Institut Cavanilles de Biodiversitat i Biologia Evolutiva

Programa de doctorat en *Biodiversitat*

Rotifer adaptation to environmental unpredictability

Lluís Franch Gras

Maig 2017

Directors:

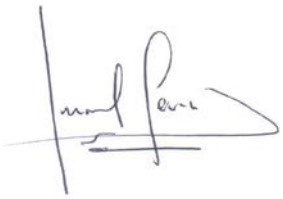
Manuel Serra Galindo

María José Carmona Navarro

Tesi presentada per LLUÍS FRANCH GRAS, per optar al grau de Doctor en Ciències Biològiques per la Universitat de València.

Signat: Lluís Franch Gras

Tesi dirigida pels Doctors en Ciències Biològiques per la Universitat de València, MANUEL SERRA GALINDO i MARÍA JOSÉ CARMONA NAVARRO.



Manuel Serra Galindo
Catedràtic d'Ecologia
Universitat de València

María José Carmona Navarro
Professora Titular d'Ecologia
Universitat de València

Aquest treball ha estat finançat per un Contracte Predoctoral del Programa *Atracció de Talents* (PREDOC13-110502) concedit a LLUÍS FRANCH GRAS per la Universitat de València, i per el projecte d'investigació (CGL2012-30779 i CGL2015-65422-P) del Ministeri d'Economia i Competitivitat del Govern d'España, cofinançat amb fons FEDER.

Agraïments

Les següents paraules són un agraïment a totes les persones o institucions que, d'alguna o altra manera, m'han donat suport per a fer possible aquesta tesi doctoral.

L'elaboració d'aquesta tesi no hauria sigut possible sense les indicacions i la supervisió dels meus directors: els Doctors Manuel Serra i María José Carmona. Els agraeix molt la confiança que van depositar en mi per a desenvolupar aquest projecte i l'estímul que això va suposar per a introduir-me en el món de la investigació. El treball de conjunt que hem realitzat durant els més de quatre anys que ha requerit l'elaboració de la tesi ha sigut molt intens, per la qual cosa he d'agrair-los també el temps que han dedicat a reflexionar amb mi sobre els resultats que anàvem obtenint i l'esforç que han realitzat per a transmetre'm les seues orientacions.

Gràcies a la Universitat de València, per finançar aquest projecte d'investigació mitjançant una ajuda 'Atracció de Talents' (PREDOC13-110502) i tres estades a l'estranger (EPREDOC14-186330; EPREDOC15-275447 i EPREDOC16-384721). Al *Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica (I+D+I)* del *Ministerio de Economia y Competitividad* per finançar aquesta recerca amb dos projectes (CGL2012-30779 i CGL2015-65422-P) cofinançats amb fons FEDER de la Unió Europea. *Thanks to the ESA (European Space Agency) and the NASA (National Aeronutics and Space Administration) for providing satellite data.* Gràcies també al '*Instituto de información Geográfica'* per facilitar-me l'accés a les seues imatges aèries. Moltes gracies als tècnics forestals de la *'Junta de Comunidades de Castilla-La Mancha'* pel seu permís per a prendre mostres de les llacunes, i també als propietaris privats que ens van permetre mostrejar. També a Font Salem per proporcionar les botelles de plàstic que necessitàvem per a créixer els

nombrosos cultius de rotífers i a Pilar Ruiz per, entre altres coses, facilitar-nos el contacte amb aquesta empresa.

Al laboratori he tingut la sort de coincidir amb Edu, qui m'ha contagiat el seu interés i passió per la investigació des del principi. Junts hem treballat moltes hores i m'ha ajudat a entendre i aplicar teories, anàlisis estadístics i moltes coses més. També vull agrair a Raquel, a Pau i a Carmen tots els moments memorables que hem passat junts. Ells han sigut els meus primers companys al laboratori i tindran sempre el meu agraïment per ensenyar-me les rutines bàsiques del laboratori i a defendre'm entre provetes, amb l'autoclau, la centrifugadora, etc. També a Ana Tortajada, amb la qual només vaig coincidir al final de la seua tesi, però en aquest breu període em va demostrar ser una gran persona amb una capacitat de superació admirable. Més endavant es va incorporar Ana Hidalgo al laboratori, amb la que he coincidit gran part de la meua tesi doctoral. Ha sigut una meravella tindre't al laboratori, la teua ajuda ha sigut valuosíssima durant tot aquest període. Després va arribar, al laboratori, Eva, de la qual sempre m'han fascinat moltes coses, però que en aquest àmbit em va mostrar una atenció als detalls infinita i una capacitat d'anàlisi i de treball impressionants. Junts al laboratori hem format un gran equip! Ja més recentment han arribat Tono, Irina, Nohemi, Nora, etc. una nova tanda d'Ecòlegs als quals els desitge la millor de les sorts. I per descomptat, no se me'n pot oblidar l'ajut del personal de l'Institut Cavanilles que he tingut per alimentar els meus clons de rotífers: Yolanda, Julia, Mª Angeles, Laura, Jorge, Rafa, etc. i també a Alejandra per ajudar-me en les gestions que he hagut de fer al llarg de la tesi.

També vull agrair als diferents grups de l'Institut Cavanilles de Biodiversitat i Biologia Evolutiva per ajudar-me en diverses qüestions: el grup d'Ecologia Integrativa, el grup de Genètica Evolutiva (especialment al grup dels pugons),

també a Ximo Baixeras per la seua ajuda per fer imatges al microscopi dels rotífers.

Thanks Dr. Eric Vermote for welcoming me at the University of Maryland and for allowing me to work in your fantastic group. També a la Doctora Belen Franch, pel seu ajut dins i fora de la ciència, que encara avui continua i esperem que per molts anys. Qui ens havia de dir que ens retrobaríem tan lluny de la nostra terra? També agrair al "*Department of Geographical Sciences*" i a tota la gent en la qual em vaig trobar allí i que em va fer gaudir de la meua estada.

En dues ocasions vaig visitar Hull i vaig tindre el plaer d'investigar en la Dra. Africa Gómez. Ella em va aportar una forma diferent de veure la recerca, estimulant la meua inquietud per les preguntes que ens fèiem en tot moment. *Thanks Dr. Christoph Hahn for guiding me through many software, programming languages... and of course to Larkins'.* De fet, guarde en la memòria grans moments de la meua estada a Hull i en gran mesura va ser per la fantàstica gent que vaig trobar allí: Dave, Lori, Domino, Berndt, Amir, Graham, James, Marco, David, Joe, Rob, Hayley, Harriet, Peter, Rosie, Lynsey, Callum, etc.

Gràcies a la meua família. Especialment als meus pares, que sempre han estat ahí, incondicionalment. Tots dos m'heu ensenyat a entendre des de les coses més senzilles a les més complicades en tots els àmbits de la vida. I també a no rendir-me i perseverar quan no s'entén res, perquè amb massa freqüència el món no hi ha qui l'entenga...

Eva, no és la primera línia que et dedique en estes pàgines, i això diu molt de tu. El gran equip que formem transcendeix al laboratori. Gràcies pel teu suport, per la teua paciència i per tot el que m'has ajudat durant tot aquest temps. Ets a tots llocs i res podria ser més agradable...

Rotifer adaptation to environmental unpredictability

"Change is the only constant in life"

Heraclitus

Index

Summary in English

Introduction

The premise that the environment is constant was assumed in the seminal mathematical models developed in early population ecology, and adopted in the empirical studies framed in this theoretical work. Nevertheless, this conception progressively changed during the second half of the 20th century and today understanding how environmental parameters vary, and how organisms respond to the degree of unpredictability associated with such variation is a central theme in both fundamental and applied evolutionary ecology. Environmental unpredictability is a phenomenon that occurs in any habitat, and is an important selection pressure for the evolution of adaptive responses in the organisms.

Quantifying the degree of predictability of habitat features is not simple. It needs a sophisticated, robust methodology, as predictability estimations cannot be based on mean values, but on variances. Moreover, this characterization needs assumptions —as explicit as possible– regarding the relevant time-scales that can be associated with predictability, these assumptions not being straightforward. Nevertheless, one of the most important constraints on fluctuation analysis is the need of long time-series data. Obtaining long time series is

regarded as costly, and these series are consequently scarce. Several studies show that remote sensing can offer solutions to a wide variety of problems in nearly all fields of environmental research. Remote sensing technology, has been highly developed in the last century and its use in several research areas has not reached yet its full potential. The information gathered by several satellites in recent years can provide the long time-series data needed to quantify unpredictability. This is the case of the Landsat Earth-observing satellites, which provide remote sensing data from 1972 to the present and have moderate spatial and intermediate temporal resolutions. Since 2008, the satellite scenes from Landsat satellite series are freely available from the United States Geological Survey (USGS), allowing for the scientific community to gather a great deal of information free of cost. The data obtained from satellites are the reflectance from various wavelength regions of the electromagnetic spectrum (Chapter 2). These reflectances can be processed and analysed to obtain information about relevant environmental features. All of this makes the satellite remote sensing data especially helpful to expand temporally and spatially our knowledge of the focus environments.

Understanding how organisms adaptively respond to environmental fluctuations —which are ubiquitous– and their degrees of predictability is a fundamental question in

evolutionary biology that has motivated a growing body of research. There are several ways by which organisms evolutionarily respond to unpredictable variation. The best known of these ways is through natural selection that acts recurrently on the heritable variation among individuals; this process is named adaptive tracking, and implies genetic evolution. Interestingly, as adaptation consumes time, it is not expectable the population to be adapted as far as unpredictable fluctuations continue. Adaptive phenotypic plasticity, another mode of evolutionary response, occurs when individuals modify their phenotypes according to environmental conditions without changing their genetics, and do it in a way that increases their performance. Bet hedging constitutes a third way of adaptive response and occurs when a genotype increases the geometric mean of fitness at the cost of a decrease in the arithmetic mean, by reducing fitness variance. Selection increasing geometric fitness is expected because long-term fitness is related to its geometric mean, rather than to its arithmetic mean. There are two main modes of bet hedging: diversified and conservative. Diversified bet hedging is a transgenerational effect that occurs when a single genotype produces different phenotypes in its offspring in advance of future unpredictable conditions. In contrast, under conservative bet hedging, an individual sacrifices expected (arithmetic) fitness to reduce temporal variance in reproductive success by employing a

single low-risk phenotype across all possible environmental future scenarios. These three kinds of responses are non-exclusive, and combinations of them may occur simultaneously in natural populations. It is worth noting that, although bet hedging is well developed theoretically in biology, empirical evidence is still scarce.

The ponds located in the Mediterranean region are characterized by having a high degree of seasonality and uncertainty at various temporal scales. This makes them an ideal study framework to investigate adaptation to environmental uncertainty. The Iberian Peninsula due to its climatic, orographical, geological and hydrological conditions, holds a rich and diverse collection of inland and coastal, saline shallow water bodies. The ponds studied in this thesis are located in an area of approximately 800 km^2. It is an endorheic territory; that is, with no defined hydrographic network, so that water tends to accumulate in the common terrain depressions. As a result of a long-term erosion of the surrounding terrains by water, salinity increased in the ponds and can now reach values as high as 40 g L^{-1}. The climate in the study region is semiarid, with a mean average annual rainfall of ca. 343 mm and a mean temperature of approx. 14 ºC. Dry periods occur between June and September, when temperatures may exceed 40 ºC and rainfall is scarce (typically< 8 mm in July). Most precipitation occurs as heavy

rains in Spring (April-May) and Autumn (October-December). The biological communities confined in this kind of non-permanent ponds are expected to be adapted and strongly reliant on patterns of pond inundation. The case study includes 20 Mediterranean saline ponds and lakes presented in Chapter 2. The focus variable is the water-surface area (*A*), which is an ecologically relevant factor in lentic water bodies that positively correlates with species diversity according to the predictions of island-biogeography theory. An increase of the pond's *A* and its correlated increase in average pond depth will result in an increase of the habitat size for aquatic organisms, thus affecting the 'colonization *vs.* extinction' balance and the habitat heterogeneity with positive effects on species persistence, complexity of the community and species diversity. Fluctuations in *A* have multiple effects on the organisms above and below the waterline, for aquatic flora and fauna —e.g. macrophytes, zooplankton and fish–. These fluctuations can strongly affect habitat conditions through variations in the physicochemical parameters —e.g. temperature, light, nutrient or solute concentration–. Solute concentration is particularly relevant in saline ponds and lakes, which are a significant, geographically widespread part of the world's inland aquatic ecosystems. Migratory flyways and nesting places of many waterfowls depend on the dynamics of inundation, particularly if birds are dependent on the ponds' food webs. Migratory patterns of

waterfowls, lifecycles of many short-lived animals and rainy patterns occur at an annual time scale or shorter. Accordingly, ecological unpredictability in these systems can be conceived as departures from an average seasonal variation —i.e. the main source of environmental unpredictability is the inter-annual variation of the within-year variation–. Episodic, *in situ* observations during the last decade showed that some ponds can frequently dry out, then developing a thick salt crust, and fill up again. Others, instead, are known to be flooded for years through several dry seasons. This suggests that the ponds present different hydrological patterns, although an accurate quantitative characterization of their hydrological features was not available before this thesis.

