysis. In total, 683 clones were picked and screened for bacterial inserts using PCR primers T7 (5′-TAATACGACTCACTATAGGG-3′) and T3 (5′-ATTAACCCTCACTAAAGGGA-3′). Bacterial inserts were then sequenced using the same methods as described above.

Chemotaxis assays

Chemotaxis assays were modified from previous studies (Zhang et al., 2005; Hong and Sommer, 2006b). Briefly, 25 μl of overnight bacterial suspension was placed 0.5 cm away from the edge of a 9 cm Petri dish filled with NGM medium. The same amount of *E. coli* OP50 was placed on the opposing side and acted as the counter attractant. Approximately 50–200 J4/adult stage *Pristionchus* individuals were placed between the two bacterial spots. All nematodes used were previously fed on *E. coli* OP50. Plates were then sealed with Parafilm® and stored at room temperature in the dark. After 24 h the number of nematodes found in each bacterial spot was recorded. A chemotaxis index was used to score the response of the nematodes, which consisted of: number of nematodes in the test bacteria – numbers of nematodes in control bacteria/total number of nematodes counted (Zhang et al., 2005). This gave a chemotaxis score ranging from –1.0 (total revulsion from test bacteria) to 1.0 (total attraction towards test bacteria). A score of around 0 means there were equal numbers of nematodes in each bacterial spot. Five plates were used per replicate, and the procedure was repeated five times for each bacterium (a total of 25 individual assays).

Chemotaxis experiments were as follows: (1) *E. coli* OP50 *versus* soil-derived bacteria; (2) *E. coli* OP50 *versus* insect-derived bacteria; (3) *B. thuringiensis* or *Bacillus* sp. 1 *versus* insect-associated bacteria – this was used to examine the effect of removing *E. coli* from the analysis and using more ecologically relevant controls; (4) *E. coli* OP50 *versus* insect and human pathogenic bacteria; (5) *P. luminescens versus* human and insect pathogenic bacteria.

Survival of *P. pacificus* exposed to bacteria

Liquid cultures of all insect- and soil-derived bacteria as well as human and insect pathogens were grown overnight at 30°C. Bacterial suspensions (200 μl) were spread evenly on 6 cm Petri dishes with NGM medium and incubated overnight. Twenty J4 *P. pacificus* were added to each plate and stored at 25°C. Survival of worms was

monitored daily for 8 days. Nematodes were transferred every 2 days to fresh plates to prevent misidentification of original worms from offspring. Mortality was determined by prodding worms with a metal pick and nematodes that did not respond were considered dead. One hundred *P. pacificus* were exposed to each of the soil-derived and insect-derived bacteria as well as the human pathogens (*P. aeruginosa* and *S. aureus*), the insect pathogens (*X. nematophila* and *Xenorhabdus* sp. and *P. luminescens*) and *E. coli* OP50 was used as the control. *C. elegans* was also exposed to each bacterium as a comparison to *P. pacificus*. Nematodes were only exposed to bacteria at the phase 1 stage which is the most virulent and were transferred to fresh plates every 2 days to ensure nematodes would be only exposed to phase 1. Nematodes were well fed on *E. coli* OP50 before addition to pathogenic bacteria to avoid starving. Any mortality observed was due to bacterial pathogenicity and no starvation or bagging behaviour was observed.

Fecundity experiments

Twenty microlitres of overnight bacteria suspension was placed on separate Petri dishes with NGM agar and left to dry. Five single virgin hermaphrodites of *P. pacificus* were individually placed on separate dishes in the bacterial spot and were stored at 25°C. The numbers of live offspring produced by each worm was recorded daily. Worms were transferred to fresh plates daily for 4–5 days.

Defecation and residence time assays

Liquid cultures of *E. coli* OP50, *Xenorhabdus* sp. and *Bacillus* sp. 2 were cultured overnight at 30°C. Bacteria were mixed with red fluorescent 0.5 μm carboxylate-modified polystyrene latex beads (Sigma Aldrich, St Louis, MO, USA) to a final bead concentration of 0.8%. Standard 6 cm NGM agar plates were seeded with 20 μl of the bacteria–bead mixture and when dry, ten J4 *P. pacificus* or *C. elegans* larvae were added to separate dishes and allowed to feed overnight at 20°C. Cycle length was recorded by timing the intervals between defecations using a dissecting microscope with a halogen light source. For each species, ten defecations of ten individual worms were recorded.

Residence time was assessed by allowing worms to feed overnight on the bacteria–bead mixture (using methods as described above). Individual worms were transferred to fresh bacterial plates and the total time taken to clear the intestine of the fluorescent bead mixture was recorded as well as the total number of defecations. We took fluorescent images of worm faeces to determine the number of defecations with beads still present (until no further fluorescence could be observed). Ten worms were analyzed for each treatment.

Statistical analysis

Chemotaxis scores, fecundity and survival (after 7 days) were compared using one way analysis of variance (ANOVA) and differences between treatments were determined using the Bonferroni multiple comparison test using Small Stata, 9.2 (StataCorp, College Station, TX, USA). Mean defecation cycles and residence times were compared using Student's *t*-test for means. The effect of bacteria over time was analyzed using Kaplan–Meier and log rank tests.

RESULTS

Meta-genomic sequencing of individual *Pristionchus* reveals a diverse bacterial flora

A total of eight *Pristionchus* individuals (four *P. entomophagus*, four *P. lheritieri*) were isolated from soil samples from the Schönbuch forest, Tübingen (Germany). From those animals, we sequenced a total of 683 bacterial clones and identified 292 unique sequences with at least 40 different bacterial 16S sequences per individual nematode. The majority of these bacteria belonged to the Pseudomonadales, Burkholderiales, Flavobacteria and Xanthomonadales and also included the plant pathogens *Erwinia* and *Agrobacterium* and human pathogens such as *Bordetella*, *Burkholderia* and *Microbacterium* (see Tables S1 and S2 in supplementary material). Other spore-forming bacteria, such as *Bacillus* sp., cannot be detected with this approach because bacterial spores escape single worm lysis. This non-saturated bacterial sequencing approach clearly indicates that *Pristionchus* nematodes ingest an enormous diversity of bacteria. However, from this type of analysis it remains unknown if the nematodes can digest all of these bacteria as a food source.

A set of 23 bacteria from the *Pristionchus* intestine

We grew beetle- and soil-derived *Pristionchus* nematodes on rich medium and isolated a total of 23 bacterial strains (Fig. 2). Specifically, we obtained bacteria derived from each nematode–beetle system, i.e. bacteria from *P. pacificus* from the oriental beetle (collected in Carver, MA, USA), *P. maupasi* from the cockchafer (Karlsruhe, Germany) and *P. entomophagus* from dung beetles (Tübingen, Germany). Together, these bacterial isolates represent all major groups previously identified in the metagenomic sequencing approach. Bacterial isolates were sequenced for species designation and represent single species isolates.

Pristionchus prefers soil-derived bacteria to *E. coli* in chemotaxis assays

In chemotaxis experiments using soil-derived bacteria (from adult *P. entomophagus*) we found that *Pristionchus* species were equally attracted to all available *Pseudomonas* strains ($P>0.05$; Fig. 3A). *Pristionchus* species were weakly attracted to the *Pseudomonas* strains and scored chemoattration indices of 0.1–0.3. By contrast, *Bacillus* sp. 1 was highly unattractive to the *Pristionchus* nematodes. More specifically, *P. pacificus* was significantly more averse to *Bacillus* sp. 1 than *P. maupasi* and *P. entomophagus* ($P<0.001$). When exposed to *Bacillus* sp. 2, the response of the *Pristionchus* nematodes varied ($P<0.001$). *P. entomophagus* was more attracted to this bacterium than *P. pacificus* and *P. maupasi,* which showed aversion.

Chemoattraction of *Pristionchus* species exposed to beetle-derived bacteria

Next, we analyzed the chemotaxis behaviour of three *Pristionchus* species exposed to beetle-derived bacteria from three settings. In the first system (cockchafer and *P. maupasi*) we found no significant differences between the chemotaxis index scores of *P. pacificus*, *P. maupasi* and *P. entomophagus* when exposed to *Pseudomonas fluorescens*, *Pseudomonas aurantiaca*, *Bacillus megaterium*, *Stenotrophomas maltophila*, *Serratia* sp. or *Bacillus* sp. ($P>0.05$; Fig. 3B). However, when exposed to *Bacillus thuringiensis*, *Enterobacter* sp. and *Pseudomonas chlororaphis* there were significant differences between each nematode species ($P<0.05$). Most strikingly, *P. pacificus* was strongly averse to *B. thuringiensis*. From the second system (dung beetle and *P. entomophagus*), we isolated five bacteria that are usually found in animal and insect faeces (Holt et al., 2000). *P. entomophagus* was significantly more attracted to *Serratia* sp., *Enterobacter amnigenus* and *Proteus vulgaris* than *P. pacificus* and *P. maupasi* ($P<0.05$) (Fig. 4A). By contrast, *P. maupasi* and *P. pacificus* responded poorly to the *Serratia* sp., *Ochrobactrum* sp., *Enterobacter* sp. and *P. vulgaris*.

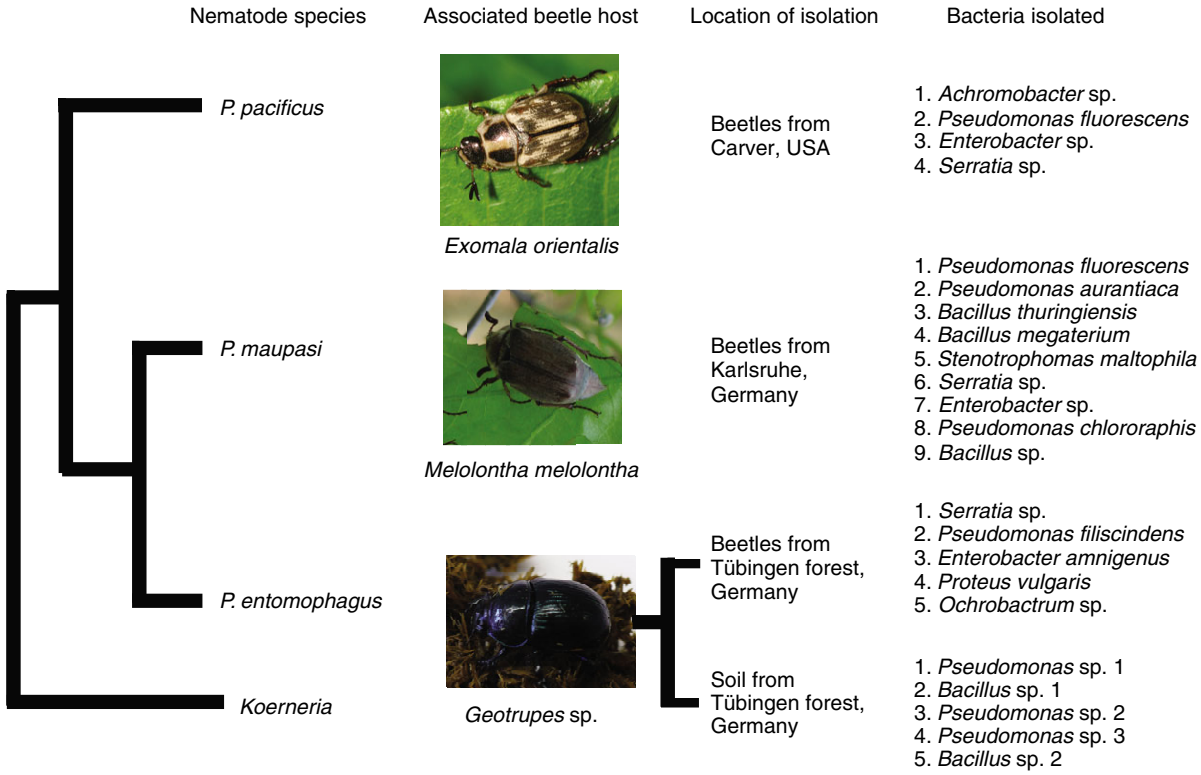


Fig. 2. Tritrophic interactions between *Pristionchus* nematodes, beetle and bacteria. Phylogeny of *Pristionchus* nematodes with associated beetle hosts detailing bacterial species isolated and used in this study.

In the third system (oriental beetle and *P. pacificus*) four bacterial species were isolated from *P. pacificus* emerging from oriental beetles including *Serratia* sp., *Achromobacter* sp., *Pantoea agglomerans* and *P. fluorescens*. These bacteria are commonly found in soil, water and the digestive tract of insects (Holt et al., 2000). All nematodes responded poorly to oriental beetle-derived bacteria, except for *P. fluorescens*, which was similarly attractive to all *Pristionchus* species (Fig. 4B). Taken together, the chemotaxis studies indicate that (1) *Pristionchus* nematodes never favour one bacterial strain completely over all others, (2) they strongly avoid certain *Bacillus* strains and (3) there is only a low *Pristionchus* species-specificity of bacteria in the chemotaxis assays.

Survival and fecundity of *P. pacificus* exposed to insect- and soil-derived bacteria

P. pacificus grown on *Bacillus* sp. 1 and *Bacillus* sp. 2 (from soil-derived nematodes) reduced brood size significantly compared to *E. coli* OP50 ($P<0.001$; Fig. 3A). This is also true for *B. thuringiensis*, *P. aurantiaca* and *Serratia* sp. from cockchafers ($P<0.05$; Fig. 3B), *Ochrobactrum*, *P. vulgaris* and *Serratia* isolated from the dung beetle ($P<0.05$; Fig. 4A) and *S. marcescens*, *P. agglomerrans* and *Achromobacter* sp. from oriental beetle ($P<0.05$; Fig. 4B).

The only bacteria that affected the survival of *P. pacificus* were *Serratia* sp. from cockchafers, and *S. marscens*, *P. agglomerrans* and *Achromobacter* sp. from the oriental beetle ($P<0.05$).

P. pacificus avoids *Bacillus* species

Next, *P. pacificus* was given the choice of either *E. coli* OP50, *B. thuringiensis* or *Bacillus* sp. 1 as a counter attractant to a range of insect-associated bacteria as an attractant. The nematode consistently avoided both *Bacillus* sp. 1 and *B. thuringiensis* and scored chemotaxis indices between 0.96–0.99 and 0.85–0.96, respectively (Fig. 5). The response to the *Bacillus* species was significantly different from the response to the *E. coli* OP50 control ($P<0.001$). These results demonstrate the revulsion *Pristionchus* has to these *Bacillus* species and provides a first indication that these nematodes might be able to avoid bacteria that have a potential harmful effect on their fitness.

P. pacificus is susceptible to insect but not human pathogenic bacteria

We assessed survival of *P. pacificus* and *C. elegans* to human and insect pathogens. *P. pacificus* was highly susceptible to *P. luminescens*, *X. nematophila* and *Xenorhabdus* sp. but was not susceptible to *S. aureus* and *P. aeruginosa* (Fig. 6A). After 24 h, only 1±0.55 *P. pacificus* were alive compared with 12±0.87 and 19±0.19 *P. pacificus* exposed to *X. nematophila* and *Xenorhabdus* sp., respectively. *P. luminescens* is therefore the most pathogenic of the entomopathogenic bacteria to *P. pacificus*. The number of surviving *P. pacificus* exposed to *S. aureus* and *P. aeruginosa* was similar to the numbers surviving *E. coli* OP50 exposure ($P>0.05$). *C. elegans* was highly susceptible to both human and insect pathogens ($P<0.001$; Fig. 6B) which is in stark contrast to the immunity *P. pacificus* has to *S. aureus* and *P. aeruginosa*.

In chemotaxis assays *P. pacificus* was exposed to either *E. coli* OP50 (as counter-attractant) or *X. nematophila*, *Xenorhabdus* sp., *P. aeruginosa* or *S. aureus* in separate experiments. We found that *P. pacificus* strongly avoided the killer species of bacteria (*X. nematophila* and *Xenorhabdus* sp.) and the human pathogens (*S. aureus* and *P. aeruginosa*; Fig. 6C). There was no difference between the responses of the nematode to these bacteria ($P>0.05$).

