# Population history and genomic diversity of Lampyris noctiluca and Luciola italica in the canton Vaud

Autor(en): **Duchen, Pablo / Höhna, Sebastian / Catalán, Ana**

Objekttyp: **Article**

Zeitschrift: **Bulletin de la Société Vaudoise des Sciences Naturelles**

Band (Jahr): **101 (2022)**

PDF erstellt am: **12.07.2024**

Persistenter Link: https://doi.org/10.5169/seals-1003693

**Nutzungsbedingungen**

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.
Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.
Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

**Haftungsausschluss**

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Ein Dienst der *ETH-Bibliothek*
ETH Zürich, Rämistrasse 101, 8092 Zürich, Schweiz, www.library.ethz.ch

**http://www.e-periodica.ch**

# Population history and genomic diversity of *Lampyris noctiluca* and *Luciola italica* in the canton Vaud

Pablo DUCHEN[1], Sebastian HÖHNA[2], Ana CATALÁN[1,3]

**Duchen P., Höhna S. & Catalán A., 2022. Population history and genomic diversity of *Lampyris noctiluca* and *Luciola italica* in the canton Vaud. *Bulletin de la Société Vaudoise des Sciences Naturelles* 101: 31-38.**

## Abstract

We studied and inferred the population history and genomic diversity of *Lampyris noctiluca* and *Luciola italica* in the canton Vaud. To do this, we used full genome data from one population of *L. noctiluca* and one population of *L. italica*. We found that genomic diversity in both species is reduced compared to other species of the same order, with *L. italica* showing the lowest diversity. Genome-wide inference of population size changes showed that both species experienced a sharp population decline in the last 500k years, with populations starting to recover only very recently. This population decline has been captured by genome-wide levels of Tajima's *D* (a statistic that measures departures from neutral evolution through both selection and demographic changes), where both species showed an average value that is positive, hinting at population contraction as a demographic explanation. We argue that our results constitute an important first step to characterize populations in terms of diversity and population trends, thus relevant for future conservation plans.

***Keywords:*** full genome data, population decline, PSMC.

**Duchen P., Höhna S. & Catalán A., 2022. Histoire des populations et diversité génomique de *Lampyris noctiluca* et *Luciola italica* dans le canton de Vaud. *Bulletin de la Société Vaudoise des Sciences Naturelles* 101: 31-38.**

## Résumé

Nous avons étudié et déduit l'histoire des populations et la diversité génomique de *Lampyris noctiluca* et *Luciola italica* dans le canton de Vaud. Dans ce but, nous avons utilisé les données complètes du génome d'une population de *L. noctiluca* et d'une population de *L. italica*. Nous avons constaté que la diversité génomique des deux espèces est réduite par rapport aux autres espèces du même ordre, *L. italica* présentant la plus faible diversité. L'inférence des changements de taille de population (au niveau génomique) a montré que les deux espèces ont subit une forte réduction de population au cours des 500 000 dernières années, celles-ci n'ayant commencé à se rétablir que très récemment. Ce déclin de la population peut être constaté par les valeurs du Tajima's *D* à l'échelle du génome, où les deux espèces présentent une valeur moyenne positive. Nous pensons que nos résultats constituent une première étape importante pour caractériser les populations en termes de diversité et de tendances démographiques, ce qui est indispensable pour les futurs plans de conservation.

***Mots-clés:*** déclin de la population, données génomiques complètes, PSMC.

---

[1] Institute for Organismal and Molecular Evolutionary Biology, Johannes Gutenberg University of Mainz, Germany.

[2] Department of Earth and Environmental Sciences, Paleontology & Geobiology, Ludwig-Maximilians University of Munich, Germany.

[3] Division of Evolutionary Biology, Ludwig-Maximilians University of Munich, Germany.

Correspondance: pduchnbo@uni-mainz.de, ana.catalan@gmail.com

© Société Vaudoise des Sciences Naturelles, 2022. Tous droits de reproduction réservés. CODEN: BSVAA6 ISSN: 0037-9603

## INTRODUCTION

Insect biodiversity plays an essential role in the ecosystem. As in most other species, this biodiversity is ultimately built upon the genetic diversity within each species which is maintained by natural selection and demographic processes. For instance, genetic diversity is gained when populations grow but can be also drastically lost when populations reduce in size. Therefore, characterizing the genetic diversity and demographic history of a species, in a location specific manner, provides us with information about population trends, which is relevant for conservation.