The monogonont rotifer *Brachionus plicatilis* is one of the zooplankton species that frequently inhabits the salt ponds of the Mediterranean region. This species has a type of reproduction called cyclic parthenogenesis, in which proliferation by ameiotic parthenogenesis (asexual phase) combines with occasional bouts of male production and sexual reproduction; the latter resulting in diapausing egg production (sexual phase). Populations of *B. plicatilis* in temperate regions are temporal —i.e. they are not active all year-round, and they colonize the water column during the so-called (planktonic) growing season–. Typically, each

growing season, the active population is initiated by the hatching of diapausing eggs from the pond sediment. Hatchlings are asexual females that produce subitaneous (i.e. non-dormant) eggs, which hatch into genetically identical daughters via parthenogenesis, thus producing clones. Sexual reproduction is induced by a chemical signal produced by the rotifers that accumulates in the medium as population density increases. Once a population density threshold is reached, asexual females are triggered to produce sexual daughters as some fraction of their offspring. Thus, sexual and asexual reproduction, typically overlap. Sexual females produce meiotic haploid eggs that develop into haploid males if they remain unfertilized and into diploid diapausing eggs if they are fertilized. Males are dwarf (100 μm length compared to 300 μm of females), do not feed and have shorter lifespan than females. Unlike asexual eggs, the sexually produced eggs are dormant embryos; therefore, sex is associated with diapause. Diapausing eggs settle in the sediment and remain dormant for a period of variable duration. Under suitable conditions, diapausing egg hatching is induced, and a new growing season starts, although a —relatively short– refractory period for hatching induction has been described in some strains. Additionally, not all diapausing eggs hatch in the season following their production. The unhatched eggs often show prolonged diapause and accumulate in the sediment forming

diapausing egg banks where they can remain viable for decades or even centuries. This life cycle is considered an adaptation to temporary habitats.

Unpredictability in natural populations can act on several organism traits, especially in those species with complex life histories. Life history theory has been developed over the last 50 years and studies the patterns of organismal growth, maturation, reproduction and survival. The study of these life-history traits is relevant, since they can interact to reduce the impact of environmental variability. A fascinating way to overcome unsuitable environmental conditions is the diapause. The diapause is a period of suspended development which is wide-spread in the animal kingdom. Diapause starts before the advent of adverse conditions, and its duration —which is mostly regulated by internal physiological factors– can exceed the duration of adverse conditions, what gives rise to many ecological and evolutionary questions. Leaving or exiting diapause may have important consequences on the dynamics of the rotifer populations dwelling in temporary habitats. First, a trade-off exists between the assignation of resources to current proliferation or to resistant stages that could resume proliferation in the future. The production of diapausing stages should be initiated beforehand the beginning of the unsuitable conditions, given that it is necessary for the survival

of the genotype. Nevertheless, if the favourable conditions prevail, an early investment in diapause can be sub-optimal, since it results in a reduction of the current population growth rate. Second, a similar compromise happens with the exit of diapause. Exiting diapause may imply to exploit a suitable environment, but may be inconvenient if the habitat becomes unsuitable too early and the genotype does not produce a new cohort of diapausing stages. All of this produces contrasting selective pressures that can lead to the appearance of bet-hedging strategies in fluctuating environments. In facultative sexual rotifers this situation becomes more complex ––and potentially more interesting– since the production of diapause stages (in form of diapausing eggs) is associated to sexual reproduction. Then, diapausing egg production, as well as having associated the costs of diapause, has additional costs derived from sexual reproduction.

Mapping genotype on phenotype and phenotype on environment is of fundamental importance for understanding how natural populations adapt to their local environments. However, the genomic signatures of the adaptive evolutionary responses to local unpredictability, especially those related to bet-hedging strategies, are little known. The development of next generation sequencing (NGS) has provided new methods for massive DNA sequencing (see

Chapter 2) that can help to get insights into these questions. NGS makes sequencing cheaper, faster and, more importantly, allows genomic studies to be performed in practically any organism. Until recently, most of our knowledge in genomics —even in deep genetics– was based on model organisms. This is not completely satisfactory in biology, especially for comparative, evolutionary and biodiversity approaches, where information on a few model organisms is not representative. This is changing with the development NGS tecnologies. For instance, new, extensive information on the genetic correlates of adaptation to different environments can be directly disentangled at the genome level on species with limited genomic resources (i.e. non-traditional genetic model organisms). Again, all of this allows the scientific community to approach new problems what helps to close the —everyday smaller– breach between ecology and evolution.

Objectives

This thesis takes advantage of the mentioned technical advances in methods and adopts the theoretical framework of evolutionary ecology in order to address local adaptation to environmental uncertainty levels, using rotifers as model organisms. This work also takes advantage of the accumulated knowledge on the monogonont rotifer *B. plicatilis*, in particular on their population biology in eastern

Spain. As explained in more detail in Chapter 2, this rotifer has been considered a model organism in evolutionary and ecological studies. The main goals of this thesis are: (1) To study variation in diapause-related life-history traits among populations of the rotifer *B. plicatilis* inhabiting ponds with different degree of environmental unpredictability. (2) To unveil, at a genome wide scale, genotypes correlated to adaptation to local environmental unpredictability by using genomic technologies. For the consecution of objectives (1) and (2), an additional objective with intrinsic methodological relevance was needed: (3) To quantify the degree of environmental unpredictability based on satellite imagery data from Mediterranean ponds.

Outline of this thesis: methodology and main results

The thesis, in the parts following the general introduction (Chapter 1), is organized as follows.

In Chapter 2 describes the methodological context. It gives an overview of methodological topics that the reader might not be familiar with. It describes the species and the habitats studied in this thesis, and it provides an overview of the remote sensing and molecular techniques used. First, the main biological features of the study organisms, the rotifer *B. plicatilis* are described. Second, the main ecogeographic characteristics of the study area are described. Third, a global

description of the study system is provided (i.e. *B. plicatilis* populations in eastern Spain). Fourth, information about the main features of the remote sensing data used and a description of how they are obtained is supplied. Finally, the main molecular techniques used throughout the thesis are introduced (genome sequencing and genotyping by sequencing, GBS). In this last section a description of how these techniques are performed and a glossary of terms are provided.

The research described in Chapter 3 takes advantage of the free access to long-term recording of remote sensing data (27 years, Landsat TM/ETM+) in order to assess a set of environmental models for estimation of environmental predictability in saline ponds. These ponds in the Mediterranean region can develop salt-crusted areas that make it difficult to distinguish between soil and water. This challenge was addressed by a novel pipeline that combines water indices and the short infrared band, as salt filter. Once this long time series of data was obtained, the predictable and unpredictable components of variation in *A* were extracted using two different approaches. The first approach, based on Colwell's predictability metrics, transforms the focus variable into a nominal one. This discretization is a key step, and here different models were considered based on how *A*'s variation could be relevant from the point of view of the focus

organism. As a second approach, general additive model fitting were developed as a new metric for quantifying predictability, this metric parallels Colwell's but it is based on a regression model. The similarity and divergence from the different predictability indices were analysed in order to determine their sensitivity to the assumptions needed to quantify predictability. Both approaches were found to extract meaningful information about the degree of predictability of the studied ponds. Interestingly, some model assumptions were found to have negligible effects, while some others can be associated to ecological features of species whose predictability needs to be assessed. The methodology described here is applicable to a wide variety of systems and will be valuable for quantifying and characterizing predictability, which is essential within the expected global increase in the unpredictability of environmental fluctuations.

In Chapter 4, rotifers were used as model organisms to investigate adaptation to environmental fluctuations. Two diapause-related traits in rotifers are studied using clones from nine *B. plicatilis* natural populations that vary in the degree of environmental unpredictability (Chapter 3). There are two key life-history traits in the life cycle of monogonot rotifers (see Chapter 2) that are associated with the entrance and exit from diapause. One is the propensity for sex, which is inversely

related to the density threshold for sex initiation and is a proxy of the timing of diapausing egg production. Other is the diapausing egg hatching fraction, which is inversely related to the diapause duration. Both traits have been proposed as instances of bet-hedging strategies that might interact to reduce the risks associated with environmental unpredictability. The propensity for sex has been proposed as a case of conservative bet hedging in rotifer populations that inhabit unpredictable habitats. When there is uncertainty regarding the onset of unsuitable periods, a low-risk strategy might be to produce diapausing eggs as soon as possible to avoid an unexpectedly short growing season. However, if the growing season is not short, early investment in sex and diapause will reduce the rate of clonal proliferation, thereby resulting in lower fitness. These considerations lead to the prediction that the propensity for sex will increase with increasing unpredictability. However, in testing this prediction, a confounding factor should be considered: if the growing seasons are predictably short, a high propensity for sex is also expected to evolve. In contrast, diapausing egg hatching fraction has been proposed in several theoretical studies as a form of diversified bet hedging. As far as rotifers cannot predict whether a particular growing season will be sufficiently long to complete the life cycle and ensure diapausing egg production, intermediate hatching rates are expected in habitats with both long growing seasons (where, ideally, all of

the eggs should hatch; i.e. good seasons) and unexpectedly short growing seasons (where, ideally, no egg should hatch; i.e. bad seasons). According to bet-hedging theory, the optimal hatching fraction should equal the frequency of good seasons. For example, in this 'good *vs*. bad growing season' scenario, a hatching fraction around 0.5 (i.e. intermediate) would be expected in a completely unpredictable habitat, since the frequency of good seasons would be somewhere around 0.5. Thus, the hypotheses to be tested are that the level of environmental unpredictability is directly related to the propensity for sex and inversely related to the hatching fraction of diapausing eggs. Genetic variation in the propensity for sex was studied by conducting 810 bioassays (9 populations × 30 clones × 3 replicates). Estimation of the diapausing egg hatching fraction was performed by using a subset ten clones that were used in the propensity-for-sex experiment. In total, 8,640 bioassays (96 diapausing eggs × 10 clones × 9 populations) were performed. Here, significant levels of genetic variation within populations for both traits are found. Moreover, as predicted, a positive correlation between pond unpredictability —quantified in Chapter 3– and the propensity for sex was found. This correlation suggests a conservative bet-hedging strategy that provides protection against growing seasons unexpectedly short. In contrast, the hatching fraction of diapausing eggs was found not to be related to the level of environmental predictability. This results

highlight the ability of rotifer populations to locally adapt to time-varying environments.

In Chapter 5 an integrative approach is used, combining environmental (Chapter 3), phenotypic (Chapter 4) and genomic data. Genotyping by sequencing (GBS) data were obtained for the set of 270 clones of the rotifer *B. plicatilis* from the nine populations in Eastern Spain used in Chapter 3. Moreover, 11.7 Gb of Illumina data is generated to assemble and annotate a draft genome for *B. plicatilis*. As a result of GBS data analysis, 4543 high quality SNPs are discovered and genotyped in a reduced in size —but dense in gene regions– genome. By using these SNPs, relevant information about the genetic structure of the natural populations can be obtained, and also, genome locations with signatures of selection can be identified. There are several methods that aim to detect loci under selection. A first group of methods, based on F_{ST} values, use allele frequencies to identify those SNPs that present significatively higher —i.e. divergent selection– or lower —i.e. balancing/purifying selection– differentiation among populations than expected under a model of neutral evolution. These methods, however, have been described to have low power, especially when a high differentiation in neutral markers exists, as reported in *B. plicatilis* populations. Recently, another group of outlier detection methods has been developed. It is based on correlating allele frequencies

to a set of environmental or phenotypic factors (i.e. genome-wide association methods, GWAS). These correlation-based models have been reported to have higher power than the differentiation-based models, but also to show higher false positive rates, which is the main drawback of the outlier analysis approaches currently available. It is worth noting that all of these methods identify markers that are not necessarily under direct selection, but instead some of them can be neutral markers 'hitchhiking' with the true locus (or loci) due to genetic linkage. Nevertheless, these techniques can be used to shed light into putative genes under selection or involved in adaptive population divergence. After applying these analyses, more than 90 SNPs showed F_{ST} values indicating that they are putatively under selection, with signatures of diverging or balancing/purifying selection. Moreover, more than 160 SNPs were found to be correlated to a set of environmental or phenotypic factors revealing signatures of local adaptation to environmental predictability. Remarkably, most of these SNPs were found to be located within genes annotated in the genome, and putative functions were associated to them.

Chapter 6 discusses in general terms the main results obtained in this thesis, proposes prospective research and sets out the most important conclusions.

Conclusions

The main conclusion of this thesis are: **(1)** The scenes from the satellites Landsat 5 and 7 —after applying a combination of procedures to discriminate water from background– provided a long time series (27 years) of the variation in water-surface area of a group of twenty Mediterranean saline water bodies. **(2)** Following the conception that predictability depends on the point of view of the focus organism and using Colwell metrics, different models for predictability estimation were developed here. Furthermore, GAM fitting was developed in this thesis as an alternative continuous approach for measuring predictability. All these models were assessed by considering how the variation in water-surface area could be relevant for the focus organism. **(3)** The application of the predictability metrics allowed quantifying predictability in a group of Mediterranean ponds. These ponds showed a wide range of predictability. **(4)** *A posteriori* classification of the models for predictability estimation showed that some assumptions had negligible effects, while others can be associated with the species assemblages for which predictability needs to be assessed. **(5)** *Brachionus plicatilis* populations inhabiting a set of nine Mediterranean saline ponds showed significant levels of within-population genetic variation for propensity for sex and for hatching fraction of diapausing eggs. **(6)** The propensity for sex in rotifer populations, and hence the early investment in diapause,

decreased with environmental predictability, while the relationship of that trait with hydroperiod length was relatively weak. This suggests a conservative, bet-hedging strategy that provides protection against unexpectedly short growing seasons. **(7)** Diapausing egg hatching fractions had intermediate values (from 44 to 88%) in all the studied populations, but hatching fractions were neither related to the level of environmental predictability nor to hydroperiod length. **(8)** Rotifer populations are able to locally diverge in diapause-related traits within a small geographical range (240 km^2) despite their potential for widespread genetic exchange through the passive dispersal of diapausing eggs. **(9)** The *B. plicatilis* genome was assembled in this thesis. Its structural annotation yielded 54,725 predicted genes. Functions were tentatively assigned to 30% of them. **(10)** Genotyping by sequencing (GBS) and the subsequent bioinformatics analyses provided a large number (4,543) of high quality single nucleotide polymorphisms (SNPs). **(11)** A number of SNPs —most of them located within genes– showed higher between-population differentiation than expected by chance and were correlated with life-history traits and environmental factors, so that they are candidates for diversifying selection for local adaptation. **(12)** Unexpectedly, a large set of SNPs, more than half of them located within genes were found to present signatures of balancing/purifying selection in *B. plicatilis*. This finding

requires further research. **(13)** A number of genes were identified as strong candidates to be part of the genomic basis of local adaptation to fluctuating environments. These genes constitute a database for future studies

Resum en Valencià

Introducció

La premissa que proposa que l'ambient és constant va ser assumida en els models matemàtics seminals desenvolupats en l'ecologia de poblacions clàssica, i també, en conseqüència, va ser adoptada en els estudis empírics emmarcats en aquests treballs teòrics. No obstant això, aquesta concepció va canviar progressivament durant la segona meitat del segle XX i, avui en dia, comprendre com varien els paràmetres ambientals, i com els organismes responen al grau d'impredictibilitat associat a aquesta variació, és un tema central en l'ecologia evolutiva, tant fonamental com aplicada. La incertesa ambiental és un fenomen que es produeix en qualsevol hàbitat, i és una pressió de selecció important per a l'evolució de respostes adaptatives en els organismes.