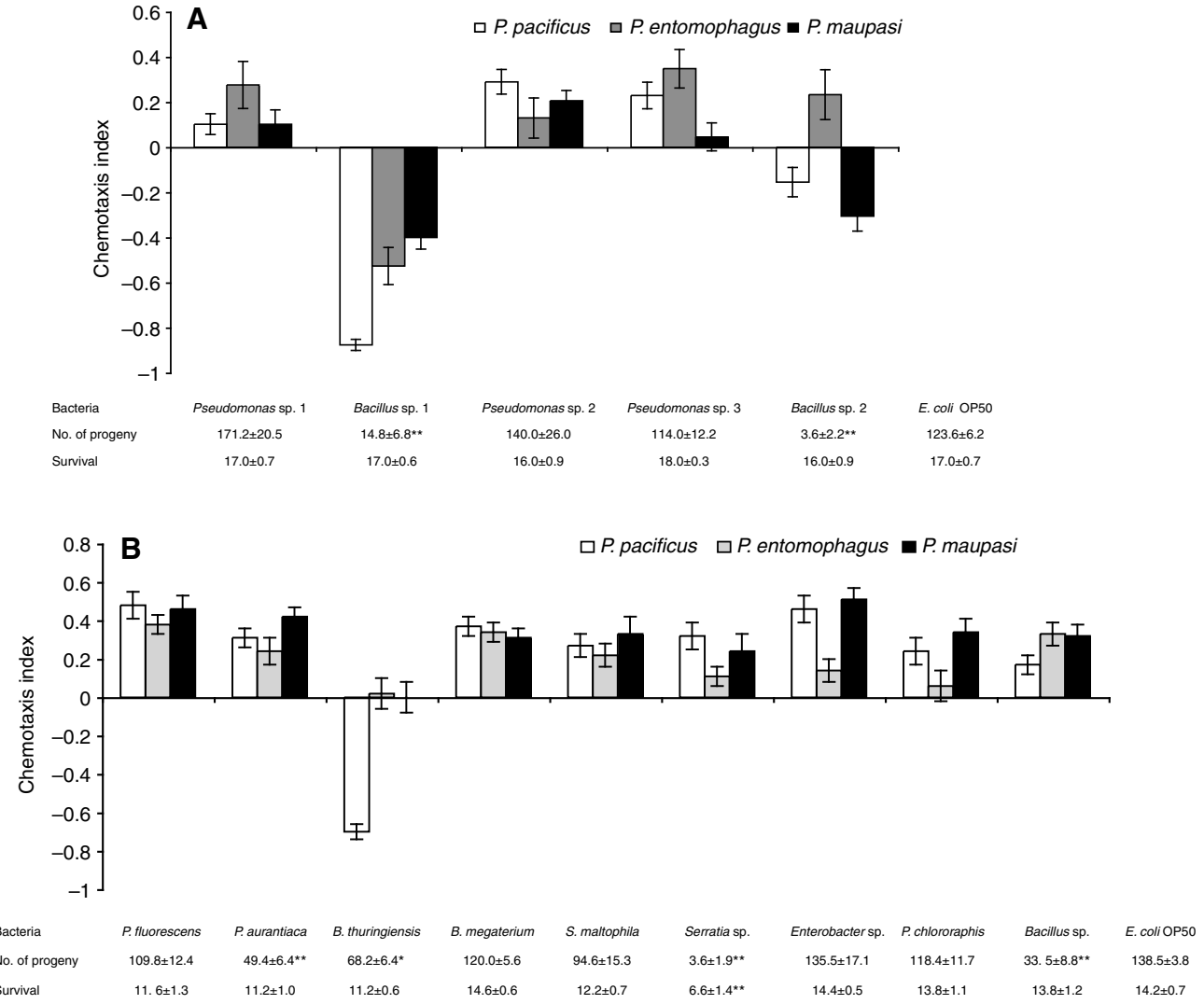


Fig. 3. Chemotaxis results of *Pristionchus* exposed to soil- and insect-derived bacteria. (A) Chemotactic response of *P. pacificus*, *P. entomophagus* and *P. maupasi* after 24 h exposure to five different soil-derived bacteria isolated from adult *P. entomophagus*. Significant differences between number of live progeny and survival are shown for *P. pacificus* only (**P*≤0.05 and ***P*≤0.001 using one way analysis of variance). Bars represent ±1 s.e.m.
(B) Chemotactic response of *P. pacificus*, *P. entomophagus* and *P. maupasi* exposed to cockchafer-associated bacteria (isolated from *P. maupasi*). Significant differences between number of live progeny and survival are shown for *P. pacificus* only (**P*≤0.05 and ***P*≤0.001 using one way analysis of variance). Bars represent ±1 s.e.m.

When *E. coli* OP50 was replaced with *P. luminescens* as the counter-attractant, *P. pacificus* was attracted to the human pathogens. The nematode found it difficult to distinguish between *X. nematophila* and *P. luminescens*, as well as *Xenorhabdus* sp. and *P. luminescens*, which was expected as the nematodes were faced with a choice of two unpleasant toxic species (Fig. 6B).

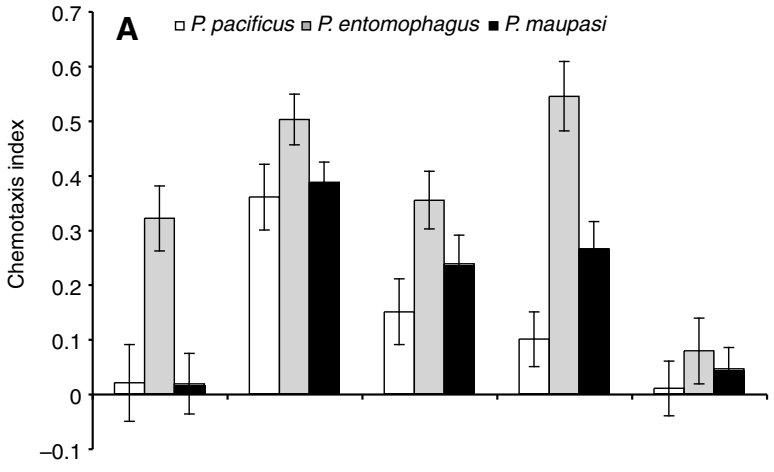
Morphological differences in nematodes influence defecation cycle and residence time of *P. pacificus*

We wanted to determine whether the absence of the grinder in *P. pacificus* affected defecation time and residence time compared with those in *C. elegans*. When *P. pacificus* is fed on *E. coli* OP50 the mean defecation time was 106±6.7 s compared to 48±1.9 s for *C. elegans* ($P<0.001$; Fig. 7A). When *P. pacificus* was fed on *Bacillus* sp. 2 the defecation cycle was significantly longer than when fed *E. coli* OP50 ($P<0.001$). There was no difference in defecation cycle when fed toxic *Xenorhabdus* sp. or *E. coli* OP50 ($P>0.05$). Conversely, the defecation cycle of *C. elegans* did not alter when fed *Bacillus* sp. 2 but was longer when fed *Xenorhabdus* ($P<0.05$). Videos demonstrating physiological differences in defecation cycles, as well as step-by-step muscle contractions can be seen in Movie 1 and Fig. S1 in supplementary material.

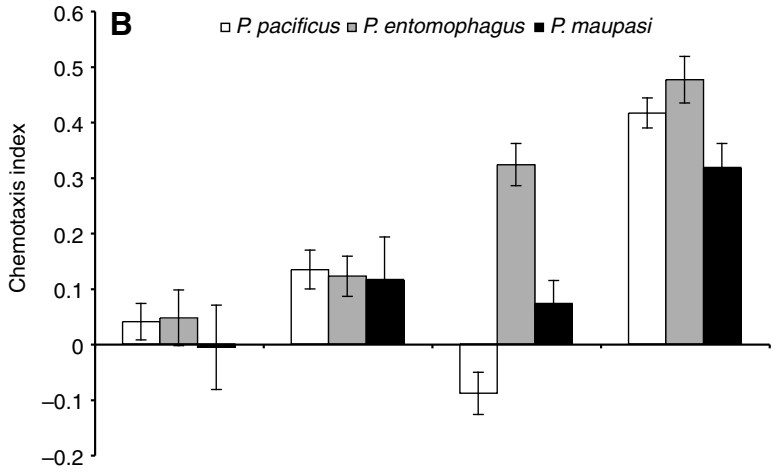
We defined the residence time as the total time recorded for a feeding worm to clear the gut of the OP50–fluorescent bead mixture after being transferred to a bacterial spot without beads. When *E. coli* OP50 was replaced with *Bacillus* sp. 2 the residence time for *P. pacificus* was significantly longer ($P<0.05$), whereas the residence time for *C. elegans* did not differ significantly ($P>0.05$; Fig. 7B).

DISCUSSION

We began an analysis of the tritrophic interactions of *Pristionchus* nematodes with bacteria associated with beetles and soil. We identified bacteria from *Pristionchus* nematodes using metagenomics and microbiological techniques and investigated the



Bacteria	*Serratia* sp.	*P. filicindens*	*E. amnigenus*	*P. vulgaris*	*Ochrobactrum* sp.	*E. coli* OP50
No. of progeny	95.0±10.8*	121.4±10.8*	135.4±7.1	84.0±12.1*	75.8±7.3**	164.7±15.4
Survival	16.8±1.1	18±1.5	19.2±0.5	18.2±0.7	17.0±0.8	18.2±0.7



Bacteria	*P. agglomerans*	*Achromobacter* sp.	*S. marcescens*	*P. fluorescens*	*E. coli* OP50
No. of progeny	21.8±7.4**	51.2±4.4**	52.0±11.1**	106.0±8.6	123.2±8.1
Survival	1.8±0.6**	9.2±1.7**	7.6±0.4**	15.6±1.2	17.8±0.9

Fig. 4. Chemotaxis results of *Pristionchus* exposed to insect-derived bacteria. (A) Chemotactic response of *P. pacificus*, *P. entomophagus* and *P. maupasi* exposed to dung beetle-associated bacteria (isolated from *P. entomophagus*) after 24 h. Significant differences between number of live progeny and survival are shown for *P. pacificus* only (*$P \leqslant 0.05$ and **$P \leqslant 0.001$ using one way analysis of variance). Bars represent ±1 s.e.m. (B) Chemotactic response of *P. pacificus*, *P. entomophagus* and *P. maupasi* exposed to oriental beetle-associated bacteria (isolated from *P. pacificus*) after 24 h. Significant differences between number of live progeny and survival are shown for *P. pacificus* only (*$P \leqslant 0.05$ and **$P \leqslant 0.001$ using one way analysis of variance). Bars represent ±1 s.e.m.

effect these bacteria have on a range of *Pristionchus* species using chemotaxis, survival and fecundity assays. We assessed survival of *P. pacificus* to human and insect pathogens and examined whether physiological functions are affected by bacterial species.

The interactions between *Pristionchus* and bacteria must be considered in the context of the necromenic life-style of these nematodes. On the death of the beetle infected with *Pristionchus* dauer larvae, bacteria and fungi rapidly colonize the beetle's body. This in turn creates a toxic 'beetle soup' consisting of large numbers of competing microorganisms originally present in the beetle intestine and the local environment. Around this time *Pristionchus* is thought to exit from the resistant dauer stage and develop into the J4 stage and then into adult form, with the main purpose to feed and reproduce. By distinguishing between an array of pathogenic and non-pathogenic bacteria and choosing the correct species to feed on, *Pristionchus* can lower the possibility of eating bacteria that cause low brood size, slow development and mortality. Once the food supply is depleted, the remaining nematodes turn into dauer larvae and search for potential beetle hosts in the surrounding soil. It is not known whether *Pristionchus* retains the bacteria it encounters before turning into the dauer form. Insect parasitic nematodes such as *S. carpocapsae* store their symbiotic bacteria *X. nematophila* in an intestinal vesicle (Martens et al., 2003) whereas *H. bacteriophora* stores *P. luminescens* in the intestine (Ciche and Ensign, 2003). Fedorko and Stanuszek (Fedorko and Stanuszek, 1971) observed *P. uniformis* dauer larva carrying bacteria cells in the gut but this observation has yet to be confirmed for *Pristionchus* species living on scarab beetles.

In general, nematodes show an enormous range of feeding strategies that allows them to occupy innumerable ecological niches (Munn and Munn, 2002). Predatory and plant parasitic nematodes have stylets, whereas soil-dwelling rhabditids, such as *C. elegans*, have a grinder in the terminal bulb of the pharynx. In contrast to *C. elegans, P. pacificus* lacks a grinder. Under laboratory conditions, bacterial lysis in the gut is incomplete and bacteria can survive the passage through the *P. pacificus* gut. We speculate that in order to gain any nutrition from these bacteria *Pristionchus* will have to increase time taken for digestion. Also, *Pristionchus* might be actively involved in bacterial dissemination in the wild, a claim that has already been made several decades ago (Chantanao and Jensen, 1968; Poinar, 1983).

The metagenomic analysis revealed that *Pristionchus* harbours a huge diversity of bacteria within its gut and on its cuticle including plant pathogenic and opportunistic human pathogens. Although mechanisms of immunity have been discovered in *C. elegans* in response to human pathogenic bacteria, *C. elegans* is susceptible to a range of naturally occurring Gram-negative and -positive bacteria and fungi (Ewbank, 2002). We have shown that *Pristionchus* associates with many different pathogenic bacteria in nature and *P. pacificus* is resistant to a number of bacteria that *C. elegans* is susceptible to, such as *S. aureus*, *P. aeruginosa* and *P. fluorescens*. Although future studies will reveal the exact molecular basis behind this resistance, the analysis of the *P. pacificus* genome provides the first insight into potential underlying mechanisms.

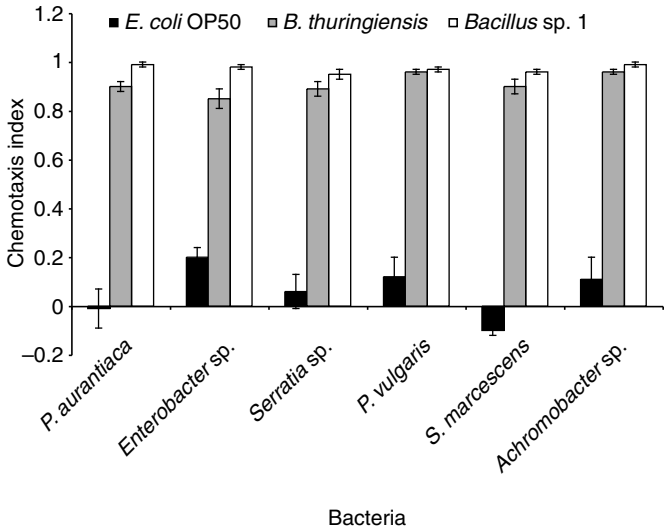


Fig. 5. Chemotaxis results of *P. pacificus* exposed to *Bacillus* species. Chemotactic response of *P. pacificus* exposed to a selection of cockchafer-, dung beetle- and oriental beetle-associated bacteria as attractants when assayed using *E. coli* OP50, *B. thuringiensis* and *Bacillus* sp. 1 as counter attractants. Bars represent ±1 s.e.m.

When compared to *C. elegans*, the *P. pacificus* genome shows a large expansion of genes encoding cytochrome P450 enzymes, glucosyl transferases, ABC transporters and other proteins that are thought to be involved in the degradation of xenobiotic compounds (C. Dieterich, S. W. Clifton, L. Schuster, A. Chinwalla, K. Delehaunty, I. Dinkelacker, R. Fulton, J. Godfrey, P. Minx, M. Mitreva et al., manuscript in revision). Future research will focus on genetic studies to elucidate methods of innate immunity, pathogenicity of associated bacteria and methods of detection and avoidance when in contact with pathogenic bacteria.

This study is one of the first to utilize both microbiological and metagenomic techniques to isolate bacteria from nematodes. Studies of other nematodes such as the slug parasitic nematode *P. hermaphrodita* have isolated 13 bacterial species from dauer larvae, culture medium and infected slugs (Wilson et al., 1995a; Wilson et al., 1995b). Bacterial communities associated with invertebrates have also been assessed using metagenomic tools. For example the bacterial communities in the hindgut paunch of a wood-feeding termite recorded 1750 bacterial 16S rRNA gene sequences that represented 12 phyla and 216 phylotypes (Warnecke et al., 2007). We did not expect to find a wealth of bacteria present in the nematode gut and on the cuticle. The function and relationship of many of these bacteria to *Pristionchus* remains unknown but from these results it can be seen that the nematodes associate with a huge diversity in nature.

In chemotaxis assays *P. pacificus* avoids species that cause ill health but the mechanism for sensing the causal properties of these bacteria currently remains unknown. We exploited this behaviour by using different *Bacillus* strains as a control and making other bacteria more attractive. Thus, the *P. pacificus* response to bacteria in chemotaxis assays differs from that of *C. elegans* in the type of

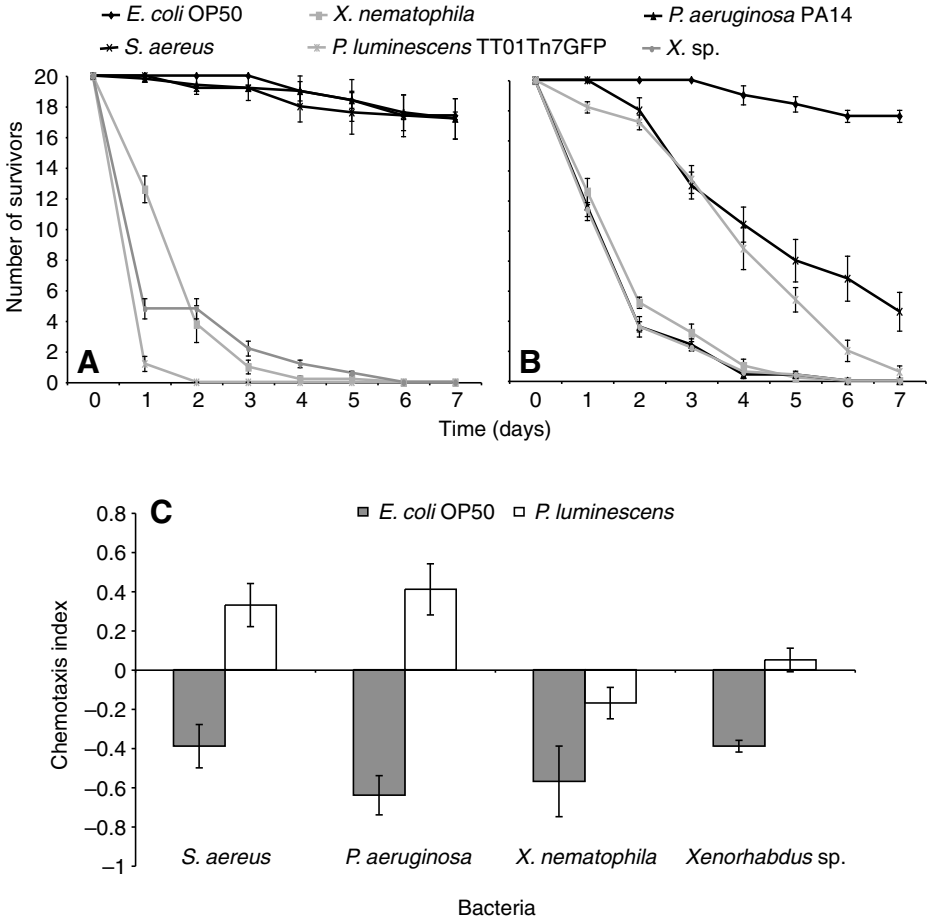


Fig. 6. Survival and chemotactic response of *P. pacificus* exposed to pathogenic bacteria. Survival of *P. pacificus* (A) and *C. elegans* (B) exposed to human and insect pathogens for 7 days. Bars represent ±1 s.e.m. (C) Chemotactic response of *P. pacificus* exposed to human and insect pathogens with *E. coli* OP50 and *P. luminescens* as the control. Bars represent ±1 s.e.m.