Fireflies are insects that play a significant role in nature. Fireflies are conspicuous beetles that produce bioluminescence, which is used as a toxicity warning signal in larvae and adults, as well as for finding mates and prey (LEWIS *et al.* 2004, LEWIS & CRATSLEY 2008). Additionally, fireflies are beneficial in agriculture due to their feeding preferences (FU & BENNO MEYER-ROCHOW 2013), thus ensuring the stability of the food web, and they are indicators of water quality and degree of contamination (FAUDZI *et al.* 2021).

In Europe, two firefly species show an ample geographical distribution: *Lampyris noctiluca* (NOVAK 2018) and *Lamprohiza splendidula* (SCHWALB 1961, GUZMÁN-ÁLVAREZ & DE COCK 2011). *L. noctiluca* is the most widely distributed European firefly, ranging from the Iberian and Balkan peninsula to Finland (NOVAK 2018), while *L. splendidula*'s distribution ranges from the Balkan peninsula to northern Germany (SCHWALB 1961, GUZMÁN-ÁLVAREZ & DE COCK 2011). In Switzerland, in addition to *L. noctiluca* and *L. splendidula*, we can find *Luciola italica* and *Luciola lusitanica* (GURCEL *et al.* 2020). In the case of *L. italica*, its habitat was, until recently, restricted from the alps towards the south. However, after its introduction in the canton Vaud in the late 1930s (FAES 1941), *L. italica* started migrating and establishing populations north of the alps (NEUMEYER 1991).

In the present study, we focused on the diversity and population history of *L. noctiluca* and *L. italica* from the canton Vaud. At the genomic level, comparing both species will help us look at differences between a well-established species (*L. noctiluca*) and a comparatively new colonizer (*L. italica*). At an ecological level, looking at both the current genomic diversity and population trends will help us establish a basis for the current ecological status of each species. Even though, the ecology and behavior of European fireflies is well documented (SCHWALB 1961, DE COCK 2009), genomic resources are lacking for these species. Such genomic resources are of uttermost importance for conservation, ecological, and adaptive studies.

## MATERIALS AND METHODS

### Sample collection

Whole firefly individuals were collected in Lausanne and surroundings for both *Lampyris noctiluca* and *Luciola italica* (figure 1). Specifically, after checking for species incidence in the *Centre Suisse de Cartographie de la Faune* (www.cscf.ch) and iNaturalist (https://www.inaturalist.org/), we searched for *L. noctiluca* in the area north of Belmont-sur-Lausanne, and *L. italica* across the northern shore of Lake Geneva, within the area of Lausanne. Sample collection began in late June 2020 and lasted until late July 2020.

## Genomic DNA extraction

We preserved all collected specimens in ethanol 96 % to ensure DNA stability for DNA extraction. We tested four extraction methods to optimize DNA extraction: two for high molecular DNA (MagAttract HMW DNA and Genomic- tip 500/G, Qiagen), and two for regular DNA extraction (NEB Monarch No. T3010 and Qiagen DNAeasy No. 69504). With the Magattract method we extracted the highest and longest DNA, with fragments longer than 50 kb. Genomic DNA quality and integrity was checked by gel electrophoresis and Qubit. The amount of total extracted DNA was ~10-15 µg from *L. noctiluca* and ~8-10 µg from *L. italica*.

## Sequencing

### Whole genome sequencing

Long DNA reads were produced with the Nanopore PromethIon platform using one flongle cell for each firefly species aiming at 60-80x coverage. Short read 150 bp paired-end Illumina at ~60 x coverage data was additionally produced from the same individual for each species. Nanopore bases were called with Guppy v4.0.11. For *L. noctiluca* a hybrid genome assembly was done with MaSuRCA v4.0.5 (Zimin *et al.* 2013). Two rounds of haplotype purging were done with Purge_dups v1.2.5 (Guan *et al.* 2020). Sequences not belonging to the class Insecta were identified using Blobtools v1.1.1 (Laetsch *et al.* 2017) and removed from the assembly. No genome polishing was performed for *L. noctiluca*, as polishing lead to worse BUSCO scores. Genome statistics were calculated with Quast v5.0.2 (Gurevich *et al.* 2013) and genome completeness was assessed with BUSCO v5.2.2 (Simão *et al.* 2015) using the dataset Insecta. *L. italica* genome was assembled with Flye v2.6 (Kolmogorov *et al.* 2019) and polished once with Medaka v1.7.2 (github.com/nanoporetech/medaka) and once with Hapo-G v1.3.2 (Aury & Istace 2021). Genome purging, identification of contaminant contigs and genome statistics including the assessment of genome completeness were done as in *L. noctiluca*.