Quantificar el grau de predictibilitat dels paràmetres de l'hàbitat no és simple. Es necessita una metodologia sofisticada i robusta, ja que les estimacions de predictibilitat no poden basar-se en valors mitjans, sinó en les variàncies. D'altra banda, aquesta caracterització requereix supòsits -tan explícits com siguen possibles- pel que fa a les escales

temporals rellevants associades amb la predictibilitat, i aquests supòsits no solen ser senzills. No obstant això, una de les limitacions més importants en l'anàlisi de les fluctuacions ambientals és la necessitat de dades de sèries temporals llargues. Obtenir sèries temporals llargues és costós, i en conseqüència, aquest tipus de sèries són escasses a la literatura científica. Diversos estudis mostren que la teledetecció pot oferir solucions a una àmplia varietat de problemes en gairebé tots els camps de la investigació. La teledetecció s'ha desenvolupat molt durant el segle passat i el seu ús en diverses àrees d'investigació no ha assolit encara el seu ple potencial. La informació recollida per diversos satèl·lits en els últims anys pot proporcionar les sèries temporals llargues necessàries per a quantificar el grau de predictibilitat ambiental. Aquest és el cas dels satèl·lits Landsat d'observació terrestre, que proporcionen dades de teledetecció des de 1972 fins al present i tenen resolucions temporals intermèdies i resolucions espacials moderades. Des de l'any 2008, les imatges de la sèrie de satèl·lits Landsat són de lliure accés des de la pàgina web de la *United States Geological Survey (USGS)*, el que permet a la comunitat científica reunir una gran quantitat d'informació de forma gratuïta. Els satèl·lits obtenen dades de reflectància de diverses regions de l'espectre electromagnètic (Capítol 2). Aquestes reflectàncies poden ser processades i analitzades per obtenir informació sobre les característiques ambientals

pertinents. Tot això fa que les dades obtingudes a través de satèl·lits siguen especialment útils per a expandir temporalment i espacialment el nostre coneixement sobre els paràmetres ambientals d'interés.

Comprendre com els organismes responen de forma adaptativa a les fluctuacions ambientals -que són ubiqües- i al seu grau de predictibilitat és una qüestió fonamental en biologia evolutiva, la qual cosa ha motivat un creixent volum d'investigació en els darrers anys. Hi ha diverses formes en què els organismes responen evolutivament a la variació impredictible. La més coneguda és a través de la selecció natural actuant de forma recurrent en la variació hereditària entre els individus; aquest procés s'ha denominat *adaptive tracking*, i implica l'evolució genètica. Cal remarcar que l'adaptació sol consumir temps i, per tant, no s'esperaria que les poblacions s'adaptaren si les fluctuacions impredictibles són recurrents. La plasticitat fenotípica és una altra forma de resposta evolutiva. Es produeix quan els individus modifiquen els seus fenotips d'acord a les condicions ambientals sense canviar la seua dotació genètica, i ho fan de forma que augmenta la seua eficàcia biològica. El *bet hedging* constitueix una tercera forma de resposta adaptativa, i es produeix quan un genotip augmenta la mitjana geomètrica de la seua eficàcia biològica a costa d'una disminució en la mitjana aritmètica d'aquesta, mitjançant una reducció de la

variància de l'eficàcia biològica. Hi ha dues formes principals de *bet hedging*: diversificador i conservatiu. El *bet hedging* diversificador és un efecte intergeneracional que es produeix quan un sol genotip produeix diferents fenotips en la seua descendència abans que ocórreguen les futures condicions impredictibles. Per contra, sota *bet hedging* conservatiu, l'individu presenta un únic fenotip de baix risc en tots els possibles escenaris futurs de l'ambient, la qual cosa suposa sacrificar l'eficàcia biològica (aritmètica) esperada a canvi de reduir la variació temporal en l'èxit reproductiu. Aquests tres tipus de resposta no són mútuament excloents, i combinacions d'elles poden ocórrer simultàniament en les poblacions naturals. Caldria assenyalar que, tot i que el *bet hedging* ha sigut àmpliament desenvolupat teòricament en biologia, l'evidència empírica hi és encara escassa.

Les llacunes situades a la regió mediterrània es caracteritzen per tindre un alt grau d'estacionalitat i d'impredictibilitat en vàries escales temporals. Això fa que aquestes siguen un sistema d'estudi ideal per investigar l'adaptació a la impredictibilitat ambiental. La península Ibèrica, a causa de les seues condicions climàtiques, orogràfiques, geològiques i hidrològiques, posseeix una rica i diversa col·lecció de cossos d'aigua somers i salins. Les llacunes estudiades en aquesta tesi es troben en una àrea d'aproximadament 800 km^2. És una zona endorreica; és a dir, que no presenta una xarxa

hidrogràfica definida, de forma que l'aigua tendeix a acumular-se en les freqüents depressions del terreny que hi han a la zona. Com a resultat de l'erosió dels terrenys circumdants per l'acció de l'aigua, la salinitat augmenta en les llacunes, podent arribar a valors tan alts com 40 g L^{-1}. El clima a la regió d'estudi és semiàrid, amb una precipitació mitjana anual de ca. 343 mm i una temperatura mitjana d'aproximadament 14 ºC. Els períodes secs de l'any es produeixen entre juny i setembre, quan les temperatures poden superar els 40 ºC i les precipitacions són escasses (típicament < 8 mm al juliol). La major part de les precipitacions es produeixen en fortes pluges a la primavera (abril-maig) i a la tardor (octubre-desembre). S'espera que les comunitats biològiques confinades en aquests tipus de llacunes no permanents estiguen adaptades als patrons d'inundació de les llacunes, sent fortament dependents dels mateixos. El cas d'estudi inclou 20 llacunes salines mediterrànies que es presenten al Capítol 2. La variable d'interés és l'àrea que ocupa la superfície d'aigua (*A*), que és un factor ecològic rellevant en els cossos d'aigua lèntics. La *A* està positivament correlacionada amb la diversitat d'espècies d'acord amb les prediccions de la teoria de biogeografia d'illes. Un augment de *A* a la llacuna i el seu augment correlacionat en la profunditat mitjana es traduirà en un increment de la grandària de l'hàbitat per als organismes aquàtics, la qual cosa afecta l'equilibri

'colonització *vs.* extinció' i també a l'heterogeneïtat de l'hàbitat. Tot això té efectes positius sobre la persistència d'espècies, la complexitat de la comunitat i la diversitat d'espècies. Les fluctuacions en *A* tenen múltiples efectes en els organismes sobre i sota de la interfase aquàtica, tant per a la flora com per a la fauna aquàtica -e.g. macròfits, espècies del zooplàncton i peixos-. Aquestes fluctuacions poden afectar fortament les condicions de l'hàbitat a través de les variacions en els paràmetres fisicoquímics -e.g. temperatura, llum, nutrients o de la concentració de solut-. La concentració de solut és particularment rellevant en les llacunes salines, que són una part important de l'extensió geogràfica dels ecosistemes aquàtics continentals del món. Tant les rutes migratòries com els llocs de nidificació de moltes aus aquàtiques depenen de la dinàmica d'inundació de les llacunes, sobretot si les aus depenen de les xarxes tròfiques de les llacunes. A aquestes llacunes, tant els patrons migratoris de les aus aquàtiques, com els cicles de vida de molts animals de vida curta, com els patrons de pluja, es produeixen en una escala temporal anual o més curta. D'acord amb això, la impredictibilitat ecològica en aquests sistemes pot ser concebuda com desviacions de la variació mitjana estacional. És a dir que la principal font d'impredictibilitat ambiental és la variació inter-anual de la variació intra-anual. Observacions episòdiques *in situ* durant l'última dècada han mostrat que algunes llacunes s'assequen amb freqüència,

desenvolupant, a continuació, una crosta de sal gruixuda. D'altres, en canvi, se sap que es mantenen inundades durant anys a través de diverses estacions seques. Això suggereix que les llacunes presenten diferents patrons hidrològics, tot i que, abans d'aquesta tesi, no es disposava d'una caracterització quantitativa precisa de les seues característiques hidrològiques.

El rotífer monogonont *Brachionus plicatilis* és una de les espècies de zooplàncton que habita amb freqüència les llacunes salines de la regió mediterrània. Aquesta espècie té un tipus de reproducció anomenat partenogènesi cíclica, en la qual la proliferació per partenogènesi ameiòtica (fase asexual) es combina ocasionalment amb la producció ocasional de mascles i reproducció sexual. Aquesta última fase de reproducció sexual resulta en producció d'ous de diapausa. Les poblacions de *B. plicatilis* en regions temperades són temporals –és a dir que els rotífers no estan actius durant tot l'any, i colonitzen la columna d'aigua durant l'anomenada estació de creixement (planctònica)-. Al començament de cada estació de creixement, la població activa s'inicia a través de l'eclosió dels ous de diapausa al sediment de la llacuna. L'eclosió dóna lloc a femelles asexuals que produeixen ous que no són de diapausa, és a dir que fan eclosió al cap de poc de temps, els quals donen lloc a filles genèticament idèntiques a través de

partenogènesi, produint així clons. La reproducció sexual és induïda per un senyal química produïda pels rotífers que s'acumula en el medi a mesura que augmenta la densitat de la població. Un cop s'ha superat un llindar de densitat poblacional, les femelles asexuals produeixen una fracció de la seua descendència com a filles sexuals. Per tant, la reproducció sexual i asexual, generalment estan superposades temporalment. Les femelles sexuals a través de meiosi produeixen ous haploides, que es converteixen en mascles haploides si els ous no són fertilitzats i en ous de diapausa (diploides) si són fertilitzats. Els mascles són molt més menuts que les femelles (longitud del mascle: 100 μm; longitud de la femella: 300 μm), no s'alimenten i tenen una vida molt més curta que les femelles. A diferència dels ous asexuals, els ous produïts sexualment són embrions latents; per tant, el sexe s'associa amb diapausa. Els ous de diapausa es dipositen en el sediment i romanen en estat latent durant un període de durada variable. En condicions adequades, s'indueix l'eclosió dels ous de diapausa i s'inicia una nova estació de creixement. En algunes soques s'ha descrit un període refractari curt per a la inducció de l'eclosió. A més, no tots els ous de diapausa fan eclosió en l'estació de creixement següent a la seua producció. Els ous que no fan eclosió sovint mostren una diapausa llarga i s'acumulen en els sediments formant bancs d'ous de diapausa, on poden romandre viables durant dècades o fins i tot segles. Aquest

cicle de vida es considera una adaptació als hàbitats temporals.

La impredictibilitat ambiental a les poblacions naturals pot actuar sobre diversos trets de la història vital de l'organisme, especialment en aquelles espècies amb cicles biològics complexos. La teoria sobre els trets de la història vital s'ha desenvolupat durant els últims 50 anys i estudia els patrons de creixement, maduració, reproducció i supervivència dels organismes. L'estudi dels trets de la història vital és rellevant, ja que aquests poden interactuar per reduir l'impacte de la variabilitat ambiental. Una fascinant manera de superar les condicions ambientals desfavorables és la diapausa. La diapausa és un període en el qual el desenvolupament embrionari entra en suspensió i que és molt freqüent al regne animal. La diapausa comença abans de l'adveniment de les condicions adverses, i la seua durada -que està regulada principalment per factors interns fisiològics- pot superar la durada de les condicions adverses. Això dóna lloc a moltes preguntes ecològiques i evolutives. De fet, abandonar la diapausa pot tenir importants conseqüències sobre la dinàmica de les poblacions de rotífers que habiten en hàbitats temporals. En primer lloc, hi ha un compromís entre l'assignació de recursos a la proliferació actual o per a la producció de formes resistents, que podrien reprendre la proliferació en el futur. La producció de les formes de

diapausa s'ha d'iniciar amb anterioritat al començament de les condicions desfavorables, ja que és l'única forma de supervivència del genotip davant d'aquestes condicions. No obstant això, si les condicions favorables s'allarguen en el temps, una inversió prematura en diapausa pot ser subòptima, ja que resulta en una reducció de la taxa de creixement poblacional actual. En segon lloc, un compromís similar ocorre amb la sortida de la diapausa. Abandonar la diapausa pot implicar explotar un ambient adequat, però pot ser un inconvenient si l'hàbitat es torna inadequat massa aviat i el genotip no produeix una nova cohort de formes de diapausa. Tot això produeix pressions selectives contraposades que poden conduir a l'aparició d'estratègies de *bet hedging* en ambients fluctuants. Als rotífers que són sexuals facultatius, aquesta situació es torna més complexa i, potencialment, més interessant, ja que la producció de formes de diapausa (en forma d'ous de diapausa) està associada a la reproducció sexual. Llavors, a aquests tipus de rotífers, la producció d'ous de diapausa té associat, als costos de diapausa, els costos addicionals derivats de la reproducció sexual.