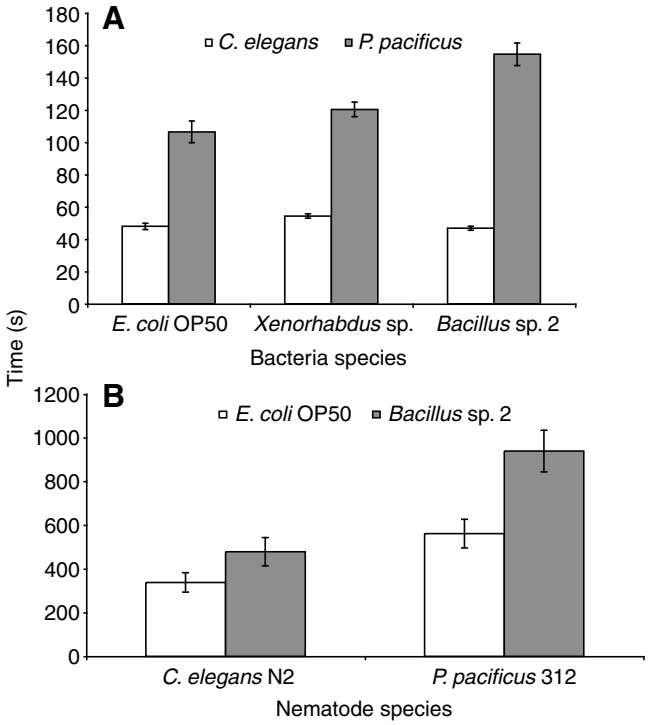


Fig. 7. Mean defecation cycle and residence time of *P. pacificus* and *C. elegans*. (A) Mean defecation time (seconds) of *C. elegans* and *P. pacificus* exposed to *E. coli* OP50, *Xenorhabdus* sp. and *Bacillus* sp. 2. Error bars represent ±1 s.e.m. (B) Mean residence time (seconds) of *E. coli* OP50 and *Bacillus* sp. 2 mixed with fluorescent beads, in *C. elegans* and *P. pacificus*. Bars represent ±1 s.e.m.

attraction (or repulsion). Similar differences have also been observed for the *P. pacificus* and *C. elegans* responses to pure insect-associated chemicals (Hong and Sommer, 2006b). Most likely, all of these differences are part of the adaptive forces that shape the tripartite interactions between bacteria, nematodes and beetle hosts in the wild.

C. elegans protects itself from potential pathogenic bacteria by avoidance behaviour and innate immunity pathways (Pujol et al., 2001; Nicholas and Hodgkin, 2004; Kurz and Ewbank, 2003). Avoidance behaviour of *C. elegans* has been recorded for *B. thuringiensis*, *S. marcescens*, *P. aeruginosa*, *P. luminescens* and *M. nematophila* using different molecular mechanisms (Pujol et al., 2001; Pradel et al., 2007; Yook and Hodgkin, 2007; Zhang et al., 2005; Beale et al., 2006; Schulenburg and Müller, 2004; Sicard et al., 2007; Hasshoff et al., 2007). In our experiments, bacteria that cause mortality to *P. pacificus* tend to score low in the chemotaxis index, ranging from –0.09±0.04 for *Serratia* sp. (from the cockchafer) to 0.37±0.05 for *S. marcescens* (from the oriental beetle). Unlike previous studies (Zhang et al., 2005) that demonstrated that *C. elegans* grown on a mixture of OP50 and *P. aeruginosa* or *S. marcescens* would avoid the two human pathogens in chemotaxis experiments, we have shown that OP50-raised *Pristionchus* species avoid *Bacillus* species without any training. Nematodes exposed to *Bacillus* species, particularly *Bacillus* sp. 1 and *B. thuringiensis* had significantly lower brood size than those grown on *E. coli* OP50. Previous studies have demonstrated that exposure of *P. pacificus* to purified Cry 5B crystal protein from *B. thuringiensis* causes a significant reduction in brood size and affects development (Wei et al., 2003).

P. pacificus is highly susceptible to *P. luminescens* and *X. nematophila* and *Xenorhabdus* sp. These bacteria are commonly found in entomopathogenic nematodes (*Steinernema* and *Heterorhabditis*) and are responsible for causing insect mortality 24–48 h after nematode penetration (Forst et al., 1997). As *Pristionchus* has a strong relationship with a number of beetle hosts that entomopathogenic nematodes can also infect, e.g. *Steinernema scarabaei,* which has been isolated from oriental beetles (Stock and Koppenhöfer, 2003), the chances of co-infection with *Pristionchus* and entomopathogenic nematodes and their associated bacteria are high. Our studies suggest that *P. pacificus* can recognize and avoid highly pathogenic bacteria such as *X. nematophila*. The ecological interaction between *Pristionchus* and entomopathogenic nematodes and bacteria clearly warrants further research.

P. entomophagus was significantly more attracted to dung beetle bacteria than to *P. maupasi* and *P. pacificus*. This was the only nematode–beetle system for which bacterial specificity was recorded, as generally the three *Pristionchus* species tested responded similarly when exposed to bacteria isolated from the respective beetle host. From this study the reasons behind nematode–beetle host specificity has not be discovered but it is not due to bacteria harboured in the beetle gut. As *Pristionchus* nematodes show a high species specificity with cockchafers, dung beetles and Colorado potato beetles (Herrmann et al., 2006a; Herrmann et al., 2006b) and display unique chemoattraction profiles towards insect pheromones and plant volatiles (Hong and Sommer, 2006b) other reasons apart from bacteria species must be considered. The behavioural response of the entomopathogenic nematode *S. carpocapsae* is correlated with nematode-induced mortality and number of infective juveniles produced on each host species (Lewis et al., 1996; Lewis, 2002). Other reasons for host specificity may include increased production of males, higher reproduction rate or better health. Also as well as bacteria living in the beetle, fungi and other parasitic or phoretic nematodes are present, perhaps these organisms contribute to nematode–beetle specificity. Further studies are needed to elucidate the exact mechanism behind these complex nematode–beetle associations.

Taken together, we have analyzed the tritrophic interactions of *Pristionchus* nematodes with various bacteria from soil and beetles. We found a range of different interactions from bacterial dissemination by the worm to reduction in brood size, longer defecation cycles and nematode mortality caused by certain bacterial strains. *Pristionchus* can recognize, respond to and avoid bacteria that cause poor health. The ability to discriminate between bacteria is important for success and survival in the soil ecosystem. Given the genetic and genomic toolkit available in *P. pacificus*, this nematode interaction with its living environment can in the future be investigated at the molecular level to provide mechanistic insight into the adaptive and non-adaptive forces that shape this nematodes ecosystem.

We are grateful to Elissa Hallem and Todd Ciche for *P. luminescens*, Heidi Goodrich Blair for *X. nematophila*, Jane Roberts and Jeremy Pearce at Becker Underwood, UK for strains of entomopathogenic bacteria. We thank members of the Sommer lab for discussions and Drs R. Hong and D. Bumbarger for comments of the manuscript. Also we thank Paul Robbins for collection of oriental beetles. This research was funded by the Max-Planck Society.

REFERENCES

Beale, E., Li, G., Tan, M. W. and Rumbaugh, K. P. (2006). *Caenorhabditis elegans* senses bacterial autoinducers. *Appl. Environ. Microbiol.* **72**, 5135-5137.
Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71-94.
Chantanao, A. and Jensen, H. J. (1969). Saprozoic nematodes as carriers and disseminators of plant pathogenic bacteria. *J. Nematol.* **1**, 216-218.

Chiang, J. T. A., Steciuk, M., Shtonda, B. and Avery, L. (2006). Evolution of pharyngeal behaviors and neuronal functions in free-living soil nematodes. *J. Exp. Biol.* **209**, 1859-1873.
Ciche, T. A. and Ensign, J. C. (2003). For the insect pathogen *Photorhabdus luminescens*, which end of a nematode is out? *Appl. Environ. Microbiol.* **69**, 1890-1897.
Ewbank, J. J. (2002). Tackling both sides of the host-pathogen equation with *Caenorhabditis elegans*. *Microbes Infect.* **4**, 247-256.
Faegri, A., Torsvik, V. L. and Goksoyr, J. (1977). Bacterial and fungal activity in soil: separation of bacteria and fungi by a rapid fractionated centrifugation technique. *Soil Biol. Biochem.* **9**, 105-112.
Fedorko, A. and Stanuszek, S. (1971). *Pristionchus uniformis* sp. n. (Nematoda, Rhabditida, Diplogasteridae), a facultative parasite of *Leptinotarsa decemlineata* Say and *Melolontha melolontha* L. in Poland. Morphology and biology. *Acta Parasitol. Pol.* **19**, 95-112.
Floyd, R., Abebe, E., Papert, A. and Blaxter, M. (2002). Molecular barcodes for soil nematode identification. *Mol. Ecol.* **11**, 839-850.
Forst, S., Dowds, B., Boemare, N. and Stackebrandt, E. (1997). *Xenorhabdus* spp. and *Photorhabdus* spp.: bugs that kill bugs. *Annu. Rev. Microbiol.* **51**, 47-72.
Fürst von Lieven, A. and Sudhaus, W. (2000). Comparative and functional morphology of the buccal cavity of Diplogastrina (Nematoda) and a first outline of the phylogeny of the taxon. *J. Zoolog. Syst. Evol. Res.* **38**, 37-63.
Han, Z. M., Hong, Y. D. and Zhoa, B. G. (2003). A study on pathogenicity of bacteria carried by pine wood nematodes. *J. Phytopathol.* **151**, 683-689.
Hasshoff, M., Böhnisch, C., Tonn, D., Hasert, B. and Schulenburg, H. (2007). The role of *Caenorhabditis elegans* insulin-like signaling in the behavioral avoidance of pathogenic *Bacillus thuringiensis*. *FASEB J.* **21**, 1801-1812.
Herrmann, M., Mayer, W. and Sommer, R. J. (2006a). Nematodes of the genus *Pristionchus* are closely associated with scarab beetles and the Colorado potato beetle in Western Europe. *Zoology.* **109**, 96-198.
Herrmann, M., Mayer, W. and Sommer, R. J. (2006b). Sex, bugs and Haldane's rule: the nematode genus *Pristionchus* in the United States. *Front. Zool.* **3**, 15-31.
Herrmann, M., Mayer, W. M., Hong, R. L., Kienle, S., Minasaki, R. and Sommer, R. J. (2007). The nematode *Pristionchus pacificus* (Nematoda: Diplogastridae) is associated with the oriental beetle *Exomala orientalis* (Coleoptera: Scarabaeidae) in Japan. *Zool. Sci.* **24**, 883-889.
Holt, J. G., Krieg, N. R., Sneath, P. H. A., Staley, J. T. and Williams, S. T. (ed.) (2000). Gram negative aerobic/microaerophilic rods and cocci. In *Bergey's Manual of Determinative Bacteriology* (9th edn), pp. 71-102. Baltimore: Williams and Wilkins.
Hong, R. and Sommer, R. J. (2006a). *Pristionchus pacificus*: a well-rounded nematode. *BioEssays* **28**, 651-659.
Hong, R. and Sommer, R. J. (2006b). Chemoattraction in *Pristionchus* nematodes and implications for insect recognition. *Curr. Biol.* **16**, 2359-2365.
Kiontke, K. and Sudhaus, W. (2006). Ecology of *Caenorhabditis* species. In *WormBook* (ed. The *C. elegans* Research Community), doi/10.1895/wormbook.1.37.1, http://www.wormbook.org.
Kurz, C. L. and Ewbank, J. J. (2003). *Caenorhabditis elegans*: an emerging genetic model for the study of innate immunity. *Nat. Rev. Genet.* **4**, 380-390.
Lane, D. J. (1991). 16S/23S rRNA sequencing. In *Nucleic Acid Techniques in Bacterial Systematics* (ed. E. Stackebrandt and M. Goodfellow), pp. 184-189. Chichester: John Wiley.
Lewis, E. E. (2002). Behavioral ecology. In *Entomopathogenic Nematology* (ed. R. Gaugler), pp. 205-224. Wallingford: CABI.
Lewis, E. E., Ricci, M. and Gaugler, R. (1996). Host recognition behavior predicts host suitability in the entomopathogenic nematode *Steinernema carpocapsae* (Rhabditida: Steinernematidae). *Parasitology* **113**, 573-579.
Martens, E. C., Heungens, K. and Goodrich-Blair, H. (2003). Early colonization events in the mutualistic association between *Steinernema carpocapsae* nematodes and *Xenorhabdus nematophila* bacteria. *J. Bacteriol.* **185**, 3147-3154.
Mayer, W., Herrmann, M. and Sommer, R. J. (2007). Phylogeny of the nematode genus *Pristionchus* and implications for biodiversity, biogeography and the evolution of hermaphroditism. *Evol. Biol.* **7**, 104.
Munn, E. A. and Munn, P. D. (2002). Feeding and digestion. In *The Biology of Nematodes* (ed. D. H. Lee), pp. 211-232. London: Taylor and Francis Group.
Nicholas, H. R. and Hodgkin, J. (2004). Responses to infection and possible recognition strategies in the innate immunity system of *Caenorhabditis elegans*. *Mol. Immunol.* **41**, 479-493.
Nussbaumer, A. D., Bright, M., Baranyi, C., Beisser, C. J. and Ott, J. A. (2004). Attachment mechanism in a highly specific association between ectosymbiotic bacteria and marine nematodes. *Aquat. Microb. Ecol.* **34**, 239-246.
Poinar, G. O. (1983). *The Natural History of Nematodes*. Englewood Cliffs, NJ: Prentice Hall.
Pradel, E., Zhang, Y., Matsuyama, T., Bargmann, C. I. and Ewbank, J. J. (2007). Detection and avoidance of a natural product from the pathogenic bacterium *Serratia marcescens* by *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* **104**, 2295-2300.
Pujol, N., Link, E. M., Liu, L. X., Kurz, C. L., Alloing, G., Tan, M. W., Ray, K. P., Solari, R., Johnson, C. D. and Ewbank, J. J. (2001). A reverse genetic analysis of components of the Toll signaling pathway in *Caenorhabditis elegans*. *Curr. Biol.* **11**, 809-821.
Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. New York: Cold Spring Harbor Press.
Schulenburg, H. and Müller, S. (2004). Natural variation in the response of *Caenorhabditis elegans* towards *Bacillus thuringiensis*. *Parasitology* **128**, 433-443.
Sicard, M., Hering, S., Schulte, R., Gaudriault, S. and Schulenburg, H. (2007). The effect of *Photorhabdus luminescens* (Enterobacteriaceae) in the survival, development, reproduction and behavior of *Caenorhabditis elegans* (Nematoda: Rhabditidae). *Environ. Microbiol.* **9**, 12-25.
Sommer, R. J., Carta, L. K., Kim, A. S. Y. and Sternberg, P. W. (1996). Morphological, genetic and molecular description of *Pristionchus pacificus* sp. n. (Nematoda: Neodiplogastridae). *Fundam. Appl. Nematol.* **19**, 511-521.
Stock, S. P. and Koppenhöfer, A. M. (2003). *Steinernema scarabaei* n. sp. (Rhabditida: Steinernematidae), a natural pathogen of scarab beetle larvae (Coleoptera: Scarabaeidae) from New Jersey, USA. *Nematology* **5**, 191-204.
Tan, L. and Grewal, P. S. (2001). Pathogenicity of *Moraxella osloensis*, a bacterium associated with the nematode *Phasmarhabditis hermaphrodita*, to the slug *Deroceras reticulatum*. *Appl. Environ. Microb.* **67**. 5010-5016.
Tan, L. and Grewal, P. S. (2002). Endotoxin activity of *Moraxella osloensis* against the grey garden slug, *Deroceras reticulatum*. *Appl. Environ. Microb.* **68**, 3943-3947.
Taylor, M. J., Bandi, C. and Hoerauf, A. (2005). *Wolbachia* bacterial endosymbionts of filarial nematodes. *Adv. Parasit.* **60**, 245-282.
The *C. elegans* Research Community (ed.) (2005). Wormbook. http://www.wormbook.org.
Warnecke, F., Luginbühl, P., Ivanova, N., Ghassemian, M., Richardson, T. H., Stege, J. T., Cayouette, M., McHardy, A. C., Djordjevic, G., Aboushadi et al. (2007). Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* **450**, 560-565.
Wei, J. Z., Hale, K., Carta, L., Platzer, E., Wong, C., Fang, S. C. and Aroian, R. V. (2003). *Bacillus thuringiensis* crystal proteins that target nematodes. *Proc. Natl. Acad. Sci. USA* **100**, 2760-2765.
Weischer, B. and Brown, D. (2000). *An Introduction to Nematodes: General Nematology, A Students Textbook*. Sofia: Pensoft.
Wilson, M. J., Glen, D. M., George, S. K. and Pearce, J. D. (1995a). Selection of a bacterium for the mass production of *Phasmarhabditis hermaphrodita* (Nematoda: Rhabditidae) as a biological control agent for slugs. *Fundam. Appl. Nematol.* **18**, 419-425.
Wilson, M. J., Glen, D. M., Pearce, J. D. and Rodgers, P. B. (1995b). Monoxenic culture of the slug parasite *Phasmarhabditis hermaphrodita* (Nematoda: Rhabditidae) with different bacteria in liquid and solid phase. *Fundam. Appl. Nematol.* **18**, 159-166.
Yook, K. and Hodgkin, J. (2007). Mos1 mutagenesis reveals a diversity of mechanisms affecting response of *C. elegans* to the bacterial pathogen *M. nematophilum*. *Genetics* **175**, 681-697.
Zauner, H., Mayer, W., Herrmann, M., Weller, A., Erwig, M. and Sommer, R. J. (2007). Distinct patterns of genetic variation in *Pristionchus pacificus* and *Caenorhabditis elegans*, two partially selfing nematodes with cosmopolitan distribution. *Mol. Ecol.* **16**, 1267-1280.
Zhang, Y., Lu, H. and Bargmann, C. I. (2005). Pathogenic bacteria induce aversive olfactory learning in *Caenorhabditis elegans*. *Nature* **438**, 179-184.