### Re-sequencing

For each sampled individual, Illumina paired-end 150 bp-long reads were generated with a 20 x sequencing depth. FastQC (Andrews 2010) v0.11.0 and Trim Galore! v0.6.6 (Krueger 2012) were run to assess read quality and trim low quality bases. BWA (Li & Durbin 2009) was used to map the generated reads to the corresponding reference genome. The mapped files in BAM format were curated by removing PCR duplicates with Picard v2.20.8, and low quality reads were discarded using SAMtools v1.10 (Li *et al.* 2009).

## Variant calling and genomic diversity

Variant calling was performed equally for both species. Using GATK v 4.1.9 (Auwera *et al.* 2013) SNPs and indels were called via local re-assembly of haplotypes with HaplotypeCaller. Joint genotyping of all sequenced samples was done with GenotypeGVCFs. VCFs statistics were drawn with BCFtools stats (Danecek *et al.* 2021) and GATK VariantsToTable. Quality scores thresholds were applied for minimum, and maximum read depth [20,1568], fisher strand [FS = 10], strand bias [SOR = 3], and nucleotide quality by depth [DP = 2]. Nucleotide polymorphism diversity (depicted by π per site) and Tajima's *D* was estimated with VCFtools (Danecek *et al.* 2011). In both cases, we used sliding windows of 5 000 base pairs.

## Demographic inference with PSMC

The pairwise sequentially Markovian coalescent (PSMC) model (Li & Durbin 2011) allows for the estimation of past population sizes by using a combination of coalescent theory and recombination events. Briefly, PSMC infers the time to the most recent common ancestor (TMRCA) between heterozygote alleles within a single individual by using a hidden Markov model, where the states are the TMRCA and the transitions are the recombination events. Thus, short regions with high heterozygosity correspond to ancient coalescent events, and longer regions of low heterozygosity correspond to recent coalescent events. Finally, this coalescent rate is inversely proportional to the population size. Here, we run PSMC on every sampled individual (using the previously generated BAM files as input) of *L. noctiluca* and *L. italica* and inferred past population sizes by assuming a generation time of one year and a mutation rate of 2.8 x 109, which falls within the general mutation rate observed in beetles (Pélissié *et al.* 2022).

# RESULTS

## Sample collection

In the area north of Belmont-sur-Lausanne we collected 21 individuals of *L. noctiluca*, and across the northern shore of Lake Geneva (within the area of Lausanne) we collected 24 individuals of *L. italica* (figure 1). The genomes of all collected individuals were fully sequenced, and this information used for the analyses presented below.

## *Lampyris noctiluca*

### Population history

In a previous study, we inferred that *L. noctiluca* originated ca. 20 million years ago (mya) (Höhna *et al.* 2021). The PSMC approach (figure 2A) to infer past population sizes indicates that *L. noctiluca* had a population size of ca. 300k individuals around 3 mya. This population then increased its size up to nearly one million individuals around 560 kya. From there on there was a steady population decline from one million to 30k individuals at around 10 kya. Finally, PSMC shows some signal of a population expansion starting 10 kya until the present, although the confidence observed here is not very high: briefly, even though all sampled individuals show