Mapejar els genotips als fenotips i els fenotips als ambients és d'una importància fonamental per entendre com les poblacions naturals s'adapten al seu ambient local. No obstant això, les bases genòmiques de les respostes evolutives

a l'adaptació a la impredictibilitat ambiental, especialment aquelles associades al *bet hedging*, són poc conegudes. El desenvolupament de la seqüenciació de nova generació (NGS) ha proporcionat nous mètodes de seqüenciació d'ADN (Capítol 2) que pot ajudar a profunditzar en el coneixement sobre aquestes qüestions. NGS fa la seqüenciació més barata, més ràpida i, el que és més important, permet que els estudis genòmics es realitzen pràcticament en qualsevol organisme. Fins fa poc, la major part del nostre coneixement en genòmica es basava en organismes model. Això no és completament satisfactori en biologia, especialment per als treballs de recerca comparatius, evolutius i de biodiversitat, on la informació sobre els organismes model no és representativa del conjunt. Això està canviant amb el desenvolupament de les tecnologies NGS. Per exemple, nova i àmplia informació sobre les correlacions genètiques de l'adaptació a diferents ambients es pot obtenir directament a escala genòmica en espècies amb recursos genòmics limitats (organismes no considerats model en la genètica tradicional). Un cop més, tot això permetrà a la comunitat científica abordar nous problemes, el que ajudarà a tancar la bretxa cada cop més menuda que separa l'ecologia i l'evolució.

Objectius

Aquesta tesi aprofita els esmentats avanços tècnics i adopta un marc teòric en ecologia evolutiva per tal d'abordar l'adaptació local als nivells d'impredictibilitat ambientals, utilitzant rotífers com a organismes model. Aquest treball també aprofita el coneixement acumulat, en particular sobre la seua biologia de poblacions a l'est d'Espanya, del rotífer monogonont *B. plicatilis*. Com s'explica amb més detall en el Capítol 2, aquest rotífer ha estat considerat com un organisme model en estudis evolutius i ecològics. Els principals objectius d'aquesta tesi són: (1) estudiar la variació en els trets de la història vital relacionats amb la diapausa a poblacions de rotífers de l'espècie *B. plicatilis* que habiten llacunes amb diferent grau d'impredictibilitat ambiental. (2) Descobrir, a escala genòmica, els genotips correlacionats amb l'adaptació a la impredictibilitat ambiental mitjançant l'ús de tecnologies genòmiques. Per a la consecució dels objectius (1) i (2), va ser necessari plantejar un objectiu addicional, amb rellevància metodològica intrínseca: (3) Quantificar el grau d'impredictibilitat ambiental a través d'imatges de satèl·lit de llacunes mediterrànies.

Esquema d'aquesta tesi: metodologia i resultats principals

Aquesta tesi, en les parts que segueixen a la introducció general (Capítol 1), està organitzada com s'especifica a continuació.

En el Capítol 2 es descriu el context metodològic. Es dóna una visió general de les metodologies en les quals el lector podria no estar familiaritzat. Es descriuen les espècies i els hàbitats estudiats en aquesta tesi, i es proporciona una visió general de la teledetecció i de les tècniques moleculars utilitzades. En primer lloc, es descriuen les principals característiques biològiques dels organismes de l'estudi: el rotífer *B. plicatilis*. En segon lloc, es descriuen les principals característiques eco-geogràfiques de la zona d'estudi. En tercer lloc, es proporciona una descripció global del sistema d'estudi (és a dir, les poblacions de *B. plicatilis* a l'est d'Espanya). En quart lloc, es fa una descripció de la forma en la qual les dades de teledetecció són obtingudes pels satèl·lits i es descriuen les principals característiques d'aquest tipus de dades. Finalment, s'introdueixen les principals tècniques moleculars utilitzades en aquesta tesi (seqüenciació del genoma i *Genotyping by Sequencing*, GBS). En aquesta última secció es proporciona una descripció de com es duen a terme aquestes tècniques i s'aporta un glossari de termes.

La investigació descrita en el Capítol 3 fa ús de dades a llarg termini obtingudes per satèl·lits (27 anys, Landsat TM / ETM +) per tal d'avaluar un conjunt de models ambientals i obtenir estimes de la predictibilitat ambiental en llacunes salines. Aquestes llacunes a la regió Mediterrània poden produir

zones cobertes de sal que dificulten la distinció entre el sòl i l'aigua. Aquest problema va ser abordat per una nova aproximació que combina els índexs de presència d'aigua i la banda infraroja curta, aquesta última actuant com a filtre de sal. Un cop es va obtenir aquesta llarga sèrie temporal, els components predictibles i impredictibles de la variació en *A* es van extreure utilitzant dos aproximacions diferents. La primera aproximació, basada en les mètriques de predictibilitat de Colwell, transforma una variable contínua en una variable nominal. Aquesta transformació és un pas clau, i aquí diferents models es van considerar basats en com la variació de *A* podria ser rellevant des del punt de vista de l'organisme d'interés. Com una segona aproximació, el *General Additive Model* (GAM) es va desenvolupar com una nova mètrica per quantificar el grau de predictibilitat ambiental. Aquesta mètrica es paral·lela a les mètriques de Colwell, però utilitza un model de regressió i manté la variable com a contínua durant l'anàlisi. La similitud i divergència dels diferents índexs de predictibilitat es van analitzar per determinar la seva sensibilitat als supòsits necessaris per a quantificar la predictibilitat. Les dues aproximacions van extreure informació significativa sobre el grau de predictibilitat de les llacunes estudiades. Alguns supòsits dels models van mostrar que no afectaven els resultats de predictibilitat, mentre que altres podien ser associades a les característiques ecològiques de l'espècie d'interés per a la

qual la predictibilitat és avaluada. La metodologia descrita en aquest capítol és aplicable a una àmplia varietat de sistemes, i serà de gran valor per a la quantificació i caracterització de la predictibilitat, el que és essencial atenent el predit augment global en la impredictibilitat de les fluctuacions ambientals.

En el Capítol 4, *B. plicatilis* va ser utilitzat com a organisme model per investigar l'adaptació a les fluctuacions ambientals. Dos trets relacionats amb la diapausa en rotífers s'estudiaren usant clons de nou poblacions naturals de *B. plicatilis* que varien en el grau d'impredictibilitat ambiental (Capítol 3). Hi ha dos trets de la història vital que son clau en el cicle de vida dels rotífers monogononts (Capítol 2). Aquests estan associats amb l'entrada i sortida de la diapausa. Un d'ells és la propensió al sexe, que està inversament relacionat amb el llindar de densitat per a la iniciació del sexe. Un altre és la fracció d'eclosió dels ous de diapausa, que està inversament relacionat amb la durada diapausa. Tots dos trets s'han proposat com a exemples d'estratègies de *bet hedging* que poden interactuar per reduir els riscos associats a la impredictibilitat de l'ambient. La propensió al sexe s'ha proposat com un cas de *bet hedging* conservatiu en les poblacions de rotífers que viuen en hàbitats impredictibles. Quan hi ha incertesa respecte al moment en el qual ocorren els períodes desfavorables, una estratègia de baix risc podria

ser produir ous de diapausa tan aviat com siga possible per evitar una estació de creixement inesperadament curta. No obstant això, si l'estació de creixement no és curta, una inversió prematura en sexe i diapausa redueix la taxa de proliferació clonal, produint un descens de l'eficàcia biològica. Aquestes consideracions condueixen a la predicció que la propensió al sexe augmentarà amb l'augment de la impredictibilitat. No obstant això, a l'hora de testar aquesta predicció, un potencial factor de confusió s'ha de considerar: si les estacions de creixement són predictiblement curtes, també s'espera que evolucione una alta propensió al sexe. En contrast, la fracció d'eclosió dels ous de diapausa s'ha proposat en diversos estudis teòrics com una forma de *bet hedging* diversificador. Com els rotífers no poden predir si una determinada estació de creixement serà prou llarga per a completar el seu cicle de vida i garantir la producció d'ous de diapausa, taxes d'eclosió intermèdies s'esperen en hàbitats en els quals hi ha tant estacions de creixement llargues (on, idealment, tots els ous hi haurien d'eclosionar; estacions de creixement bones) com estacions de creixement inesperadament curtes (on, idealment, cap ou hi hauria d'eclosionar; estacions de creixement males). Segons la teoria de *bet hedging*, la fracció de eclosió òptima ha de ser igual a la freqüència d'estacions de creixement bones. Per exemple, una fracció de l'eclosió al voltant de 0,5 (és a dir, intermèdia) seria d'esperar en un hàbitat

completament impredictible, ja que la freqüència de temporades bones estaria al voltant de 0,5. Per tant, les hipòtesis testades són que el nivell d'impredictibilitat ambiental està directament relacionat amb la propensió al sexe i inversament relacionada amb la fracció d'eclosió d'ous de diapausa. La variació genètica en la propensió al sexe es va estudiar mitjançant la realització de 810 bioassaigs (9 poblacions × 30 clons × 3 rèpliques). L'estimació de la fracció d'eclosió dels ous de diapausa es va realitzar mitjançant l'ús d'un subconjunt de deu clons que ja es van utilitzar en l'experiment propensió al sexe. En total, 8.640 bioassaigs (96 ous de diapausa × 9 poblacions × 10 clons) es van realitzar. Com a resultat d'aquests experiments es van trobar, a ambdós trets, nivells significatius de variació genètica a nivell intra-poblacional. D'altra banda, com es va predir, es va trobar una correlació positiva entre el grau d'impredictibilitat —quantificat en el Capítol 3– i la propensió al sexe. Aquesta correlació suggereix una estratègia de *bet hedging* conservativa que proporciona protecció contra les estacions de creixement inesperadament curtes. En contrast, es va trobar que la fracció de l'eclosió dels ous de diapausa no estava relacionada amb el nivell de predictibilitat ambiental. Tot això posa en relleu la capacitat de les poblacions de rotífers per adaptar-se localment als ambients variables en el temps.

En el Capítol 5 es fa servir una aproximació integradora en el que es fa ús de dades ambientals (Capítol 3), fenotípiques (Capítol 4) i genòmiques. Es va obtenir dades de Genotyping by Sequencing (GBS) per al conjunt de 270 clons de *B. plicatilis* de les nou poblacions a l'est d'Espanya utilitzades en el Capítol 3. D'altra banda, 11,7 Gb de dades de seqüenciació es generaren per a assemblar i anotar un genoma per *B. plicatilis.* Com a resultat de l'anàlisi de dades GBS, 4543 SNPs d'alta qualitat es descobriren i es genotiparen en un genoma d'una mida reduïda però que és dens en regions gèniques. Mitjançant l'ús d'aquests SNP, es pot obtenir informació rellevant sobre l'estructura genètica de les poblacions naturals, i també, es poden localitzar les parts del genoma que presenten evidencies d'estar sota selecció. Hi ha diversos mètodes que tenen com a objectiu detectar *loci* sota selecció. Un primer grup de mètodes, es basa en els valors de F_{ST} i utilitza les freqüències al·lèliques per tal d'identificar els SNPs que presenten valors significativament més alts -i.e. selecció divergent- o baixos -i.e. la selecció balancejant o purificadora- sota un model d'evolució neutral. Aquests tipus de mètodes, però, s'ha descrit que tenen baixa potència, especialment quan hi ha una alta diferenciació en els marcadors neutres, com s'ha descrit en les poblacions de *B. plicatilis*. Recentment, un altre grup de mètodes de detecció de *loci* sota selecció s'ha desenvolupat. Es basa en la correlació de les freqüències al·lèliques a un conjunt de

factors ambientals o fenotípics (*Genome-wide association analysis*, GWAS). Aquests mètodes basats en la correlació, s'ha descrit que tenen una potència major que els models basats en la diferenciació, no obstant també s'ha vist que poden mostrar majors taxes de falsos positius, que és el principal inconvenient dels mètodes d'anàlisi de *loci* sota selecció disponibles en l'actualitat. No obstant això, aquestes tècniques es poden utilitzar per aportar informació sobre els gens que estan potencialment sota selecció o involucrats en processos de diferenciació poblacional adaptativa. Al aplicar aquestes aproximacions, més de 90 SNPs van mostrar valors de F_{ST} indicatius d'estar potencialment sota selecció. D'altra banda, es van trobar més de 160 SNPs correlacionats a un conjunt de factors ambientals o fenotípics. Sorprenentment, la majoria d'aquests SNPs es van trobar dins de gens anotats al genoma i les seues funcions putatives van ser associades.

Al Capítol 6, es discuteix en termes generals els principals resultats obtinguts en aquesta tesi, es proposen potencials investigacions futures, i s'exposen les conclusions més importants.

Conclusions

Les conclusions principals d'aquesta tesi són: **(1)** Les imatges dels satèl·lits Landsat 5 i 7 -després de l'aplicació d'una combinació de procediments per discriminar entre l'aigua i el

fons- proporcionen una sèrie temporal llarga (27 anys) de la variació en l'àrea que ocupa la superfície d'aigua d'un grup de llacunes salades mediterrànies. **(2)** Tenint en consideració la concepció que la predictibilitat depén del punt de vista de l'organisme d'interés, i utilitzant les mètriques de Colwell, es van desenvolupar a aquesta tesi diferents models per a l'estimació de la predictibilitat. A més, l'ajustament GAM es va desenvolupar en aquesta tesi com una aproximació alternativa per a la quantificació de la predictibilitat ambiental. Tots aquests models van ser avaluats tenint en compte de quina forma la variació en l'àrea de la superfície de l'aigua podria ser rellevant per a l'organisme d'interés. **(3)** L'aplicació de les mètriques de predictibilitat va permetre la quantificació de la predictibilitat en un grup de llacunes mediterrànies. Aquestes llacunes van mostrar un rang ampli de predictibilitat. **(4)** Mitjançant una classificació *a posteriori* dels models emprats per estimar la predictibilitat, es va concloure que alguns dels supòsits no afectaven els resultats de predictibilitat, mentre que uns altres podien ser associats a les característiques ecològiques de l'espècie d'interés per a la qual la predictibilitat és avaluada. **(5)** Les poblacions de *Brachionus plicatilis* que habiten un conjunt de nou llacunes salines mediterrànies van mostrar nivells significatius de variació genètica dins de la població per la propensió al sexe i per a la fracció d'eclosió d'ous de diapausa. **(6)** La propensió al sexe en les poblacions de rotífers, i per tant la

inversió prematura en diapausa, va disminuir amb la predictibilitat ambiental, mentre que la relació d'aquest tret amb la longitud de l'hidroperíode va ser relativament feble. Això suggereix una estratègia de *bet hedging* de tipus conservatiu, que proporciona protecció contra les estacions de creixement inesperadament curtes. **(7)** Les fraccions d'eclosió d'ous de diapausa van presentar valors intermedis (del 44 al 88%) en totes les poblacions estudiades, però no es va trobar cap relació significativa respecte al nivell de predictibilitat ambiental ni respecte a la longitud del hidroperíode. **(8)** Les poblacions de rotífers, malgrat el seu potencial per a l'intercanvi genètic a través de la dispersió passiva d'ous de diapausa, son capaces de divergir localment en els trets associats a diapausa dins d'un petit rang geogràfic (240 km^2). **(9)** El genoma de *B. plicatilis* es va assemblar en aquesta tesi. La seva anotació estructural va produir 54,725 gens. S'han assignat funcions temptativament per al 30% d'ells. **(10)** El *Genotyping by sequencing* (GBS) i l'anàlisi bioinformàtic posterior proporcionaren un gran nombre (4543) de polimorfismes de nucleòtid simple (SNPs) d'alta qualitat. **(11)** Un alt nombre de SNPs -la majoria d'ells situats dins de regions gèniques- van mostrar major diferenciació genètica entre les poblacions del que es podria esperar per atzar i també es van mostrar correlacionats amb els trets d'història vital i els factors ambientals testats, pel que són candidats a estar sota la selecció diversificadora que

condueix a l'adaptació local. **(12)** De forma inesperada, un elevat nombre de SNPs, més de la meitat dels quals situats dins de regions gèniques, van mostrar senyals de selecció balancejant o purificadora en *B. plicatilis*. Aquest descobriment requereix més investigació. **(13)** Un nombre elevat de gens van ser identificats com candidats forts per ser part de la base genòmica de l'adaptació local a la fluctuació ambiental. Aquests gens poden servir com una base de dades per a futurs estudis.