Weakly deleterious mutations dominate genetic diversity in Pristionchus pacificus.

Christian Rödelsperger [*], Richard A. Neher [†], Andreas M. Weller [*], Gabi Eberhardt [*], Hanh Witte [*], Werner Mayer [*], Christoph Dieterich [‡], Ralf J. Sommer [*]

[*]Department for Evolutionary Biology, Max Planck Institute for Developmental Biology,72076, Tübingen, Germany,[†]Evolutionary Dynamics and Biophysics Group, Max Planck Institute for Developmental Biology, 72076, Tübingen, Germany, and [‡]Berlin Institute for Medical Systems Biology (BIMSB) at the Max Delbrück Center for Molecular Medicine Berlin, Robert-Rößle-Str. 10, 13125 Berlin, Germany

Submitted to Proceedings of the National Academy of Sciences of the United States of America

Genetic variation originates from mutation and recombination and is removed by selection and drift. While selection and demographic processes depend on the environment and the ecology of an organism, mutation and recombination depend primarily on intrinsic properties, most prominently the mating-system. In nematodes that reproduce by self-fertilization with occasional outcrossing, recombination is reduced, which can result in large haplotype blocks and a reduced efficacy of selection. In this study, we present whole-genome sequences of 104 strains of the self-fertilizing nematode *Pristionchus pacificus* and the draft genome of the obligate outcrossing sister species *P. exspectatus*. *P. pacificus* is ten times more diverse than *C. elegans* and exhibits substantial population structure. We find that 40% of non-synonymous variation in coding regions is weakly deleterious, which results in departures from the neutral model. We polarized variation in *P. pacificus* using the *P. exspectatus* genome and investigate the properties of site frequency spectra (SFS) of derived alleles. The SFS decay more rapidly then expected in a neutral model and show an excess of high frequency derived alleles characteristic of extensive linked selection. Consistent with this idea, we find haplotype blocks that span several megabases. Together, our findings suggest that *P. pacificus* is a diverse species with distinct subpopulations in which deleterious mutations are constantly brought to high frequency by linked selection.

Significance statement. With the sequencing of 104 strains of the nematode *Pristionchus pacificus*, a satellite model to the classical model organism *C. elegans*, we present one of the largest whole genome sequencing efforts of natural isolates in multicellular organisms. Using the draft genome of its sister species as outgroup, we identified several genomic features that are consistent with strong linked selection. Most prominently, around two thirds of non-synonymous mutations are deleterious, but may still be brought to high frequencies by linked selection. We also find haplotype blocks millions of bases in length. The overall genetic diversity is ten-fold higher than in *C. elegans* but structured into several clades. These differences highlight the importance of comparative studies of nematode genome evolution.

Introduction

A number of nematode and plant species reproduce by self-fertilization (hermaphroditism) with occasional outcrossing events via males, rather than obligatory outcrossing such as vertebrates and flies. Rare out-crossing in self-fertilizing species reduces effective recombination and can result in a decrease of neutral diversity through selection against linked deleterious variants (background selection) [1] or "hitchhiking" with linked beneficial variants [2, 3]. In addition, linkage between loci with non-neutral variation leads to Hill-Robertson interference [4] and more generally to a reduced efficacy of selection [5]. To contrast the stochastic effects of linked selection from those due to genetic drift, various forms of linked selection are called genetic draft - a term introduced by Gillespie [3]. Draft not only reduces diversity, but reshapes patterns of genetic variation in ways different from genetic drift [6, 7, 8]. A quantitative analysis of genetic variation therefore allows disentangling the roles of drift and draft. The prediction that genetic diversity should be negatively correlated with the local recombination rate has been verified in populations of *Drosophila* and *C. briggsae* [9, 10]. If linked selection has strong effects in obligate outcrossers, such as *Drosophila*, we expect even more dramatic signatures in selfing species. Indeed, Andersen et al [11] have shown that *C. elegans* has experienced chromosome wide sweeps in its recent history and exhibits very little genetic variation.

Over the last two decades, the hermaphroditic nematode *P. pacificus* has been established as a satellite model organism to *C. elegans* and a number of comparative studies have revealed divergent patterns in vulva formation, dauer development, and feeding behavior [12, 13, 14]. Similar to *C. elegans*, *P. pacificus* is a cosmopolitan nematode with a global distribution. *P. pacificus* has a clearly defined ecological niche and is found with scarab beetles in a necromenic association: nematodes infest beetles as arrested dauer larvae, wait for the insect's death to resume development and feed on the growing microbes on the carcass of the beetle. Phylogenetic studies have shown that hermaphroditism has evolved multiple times within the genera *Caenorhabditis* and *Pristionchus* [15]. Recently, a very closely related outcrossing sister species, *P. exspectatus*, has been isolated from stag beetles in Japan that still forms viable, but sterile F1 hybrids with *P. pacificus*. The close phylogenetic relationship between both species provides a unique framework for studying genome evolution and associated population genetic processes.

In this study, we present the draft genome of *P. exspectatus* and whole-genome sequencing data of 104 *P. pacificus* strains. Using *P. exspectatus* as outgroup, we investigated population structure and genetic diversity of *P. pacificus*. The majority of our *P. pacificus* isolates fall into three well-defined clades and the global diversity in *P. pacificus* is one order of magnitude higher than in *C. elegans*. We show that diversity in distinct clades is shaped by similar population genetic processes, most prominently background selection on weakly deleterious polymorphism. We find that roughly two thirds of non-synonymous polymorphisms are deleterious, yet a fraction of those nevertheless segregates at high frequency. Consistent with extensive non-neutral variation, the site frequency spectra (SFS) deviate from neutral expectations and show signs of genetic draft. The central regions of most chromosomes harbor megabase-sized haplotype blocks in high linkage disequilibrium (LD), concomitant with reduced diver-

Reserved for Publication Footnotes

sity. Our analysis suggests that linked selection is the dominant force shaping genetic diversity in *P. pacificus*.

Results

The draft genome of P. exspectatus. The outcrossing nematode *P. exspectatus* RS5522 has recently been identified as the closest known relative of the self-fertilizing hermaphrodite *P. pacificus* [16]. We sequenced an inbred strain of *P. exspectatus* to an approximate coverage of 97X using libraries with different insert sizes (Table 1) (see Supplementary Methods for details). The genome draft contains 167Mb of assembled sequence in 4,412 scaffolds spanning a total length of 177.6Mb with an N50 value of 142kb. Realignment of the sequencing data placed 95.6-97.1% of reads onto the assembly, 94.9% of mapped pairs were in correct orientation. Despite full-sibling inbreeding for 10 generations we observed extensive amounts of ambiguous bases in regions spanning 21.5Mb. In contrast to a previous analysis of assemblies of outcrossing nematodes [17], ambiguous regions were found on all chromosomes. In addition, ambiguous regions in *P. exspectatus* were associated with higher read coverage, and contained even tri-allelic positions (Supplemental Fig. 1). Specifically, we found that ambiguous regions spanning 21.5Mb (12%) of the genome contained a higher read coverage when using a threshold of 5 heterozygous base calls per 1kb to define an ambiguous region (Supplemental Fig. 1B). This finding suggests that these regions represent repeats or recent duplications rather than remaining heterozygosity not eliminated by inbreeding. Additional support for this hypothesis was obtained from the re-sequencing of the wild type, non-inbreed *P. exspectatus* strain. While 10 generation inbreeding should result in a theoretical reduction of heterozygosity to 14%, the observed reduction to only 44% indicates that 79% of ambiguous positions (nearly 10% of the total genome) are due to repeats and duplications. Thus, we estimate an actual genome size of *P. exspectatus* of around 200Mb. Whole genome alignments between *P. pacificus* and *P. exspectatus* covered 79.2 and 78.9Mb of uniquely alignable sequence, respectively, and revealed 5,221,279 substitutions and 1,161,799 indels indicating a sequence divergence of 10% that is distributed uniformly across the chromosomes.

Using RNA-seq derived gene models as training set, 24,642 complete gene models (28,236 including partial models) were predicted. Excluding genes with ambiguous gene structures due to alternative isoforms, evaluation of single-transcript RNA-seq gene models showed that 80.3% of expressed exons overlapped predicted exons on the same strand. For 57.6% of those exons, start and end positions were predicted correctly at nucleotide resolution. However, for 44% of predicted exons showed no evidence for expression indicating either false predictions or constitutively low and/or highly spatiotemporally restricted expression. Protein domain annotation using PFAM showed strong correlations between the two species (Spearman's ρ=0.77). Only two domain families (PF01498 and PF01359) were strongly expanded in *P. pacificus* relative to *P. exspectatus*; these domains correspond to DNA transposons of the mariner family suggesting an increased DNA transposon activity in *P. pacificus* following speciation. In this study, we use the *P. exspectatus* genome sequence (available at www.pristionchus.org/variome/) as outgroup for population genetic analysis of *P. pacificus*. A detailed comparison to *P. pacificus* and another closely related species, *P. arcanus*, will be presented elsewhere.

Feature	*P. pacificus*	*P. exspectatus*
Assembly size	172.5 Mb	177.6 Mb
Assembled sequence	153.2 Mb	167.0 Mb
N50 scaffold size	1.25 Mb	0.14 Mb
GC-content	42.8%	42.8%
Predicted genes	28,666	24,642
Coding sequence	27.5 Mb	27.1 Mb
Gene length	1.9 (1.0-3.5)* kb	2.5 (1.3-4.5)* kb
Transcript length	0.7 (0.4-1.2)* bp	0.8 (0.4-1.4)* kb
Exons per gene	7 (4-12)*	8 (5-14)*
Exon length	87 (65-115)* bp	86 (63-112)* bp
Intron length	119 (58-242)* bp	128 (63-253) * bp

*Numbers about gene structures denote median (first and third quartile)

Table 1. Summary of the *P. pacificus* and *P. exspectatus* genome sequences.

Sequencing of 104 natural isolates. To investigate genetic diversity and the underlying population genetic processes, we selected 104 strains of *P. pacificus* including the reference strain PS312 for second-generation sequencing. Strains were selected based on biogeography, beetle association and microsatellite patterning [18] (Fig. 1a). Sixty-one of the 104 sequenced strains are from the Island of La Réunion in the Indian Ocean, which represents a hotspot of *P. pacificus* biodiversity and has been the focus of recent population genetic studies on *P. pacificus* [18]. All strains were inbred for at least 10 generations and known to be largely homozygous at microsatellite markers. We sequenced genomic DNA of these strains on the Illumina platform to a mean coverage of 6-37 X (after alignment to the reference genome).

Individual strains showed between 23,381 and 1,402,808 single nucleotide variants (SNVs) relative to the reference genome (Supplementary Table 1). In total, we identified 7,112,381 SNV and

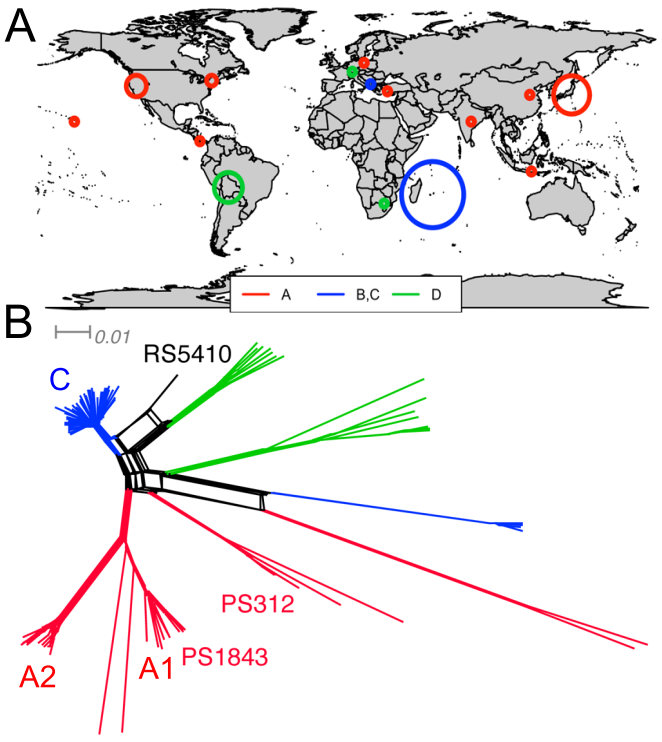


Figure 1. *(A)* Worldwide sampling of *P. pacificus* strains. Circle sizes indicate the numbers of sequenced strains. Colors denote the predominant mitochondrial and microsatellite clade per region [18]. *(B)* Population structure of *P. pacificus* natural isolates visualized as a split-network based on average number of substitutions within one million variables sites that were genotyped in all strains. The three most deeply sampled clades are used for further analysis. Clade C (N=44) is almost exclusively sampled from la Réunion, A_1 (N=15) contains strains from northern America and Asia, and A_2 (N=16) consists of south and central American strains as well as strains from the Indian ocean. Colors indicate the predominant mitochondrial and microsatellite clade per branch [18].

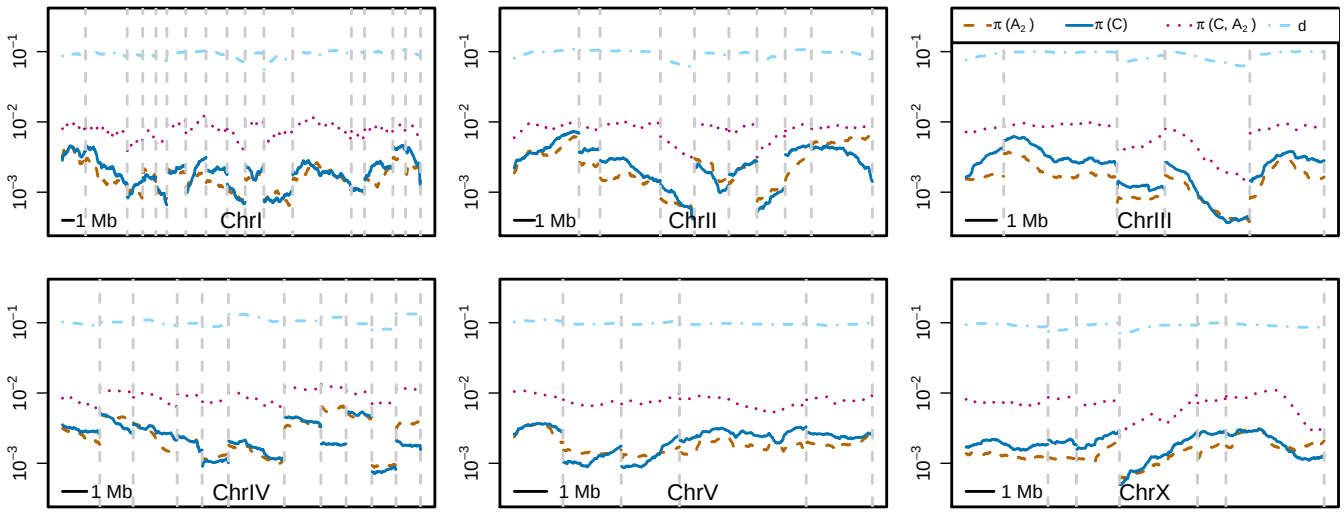


Figure 2. Nucleotide diversity π within and between clades C and A_2. π values averaged per 1Mb windows are shown. Supercontigs (>1Mb) with markers on the genetic map [24] were concatenated to visualize the chromosomal distribution. The dashed lines denote supercontig boundaries with unknown physical distance. Diversity within different clades is very well correlated suggesting that similar population genetic factors shape this variation.