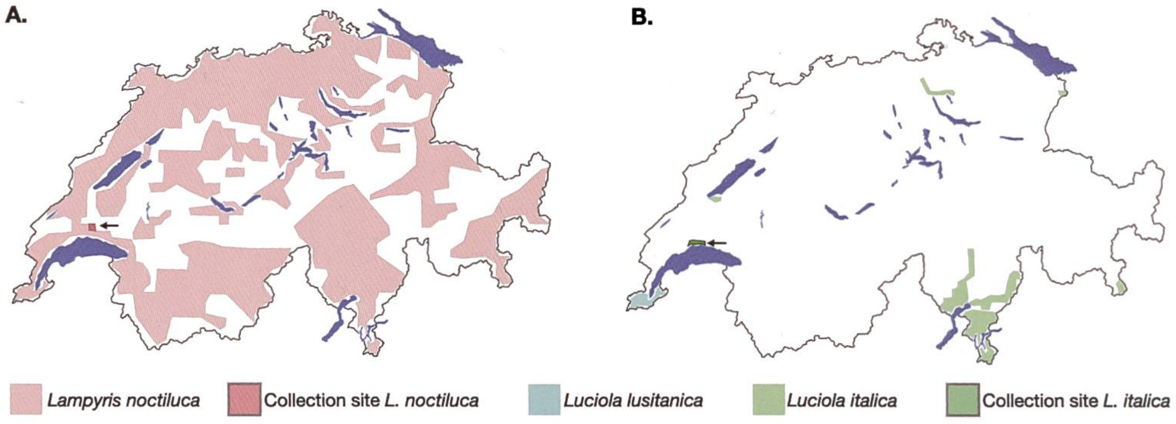

**Figure 1.** Distribution of *Lampyris noctiluca* and *Luciola italica* in Switzerland. Highlighted (and pointed with arrows) are the collection sites of the present study.

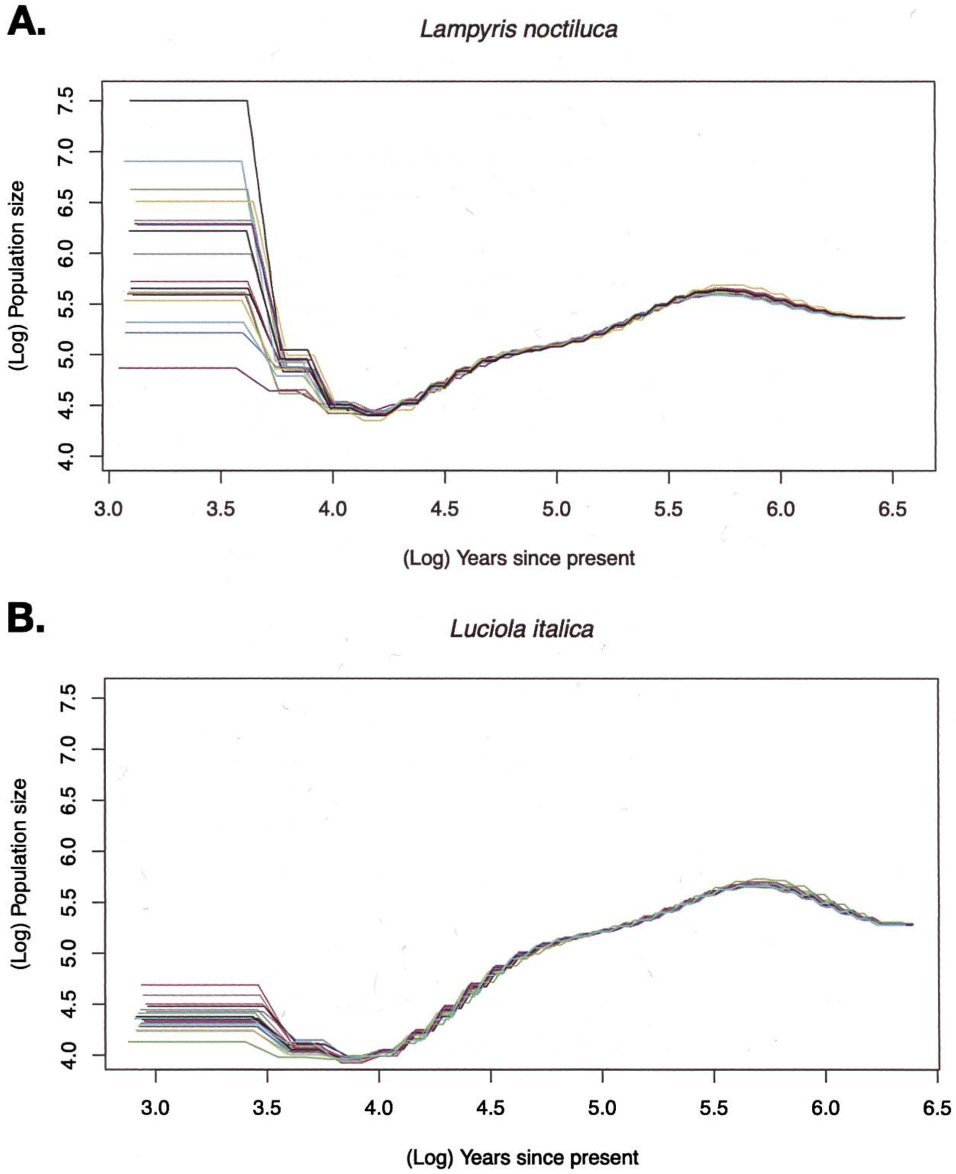

**Figure 2.** Population size changes in *Lampyris noctiluca* (A) and *Luciola italica* (B) inferred by PSMC. Each color represents a different individual among all sampled individuals.

comparable histories of population size changes until 10 kya, there are discrepancies between individuals regarding the inference of the last ten thousand years (figure 2A). This is the case because PSMC uses recombination events to estimate past population sizes, and thus, the more we go in the past the more recombination events have happened.