1

Introduction

Environmental fluctuations in population ecology

Most of the fundamental theory in population ecology assumes that organisms live in constant environments (i.e. homogeneous in space and time). For instance, the considered by most (e.g. Ginzburg 1986, Brown 1997, Berryman 1999) as the first principle of population ecology — the exponential growth model– states that 'a population will grow (or decline) exponentially as long as the environment experienced by all individuals in the population remains constant' (Turchin 2001, from Malthus 1798). The premise that the environment is constant was assumed in the seminal mathematical models developed in early population ecology, and adopted in the empirical studies framed in this theoretical work. This conception progressively changed during the second half of the 20th century (McIntosh 1991, Wiens 2000). For instance, May (1974) illustrated the need of considering habitat fluctuations by showing that even in an

exponential growth model, the incorporation of random environmental fluctuations —if they are large enough– could lead to extinction. Later, Southwood (1977) made one of the first efforts in classifying habitats with respect to the variations in space and in time. He emphasized that classifications of habitats fluctuating in time must not only depend only on the length of the favourable and unfavourable periods, but also on the generation time of the focus species. This idea is stressed in the textbook by Begon et al. (2006). These authors warn the reader that the organisms do not 'sense the environment as we do', and that we have to adopt the organism's-eye view of the environment. From the organism's-eye view, predictability is related to its ability to anticipate and adjust to a future environmental condition, thus involving a time scale (MacArthur and Levins 1964; Levins 1968). Distinction between fine-grained and coarse-grained environmental heterogeneity makes an effort in this direction (e.g. Hutchinson 1978). An instance of this is the annual temperature fluctuation, which might be experienced very differently by short-lived invertebrates —fine grain– if compared to long-lived vertebrates —coarse grain– (Crozier et al. 2008).

Inspired in the classification made by Southwood (1977), and in order to illustrate how habitats fluctuate in time, three idealized habitats are shown in Figure 1.1. The first idealized

habitat is constant, the second one varies in a periodic way, and, finally, in the third one, there is variation but it is not periodic. In terms of the degree of predictability, the first two habitats would be grouped as predictable and the latter as unpredictable.

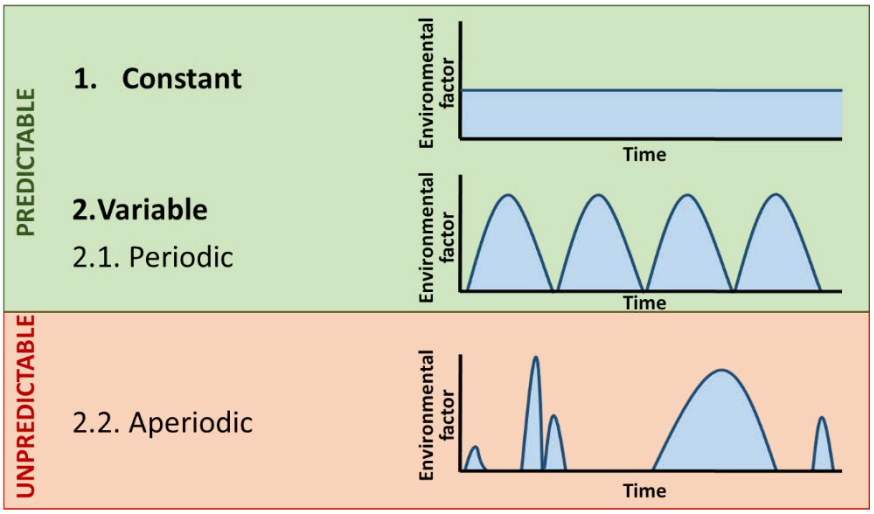


Figure 1.1. Three idealized situations showing different ways of fluctuations of the environmental factor (either biotic or abiotic) over time in a habitat with no monotonous trend (after Southwood 1977). Blue areas show factor values favouring the focus species. In terms of the degree of predictability, habitats 1 and 2.1 would be grouped as predictable and 2.2 as unpredictable.

Unlike in these idealized habitats, in nature, any habitat shows some constancy and some variation in their features; therefore dissecting the constant, periodic and aperiodic components in the habitat patterns provides a realistic view. Nevertheless, quantifying the degree of predictability of

habitat features is not simple. It needs of a sophisticated, robust methodology, as it cannot be based on mean values, but on variances. Moreover, as stated above, this characterization needs assumptions —as explicit as possible– regarding the relevant time-scales that can be associated with predictability, the assumptions not being straightforward. At last, one of the most important constraints on fluctuation analysis is the need of long time-series data. Several indices have been developed to estimate the degree of predictability from time-series data, such as Colwell's predictability index (1974), Fourier transform, and asymmetric eigenvector maps (Blanchet et al. 2008; Legendre and Gauthier 2014). Basically, they decompose the time series by periodic and stochastic variation (Sabo and Post 2008), and associate them with predictable and unpredictable fluctuations, respectively.

Adaptation to environmental unpredictability

Understanding how organisms adaptively respond to environmental fluctuations —which are ubiquitous– and their degrees of predictability is a fundamental question in evolutionary biology that has motivated a growing body of research (e.g. Crozier et al. 2008; Chevin et al. 2010; Chown et al. 2010). There are several ways by which organisms evolutionarily respond to unpredictable variation (Simons

2011). The best known of these ways is through natural selection that acts recurrently on the heritable variation among individuals; this process is named adaptive tracking, and implies genetic evolution (e.g. Lynch and Lande 1993; Tufto 2015). Interestingly, as adaptation consumes time, it is not expectable the population to be adapted as far as unpredictable fluctuations continue. Adaptive phenotypic plasticity, another mode of evolutionary response, occurs when individuals modify their phenotypes according to environmental conditions without changing their genetics, and do it in a way that their performance increases (Reed et al. 2010). Bet hedging constitutes a third way of adaptive response and occurs when a genotype increases the geometric mean of fitness at the cost of a decrease in the arithmetic mean, by reducing fitness variance (Philippi and Seger 1989). Selection increasing geometric fitness is expected because long-term fitness is related to its geometric mean, rather than to its arithmetic mean (Gillespie 1974). There are two main modes of bet hedging: diversified and conservative (Philippi and Seger 1989). Diversified bet hedging is a transgenerational effect that occurs when a single genotype produces different phenotypes in its offspring in advance of future unpredictable conditions (Childs et al. 2010). In contrast, under conservative bet hedging, an individual sacrifices expected (arithmetic) fitness to reduce temporal variance in reproductive success by employing a

single low-risk phenotype across all possible environmental future scenarios (Botero et al. 2015). These three kinds of responses are non-exclusive, and combinations of them may occur simultaneously in natural populations (Tufto 2015). Worth noting, although bet hedging is well developed theoretically in biology (Cohen 1966; Starrfelt and Kokko 2012), empirical evidence is still scarce (Simons 2011).

Life-history traits and diapause

Life history theory has been developed over the last 50 years and studies the patterns of organismal growth, maturation, reproduction and survival (Flatt and Heyland 2012). Those features —called life-history traits– can be quantified and analysed through well-developed demographic statistical methods and are considered fitness components (García-Roger 2006; Flatt and Heyland 2012). Unpredictability in natural populations can act on several organism traits, especially in those species with complex life histories (Crozier et al. 2008). In addition, life-history traits can interact to reduce the impact of environmental variability (Brown and Venable 1986; Rees 1994; Ellner et al. 1998). A fascinating way to overcome unsuitable environmental conditions is the diapause. The diapause is a period of suspended development which is wide-spread in the animal kingdom (Danks 1987). Diapause has been described as a mechanism

of dispersion through time, working as a 'time machine' (Hairston 1998).

Diapause starts before the advent of adverse conditions, and its duration —which is mostly regulated by internal physiological factors– can exceed the duration of adverse conditions (Hairston et al. 1995; Hand and Podrabsky 2000), what gives rise to many ecological and evolutionary questions. Leaving or exiting diapause may have important consequences on the dynamics of the populations dwelling in temporary habitats. First, a trade-off exists between the assignation of resources to current proliferation or to resistant stages that could resume proliferation in the future (e.g. Serra and King 1999). The production of diapausing stages should be initiated beforehand the beginning of the unsuitable conditions, given that it is necessary for the survival of the genotype. Nevertheless, if the favourable conditions prevail, an early investment in diapause can be sub-optimal, since it results in a reduction of the current population growth rate (Serra and King 1999). Second, a similar compromise happens with the exit of diapause. Exiting diapause may imply to exploit a suitable environment, but may be inconvenient if the habitat becomes unsuitable too early and the genotype does not produce a new cohort of diapausing stages (Seger and Brockman 1987; Brendonck and De Meester 2003). All of this produces contrasting selective pressures that can lead to the

appearance of bet-hedging strategies in fluctuating environments (Ellner 1997).

In facultative sexual rotifers this situation becomes more complex —and potentially more interesting– since the production of diapause stages (in form of diapausing eggs) is associated to sexual reproduction. Then, diapausing egg production, as well as having associated the costs of diapause, has additional costs derived from sexual reproduction. One of the most referred costs of sex was described by Maynard Smith (1978). It is the 'two-fold cost of sex', caused because part of the reproductive resources are allocated to males, which do not actually make any contribution to population growth. Despite rotifer males are haploids and dwarf, the two-fold cost applies as a result of sex-allocation evolution (Aparici et al. 1998; 2002). Moreover, the optimal timing to entering diapause is affected by this linkage between sex and diapause, since sexual reproduction patterns (i.e. the timing and the amount of sex; Serra and Carmona, 1993; Serra and King, 1999) are also expected to affect the genetic diversity of diapausing eggs. Several studies (Gómez and Carvalho 2000; Ortells et al. 2006) have shown that, due to clonal selection, genetic diversity is eroded during the growing season (i.e. during the lapse of time in which rotifer populations are active in the water column). As a result, diapausing eggs produced early in the growing

season are expected to harbour more diversity than diapausing eggs produced lately in the season. Therefore, in this respect environmental variability is expected to benefit lineages with early sex induction owing to the genetic variation conferred by sexual reproduction (Peck et al. 1999; Frantz et al. 2006), which allows populations to develop over a broader range of ecological conditions (Pound et al. 2002). All of this plus the existing knowledge of the high genetic differentiation among populations of the facultative sexual rotifer *Brachionus plicatilis* in the eastern Iberian Peninsula (Gómez et al. 2002a; Campillo 2009), make these populations a good model system to study the adaptation to environmental unpredictability (see Chapter 2).

Technological revolution in methodologies

Advances in science usually occur either because of the appearance of new theories that provoke a paradigm shift or by the appearance of new technologies that open new paths to obtain data. The development of this thesis makes use of two powerful technologies that can be helpful and make a difference in modern ecology.

First one is remote sensing technology, which has been highly developed in the last century (Schowengerd 2007). Nonetheless, its use in several research areas has not reached

yet its full potential, as is the case with ecology (Pettorelli et al. 2014). Several studies show that it can offer solutions to a wide variety of problems in nearly all fields of environmental research (Colwell et al. 1983; Campbell 2002; Roach et al. 2013). Obtaining long time series is regarded as costly, and these series are consequently scarce (May and McLean 2007). In this regard, the information gathered by several satellites in recent years can provide long time-series data. This is the case of the Landsat satellite series, which ranges from 1972 to the present and have moderate spatial and intermediate temporal resolutions (Frazier and Page 2000). Since 2008, the satellite scenes from Landsat satellite series are freely available from the United States Geological Survey (USGS), allowing for the scientific community to gather a great deal of information free of cost (Asner 2009). The data obtained from satellites —the reflectances from various wavelength regions of the electromagnetic spectrum (Chapter 2)– can be processed and analysed to obtain information about relevant environmental features. All of this makes the satellite remote sensing data especially helpful to expand temporally and spatially our knowledge of the focus environments.