2,119,730 indel (<100 bp) positions. We assessed the quality of variant calls by comparing the SNVs derived from short read alignments with the SNVs derived from Sanger whole genome alignments of the *P. pacificus* mapping strain PS1843 from Washington [19]. Out of 474,285 SNV calls that are covered by both platforms, 98% were in agreement, indel calls showed an agreement of 79-96% (see Methods). Figure 1b shows the phylogenetic relationship of the 104 sequenced strains as a split-network with the three most deeply sampled clades being A_1, A_2, and C covering at least 15 strains per clade.

Despite extensive inbreeding, an average of 6.8±6.2% of SNVs for all strains were called as heterozygous (Supplemental Table 1). Interestingly, we found that 14.3-35.4% of these heterozygous SNVs fall into large duplications ≥2kb (Supplemental Table 1). For the majority of the remaining heterozygous SNVs, several lines of evidence suggest that they also derive from duplications, which are however, below the limit of detection. First, a subset of 20 strains that was inbred for more than 30 generations shows similar levels of heterozygosity than the other strains. Second, the isolate RS5410 from La Réunion (Fig.1B), which was previously identified as an admixed strain [18], also shows similar levels of heterozygosity (8.1%). Third, structural variants and copy number variations in a range between 2-1000kb were detected using a method comparing differences in read depth relative to re-sequencing data of the reference strain [20]. Compared to the reference genome, an average of 6.9±% and 2.4±0.5% of the genome was predicted as deleted and duplicated, respectively. Finally, previous comparisons of predicted deletions with PCR amplification experiments for three cellulase genes in 24 strains showed perfect agreement [21]. Thus, the trend of recent, high-level duplications as found in *P. exspectatus* is also eminent in *P. pacificus* wild isolates.

Genome-wide levels of nucleotide diversity. Principal component analysis (PCA) of the SNV data revealed strong population structure with the first two principal components explaining 22.9% and 16.3% of the global variability, respectively (Supplemental Fig. 2). The clades A_1, A_2 and C are clearly separated along the first two principal components, a finding that is consistent with the split network (Fig. 1). However, further PCA within clade C, the most deeply sampled clade with a total of 44 strains, did not reveal additional clusters but only a slight signal reflecting the local geography on La Réunion Island. We analyzed the nucleotide diversity within clade C and A_2, together with average distances between these two clades (Fig. 2, see supplementary Fig. 3 for clade A_1). Within clades nucleotide diversity varies between 0.02% and 0.7% along the genome. The fluctuation of nucleotide diversity along the genome is well correlated between clades. In contrast, the comparison between clades reveals more stable divergence in the order of 0.8%. The divergence between

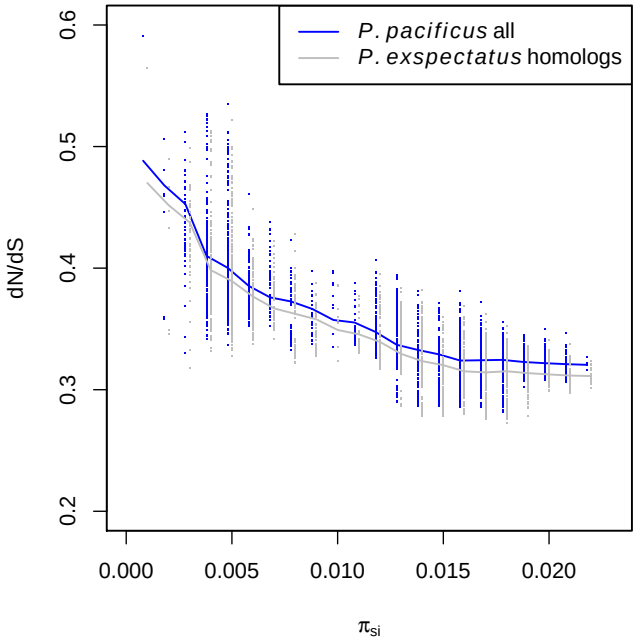


Figure 3. Across all pairwise comparisons of the 104 *P. pacificus* strains, dN/dS ratios decrease with divergence (π_{si}) from 0.5 to 0.3, suggesting that around 50% of nonsynonymous mutations have been selected against at very short time periods and were not observed as variable. In addition, around 20% of nonsynonymous variations are pruned with increasing divergence. Assuming that this drop is due to purifying selection, we conclude that 20% of non-synonymous are weakly deleterious and segregate as high frequency polymorphisms.

P. pacificus and *P. exspectatus* is very homogeneous along the chromosomes (Fig. 2). This very homogeneous divergence between the sister species suggests a constant accumulation of mutations and the absence of large-scale mutation rate variation. Therefore, the correlated variation of genetic diversity within clades that is in strong contrast to the between clade and interspecies diversity, probably reflects population genetic processes common to all clades. We speculate that the residual fluctuations in the inter-clade divergence are likely due to diversity fluctuations in the population ancestral to all clades, which are gradually swamped by accumulating divergence.

Deleterious variants at high frequencies. Restricting the analysis of the global *P. pacificus* diversity to coding regions, we find a mean silent site diversity π_{si} = 1.2±1.2% and a non-synonymous site diversity of π_{ns} = 0.3 ±0.3 % (Supplemental Figure 4). To obtain further insight into purifying selection on coding regions, we calculated the ratio of non-synonymous and synonymous differences (dN/dS) for all pairwise comparisons of the 104 strains and plot this against their genome wide distance (Fig. 3). Closely related strains within one clade show dN/dS of 0.5, suggesting that half the non-synonymous mutations were eliminated quickly. With increasing distance, the dN/dS ratio drops from 0.5 to 0.3 when comparing strains between clades. Finally, we find dN/dS=0.27 for the comparison of *P. pacificus* with *P. exspectatus* 1:1 orthologs. Since genetic distances roughly correspond to the time of separation, dN/dS values at different distances probe purifying selection on different time scales. From the observation $dN/dS \approx 0.5$ for closely related strains, we conclude that 50% of non-synonymous substitutions are so deleterious that they are rarely seen in wild isolates and elimiated quickly. In contrast, another 20% of non-synonymous substitutions is weakly deleterious and segregates as high frequency polymorphisms. Such extensive non-neutral variation distorts genealogies and biases tests for adaptive evolution [22, 23]. The most immediate consequence of deleterious mutations is a reduction of diversity by background selection [1]. Indeed, we find a strong negative correlation between the fraction of coding sequences with diversity (Spearman's ρ=-0.35 for 100 kb windows).

Site frequency spectra (SFS) and linked selection. The histogram of SNVs present in k out of n strains (SFS) provides a rich summary of genetic diversity that is informative about the demographic and evolutionary history of populations. Particularly, the polarized or un-folded frequency spectrum allows disentangling the effects of demography and selection, although it is sensitive to errors in polarization [25]. To minimize such errors, we broke the genome into 50kb intervals and inferred local genealogical trees and the ancestral sequences (see methods). Blocks of 50kb are in substantial LD for most of the genome, but the existence of many SNVs is sufficient to resolve the phylogenetic relationship and to reliably infer ancestral character states. After rooting trees with the *P. exspectatus* genome for polarization, we investigated the SFS of derived alleles within the most deeply sampled clade C (Fig. 1B).

We plotted the SFS on a double logarithmic scale such that power-laws show as straight lines (Fig. 4A). At frequencies below 10%, the observed SFS is compatible with a $1/k$ decay as expected in neutrally evolving populations of constant size [26]. At higher frequencies, however, the SFS decreases much more rapidly, before increasing again for alleles close to fixation. At intermediate derived allele frequencies between 10-40%, the SFS is compatible with $1/k^2$ indicated as dashed line. The $1/k^2$ behavior and the non-monotonicity latter is expected if the dominant force changing allele frequencies is selection at linked loci or genetic draft [8, 25]. Synonymous, non-synonymous and noncoding polymorphisms all follow very similar distributions, suggesting that the dynamics of polymorphisms is quasi-neutral in the sense that their fate is determined by the genetic background rather than their own effect on fitness. While the SFS is insensitive to the type of mutation, stratifying the SFS by the overall nucleotide diversity in surrounding regions reveals a strong dependence of the shape of the SFS on coalescence time (Fig. 4B).

Disentangling the effects of demography and genetic draft on SFS can be subtle, since exponential expansion also results in an

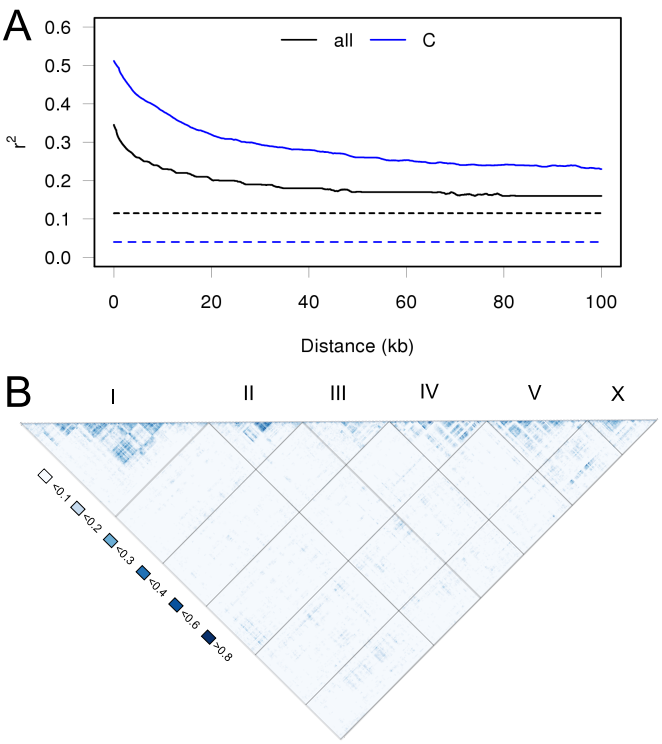


Figure 5. Linkage disequilibrium in *P. pacificus*. *(A)* LD drops over the first 20kb but stays on a relatively high level across longer distances (solid lines) and even across chromosomes (dashed lines). We attribute this high level of background LD to the strong differentiation between subclades as a result of geographic separation. Restricted to clade C, local LD is higher, while inter-chromosome LD is low. *(B)* Average LD between all 100kb windows for clade C across the genome was calculated as average r^2 for 10 pairs of biallelic SNVs with 5-95% allele frequency. LD across chromosomes is virtually absent indicating the effective reshuffling of chromosomes in outcrossing events. However, clade C shows several blocks spanning megabases in strong LD on most chromosomes.

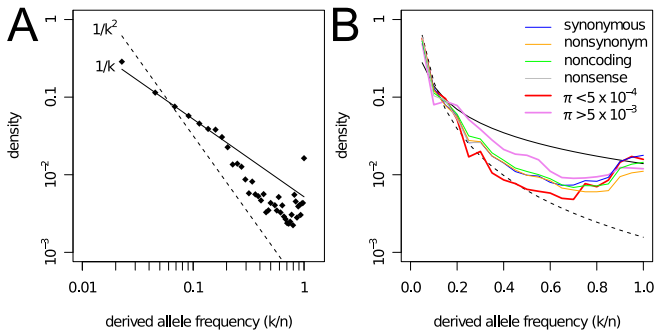


Figure 4. *(A)* Site frequency spectrum (SFS) of clade C is shown on a double logarithmic plot. If neutral diversity is shaped by constant genetic drift, one expects that the number of derived alleles present in k out of n strains decays as $1/k$, indicated as straight solid line. The dashed line indicates the corresponding expectation, $1/k^2$, under a genetic draft model (for rare alleles). At frequencies below 10%, the SFS is proportional to 1/k, at intermediate frequencies between 10% and 50%, the SFS is steeper and compatible with $1/k^2$. *(B)* The SFS decays less rapidly in regions of high than of low diversity. In constrast, distinct functional categories have almost indistinguishable SFSs.

$1/k^2$ decay. Two observations, however, argue against dominant demographic effects. First, the SFS expected under draft differs from that under population expansion at the high frequency end where the number of copies of a mutation k approaches the number of strains n: Models with draft predict an increase of SFS as k approaches n, while exponential expansion does not [25]. Consistent with draft, we observe a marked up-tick beyond 60-70% derived allele frequency. Second, demographics should affect the entire genome in a similar manner. However, we observe different SFS in regions of high and low diversity with signatures of draft being more pronounced in low diversity regions (Fig. 4B). This observation is consistent with the expectation that draft reduces coalescence time and hence diversity. Young alleles and the rare end of the frequency spectrum are compatible with genetic drift. Overall, however, *P. pacificus* is not compatible with a neutral model and strong distortions of genealogies by purifying selection [27, 8] and possibly adaptation are a likely explanation for the observed SFS.

Linkage disequilibrium and haplotype structure. Finally, we investigated the pattern of LD in *P. pacificus*. We found significant local LD that drops over a distance of approx. 20kb, but remained higher than randomly expected (Fig. 5). When all strains are compared, substantial LD persists even across chromosome borders with r^2 values of 0.12 (Fig. 5b) which likely results from population structure (Supplemental Figure 3). In contrast, within the more closely related clade C strains, we find very little LD across chromosomes (r^2 = 0.005-0.01), but blocks of local LD, the largest of which span several Mb (Fig. 5a). The lack of LD across chromosomes supports our assertion that clade C is a rather well mixed population suitable for population genetic analysis. A complementary analysis of haplotype blocks (see Methods) also revealed shared haplotypes blocks of megabase size between clade C strains (Supplemental Figure 5). While long haplotype blocks conserved across continents were found in *C. elegans*, most of shared haplotype blocks in *P. pacificus* are clade-specific. Consistent with reduction of neutral diversity through linked selection, we find a moderate anticorrelation between LD and diversity (Spearman's ρ=-0.25 for 100 kb windows).

Discussion

By sequencing 104 strains of *P. pacificus* and a closely related outgroup *P. exspectatus*, we have shown that *P. pacificus* contains extensive population structure. Individual clades show similar levels of nucleotide diversity with a genome-wide profile that is highly correlated between clades. The diversity levels within clades are comparable to levels found in *C. elegans* [11]. In contrast, the between clade diversity can help to reveal evolutionary processes of more distantly related lineages within a selfing organism. By analyzing the ratio of synonymous and non-synonymous differences between strains at various distances, we characterized the strength of purifying selection on coding regions. 50% of non-synonymous mutations are so deleterious that they are not found in a typical population sample. The estimated fraction of 50% is supported based on the finding of the accompanying manuscript showing that *P. pacificus* mutation accumulation lines exhibit a dN/dS ratio of 1 (Weller et al. 2013). Within the sampled populations, another 20% of non-synonymous variants is weakly selected against and pruned only over time scales on the order of the separation between clades, which corresponds to roughly 1% divergence at silent sites. Combined with our finding that nucleotide diversity anticorrelates strongly with the fraction of the genome that is coding, we conclude that background selection plays an important role in shaping *P. pacificus* diversity.

Recent theoretical work has shown that background selection cannot be fully described by a reduced effective population size but results in substantial distortions of genealogies [27, 22]. Purifying selection on weakly deleterious mutations, in particular, manifests itself in SFS with a steeper decay and an up-tick at high derived allele frequency, both of which are characteristic signatures of genetic draft or linked selection [25, 8]. The signature of genetic draft in the SFS of *P. pacificus* is consistent with previous studies that proposed selection at linked sites as one important factor shaping genomic diversity of self-fertilizing nematodes [10, 28, 11]. In addition to distorted SFS, we also find megabase scale haplotype blocks in strong linkage disequilibrium, which also suggests ample linked selection. Whether on top of purifying selection adaptive substitutions or fluctuating selection play a prominent role remains currently unclear. McDonald-Kreitman type tests[29] that have been used to quantify adaptive evolution in *Drosophila* [30] are very vulnerable to segregating deleterious mutations[23] and hence not applicable to *P. pacificus*. Furthermore, those tests make the assumption that synonymous mutations are neutral, which remains unlikely in species with a compact genome [31]. Methods used in *Drosophila* to detect more recent adaptations based on signatures of hitch-hiking [32, 33] are also inapplicable in a organisms with long range LD. Nevertheless, the large number of weakly deleterious polymorphims as observed here, implies frequent fixation of the latter and requires compensatory mutations. We therefore suggest that *P. pacificus* is in a dynamic balance between deleterious and beneficial mutations [34]. In addition, the environment faced by *P. pacificus* is most likely changing on time scales shorter than the divergence between clades and mutations that are deleterious now might have been beneficial in the past.

P. pacificus and *C. elegans* occupy distinct ecological niches, which has resulted in substantially different natural histories. It has been speculated that the reduction of diversity observed in *C. elegans* which was caused by recent migration patterns and strong selective sweeps, might be linked to the human dispersal within the last centuries [11]. Our findings of higher diversity, strong population subdivision, recent transposon activity, large haplotype blocks, and background selection are in accordance with the predicted consequences of a preferentially selfing mode of reproduction [35]. Thus, our genome-wide analysis of *P. pacificus* populations describes a complementary picture to the one obtained from *C. elegans* and highlights the importance of comparative studies for nematode genome evolution. Finally, our catalogue of natural variation will form the basis for further studies associating phenotypic variability to the genome of *P. pacificus*.