### Genomic diversity and neutrality

Genome-wide nucleotide diversity (depicted by $\pi$ per site) ranges from 0.00063 to 0.0068, with mean 0.0027 (figure 3A). Likewise, genome-wide Tajima's D values range from -1.34 to 1.06, with mean at 0.14 (figure 3B). Tajima's *D* statistic measures departures from neutrality, with positive values indicating balancing selection or sudden population contractions. In the case of *L. noctiluca*, we assume that the positive values are indicative of past population decline and recent balancing selection.

## ***Luciola italica***

### Population history

The PSMC approach (figure 2B) to infer past population sizes indicates that *L. italica* had a population size of ca. 160k individuals around 3 mya. This population then increased its size up to 560k individuals around 500 kya. From there on there was a steady population decline down to 10k individuals at around 6.3 kya. Finally, PSMC shows signal of a slight population expansion (up to 30k individuals) starting 6.3 kya until the present, although (as in the case of *L. noctiluca*) the confidence of this last part is reduced.

### Genomic diversity and neutrality

As above, genome-wide nucleotide diversity ranged from 0.00000212 to 0.0107, with mean 0.0022 (figure 3A). Likewise, genome-wide Tajima's *D* values ranged from -2.24 to 3.97, with mean at 0.67 (figure 3B). In the case of *L. italica*, we have shown that this population had experienced a sharp population decline, thus, we assume that the positive Tajima's *D* values can explain part of this.

## DISCUSSION

Both *L. noctiluca* and *L. italica* populations have been declining over the past several thousand years (figure 2). This population decline is also reflected in the average positive Tajima's *D* values computed genome-wide (figure 3B). The population decline seems to be stronger in *L. italica* than in *L. noctiluca*, which is also reflected by the average higher Tajima's *D* values in *L. italica* (figure 3B). This distribution is wider for *L. italica* than for *L. noctiluca*, which shows the presence of loci representing recent population growth, as well as loci with signs of population contractions and balancing selection. It is known that *L. italica* has been recently introduced in canton Vaud in the late 1930s (Faes 1941) and started migrating and establishing populations north of the alps (Neumeyer 1991). On the contrary, the distribution of Tajima's *D* in *L. noctiluca* is narrower and centered closer to zero, indicating a relatively more stable population.