Second one is next generation sequencing (NGS), a new method for massive DNA sequencing (see Chapter 2) with a rapid development during the last decade. NGS makes

sequencing cheaper, faster and, more importantly, allows genomic studies to be performed in practically any organism. Until recently, most of our knowledge in genomics —even in deep genetics– was based on model organisms. This is not completely satisfactory in biology, especially for comparative, evolutionary and biodiversity approaches, where information on a few model organisms is not representative. This is changing with NGS. For instance, new, extensive information on the genetic correlates of adaptation to different environments can be directly disentangled at the genome level on species with limited genomic resources (i.e. non-traditional genetic model organisms, but see Stapley et al. 2010; Klepsatel and Flatt 2011). Again, all of this allows the scientific community to approach new problems, in this case by mapping genotype on phenotype and phenotype on environment what helps to close the —everyday smaller– breach between ecology and evolution.

Objectives and outline of this thesis

This thesis takes advantage of the mentioned technical advances in methods and adopts the theoretical framework of evolutionary ecology explained above in order to address local adaptation to uncertainty levels, using rotifers as model organisms. This work also takes advantage of the accumulated knowledge on the monogonont rotifer *B. plicatilis*, in particular on their population biology in eastern

Spain. As explained in more detail in Chapter 2, this rotifer has been considered a model organism in evolutionary and ecological studies.

The main goals of this thesis are:

(1) To study variation in diapause-related life-history traits among populations of rotifer *B. plicatilis* inhabiting ponds with different degree of environmental unpredictability.

(2) To unveil, at a genome wide scale, genotypes correlated to adaptation to local environmental unpredictability by using genomic technologies.

For the consecution of objectives (1) and (2), an additional objective with intrinsic methodological relevance was needed:

(3) To quantify the degree of environmental unpredictability based on satellite imagery data from Mediterranean ponds.

The thesis, in the parts following this general introduction (Chapter 1), is organized in the following way:

Chapter 2 describes the methodological context. It is an introduction to the species and the habitats studied in this

thesis, as well as an overview of the main remote sensing and molecular techniques used. Thus, Chapter 2 is an extension of the general introduction of the thesis, but it focuses on materials and methods. The aim of this chapter is to provide some background to help readers who are not familiar with some aspects developed in further chapters of this thesis.

In Chapter 3, a remote sensing data set spanning 27 years (Landsat TM/ETM+) is used to obtain a time series of the water-surface area in twenty Mediterranean ponds. This time series is utilised to assess a set of environmental models for estimation of environmental predictability.

In Chapter 4, two diapause-related life-history traits are studied in rotifers using clones from nine *B. plicatilis* natural populations that, according to results in Chapter 3, vary in the degree of environmental unpredictability.

Chapter 5 describes and analyses genotyping by sequencing (GBS) data obtained from all the rotifer clones used in Chapter 4 in order to detect genetic variants at the genome level. Environmental, phenotypic and genomic data are integrated in order to study the genetic correlates of local adaptation to environmental unpredictability, and an effort to identify responsible genes for that adaptation is made.

Chapter 6 discusses in general terms the main results obtained in this thesis, proposes prospective research and sets out the most important conclusions.

2

The methodological context

Summary

This chapter gives an overview of methodological topics that the reader might not be familiar with. The chapter describes the species and the habitats studied in this thesis and provides an overview of the remote sensing and molecular techniques used. First, the main biological features of the study organisms, the rotifer Brachionus plicatilis *are described. This zooplankter has a life cycle called cyclical parthenogenesis, which includes both sexual and asexual (i.e. parthenogenetic) reproduction. This cycle is a key feature for the persistence of rotifer populations dwelling in temporary environments, since the product of the sexual reproduction are diapausing eggs, the stages able to resist adverse conditions (e.g. pond desiccation, freezing). Second, the main ecogeographic characteristics of the study area are described. Third, a global description of the study system is provided (i.e.* B. plicatilis *populations in eastern Spain). Fourth, information about the main features of the remote sensing data used and a description of how they are obtained are supplied. Finally, the main molecular techniques used throughout the thesis are introduced (genome sequencing and genotyping by sequencing, GBS). In this last section a description of how these techniques are performed and a glossary of terms are provided.*

The study organism: *Brachionus plicatilis*

In this thesis a microscopic aquatic invertebrate, the monogonont rotifer species *Brachionus plicatilis* Müller*, 1786* (hereafter *B. plicatilis*) is used as model organism. Rotifers are bilateral metazoans that make up what was classically regarded as a phylum (see below) including three classes: Bdelloidea, Monogononta and Seisonidea. The term Rotifera (L., *rota* + *ferre*, wheel-bearers) refers to the ciliary corona present in the apical end of their body, which they use in locomotion and food gathering (Wallace and Snell 1991; Wallace and Smith 2009). The metachronous beating of cilia makes the corona to be perceived as rotating due to a visual effect.

More than 2,000 rotifer species —most of them, monogononts– have been described, the species being generally aquatic and free-living. Although most of the species inhabit freshwater bodies, marine and brackish species exists (Wallace and Snell 1991). Indeed, some of these species have a wide biogeographical range (Koste, 1978), being one of the major groups of zooplanktonic organisms in continental water bodies along with copepods and cladocerans. As a consequence of their typically high reproductive rates, they can reach large population sizes constituting a substantial part of continental zooplankton, and playing a critical role in aquatic food webs

(Starkweather, 1987; Walz, 1997; Armengol et al. 2001). Rotifer size ranges from 40 µm to 3,000 µm, although most of them measure from 100 to 500 µm (Hickman et al. 1997). Given their minute size, their diet consist of matter small enough to fit through their tiny mouths. Most rotifers may be described as euryphagous, and typically feed on bacteria, algae, protozoa, and yeast, as well as organic detritus (Wallace and Smith 2009).

Unlike in the classical classification, Rotifera is no longer considered a monophyletic group formed by the classes Bdelloidea, Monogononta and Seisonidea. Recently, the Acanthocephala, that were thought to be a discrete phylum, have been proposed to be highly modified rotifers (Welch 2000), a sister group of Seisonidea (Sielaf et al. 2016). Thus, in some recent revisions, Rotifera and Acanthocephala are placed together in a new larger taxa called Syndermata (Welch 2000; Fontaneto and Jondelius 2011; Wey-Fabrizius et al. 2014). The type of reproduction is different among the four different groups of Syndermata: Seisonids and Acantocephalans are obligate sexuals, Bdelloids obligate asexuals and Monogononts cyclical parthenogens (i.e. they combine asexual and sexual reproduction in their life cycle).

The species *B. plicatilis* (Figure 2.1) belongs to the *Brachionus plicatilis* species complex, a complex of cryptic species (i.e.

distinct but morphologically similar species classified as a single species by classical, morphological based taxonomy), which includes at least fifteen species (Gómez et al. 2002b; Gómez 2005; Suatoni et al. 2006; Mills et al. 2016). In the past, when the existence of the complex was ignored, the name *B. plicatilis* was used for any specimen belonging to the species complex, which introduces ambiguity in assigning specific past-reported features to the biological species in the complex. Then, the term *B. plicatilis sensu stricto* has been used to clarify that the biological species was referred to. In this thesis, the name *B. plicatilis* (i.e. dropping '*sensu stricto*') will term the biological species as re-described in Ciros et al. (2001) and an appropriate disambiguation will be used if needed.

The *B. plicatilis* complex of species is a very diverse and widespread taxon with affinity to saline habitats (Mills et al. 2016). It is the most studied monogonont cryptic species taxon and species from this complex have been used in a wide variety of studies (Ricci et al, 2000; Kostopoulou et al. 2012), such as ecotoxicology (e.g. Snell and Carmona 1995; Snell and Janssen 1995; Snell and Joaquim-Justo 2007; Dahms et al. 2011), aquaculture (e.g. Lubzens et al. 1989; 2001; Hagiwara et al. 1997), osmoregulation (Lowe et al. 2005), ageing (Snell 2014; Snell et al. 2015) and, with special relevance for this thesis, studies on evolutionary and population ecology (e.g.

Gómez et al. 2002a; Ortells et al. 2006; Campillo et al. 2009; Carmona et al. 2009; Alcántara-Rodríguez et al. 2012; Fontaneto et al. 2012; Montero-Pau et al. 2016).

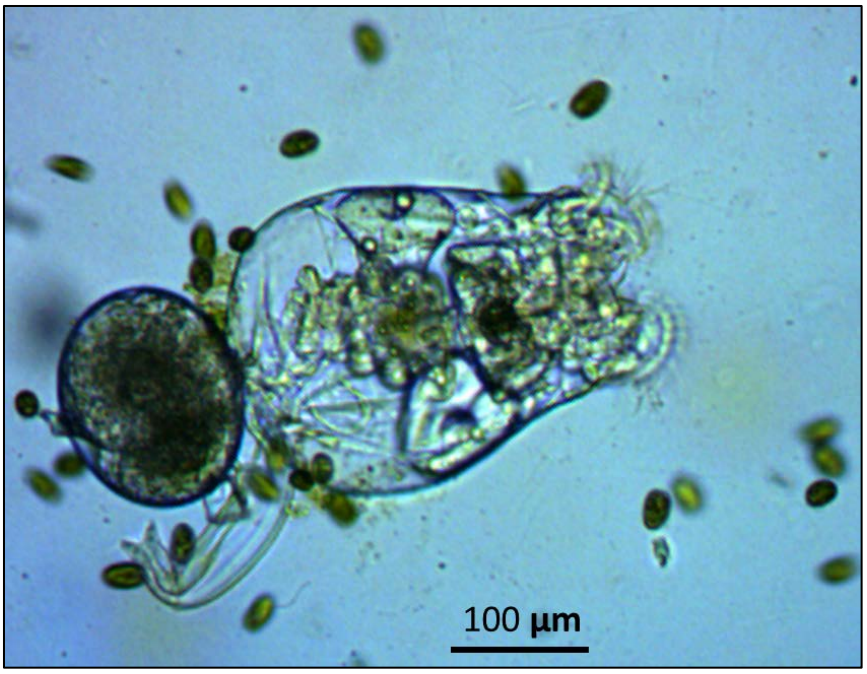


Figure 2.1. Microphotography of a *B. plicatilis* female bearing an asexual egg

The biological cycle of *B. plicatilis* is the typical one of monogonont rotifers belonging to genus *Brachionus* (Figure 2.2) (e.g. Gilbert 1974; Wallace and Smith 2009; Wallace et al. 2015). Proliferation by ameiotic parthenogenesis (asexual phase) combines with occasional bouts of male production and sexual reproduction; the latter resulting in diapausing egg production (sexual phase). Populations of *B. plicatilis* in

temperate regions are temporal; i.e. they are not active all year-round, and they colonize the water column during the so-called (planktonic) growing season. Typically, each growing season, the active population is initiated by the hatching of diapausing eggs from the pond sediment (Pourriot and Snell 1983) (Figure 2.2). Hatchlings are asexual females that produce subitaneous (i.e. non-dormant) eggs, which hatch into genetically identical daughters via parthenogenesis, thus producing clones. Sexual reproduction is induced by a chemical signal produced by the rotifers that accumulates in the medium as population density increases (Snell and Boyer 1988; Carmona et al. 1993; 1995; Stelzer and Snell 2003; Snell et al. 2006). Once a population density threshold is reached (Carmona et al. 2011), asexual females are triggered to produce sexual daughters as some fraction of their offspring (e.g. Gilbert 1974; Pourriot and Snell 1983; Schröder 2005). Thus, sexual and asexual reproduction, typically overlap. Sexual females produce meiotic haploid eggs that develop into haploid males, if they remain unfertilized, and into diploid diapausing eggs, if they are fertilized. Males are dwarf (100 μm length compared to 300 μm of females, Ciros-Pérez et al. 2001), do not feed and have shorter lifespan than females (Wallace et al. 2006).

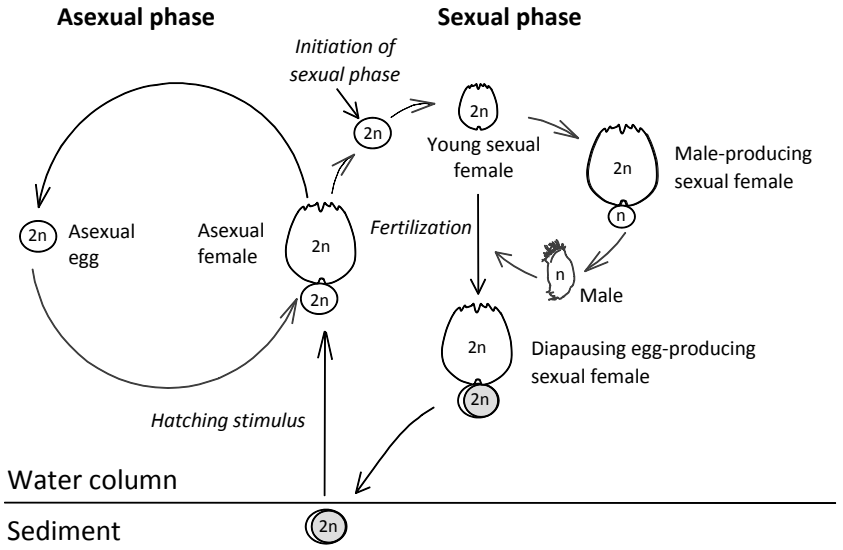


Figure 2.2. Life cycle of *B. plicatilis* (modified from Tortajada 2016).

Unlike asexual eggs, the sexually produced eggs are dormant embryos; therefore, sex is associated with diapause. Diapausing eggs settle in the sediment and remain dormant for a period of variable duration. Under suitable conditions, diapausing egg hatching is induced, and a new growing season starts, although a —relatively short– refractory period for hatching induction has been described in some strains (Hagiwara and Hino 1989; Martínez-Ruiz and García-Roger 2014). Additionally, not all diapausing eggs hatch in the season following their production (Martínez-Ruiz and García-Roger 2014; Tarazona et al. 2017). The unhatched eggs often show prolonged diapause and accumulate in the sediment forming diapausing egg banks where they can remain viable

for decades or even centuries (Marcus et al. 1994; Kotani et al. 2001; García-Roger et al. 2006). This life cycle is considered an adaptation to temporary habitats (Serra and King 1999).