Materials and Methods

Genomic library preparation. For preparation of genomic DNA, the MasterPure DNA purification Kit from Epicentre was used, resulting in high yields of clean DNA. DNA was quantified by Qubit measurement and diluted with TE to 20 ng/μl in a total volume of 55 μl. Genomic libraries were generated using the TruSeq DNA Sample Preparation Kit / v2 from Illumina. DNA was sheared with the given settings using the Covaris S2 System. Following the protocol, end repair, adenylation and index adapter ligation were performed. After running samples on a 2% agarose gel for 90 minutes, gel slices ranging in size from 400-500 bp were excised resulting in an insert size of approximately 300-400 bp. After fragments were amplified by PCR, libraries were validated on a Bioanalyzer DNA 1000 chip. All libraries were diluted to a concentration of 10 nM in 0,1% EB-Tween and pooled to 1, 4, 8, and 12-plexes). For mate pair libraries, clean genomic DNA was prepared using the Genomic-tip 100/G Kit from Qiagen. DNA was quantified by Qubit measurement, diluted with Tris-HCl to 150 ng/μl in a total volume of 70 μl and was sheared with the given settings (using SC 13) with the Hydroshear. For *P. exspectatus*, 3 kb and 5 kb Mate Pair libraries were generated using the Mate Pair Library v2 Kit for 2-5 kb Libraries from Illumina. Details about the genome assembly and annotation are described in the Supplementary Methods.

Alignment and variant calling. Low quality bases in the first 36bp of raw reads with a quality below 20 (error probability=1%) were masked and reads were trimmed at the first occurrence of a low quality (<20) base in the rest of the read. Reads were aligned to the Hybrid1 genome assembly of the *P. pacificus* PS312 strain (California) using stampy (version 1.0.12) [36]. Duplicate reads were removed and reads were locally realigned using GATK (version 2.1-13). SNVs and small indels were called using samtools (0.1.18)[37] excluding positions with >

100x coverage. Variants with a quality score below 20 were excluded. Large structural variations were called using cnv-seq [20]. The previously published low coverage Sanger assembly of the *P. pacificus* PS1843 (Washington) strain [19] was used to compare variant calls between the Sanger and Illumina platform. The Sanger assembly was aligned to the PS312 Hybrid1 reference genome using mugsy (version v1r2.2)[38]. Based on homozygous 474,285 SNV positions that were covered by Sanger data, the genotyping accuracy was defined as the fraction of SNVs, called by both platforms among all Illumina variant calls. For 938,544 SNVs obtained from the whole genome alignment of the PS1843 Sanger assembly, only 53% of variant calls agreed with the Illumina data. We attribute this high number of putative false positive SNVs to the low coverage (1x) of the Sanger data. 79% of Illumina-based indel calls were found to be in agreement with the Sanger data based on comparison of the equivalent indel region[39], however, 96% of Illumina-based indel calls overlapped with an Sanger-based indel in up to 10bp distance.

Population genetic analysis. For visualization of phylogenetic relationships one million SNVs with genotypes in all strains were randomly selected and concatenated as input for SplitsTree4 (version 4.12.6) [40]. Principal component analysis was done with EIGENSOFT (version 3.0) [41]. π_{si},π_{ns} and dN/dS values were normalized using the numbers of silent and nonsynonymous sites. Chromosome-wide plots were generated by concatenating Contigs (>1Mb) with respect to marker positions on the genetic map [24]. Contigs spanning $\sim$40% of the total assembly were excluded from these plots as they lack any genetic marker or were shorter than 1Mb. Shared haplotype blocks were identified by the program GERMLINE (version 1-5-1)[42]. The ancestral state of SNPs was inferred by building a maximum likelihood tree [43] of all clade C strains and the *P. exspectatus* sequence in non-overlapping windows of 50kb (20kb and 100kb yield similar results). We inferred the most likely ancestral sequences of all nodes of the tree using [44] and the substitution model FN81. For each of the resulting trees, the sequence at the node that is the parent of *P. exspectatus* was used as the ancestral sequence of clade C.

Data availability. All reads were submitted to the NCBI Sequence Read Archive. Variant calls, and the *P. exspectatus* genome assembly are available at http://www.pristionchus.org/variome/ .

1. Charlesworth, B, Morgan, M. T, & Charlesworth, D. (1993) The effect of deleterious mutations on neutral molecular variation. *Genetics* 134, 1289–303.
2. Maynard Smith, J & Haigh, J. (1974) The hitch-hiking effect of a favourable gene. *Genet Res* 23, 23–35.
3. Gillespie, J. H. (2000) Genetic drift in an infinite population. The pseudohitchhiking model. *Genetics* 155, 909–19.
4. Hill, W. G & Robertson, A. (1966) The effect of linkage on limits to artificial selection. *Genet Res* 8, 269–94.
5. Barton, N. H & Charlesworth, B. (1998) Why sex and recombination? *Science* 281, 1986–90.
6. Barton, N. (1998) The effect of hitch-hiking on neutral genealogies. *Genet Res* 72, 123–133.
7. Braverman, J. M, Hudson, R. R, Kaplan, N. L, Langley, C. H, & Stephan, W. (1995) The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. *Genetics* 140, 783–96.
8. Neher, R. A & Shraiman, B. I. (2011) Genetic draft and quasi-neutrality in large facultatively sexual populations. *Genetics* 188, 975–996.
9. Begun, D. J, Holloway, A. K, Stevens, K, Hillier, L. W, Poh, Y.-P, Hahn, M. W, Nista, P. M, Jones, C. D, Kern, A. D, Dewey, C. N, Pachter, L, Myers, E, & Langley, C. H. (2007) Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. *PLoS Biol* 5, 2534–2559.
10. Cutter, A. D & Choi, J. Y. (2010) Natural selection shapes nucleotide polymorphism across the genome of the nematode *Caenorhabditis briggsae*. *Genome Research* 20, 1103–1111. PMID: 20508143.
11. Andersen, E. C, Gerke, J. P, Shapiro, J. A, Crissman, J. R, Ghosh, R, Bloom, J. S, Félix, M.-A, & Kruglyak, L. (2012) Chromosome-scale selective sweeps shape *Caenorhabditis elegans* genomic diversity. *Nat Genet* 44, 285–290.
12. Kienle, S & Sommer, R. J. (2013) Cryptic variation in vulva development by cis-regulatory evolution of a hairy-binding site. *Nat Commun* 4, 1714.
13. Sinha, A, Sommer, R. J, & Dieterich, C. (2012) Divergent gene expression in the conserved dauer stage of the nematodes *Pristionchus pacificus* and *Caenorhabditis elegans*. *BMC Genomics* 13, 254.
14. Bumbarger, D. J, Riebesell, M, RÃ¶delsperger, C, & Sommer, R. J. (2013) System-wide rewiring underlies behavioral differences in predatory and bacterial-feeding nematodes. *Cell* 152, 109–119.
15. Denver, D. R, Clark, K. A, & Raboin, M. J. (2011) Reproductive mode evolution in nematodes: insights from molecular phylogenies and recently discovered species. *Mol Phylogenet Evol* 61, 584–592.
16. Kanzaki, N, Ragsdale, E. J, Herrmann, M, Mayer, W. E, & Sommer, R. J. (2012) Description of three pristionchus species (Nematoda: Diplogastridae) from Japan that form a cryptic species complex with the model organism *P. pacificus*. *Zoolog Sci* 29, 403–417.
17. Barrière, A, Yang, S.-P, Pekarek, E, Thomas, C. G, Haag, E. S, & Ruvinsky, I. (2009) Detecting heterozygosity in shotgun genome assemblies: Lessons from obligately outcrossing nematodes. *Genome Res* 19, 470–480.
18. Morgan, K, McGaughran, A, Villate, L, Herrmann, M, Witte, H, Bartelmes, G, Rochat, J, & Sommer, R. J. (2012) Multi locus analysis of *Pristionchus pacificus* on la Reunion island reveals an evolutionary history shaped by multiple introductions, constrained dispersal events and rare out-crossing. *Mol Ecol* 21, 250–266.
19. Dieterich, C, Clifton, S. W, Schuster, L. N, Chinwalla, A, Delehaunty, K, Dinkelacker, I, Fulton, L, Fulton, R, Godfrey, J, Minx, P, Mitreva, M, Roeseler, W, Tian, H, Witte, H, Yang, S.-P, Wilson, R. K, & Sommer, R. J. (2008) The *Pristionchus pacificus* genome provides a unique perspective on nematode lifestyle and parasitism. *Nat Genet* 40, 1193–1198.
20. Xie, C & Tammi, M. T. (2009) Cnv-seq, a new method to detect copy number variation using high-throughput sequencing. *BMC Bioinformatics* 10, 80.
21. Mayer, W. E, Schuster, L. N, Bartelmes, G, Dieterich, C, & Sommer, R. J. (2011) Horizontal gene transfer of microbial cellulases into nematode genomes is associated with functional assimilation and gene turnover. *BMC Evol Biol* 11, 13.
22. Seger, J, Smith, W, Perry, J, Hunn, J, Kaliszewska, Z, Sala, L, Pozzi, L, Rowntree, V, & Adler, F. (2010) Gene genealogies strongly distorted by weakly interfering mutations in constant environments. *Genetics* 184, 529.
23. Messer, P. W & Petrov, D. A. (2012) The McDonald-Kreitman test and its extensions under frequent adaptation: Problems and solutions. *arXiv:1211.0060*.
24. Srinivasan, J, Sinz, W, Jesse, T, Wiggers-Perebolte, L, Jansen, K, Buntjer, J, van der Meulen, M, & Sommer, R. J. (2003) An integrated physical and genetic map of the nematode *Pristionchus pacificus*. *Mol Genet Genomics* 269, 715–722.
25. Neher, R. A & Hallatschek, O. (2012) Genealogies of rapidly adapting populations. *Proc Natl Acad Sci USA*.
26. Wakeley, J. (2008) *Coalescent theory*. (Roberts & Company).
27. Walczak, A. M, Nicolaisen, L. E, Plotkin, J. B, & Desai, M. M. (2012) The structure of genealogies in the presence of purifying selection: A "fitness-class coalescent". *Genetics* 190, 753–779.
28. Rockman, M. V & Kruglyak, L. (2009) Recombinational landscape and population genomics of *Caenorhabditis elegans*. *PLoS Genet* 5, e1000419.
29. McDonald, J. H & Kreitman, M. (1991) Adaptive protein evolution at the adh locus in *Drosophila*. *Nature* 351, 652–4.
30. Sella, G, Petrov, D. A, Przeworski, M, & Andolfatto, P. (2009) Pervasive natural selection in the *Drosophila* genome? *PLoS Genet* 5, e1000495.
31. Lawrie, D. S, Messer, P. W, Hershberg, R, & Petrov, D. A. (2013) Strong purifying selection at synonymous sites in *D. melanogaster*. *PLoS Genet* 9, e1003527.
32. Andolfatto, P. (2007) Hitchhiking effects of recurrent beneficial amino acid substitutions in the *Drosophila* melanogaster genome. *Genome Res* 17, 1755–62.
33. Macpherson, J. M, Sella, G, Davis, J. C, & Petrov, D. A. (2007) Genomewide spatial correspondence between nonsynonymous divergence and neutral polymorphism reveals extensive adaptation in Drosophila. *Genetics* 177, 2083–99.
34. Goyal, S, Balick, D. J, Jerison, E. R, Neher, R. A, Shraiman, B. I, & Desai, M. M. (2012) Dynamic mutation selection balance as an evolutionary attractor. *Genetics* 191, 1309–1319.
35. Charlesworth, D & Wright, S. I. (2001) Breeding systems and genome evolution. *Curr Opin Genet Dev* 11, 685–690.
36. Lunter, G & Goodson, M. (2011) Stampy: a statistical algorithm for sensitive and fast mapping of illumina sequence reads. *Genome Res* 21, 936–939.
37. Li, H, Handsaker, B, Wysoker, A, Fennell, T, Ruan, J, Homer, N, Marth, G, Abecasis, G, Durbin, R, & Subgroup, . G. P. D. P. (2009) The sequence alignment/map format and samtools. *Bioinformatics* 25, 2078–2079.
38. Angiuoli, S. V & Salzberg, S. L. (2011) Mugsy: fast multiple alignment of closely related whole genomes. *Bioinformatics* 27, 334–342.
39. Krawitz, P, Rödelsperger, C,Jäger, M, Jostins, L, Bauer, S, & Robinson, P. N. (2010) Microindel detection in short-read sequence data. *Bioinformatics* 26, 722–729.
40. Huson, D. H & Bryant, D. (2006) Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* 23, 254–267.
41. Patterson, N, Price, A. L, & Reich, D. (2006) Population structure and eigenanalysis. *PLoS Genet* 2, e190.
42. Gusev, A, Lowe, J. K, Stoffel, M, Daly, M. J, Altshuler, D, Breslow, J. L, Friedman, J. M, & Pe'er, I. (2009) Whole population, genome-wide mapping of hidden relatedness. *Genome Res* 19, 318–326.
43. Price, M. N, Dehal, P. S, & Arkin, A. P. (2009) Fasttree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 26, 1641–50.
44. Knight, R, Maxwell, P, Birmingham, A, Carnes, J, Caporaso, J. G, Easton, B. C, Eaton, M, Hamady, M, Lindsay, H, Liu, Z, Lozupone, C, McDonald, D, Robeson, M, Sammut, R, Smit, S, Wakefield, M. J, Widmann, J, Wikman, S, Wilson, S, Ying, H, & Huttley, G. A. (2007) Pycogent: a toolkit for making sense from sequence. *Genome Biol* 8, R171.

The genome of Pristionchus pacificus is shaped by opposing forces of biased mutations and counterbalancing selection.

Andreas M. Weller[1], Christian Rödelsperger[1], Gabi Eberhardt[1], Ruxandra I. Molnar[1] and Ralf J. Sommer[1]

[1] *Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany*

(Dated: August 6, 2013)

Base substitution mutations are a major source of genetic novelty and mutation accumulation line (MAL) studies revealed a nearly universal AT bias in de novo mutation spectra. While a comparison of de novo mutation spectra with the actual nucleotide composition in the genome suggests the existence of general counterbalancing mechanisms, little is known about the evolutionary and historical details of these opposing forces. Here, we correlate MAL derived mutation spectra with patterns observed from population resequencing. Variation observed in natural populations has already been subject to evolutionary forces. Distinction between rare and common alleles, the latter of which are close to fixation and of presumably older age, can provide insight into mutational processes and their influence on genome evolution. We provide a genome-wide analysis of de novo mutations in 22 MALs of the nematode *Pristionchus pacificus* and compare the spectra with natural variants observed in resequencing of 104 natural isolates. MALs show an AT bias of 5.3, one of the highest values observed to date. In contrast, the AT bias in natural variants is much lower. Specifically, rare derived alleles show an AT bias of 2.4, whereas common derived alleles close to fixation show no AT bias at all. These results indicate the existence of a strong opposing force and they suggest that the GC content of the P. pacificus genome is in equilibrium. We discuss GC biased gene conversion as potential mechanism acting against AT biased mutations. This study provides insight into genome evolution by combining MAL studies with natural variation.

I. INTRODUCTION

Base-substitution mutations are a major source of genetic novelty and provide the raw material on which evolutionary forces will act. An accurate picture of the processes generating de novo mutations and their frequencies is fundamental for the understanding of evolution (LYNCH, 2007), as well as for mutational processes causing diseases (LU *et al.*, 2011). Mutation accumulation (MA) line analysis became the method of choice to study the frequencies and spectra of de novo mutations (MUKAI, 1964). Most recently, MA lines have been used in model species in combination with whole genome sequencing. Among multicellular organisms, this included *Caenorhabditis elegans* (DENVER *et al.*, 2009, 2012) and *C. briggsae* (BAER *et al.*, 2005; HOWE *et al.*, 2010), *Drosophila melanogaster* (KEIGHTLEY *et al.*, 2009), *Daphnia pulex* (SEYFERT *et al.*, 2008), and *Arabidopsis thaliana* (OSSOWSKI *et al.*, 2010). These studies showed that the base substitution mutation rate (μ_{bs}, measured in mutations per site per generation), while stable to one order of magnitude, is positively correlated with genome size, but negatively correlated with the effective population size across organisms (LYNCH, 2007). In contrast, the mutation spectra vary considerably between species. In all studies done so far, de novo mutations from G/C to A/T outweigh mutations from A/T to G/C, a phenomenon that has been described as AT bias (LYNCH, 2007). Such a mutational bias creates a pressure towards higher AT content of genomes. However, the actual AT content of most genomes is far from what would be predicted by these mutational forces alone, indicating the existence of other forces that operate in the opposite direction (LYNCH, 2007).