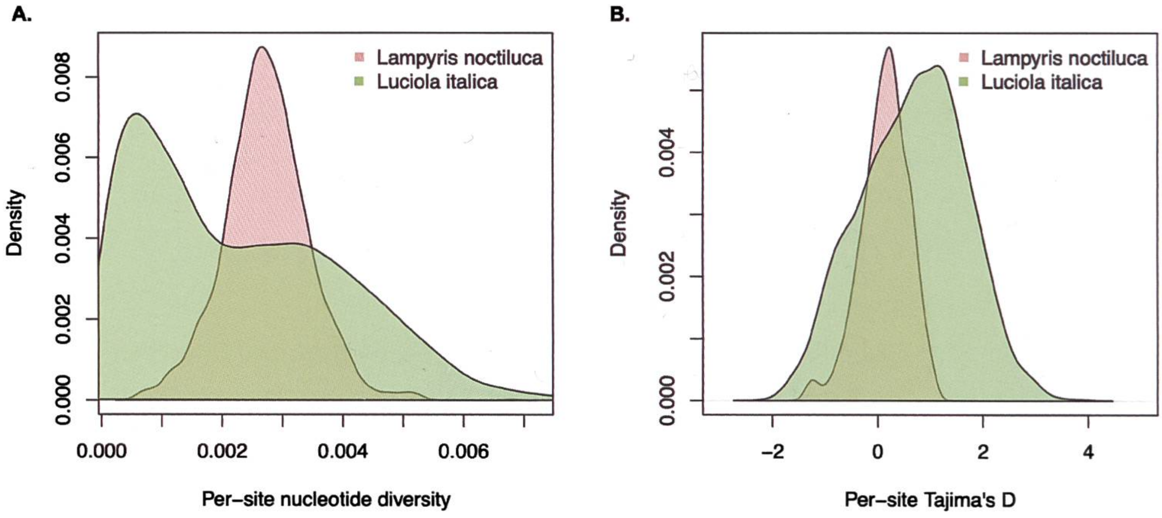

**Figure 3.** Distribution of genome-wide values of π (A) and Tajima's *D* (B) per site.

Genomic diversity levels of both *L. noctiluca* and *L. italica* are lower than the levels observed for other insects of the same order. For instance, average diversity in the beetles *Leptinotarsa decemlineata* and *Tribolium castaneum* can be three to four times larger (Diome *et al.* 2013, Crossley *et al.* 2019). Additionally, diversity levels in other insects, such as the domesticated silkmoth *Bombyx mori* range from 0.0006 to 0.02685 (mean 0.011) and in the wild *Bombyx mandarina* from 0.0052 to 0.0312 (mean 0.017) (Guo *et al.* 2011), meaning a six-fold difference in diversity compared to the fireflies of the present study. Here, maximum diversity levels do not go above 0.01. The case of *L. italica* is the most relevant in this case, because this species has a great proportion of diversity values lower than 0.001 (figure 3A), which is much lower than the diversity observed in *L. noctiluca*, and the other insect species. We attribute these low diversity values to the overall population decline that both species experienced in the past 500k years, plus the current habitat threats posed by human activities. Additionally, *L. italica* has been only introduced in the canton Vaud a few decades ago (Faes 1941), generating a bottleneck effect. Thus, it is a possibility that it takes time for this species to build up genomic diversity again.

In summary, we have calculated genome-wide diversity levels of *L. noctiluca* and *L. italica* from the canton Vaud and contrasted those to genome-wide Tajima's *D* values to gain insights on the demographic history of these two firefly species. Then, we have used single nucleotide polymorphisms as input for the PSMC algorithm to infer past population size changes in both species. With all these results we were able to characterize population trends and current diversity, which are determinant for conservation strategies. Finally, we would like to emphasize the importance and potential of genomic studies in advancing the knowledge of ecologically important species, with the present study being just an example of one such application.

## ACKNOWLEDGEMENTS

We would like to thank the Société Vaudoise des Sciences Naturelles (SVSN) and the Société Academique Vaudoise (SAV) for funding and supporting this project at all stages. Anne Freitag for support at the Musée de zoologie Lausanne. Yannick Chittaro, Stève Breitenmoser, and Vivien Cosandey for advice on sampling locations. Emily Williams for help during field work, and Gabriele Kumpfmüller at the LMU for her lab expertise.

## REFERENCES

Andrews S., 2010. FastQC: a qualit control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Aury J.M. & Istace B., 2021. Hapo-G, haplotype-Aware polishing of genome assemblies with accurate reads. *NAR Genomics and Bioinformatics* 3: 1–9.

Auwera G.A., Carneiro M.O., Hartl C., Poplin R., del Angel G., Levy-Moonshine A., Jordan T., Shakir K., Roazen D., Thibault J., Banks E., Garimella K. V., Altshuler D. & Gabriel S., DePristo M.A., 2013. From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Current Protocols in Bioinformatics* 43.

Crossley M.S., Rondon S.I. & Schoville S.D., 2019. Patterns of genetic differentiation in Colorado potato beetle correlate with contemporary, not historic, potato land cover. *Evolutionary Applications* 12: 804–814.

Danecek P., Auton A., Abecasis G., Albers C.A., Banks E., DePristo M.A., Handsaker R.E., Lunter G., Marth G.T., Sherry S.T., McVean G., Durbin R. & Group 1000 Genomes Project Analysis., 2011. The variant call format and VCFtools. *Bioinformatics* 27: 2156–2158.

Daneck P., Bonfield J.K., Liddle J., Marshall J., Ohan V., Pollard M.O., Whitwham A., Keane T., McCarthy S.A., Davies R.M. & Li H., 2021. Twelve years of SAMtools and BCFtools. *Gigascience* 10: p.giab008.