The study region. Water bodies in inland eastern Spain

Due to its climatic, orographical, geological and hydrological conditions, the Iberian Peninsula holds a rich and diverse collection of inland and coastal, saline shallow water bodies (Alonso 1998). The environmental conditions of many of those ponds in Spain have been also explored in several studies, existing a deep knowledge of co-occurring phyto- and zooplankton species (Carrillo et al. 1987; García et al. 1997; Antón-Pardo and Armengol-Díaz 2010), hydrophilic plants (Cirujano 1981; 1988), nutrient dynamics (García-Ferrer et al. 2003; De Vicente et al. 2006) or geology (De la Peña and Marfil 1986). Interestingly, the Iberian Peninsula was one of the main European glacial refugia throughout the Pleistocene ice ages (Gómez et al. 2000; Gómez et al. 2007).

The study ponds are located in an area of approximately 800 km^2 (Figure 2.3), and the region is endorheic territory; that is, with no defined hydrographic network, so that water tends to accumulate in the common terrain depressions. As a result of a long-term erosion of the surrounding terrains by water, salinity increased in the ponds and can now reach values as

high as 40 g L^{-1} (Cirujano et al. 1988; de la Peña and Marfil 1986). Most of these ponds and their surroundings have been declared Site of Community Importance in the Natura 2000 network (ES4210004). Twenty of the ponds within this area are studied in Chapter 3 (Table 2.1). *B. plicatilis* was found at a subset of nine of these ponds. The corresponding populations are investigated in Chapters 4 and 5. The above-mentioned twenty ponds are shallow (maximum depth around 1 m), variable in their mean area (0.00013-1.19 km^2) and brackish or saline (salinity ranging 6.5-40 g L^{-1}). Precipitation is the main water inflow, but some ponds are also connected to groundwater (Gómez-Alday et al. 2014).

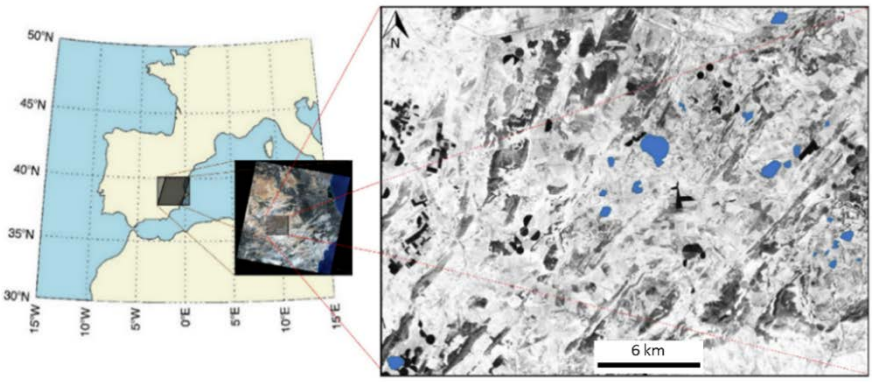


Figure 2.3. Location of the study region (38°55.4' to 38°41.803'N and 1°47.32' to 1°24.26'W; left) and Landsat 5 scene showing the area where the ponds (highlighted in blue) included in this study are located (right).

Table 2.1. Ponds considered in Chapter 3. Pond location coordinates according to Datum WGS84.

Pond name	Acronym	Pond location
Pétrola*	PET	38°50'16.82"N, 1°33'49.22"W
Salobralejo*	SAL	38°54'52.11"N, 1°28'6.95"W
Ontalafia	ONT	38°43'21.23"N, 1°46'3.91"W
Hoya Grande	HYG	38°49'35.17"N, 1°28'31.17"W
El Saladar	SLD	38°47'21.72"N, 1°25'8.00"W
Atalaya de los Ojicos*	ATA	38°46'20.97"N, 1°25'49.12"W
Horna	HOR	38°50'0.77"N, 1°36'3.87"W
Hoya Rasa*	HYR	38°47'6.06"N, 1°25'37.56"W
Casa Villora	CVI	38°48'11.47"N, 1°36'18.10"W
Hoya Redonda	HRE	38°49'5.88"N, 1°34'49.96"W
Hoya del Norte	HYN	38°50'17.10"N, 1°27'23.08"W
Hoya Chica*	HYC	38°49'46.22"N, 1°27'49.74"W
La Campana*	CAM	38°51'29.06"N, 1°29'36.97"W
Mojón Blanco	BLA	38°47'49.95"N, 1°25'55.47"W
Hoya del Monte*	HMT	38°50'44.87"N, 1°26'38.70"W
Casa Villora2	CVI2	38°49'1.33"N, 1°36'37.03"W
Hoya de las Ánades	HYA	38°51'44.84"N, 1°32'38.59"W
Hoya Yerba*	HYB	38°46'46.02"N, 1°26'6.60"W
Hoya Elvira	HYE	38°46'42.13"N, 1°26'43.36"W
Hoya Turnera*	HTU	38°46'31.19"N, 1°24'37.41"W

* Ponds where *B. plicatilis* populations were observed (studied in Chapter 4 and 5)

The climate in the study region is semiarid, with a mean average annual rainfall of ca. 343 mm and a mean temperature of approx. 14 ºC (from data daily collected at the AB07 meteorological station in Pozo Cañada, Ministry of Agriculture, Food and Environment of Spain). Dry periods occur between June and September, when temperatures may exceed 40 ºC and rainfall is scarce (typically < 8 mm in July). Most precipitation occurs as heavy rains in Spring (April-May) and Autumn (October-December; Gómez-Alday et al. 2014).

Episodic, in situ observations during the last decade showed that some ponds can frequently dry out (Cirujano 1988; López Donate et al. 2004), then developing a thick salt crust, and fill up again (Figure 2.4). Others, instead, are known to be flooded for years through several dry seasons. This suggests that the ponds present different hydrological patterns, although an accurate quantitative characterization of their hydrological features was not available before this thesis.

Figure 2.4. Aerial images of PET (above) and CAM (below) ponds, illustrating the temporal variation in water-surface area (green-black) and salt crust (white).

Studies on rotifer population biology in eastern Spain

The best known system for rotifer population biology is eastern Spain —including the region studied here and others–. The studies performed in this part of Spain contributed to unmask the existence of cryptic species in the *B. plicatilis* complex and

their phylogenetic relationships (Gómez and Serra 1995; Gómez et al. 1995; Gómez and Snell 1996; Serra et al. 1998; Ortells et al. 2000; Ciros-Pérez et al. 2001; Fontaneto et al. 2007). Moreover, the intensive sampling campaigns performed on lakes and ponds in eastern Spain (Ortells 2002; Lapesa 2004; García-Roger 2006; Campillo et al. 2009) resulted in information on the geographical distribution and phylogeography of the *B. plicatilis* species complex, showing that there are several clades in the Iberian Peninsula according to mitochondrial sequences (Gómez et al. 2000; Campillo et al. 2011).

Six species of the complex have been found inhabiting inland salt lakes and coastal lagoons of the region (Gómez et al. 2002b, 2007) and subsets of these species commonly co-occur in these habitats (Ortells et al. 2003; Gómez et al. 2005; Lapesa et al. 2004; Montero-Pau et al. 2011). This information has been used to unveil their ecological requirements (e.g. salinity, temperature), to describe their seasonal specialization, and to characterize their abiotic and biotic niche differentiation (Gómez et al. 1997; Ciros-Pérez et al. 2001; Ortells et al. 2003; Gómez et al. 2007; Montero-Pau et al. 2011; Gabaldón et al. 2013; Gabaldón et al. 2016). Temperature and salinity play a role in niche differentiation of some species (Gómez et al. 1997; Montero-Pau et al. 2011; Gabaldón et al. 2015a) and a differential competitive ability

and vulnerability to predation has been reported in species differing in body size (Ciros-Pérez et al. 2001; 2004). But, in species of the complex that are extremely similar in size, the coexistence is favoured by both differences in the response to fluctuating abiotic conditions and in life-history traits related to diapause (Montero-Pau et al. 2011; Gabaldón et al. 2015b).

Interestingly, *B. plicatilis* populations show strong genetic differentiation in eastern Spain (Gómez et al. 2002a; Campillo et al. 2009; Montero-Pau et al. 2016) that has been associated with signatures of local adaptation (Campillo 2010). This interpopulation differentiation is found even among closely located populations and despite their potential high passive dispersal via diapausing eggs (Frish et al. 2007), which suggest a low level of effective gene flow among them (Gómez et al. 2007). The discordance between migration and gene flow, has been mainly explained on the basis of long-lasting founder effects (Gómez and Carvalho 2000; Gómez et al. 2000; Campillo 2010; Montero-Pau 2012) according to what has been hypothesized in other zooplankters (Boileau et al. 1992; De Meester 2000). However, rotifer population differentiation can be promoted or prevented by the effects of outcrossing on fitness. These rotifer populations have been shown to be affected by inbreeding and outbreeding in experimental crosses (Tortajada et al. 2009; 2010); whereas outbreeding depression seems not to constitute a barrier to

gene flow, inbreeding depression in small populations recently founded could give new migrants a fitness advantage and favor gene flow into the populations (Tortajada 2016).

Rotifer populations inhabiting water bodies in eastern Spain have also been found to harbour high degree of intrapopulation genetic variation (Gómez et al. 2002a; Campillo et al. 2009; Carmona et al. 2009; Gabaldón and Carmona 2015), which shows the potential of these populations to evolve. As a result, short-term micro-evolutionary adaptation to local selection conditions can be expected in these rotifers (Declerk and Papakostas 2016).

All these accumulated information on population biology makes the *B. plicatilis* populations in eastern Spain a good model system to study the adaptation to environmental unpredictability.

Remote sensing techniques

Sensors in satellites is one of the several ways of obtaining remote sensing data from the earth surface. This will be showed in Chapter 3, where scenes from the satellites Landsat 5 and 7 are exploited. Hereafter, by default, the explanations will assume these two satellites are the relevant ones. The

sensors on-board of the satellites store the information in scenes that are composed of a matrix of pixels (Figure 2.5). A pixel is the smallest unit of a scene. In a Landsat scene, the pixel is square and integrates the information from a georeferenced earth-surface area. The side length of a pixel is known as spatial resolution. In Landsat 5 and 7 this resolution ranges between 15 and 120 m, depending on the spectral band (see below). Apart from the spatial resolution, another important factor is the time lasted between the acquisitions of two scenes at the same area (i.e. the temporal resolution), what in the satellites Landsat 5 and 7 is 16 days.

When acquiring a scene, Landsat 5 or 7 satellites record the reflected radiance from the Earth surface in various wavelength regions of the electromagnetic spectrum (EMS). Each region of the electromagnetic spectrum that a sensor captures is a spectral band and, in Landsat 5 and 7 satellites, there are 7 and 8 spectral bands respectively (see Table 2.2).

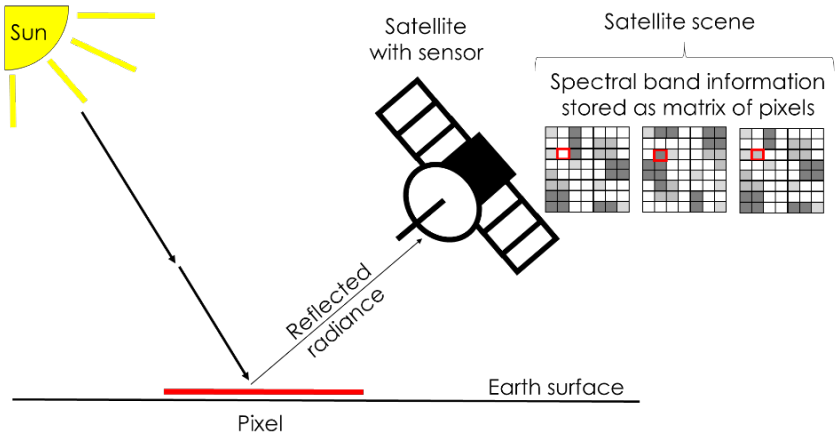


Figure 2.5. Simplified diagram showing acquisition of satellite remote sensing data. The three squared matrices on the upper-right corner are composed of 8 x 7 pixels, each matrix corresponding to a spectral band and all together composing a satellite scene.

The data gathered by the satellite has to be processed in order to obtain the proportion of light reflected from each spectral band (i.e. reflectivity) in every pixel. This information can be used to differentiate different types of surfaces in the Earth, if the different materials produce different proportion of light reflected in one or more regions of the EMS (i.e. their spectral signature). Thereby, single or multiple spectral band combinations can be used in order to determine the type of materials in the Earth that the pixel represents (e.g. vegetation, soil, water, salt). This and the long-term time series gathered by satellites, makes this information very useful in order to characterize the variation of the relevant features of natural habitats.

Table 2.2. Spectral band wavelength and spatial resolution of the satellites Landsat 5 and 7.

Band number	Landsat 5		Landsat 7	
	Wavelength (µm)	Spatial resolution (m)	Wavelength (µm)	Spatial resolution (m)
1	0.45-0.52	30	0.45-0.51	30
2	0.52-0.60	30	0.52-0.60	30
3	0.63-0.69	30	0.63-0.69	30
4	0.76-0.90	30	0.77-0.90	30
5	1.55-1.75	30	1.55-1.75	30
6	10.41-12.5	120	10.4-12.5	60
7	2.08-2.35	30	2.08-2.35	30
8	-	-	0.52-0.9	15

Molecular techniques

In Chapter 5 two major molecular techniques are used: (1) genome sequencing and (2) genotyping by sequencing (GBS). Box 2.1 defines a set of general terms associated to the molecular techniques employed in Chapter 5. It follows an outline of the general procedures used in this thesis, and —for the sake of providing a global synopsis– it ignores relevant details that are given in the Chapter 5.

Box 2.1. Glossary

Adapters: specific oligonucleotides added to the *problem DNA fragments* that match other oligonucleotides and are necessary, in most *sequencing platforms*, for *DNA sequencing*.

Barcode: in the context of *NGS*, a known, specified sequence of nucleotides added as a tag at the beginning of a *problem DNA fragment*, the tag being used to identify the fragment of origin when performing multiplex sequencing.

Contig: a contiguous genomic sequence obtained from the *problem DNA* in which the order and identity of all bases is known to a high confidence level.