The comparison of MA line derived mutational spectra with genome composition in organisms covering all domains of life has been most powerful in providing insight into genome evolution. One additional fruitful line of investigation will be the correlation of MA line derived patterns of mutation frequencies and spectra with those patterns observed from natural variation studies. Variation observed in natural populations has already been affected by adaptive and non-adaptive forces and can therefore capture the result of mutational processes and evolutionary forces more accurately than a single reference genome sequence of the organism in question. Such studies will be most informative for species with a comprehensive collection of wild isolates that cover the complete geographic range and micro-evolutionary time frame of diversification of the organisms. Specifically, the comparison of MA line derived patterns with those of rare and common alleles, the latter of which are close to fixation and thus of presumptive older age, might provide valuable new insight into mutational processes and their influence on genomes and evolution.

One candidate for such studies is the nematode *Pristionchus pacificus*. This species has recently been developed as a model system for integrative studies in evolutionary biology (SOMMER, 2009). *P. pacificus* has well-developed functional tools for genetic, genomic and transgenic studies. Detailed knowledge about its ecology based on its tight association with scarab beetles, its population genetics and natural history are available (MCGAUGHRAN *et al.*, 2013). MA lines have been generated in *P. pacificus* (MOLNAR *et al.*, 2011) and more than 600 wild isolates covering its cosmopolitan distribution are available in the laboratory. Four major clades, called A to D, can be distinguished within *P. pacificus*, with two clades (B and C) being endemic to the Mascarene Islands La Reunion and Mauritius in the Indian Ocean (MORGAN *et al.*, 2012). Of particular interest for genomic studies is *P. pacificus* clade C that underwent a massive radiation on La Reunion (MORGAN *et al.*, 2012). A subset of 104 *P. pacificus* wild isolates had their genome of 169 MB re-sequenced, which revealed strong divergence resulting in more than 6 million single nucleotide polymorphisms (SNP)

(Rödelsperger *et al.* 2013, accompanying paper). Furthermore, the patterns of divergence and global dispersal between the four *P. pacificus* clades and the repeated invasion of the geologically young volcanic island of La Reunion (HERRMANN *et al.*, 2010; MCGAUGHRAN *et al.*, 2013) indicated that the *P. pacificus* genome is shaped by weakly deleterious mutations and linked selection (Rödelsperger *et al.* 2013, accompanying paper). These studies provide a genomic and phylogenetic framework for evolutionary comparisons of mutational patterns.

Here, we re-sequenced the full genome of 22 independent MA lines and followed the mutational patterns along three different avenues of analysis. First, we studied de novo mutations in MA lines and provide estimates of μ_{bs} for *P. pacificus* as baseline for reconstruction of the evolutionary history of the species. We used this μ_{bs} estimate to calculate the time to the most recent common ancestor (TMRCA) of the *P. pacificus* clades. Second, we investigated the role of chromosome location and coding state on mutation rates. By comparing signatures of selection against non-synonymous substitutions between MA line mutations and natural variants, we can test the assumption of neutral mutation accumulation in MA lines. Finally, we studied the mutation spectrum in terms of transition - transversion ratios and AT bias. We observed strong differences in the mutation spectra between MA line derived de novo mutations and natural variants. In addition, the mutational bias in natural variants is negatively correlated with their age, suggesting the result of selection processes in nature that balance mutation spectra.

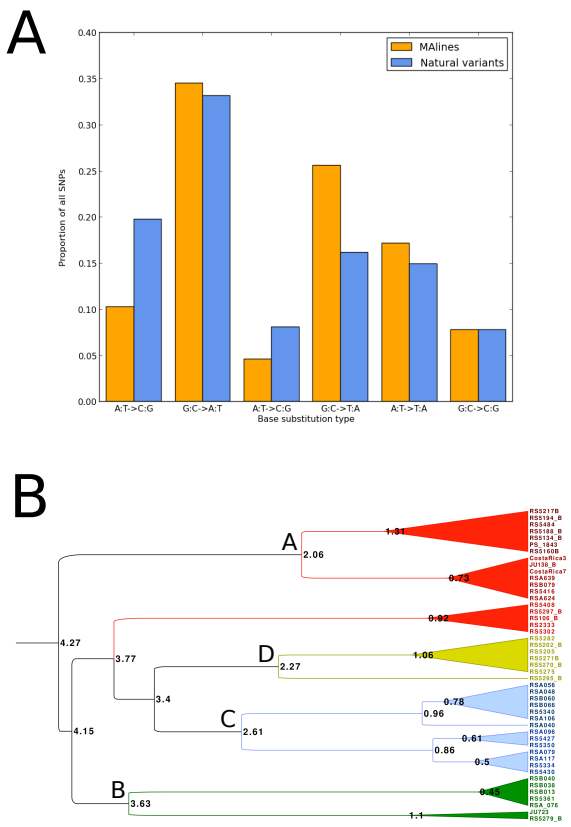


Figure 1 **(A)** Comparison between the base-substitution types in MA lines mutations and natural variants as percentage of all substitutions. **(B)** Phylogeny of *P. pacificus* derived from the mutation rate estimated in this study and 5 MB of sequence. Number at nodes represents time to the most recent common ancestor of the node in million generations.

II. RESULTS

A. Whole genome sequencing of MA lines

P. pacificus MA lines were propagated for 142 generations and first analyzed for mutation accumulation and estimates of the TMRCA in the mitochondrial genome (MOLNAR *et al.*, 2011). Here, we sequenced the nuclear genome of 22 MA lines using next generation sequencing technology (see Materials and Methods). In total, 746 million reads were generated resulting in mean genome coverage of 16x. After alignment to the reference, a total genome sequence of 49-141MB per MA line passed the quality cutoff for mutation calling. To avoid false-positive mutation calls due to assembly errors, we used a conservative pipeline that selected for high quality putative mutations and compared them between lines to filter for duplicate sites. We evaluated 21 mutations by PCR and Sanger sequencing and confirmed all of them as valid calls indicating that the number of false positive mutation calls is limited.

B. De novo mutations and mutation frequencies

In total, we identified 802 de novo mutations in the nuclear genome of these 22 MA lines (Fig.1A). Mutations are equally distributed over all chromosomes (Fig.2) indicating no evidence for mutation hot spots in *P. pacificus*. We calculated the rate of base-substitution μ_{bs} as the number of mutations divided by the number of generations and sites. Thus, the line-specific base substitution rate was derived by taking the number of sites actually considered in each line into account. The mean μ_{bs} is 2e-9 with a non-significant (χ^2 test, P = 0.27) variation among the 22 MA lines from 1.36 - 2.64e-9. By analyzing a sequence of 5 MB in 47 representative wild isolates, we used the identified μ_{bs} to calculate the TMRCA of the species. Following an approach described elsewhere (MOLNAR *et al.*, 2011), we estimated the TMRCA for all *P. pacificus* strains at 4.27e6 generations (Fig.1B).

C. Influence of coding state on mutation rate

Next, we evaluated the effect of mutations on gene function (Fig.1A and Table 1). Of the 802 mutations, 127 were found in exons, including 9 premature stop codons, 88 non-synonymous and 30 synonymous changes. This distribution does not differ from an assumed random distribution (Fishers Exact Test, P = 0.76). Five of the genes with stop codons were orphan genes without a *C. elegans* 1:1 ortholog, while three

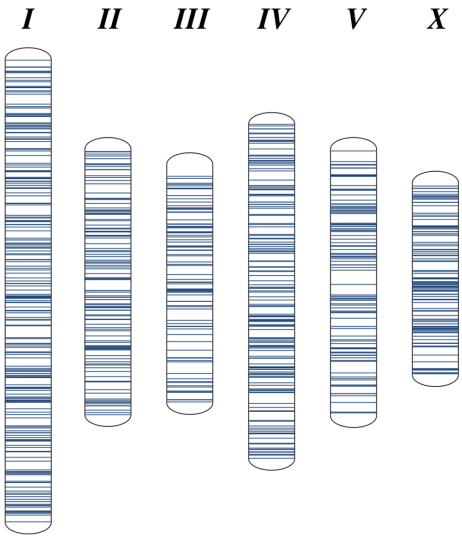


Figure 2 Chromosome location of de novo mutations in the MA lines. Each vertical bar represents a single mutation. Chromosome sizes are drawn to scale according to their genetic size.

genes have a *C. elegans* 1:1 ortholog with unknown function. The dN/dS rate ω serves as a measurement of the strength of selection acting on a sequence, with ratios >1, 1 and <1 indicating positive, neutral and purifying selection, respectively (YANG and BIELAWSKI, 2000). We used a Maximum Likelihood approach to calculate ω as 1.06 over all 118 genes with exonic mutations in the MA lines. Interestingly, variants found in the same 118 genes in the genome of the available 104 natural isolates have a ω of only 0.17. These results suggest that mutations observed in MA lines follow neutral expectations, whereas natural variation within the same genes indicates the action of purifying selection.

D. Mutation spectra and AT bias

To obtain insight into the processes generating mutations, we analyzed the mutation spectra in the 22 sequenced MA lines. We observed a ratio of transitions to transversions (Ts/Tv) in MA lines and natural populations of 0.8 and 1.1, respectively. The Ts/Tv ratio is higher in exons than in introns and intergenic regions, both in MA lines and natural variants (Fishers Exact Test, $P < 0.0001$). Next, we analyzed μ_{bs} separately for all six possible base substitution types. We found a strong bias towards G:C to T:A transversions and G:C to A:T transitions (Fig.1A) On average, we observed 3.9 mutations towards A/T for every mutation towards G/C. We calculated the AT bias by weighting for GC content in the genome following standard procedures (LYNCH, 2007). We found an AT bias of 5.3, indicating a strong decrease in GC content of the genome as a consequence of de novo mutations. This value represents one of the highest AT bias values based on whole genome sequencing known to date (Table 2).

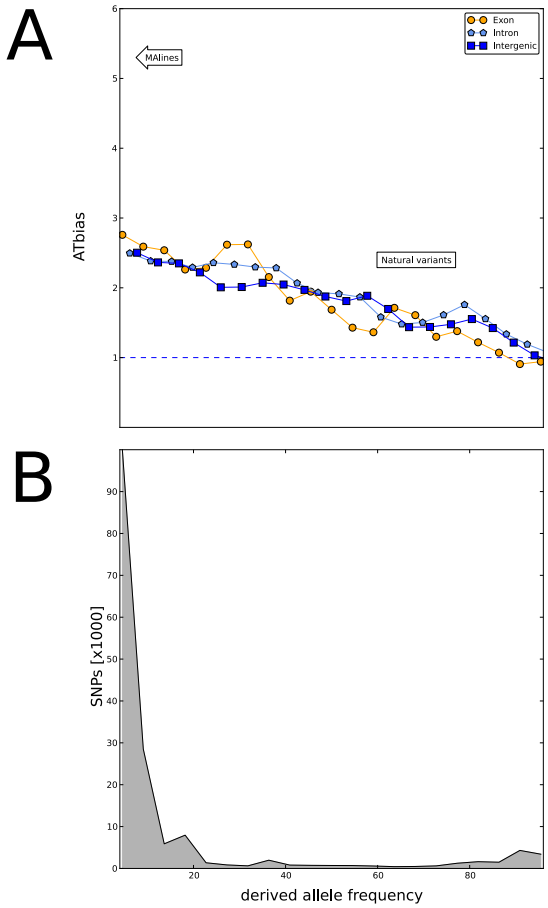


Figure 3 **(A)** AT bias in the natural variants of *P. pacificus* divided by derived allele frequencies in the population. MA line derived AT bias is included for comparison. **(B)** Number of natural variants studied per derived allele frequency.

E. AT bias and population structure

Finally, we compared the mutation spectra between MA lines and natural variants to infer the mode by which selective forces act on newly arising natural variants. In total, 655,705 natural variants were considered. An overwhelming majority of these variants were present in only a few strains (Fig.3). In contrast to the AT bias of *P. pacificus* MA line mutations, the mean AT bias of the natural variants is only 2.4. In addition, the AT bias of derived alleles is negatively correlated with their frequency among the sampled strains, i.e. alleles close to fixation show no bias compared to rare alleles (Fig.3A). The AT bias does not differ between coding and noncoding regions (Fig.3A). Together, these results suggest that a strong force exists counterbalancing the mutational AT bias.

III. DISCUSSION

In this study we analyzed base-substitution mutations in 22 MA lines and compared the mutation patterns and spectra with natural variants obtained in 104 wild isolates of *P. pacificus*. The mutation rate of *P. pacificus* is surprisingly similar to the μ_{bs} of 0.8 - 2.1e-9 reported for several Caenorhabditis species (DENVER *et al.*, 2012). These findings suggest a stable mutation rate in nematodes over an evolutionary separation time of more than 200 million years, the expected time to the TMRCA of *Pristionchus* and *Caenorhabditis* (DIETERICH *et al.*, 2008). As the mutation rate found in MA lines in *Drosophila melanogaster* (KEIGHTLEY *et al.*, 2009) is also very similar, our result is another indicator that mutation rates can be generalized even among higher animal taxa. This result lends credence to phylogenetic analysis based on molecular clocks, such as our estimation of the TMRCA of the *P. pacificus* clades, which is at the same order of magnitude as previous mtDNA-based analyses (MOLNAR *et al.*, 2011).

The finding of equal mutation rates for nematodes with very different effective population sizes (Rödelsperger *et al.* 2013, accompanying paper) and gene numbers (DIETERICH *et al.*, 2008) is unexpected under the mutation-barrier framework. The genomic content of coding DNA and the effective population size are argued to be the lower limit for an ever-decreasing mutation rate driven by natural selection, as the overwhelming majority of mutations are deleterious (LYNCH, 2010a) (SUNG *et al.*, 2012).

We did not find a deviance from random expectations between individual MA lines, chromosomes or coding states. In addition, the effects of mutations in exons were randomly distributed between stop codons, non-synonymous and synonymous mutations. A formal analysis of dN/dS further showed that there is no evidence for selection in the MA lines, while there is a clear trace of strong purifying selection in the natural variants. These results follow the a priori expectation that the benign lab environment leads to an unbiased accumulation of mutations in MA lines. Together, the ω of 1 and the random distribution of mutations are in accordance with the central premise of MA experiments, observing de novo mutations without the strong influence of selection.

The spectrum of de novo mutations in *P. pacificus* shows an AT bias of 5.3, one of the highest values observed by high-throughput sequencing to date (Table 2) (LYNCH, 2010b). This result allows two important conclusions. First, the nearly identical mutation frequencies of *P. pacificus* and *C. elegans* result from completely different mutation spectra as *C. elegans* has an AT bias of only 2.2. One possible explanation for these findings could be that *P. pacificus* and *C. elegans* MA lines have been exposed to different mutagenic forces, although the laboratory conditions of the MA lines should have been reasonably similar. Given the observed differences in the mutational forces acting on *P. pacificus* and *C. elegans*, the identical mutation frequencies might therefore result from a selection-like process. It should be noted that the proofreading activities of the DNA replication machinery has not been subject to direct investigations and therefore, it remains unknown if and to what extend proofreading activities are altered and locally optimized during the course of evolution.

The second major conclusion deriving from the high AT bias observed in MA lines is the missing correlation with the GC content of the *P. pacificus* genome. If de novo mutations were the only factor influencing nucleotide composition one would expect a GC content of only 16% for the *P. pacificus* genome (LYNCH, 2007). However, the observed GC content is 42%, is stable over all studied natural isolates of the species (Rödelsperger *et al.* 2013, accompanying paper) and is thus far from a pure mutation-based equilibrium. This strong discrepancy suggests the existence of powerful forces acting in the opposite direction. Indeed, the analysis of the AT bias in natural isolates provides strong support for the existence of such counterbalancing forces. First, the mean AT bias in natural isolates is 2.4 and already much lower than what is seen in MA lines. Derived alleles that are close to fixation and therefore have been exposed to selection for a longer time period show no AT bias at all. This observation strongly suggests that AT-driving variants are eliminated from the genome and that this tendency increases with the age of the variant. Furthermore, observing no AT bias in those variants that are closest to fixation suggest that the GC content of the *P. pacificus* genome is in an equilibrium state. This conclusion is also supported by the observations that the sister species *P. exspectatus* has a similar GC content (Rödelsperger *et al.* 2013, accompanying paper).

While empirical studies in multicellular organisms are sparse, there is ample evidence that GC biased gene conversion (gcgcBGC) represents one of the major counterbalancing forces acting opposite to the AT bias (LYNCH, 2007). gcBGC is a mechanism to resolve double-strand breaks by using a homologous region (e.g. from the sister chromatid) as a repair template. Mismatches between the sequences at heterozygous sites are non-randomly resolved in favor of G:C nucleotides, which leads to an increase of genomic GC content (LYNCH, 2007). Even though natural selection influences survival rates of organisms as a whole, while gcBGC acts through transmission ratio of individual alleles, both processes would lead to the same outcome of increased frequencies of G:C alleles over time. Given the currently available technology, it is virtually impossible to distinguish between the effects of natural selection and gcBGC (LYNCH, 2007; NAGYLAKI, 1983; WALSH, 1983). Several observations suggest the frequency of gcBGC to be magnitudes higher than mutations, with values of 10-5 in multicellular species and up to 10-2 in fungi (MARAIS, 2003). Given that the AT bias in *P. pacificus* is not different between coding and noncoding regions, we assume gcBGC to be a likely candidate for the observed phenomenon, although there are no direct observations on gcBGC in *P. pacificus*.