Diome T., Thiaw C., Ndong A. & Kane M., 2013. Genetic diversity of *Tribolium castaneum* (Herbst) population in storage infrastructures and agro ecological zones in Senegal. Int. *Journal of Biosciences* 3: 248–258.

Faes H. 1941. L'introduction de la luciole (*Luciola italica*) dans le canton de Vaud. *Bulletin de la Société Vaudoise des Sciences Naturelles* 61: 451–452.

Faudzi R., Abas A., Othman N.W. & Mazlan S.M., 2021. Effect of Water Quality on the Abundance of Firefly Populations at Cherating River, Pahang, Malaysia. *EnvironmentAsia* 14.

Fu X. & Benno Meyer-Rochow V., 2013. Larvae of the firefly *Pyrocoelia pectoralis* (Coleoptera: Lampyridae) as possible biological agents to control the land snail Bradybaena ravida. *Biological Control* 65: 176–183.

Guan D., Guan D., McCarthy S.A., Wood J., Howe K., Wang Y., Durbin R. & Durbin R., 2020. Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics* 36: 2896–2898.

Guo Y., Shen Y.H., Sun W., Kishino H., Xiang Z.H. & Zhang Z., 2011. Nucleotide diversity and selection signature in the domesticated silkworm, *Bombyx mori*, and wild silkworm, *Bombyx mandarina*. *Journal of Insect Science* 11: 155.

Gurcel K., Chittaro Y. & Sanchez A., 2020. Contribution à la connaissance des lucioles et lampyres de Suisse et observation de *Luciola lusitanica* Charpentier, 1825 à Genève. *Entomo Helvetica* 13: 81–96.

Gurevich A., Saveliev V., Vyahhi N. & Tesler G., 2013. QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* 29: 1072–1075.

Guzmán-Álvarez R.J. & De Cock R., 2011. The biology and distribution of glow-worms (Coleoptera: Lampyridae) in Spain. *Lampyrid* 1: 22–31.

Höhna S., Duchen P., Lower S.E. & Catalán A., 2021. A Time-calibrated Firefly (Coleoptera: Lampyridae) Phylogeny: Using Ultra Conserved Elements for Divergence Time Estimation. *bioRxiv.*

Kolmogorov M., Yuan J., Lin Y. & Pevzner P.A., 2019. Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnoly* 37: 540–546.

Krueger F., 2012. Trim Galore!: A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. Available at https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/

Laetsch D.R., Blaxter M.L. & Leggett R.M., 2017. BlobTools: Interrogation of genome assemblies [version 1; peer review: 2 approved with reservations]. *F1000Research* 6: 1287.

Lewis S.M. & Cratsley C.K., 2008. Flash Signal Evolution, Mate Choice, and Predation in Fireflies. *Annual Review of Entomoly* 53: 293–321.

Lewis S.M., Cratsley C.K. & Demary K., 2004. Mate recognition and choice in Photinus fireflies. *Annales Zoologici Fennici* 41: 809–821.

Li H. & Durbin R., 2011. Inference of human population history from individual whole-genome sequences. *Nature* 475: 493–496.

Neumeyer R., 1991. Ein neuer Fundort von *Luciola italica* (Col., Lampyridae) nördlich der Alpen. *Bulletin de la Société Entomologique Suisse* 64: 349–350.

Novak M., 2018. Redescription of immature stages of central European fireflies, Part 1: *Lampyris noctiluca* (Linnaeus, 1758) larva, pupa and notes on its biology (Coleoptera: Lampyridae: Lampyrinae). *Zootaxa* 4378: 516.

Pélissié B., Chen Y.H., Cohen Z.P., Crossley M.S., Hawthorne D.J., Izzo V. & Schoville S.D., 2022. Genome Resequencing Reveals Rapid, Repeated Evolution in the Colorado Potato Beetle. *Molecular Biology and Evolution* 39: msac016.

Schwalb H., 1961. Contribution to the biology of the domestic Lampyridae *Lampyris noctiluca* Geoffr. and *Phausis splendidula* Lec. and experimental analysis of their predatory and sexual behavior. *Zoologische Jahrbücher* 88: 399–550.

Simão F.A., Waterhouse R.M., Ioannidis P., Kriventseva E. V. & Zdobnov E.M., 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31: 3210–3212.

Zimin A. V., Marçais G., Puiu D., Roberts M., Salzberg S.L. & Yorke J.A., 2013. The MaSuRCA genome assembler. *Bioinformatics* 29: 2669–2677.