A possible explanation for the increased rate of both AT bias by mutation and GC-bias by gcBGC might be in an increased role of oxidative stress in *P. pacificus*. Oxidative stress boosts AT biased mutations by deamination of cytosine and guanine (TEEBOR *et al.*, 1988). At the same time, it causes double-strand breaks which might possibly be repaired via gcBGC (LETAVAYOVÁ *et al.*, 2006; LYNCH, 2010a). Thus, oxidative stress might induce opposing forces on the genomic GC content that cancel each other out, making them effec-

tively invisible if only fixed natural variants are studied. Potential support for a role of oxidative stress comes from the known ecosystem of *P. pacificus* and related nematodes. Unlike *C. elegans*, *Pristionchus* lives in tight association with scarab beetles in a necromenic interaction (HERRMANN *et al.*, 2010, 2007). Nematodes rest on the living beetle in the dauer stage (WELLER *et al.*, 2010) and only resume development on the beetle carcass to feed on all type of microbes (MCGAUGHRAN *et al.*, 2013). Consistent with this stressful environment the *P. pacificus* genome shows a strong increase in the number of enzymes potentially involved in the detoxification of xenobiotics. For example, *P. pacificus* contains 198 gene predictions encoding for cytochrome P450 enzymes, whereas *C. elegans* has only 59 such gene predictions (DIETERICH *et al.*, 2008). While the ecosystem and the genome of *P. pacificus* strongly suggest a life under oxidative stress, direct evidence for these environmental circumstances on mutational processes is not yet available and awaits future analysis.

IV. MATERIALS AND METHODS

V. MA LINE CREATION AND LIBRARY PREPARATION

The creation of mutation-accumulation lines has been described in detail by (MOLNAR *et al.*, 2011). Briefly, 100 lines were initiated from the offspring of a single individual of the strain *P. pacificus* PS312. The lines were propagated for 142 generations under benign conditions by transfering a single random hermaphrodite per generation. The 82 surviving lines were frozen in liquid nitrogen (MOLNAR *et al.*, 2011). We later thawed 22 lines and extracted DNA for Illumina sequencing using the MasterPure DNA purification kit. DNA was first sheared using the Covaris S2 system. End repair, adenylation and ligation of one of the 12 index adapters was performed using reagents and the standard protocol of the Illumina TruSeq v2 sample preparation kit as described in (Rödelsperger *et al.* 2013, accompanying paper). Samples were selected for an insert size of 300-400bp by excision from an agarose gel. Library concentration was finally determined on a Bioanalyzer DNA 1000 chip and normalized to 10nM before pooling into 12-plexes. The libraries were sequenced in-house on an Illumina HiSeq 2000 sequencer.

VI. SEQUENCING AND VARIANT CALLING

Illumina reads were aligned to the reference genome (*P. pacificus* Hybrid Assembly) by stampy (version 1.0.13) (LUNTER and GOODSON, 2011). Duplicate reads were removed using the samtools (version 0.1.17) (LI *et al.*, 2009) "rmdup" command. After quality filtering, coverage per line was between 4x and 34x. Initial SNP calling was performed using samtools and bcftools (version 0.1.17) (LI *et al.*, 2009). In order to remove low-quality mutations as well as mutations called in several of the lines (which indicated a mutation already present in the founder animal), we employed a custom Python script for further conservative filtering. Putative mutations were accepted as candidates if they A) were covered by between 5 and 30 reads, B) had a minimum phred quality score of 10 and C) had at least an FQ-value of 40 (indicating a homozygous position). Candidates were then compared to the founder animal and all other MA lines. Mutations were accepted as unique only if A) not a single read in any other line indicated the same alternative base and B) at least 10 other lines had a high-quality reference consensus at the given position. We used classic Sanger sequencing to verify the accuracy of our mutation calling. Primers for 21 randomly selected mutations were designed using a custom pipeline implementing Primer3 (UNTERGASSER *et al.*, 2007) for initial primer finding and NCBI Blast+ (JOHNSON *et al.*, 2008) to confirm the uniqueness of primer candidates in the reference sequence. After sequencing, reads were aligned by NCBI Blast+ (JOHNSON *et al.*, 2008) and mutations were called using a custom Python script.

VII. MUTATION RATE AND SPECTRUM

The mutation rate was calculated as previously described by (DENVER *et al.*, 2009) as μ_{bs} = m/(LnT) where μ_{bs} is the base substitution rate per generation, m is the number of mutations, L is the number of lines, n is the number of sites that passed our criteria for mutation calling and T is the number of generations. We used X2 tests to evaluate the randomness of mutations between lines, chromosomes and functional categories. The expected random distribution was calculated based on the null expectation that e.g the total number of mutations would be uniformly distributed among the lines in accordance with the number of sites actually considered in each line. A custom Python script identified the coding context of each mutation using the set of genes described previously (BORCHERT *et al.*, 2010). All mutations that did not fall between the boundaries of a known gene were classified as intergenic. Three mutations that hit two different genes, on the sense- and antisense strand, were counted for both hits. We followed the procedure (LYNCH, 2007) p.125f to describe the mutation spectrum. Shortly, the AT bias is the ratio of mutations from A/T to G/C to the mutations in the other direction, both weighted by the genome content of nucleotides allowing the respective mutation type. The expected equilibrium AT composition is calculated as m/(1+m) where m is the AT bias. The Ts/Tv rate is simply the ratio of the total number of transitions to transversion mutations.

VIII. TMRCA ESTIMATION

We inferred relative divergence times for the phylogeny of the 47 P. pacificus strains using BEAST (v.1.7.4) (DRUMMOND *et al.*, 2012) on a randomly chosen dataset of 5MB. We inferred the topology of the trees under the birth-death speciation model (GERNHARD, 2008), using the HKY model of sequence evolution (HASEGAWA *et al.*, 1985) and assuming a strict molecular clock on all branches with the mutation rate found in the MA lines. We ran five independent runs for 10 million generations, sampling every 1000th generation.

We verified the convergence of runs by examining the effective sample size of the likelihood and posterior probability parameters for each analysis and by visual inspection in Tracer (v.1.5) (DRUMMOND and RAMBAUT, 2007). The trees and parameter estimates from the five runs were combined using LogCombiner (v.1.7.4) (DRUMMOND *et al.*, 2012). The results were considered reliable only when the effective sampling size of all parameters was above 100. Using TreeAnnotator (v.1.7.4) (DRUMMOND *et al.*, 2012), the samples from the posterior were summarized on the maximum credibility tree, with the posterior probability limit set to 0.8 and summarizing mean node heights.

IX. ANALYSIS OF NATURAL VARIANTS

For the comparison of AT bias between MA lines and natural strains, we used a dataset of 104 re-sequenced strains (Rödelsperger *et al.* 2013, accompanying paper). As described in (Rödelsperger *et al.* 2013, accompanying paper), ancestral alleles were determined based on maximum likelihood comparison to all 104 sequenced strains and the outgroup species *P. exspectatus*. In total, 655.705 natural variants from this dataset had a derived allele present in at least one strain. The codeml program from the PAML package (YANG, 1997) was used to infer a Maximum-Likelihood estimate of the dN/dS ratio ω in de novo mutations and natural variants separately. ω was averaged over all exons of the 118 genes carrying an exonic de novo SNP in the MA lines. The same exons were used in the natural variants to obtain a comparable estimate of ω. Codeml was run assuming the F3x4 codon frequency model and a uniform ω across branches.

X. DATA AVAILABILITY

All reads were submitted to the NCBI Sequence Read Archive. The reference genome and linkage maps for *P. pacificus* are available at http://www.pristionchus.org.

Acknowledgments

We thank Dr. R. Neher, F. Zanini and Dr. A. Künstner for many fruitful discussions and for critically reading the manuscript. This study was funded by the Max Planck Society.

References

BAER, C. F., F. SHAW, C. STEDING, M. BAUMGARTNER, A. HAWKINS, *et al.*, 2005 Comparative evolutionary genetics of spontaneous mutations affecting fitness in rhabditid nematodes. Proceedings of the National Academy of Sciences of the United States of America **102**: 5785–5790.

BORCHERT, N., C. DIETERICH, K. KRUG, W. SCHÜTZ, S. JUNG, *et al.*, 2010 Proteogenomics of Pristionchus pacificus reveals distinct proteome structure of nematode models. Genome Research **20**: 837–846.

DENVER, D. R., P. C. DOLAN, L. J. WILHELM, W. SUNG, J. I. LUCAS-LLEDÓ, *et al.*, 2009 A genome-wide view of Caenorhabditis elegans base-substitution mutation processes. Proceedings of the National Academy of Sciences **106**: 16310–16314.

DENVER, D. R., L. J. WILHELM, D. K. HOWE, K. GAFNER, P. C. DOLAN, *et al.*, 2012 Variation in Base-Substitution Mutation in Experimental and Natural Lineages of Caenorhabditis Nematodes. Genome Biology and Evolution **4**: 513–522.

DIETERICH, C., S. W. CLIFTON, L. N. SCHUSTER, A. CHINWALLA, K. DELEHAUNTY, *et al.*, 2008 The Pristionchus pacificus genome provides a unique perspective on nematode lifestyle and parasitism. Nature Genetics **40**: 1193–1198.

DRUMMOND, A., and A. RAMBAUT, 2007 BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology **7**: 214+.

DRUMMOND, A. J., M. A. SUCHARD, D. XIE, and A. RAMBAUT, 2012 Bayesian Phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution **29**: 1969–1973.

GERNHARD, T., 2008 The conditioned reconstructed process. Journal of Theoretical Biology **253**: 769–778.

HASEGAWA, M., H. KISHINO, and T. YANO, 1985 Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of molecular evolution **22**: 160–174.

HERRMANN, M., S. KIENLE, J. ROCHAT, W. E. MAYER, and R. J. SOMMER, 2010 Haplotype diversity of the nematode Pristionchus pacificus on Réunion in the Indian Ocean suggests multiple independent invasions. Biological Journal of the Linnean Society **100**: 170–179.

HERRMANN, M., W. E. MAYER, R. L. HONG, S. KIENLE, R. MINASAKI, *et al.*, 2007 The Nematode Pristionchus pacificus (Nematoda: Diplogastridae) Is Associated with the Oriental Beetle Exomala orientalis (Coleoptera: Scarabaeidae) in Japan. Zoological Science **24**: 883–889.

HOWE, D. K., C. F. BAER, and D. R. DENVER, 2010 High rate of large deletions in Caenorhabditis briggsae mitochondrial genome mutation processes. Genome biology and evolution **2**: 29–38.

JOHNSON, M., I. ZARETSKAYA, Y. RAYTSELIS, Y. MEREZHUK, S. MCGINNIS, *et al.*, 2008 NCBI BLAST: a better web interface. Nucleic Acids Research **36**: W5–W9.

KEIGHTLEY, P. D., U. TRIVEDI, M. THOMSON, F. OLIVER, S. KUMAR, *et al.*, 2009 Analysis of the genome sequences of three Drosophila melanogaster spontaneous mutation accumulation lines. Genome Research **19**: 1195–1201.

LETAVAYOVÁ, L., E. MARKOVÁ, K. HERMANSKÁ, V. VLČKOVÁ, D. VLASÁKOVÁ, *et al.*, 2006 Relative contribution of homologous recombination and non-homologous end-joining to DNA double-strand break repair after oxidative stress in Saccharomyces cerevisiae. DNA Repair **5**: 602–610.

LI, H., B. HANDSAKER, A. WYSOKER, T. FENNELL, J. RUAN, *et al.*, 2009 The Sequence Alignment/Map format and SAMtools. Bioinformatics **25**: 2078–2079.

LU, P., X. HAN, J. QI, J. YANG, A. J. WIJERATNE, *et al.*, 2011 Analysis of Arabidopsis genome-wide variations before and after meiosis and meiotic recombination by re-sequencing Landsberg erecta and all four products of a single meiosis. Genome Research **22**: 508–518.

LUNTER, G., and M. GOODSON, 2011 Stampy: A statistical algorithm for sensitive and fast mapping of Illumina sequence reads. Genome Research **21**: 936–939.

LYNCH, M., 2007 *The Origins of Genome Architecture*. Sinauer Associates Inc, 1 edition.

LYNCH, M., 2010a Evolution of the mutation rate. Trends in Genetics **26**: 345–352.

LYNCH, M., 2010b Rate, molecular spectrum, and consequences of

human mutation. Proceedings of the National Academy of Sciences **107**: 961–968.

LYNCH, M., W. SUNG, K. MORRIS, N. COFFEY, C. R. LANDRY, *et al.*, 2008 A genome-wide view of the spectrum of spontaneous mutations in yeast. Proceedings of the National Academy of Sciences **105**: 9272–9277.

MARAIS, G., 2003 Biased gene conversion: implications for genome and sex evolution. Trends in Genetics **19**: 330–338.

MCGAUGHRAN, A., K. MORGAN, and R. J. SOMMER, 2013 Unraveling the evolutionary history of the nematode Pristionchus pacificus: from lineage diversification to island colonization. Ecol Evol : n/a.

MOLNAR, R. I., G. BARTELMES, I. DINKELACKER, H. WITTE, and R. J. SOMMER, 2011 Mutation Rates and Intraspecific Divergence of the Mitochondrial Genome of Pristionchus pacificus. Molecular Biology and Evolution **28**: 2317–2326.

MORGAN, K., A. MCGAUGHRAN, L. VILLATE, M. HERRMANN, H. WITTE, *et al.*, 2012 Multi locus analysis of Pristionchus pacificus on La Réunion Island reveals an evolutionary history shaped by multiple introductions, constrained dispersal events and rare out-crossing. Molecular Ecology **21**: 250–266.

MUKAI, T., 1964 The genetic structure of natural populations of Drosophila melanogaster. Genetics **50**: 1–19.

NAGYLAKI, T., 1983 Evolution of a large population under gene conversion. Proceedings of the National Academy of Sciences **80**: 5941–5945.

OSSOWSKI, S., K. SCHNEEBERGER, J. I. LUCAS-LLEDÓ, N. WARTHMANN, R. M. CLARK, *et al.*, 2010 The Rate and Molecular Spectrum of Spontaneous Mutations in Arabidopsis thaliana. Science **327**: 92–94.

SEYFERT, A. L., M. E. A. CRISTESCU, L. FRISSE, S. SCHAACK, W. K. THOMAS, *et al.*, 2008 The Rate and Spectrum of Microsatellite Mutation in Caenorhabditis elegans and Daphnia pulex. Genetics **178**: 2113–2121.

SOMMER, R. J., 2009 The future of evo-devo: model systems and evolutionary theory. Nat Rev Genet **10**: 416–422.

SUNG, W., M. S. ACKERMAN, S. F. MILLER, T. G. DOAK, and M. LYNCH, 2012 Drift-barrier hypothesis and mutation-rate evolution. Proceedings of the National Academy of Sciences **109**: 18488–18492.

TEEBOR, G. W., R. J. BOORSTEIN, and J. CADET, 1988 The Repairability of Oxidative Free Radical Mediated Damage to DNA: A Review. Int J Radiat Biol **54**: 131–150.

UNTERGASSER, A., H. NIJVEEN, X. RAO, T. BISSELING, R. GEURTS, *et al.*, 2007 Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Research **35**: W71–W74.

WALSH, J. B., 1983 Role of biased gene conversion in one-locus neutral theory and genome evolution. Genetics **105**: 461–468.

WELLER, A. M., W. E. MAYER, R. RAE, and R. J. SOMMER, 2010 Quantitative Assessment of the Nematode Fauna Present on Geotrupes Dung Beetles Reveals Species-Rich Communities with a Heterogeneous Distribution. Journal of Parasitology **96**: 525–531.

YANG, Z., 1997 PAML: a program package for phylogenetic analysis by maximum likelihood. Computer applications in the biosciences : CABIOS **13**: 555–556.

YANG, Z., and J. P. BIELAWSKI, 2000 Statistical methods for detecting molecular adaptation. Trends in Ecology & Evolution **15**: 496–503.

Table I Properties of de novo mutations in MA lines

Coding state	μ_{bs}	A:T to G:C[a]	G:C to A:T	A:T to C:G	G:C to T:A	G:C to C:G	A:T to T:A	N	Ts:Tv	AT bias[b]
Exon	2e-09	12	49	3	38	9	16	127	0.92	5.3
Intron	1.9e-09	33	108	17	79	22	73	332	0.74	5.9
Intergenic	2.1e-09	39	118	17	89	31	49	343	0.84	4.8

[a]Total number in all 22 MA lines.
[b]Weighted by nucleotide composition at the coding state considered

Table II Base substitution types, Ts:Tv ratio and AT bias across species

Species	A:T to G:C[a]	G:C to A:T	A:T to C:G	G:C to T:A	G:C to C:G	A:T to T:A	N	Ts:Tv	AT bias[b]
P. pacificus	0.105	0.344	0.046	0.258	0.078	0.173	802	0.81	5.3
C. elegans (DENVER *et al.*, 2009)	0.105	0.205	0.113	0.281	0.077	0.22	391	0.448	2.235
D. melanogaster (KEIGHTLEY *et al.*, 2009)	0.19	0.305	0.115	0.155	0.115	0.121	174	0.977	1.509
A. thaliana (OSSOWSKI *et al.*, 2010)	0.118	0.588	0.071	0.094	0.071	0.059	85	2.4	3.625
S. cerevisiae (LYNCH *et al.*, 2008)	0.103	0.286	0.057	0.312	0.182	0.06	1250	0.638	3.74

[a]Fraction of all mutations
[b]Weighted by genomic nucleotide composition