DNA sequencer or sequencing platform: Instrument that allows sequencing the *problem DNA fragment*.

DNA sequencing: process that allows determining the order of the bases in a *problem DNA fragment*.

Functional annotation (process): *to* associate biological information to the genomic elements structurally annotated; e.g., their biochemical function, biological function, if it is involved regulation and interactions, their expression, etc.

Gap: a genomic sequence with an unknown identity of their bases but in which a rough estimation of its length exists.

Gene ontology terms: descriptions of the known features of a gene according to its expression product classified in three domains of terms: (1) cellular component; (2) associated molecular function; and (3) associated biological process.

Genome assembly (process): to reconstruct an original DNA sequence for the *problem DNA* by aligning and merging *reads* obtained from the *DNA sequencer*.

***Genomic DNA library*:** in the context of this thesis, collection of genomic *problem DNA fragments* to which specified oligonucleotides (i.e. *barcodes* and/or *adapters*) are added by PCR to allow *DNA sequencing* and information recovery.

Multiplex sequencing: *to* process multiple *problem DNA fragments* at the same time in order to get their sequences.

Next generation sequencing (NGS): type of *DNA sequencing* based on modern sequencing techniques (i.e. non-Sanger) yielding millions or billions of small *reads* from the *problem DNA fragments*.

Paired end sequencing (PE): a type of *DNA sequencing* which starts on both ends of the *problem DNA fragments*.

Problem DNA fragment: fragment of DNA that is sequenced in which the order of its bases is unknown.

Read: Each of the DNA sequences obtained from the *DNA sequencer*.

***Scaffold*:** a genomic sequence composed of contigs and gaps.

Single end sequencing (SE): a type of DNA sequencing which starts from only one end (i.e. extremity) of the problem DNA fragments.

Single nucleotide polymorphism (SNP): variation in a single nucleotide locus between members of a species or paired chromosomes in an individual.

Structural annotation (process): to identify the physical regions of a genome containing a genomic element of interest such as genes, mRNA, transcript, repeat sequences, etc.

Tag: consensus sequence obtained from *reads* putatively representing the same genome locations, and that are collapsed during GBS data analysis.

Genome sequencing

Whole genome sequencing is the process of determining the DNA sequence of all the chromosomes of the genome of an organism, as well as of the DNA in mitochondria and chloroplasts, where applicable.

The first step in genome sequencing (Figure 2.6) is the extraction of high quality genomic DNA from the focus species. In the case of rotifer *B. plicatilis*, it involved to raise eight cultures of 1.5 L for 7-10 days at 25 ºC (see Chapter 5 for details) in order to get sufficient biological material (at least 10,000 individuals in each bottle) and to obtain enough DNA to sequence the genome. Once the DNA from a sample is obtained, DNA quality and quantity has to be assessed. Next step is DNA shearing (e.g. through sonication) down to ca. 550 bp. This fragmentation is not sequence-dependent. In the process of library preparation, Illumina adapters matching the oligonucleotides in the sequencing plate are added by PCR. Then a single run is performed in a NGS platform, yielding pair-end reads (Box 2.1). Finally, the raw sequence data obtained has to be analysed by using several bioinformatics analysis steps including assembly and structural and functional annotation (see Box 2.1).

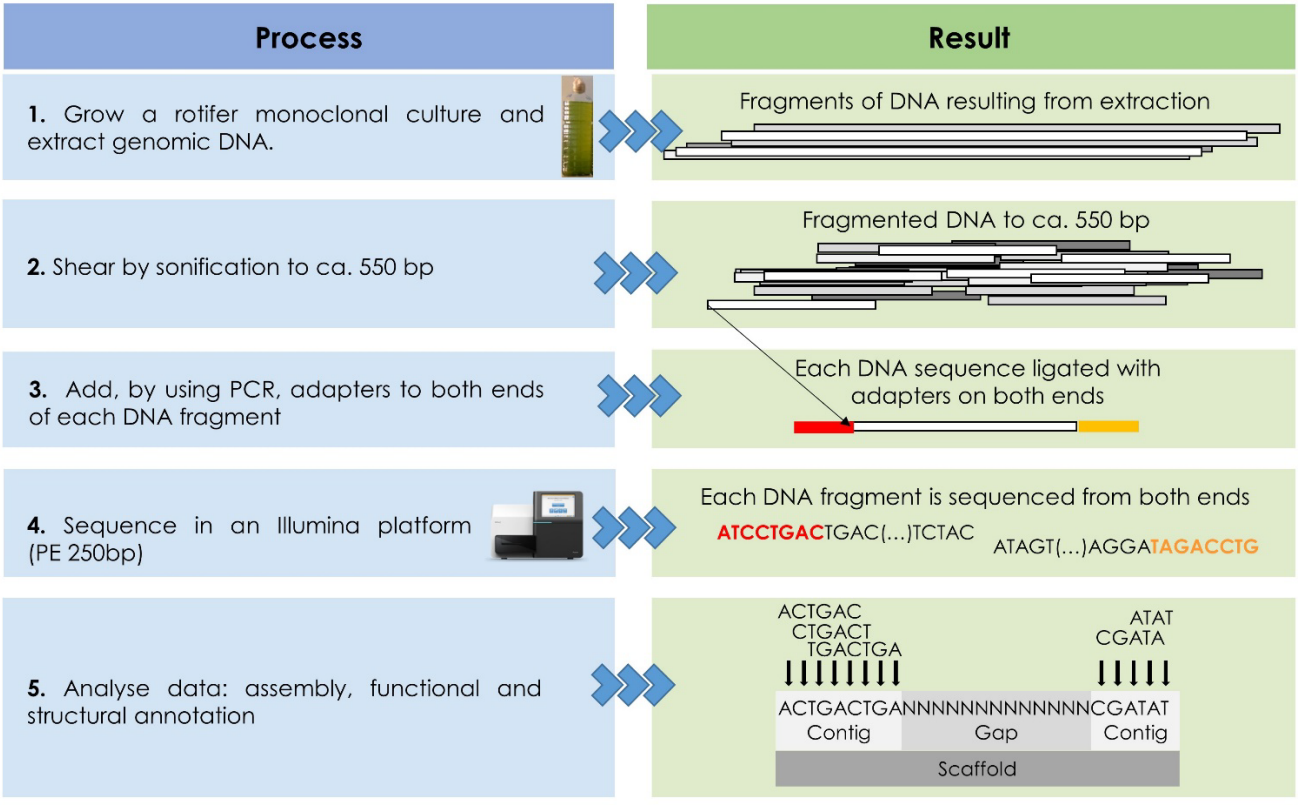


Figure 2.6. Main steps in genome sequencing (see the text in this chapter and Chapter 5 for details).

Genotyping by sequencing (GBS)

One of the approaches recently developed by using NGS is genotyping by sequencing (GBS). GBS allows to genotype a large number of single nucleotide polymorphisms (SNPs) in the genome of multiple individuals at a reduced cost. This data can be used to characterize at a genome scale the genetic structure of natural populations and to identify genomic regions putatively under selection.

For GBS, in contrast to genome sequencing, the genomic DNA extraction is performed on each of the samples separately (e.g. 270 samples corresponding to 9 populations x 30 clones of *B. plicatilis*, as studied in Chapter 4 and 5; Figure 2.7). Once the DNA from all samples is obtained, DNA quality and quantity has to be assessed. Then, the DNA libraries have to be constructed and sequenced. For doing this, the genomic DNA from each sample is digested by a restriction enzyme (e.g. ApeK1; Chapter 5). Note that this fragmentation is sequence-dependent, and only the fragment of DNA preceded by a cutting sequence of the restriction enzyme is sequenced. Once DNA is fragmented, the Illumina adapters and a specific barcode for each sample have to be added to each DNA fragment by PCR. This barcode in GBS is a sequence between 4 and 8 bp and enables to identify the original sample from which the DNA information is obtained. This barcode is essential for samples to be distinguished and

sorted during data analysis, since the DNA from the different samples are pooled to be amplified and sequenced together. Afterwards, DNA is amplified and sequenced on a NGS platform yielding single-end reads. Once the DNA is sequenced, the raw sequence data so obtained have to be analysed. This analysis starts with removing the effect of multiplexing (i.e. demultiplexing). Then, the reads (Box 2.1) are assigned to each of the samples accordingly to their barcode. Finally, further bioinformatics analysis are performed in order to discover single nucleotide polymorphisms (SNPs) (Chapter 5).

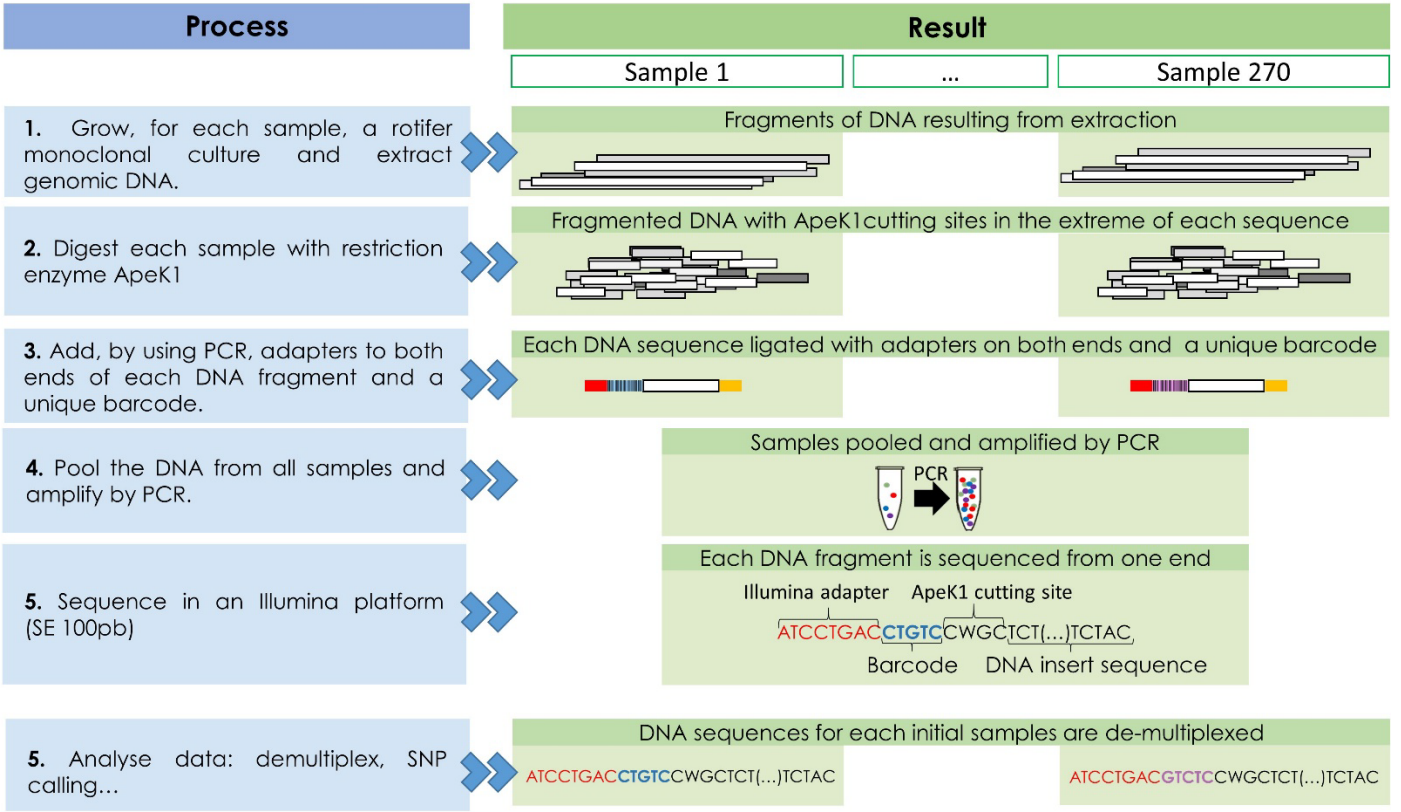


Figure 2.7. Main steps in GBS of 270 *B. plicatilis* clones (see the text in this chapter and Chapter 5 for details).

3

Quantifying unpredictability: a multiple model approach based on satellite imagery data from Mediterranean ponds

Summary

*Fluctuations in environmental parameters are increasingly being recognized as essential features of any habitat. Quantifying whether environmental fluctuations are prevalently predictable or unpredictable is remarkably relevant to understanding the evolutionary responses of organisms. However, when characterizing natural habitat's features, ecologists typically face two problems: (1) gathering long-term data and (2) handling that hard-won data. The research described in this chapter takes advantage of the free access to long-term recording of remote sensing data (27 years, Landsat TM/ETM+) in order to assess a set of environmental models for estimation of environmental predictability. The case study included the 20 Mediterranean saline ponds and lakes presented in Chapter 2. The focus variable was the water-surface area (*A*). Saline ponds can develop salt-crusted areas that make it difficult to distinguish between soil and water. A novel pipeline that combines water indices and the short infrared band, as salt filter, addressed this challenge. The predictable and unpredictable components of variation in* A *were extracted using two different approaches. The*

first approach, based on Colwell's predictability metrics, transforms the focus variable into a nominal one. This discretization is a key step, and here different models were considered based on how A's variation could be relevant from the point of view of the organisms of interest. As a second approach, general additive model fitting is introduced as a new metric for quantifying predictability. Both approaches extracted meaningful information about the degree of predictability of the studied ponds. Interestingly, some model assumptions had negligible effects, while some others could be associated to ecological features of species whose predictability needs to be assessed. The methodology described here is applicable to a wide variety of systems and will be valuable for quantifying and characterizing predictability, which is essential within the expected global increase in the unpredictability of environmental fluctuations.

Introduction

Fluctuations in environmental parameters and their potentially associated unpredictability are increasingly being recognized as essential features of any habitat (Williams and Hastings 2011) because, as introduced in Chapter 1, they are expected to influence the performance of the inhabiting organisms and affect upper levels of ecological organization (Simons 2011). Human activity often increases environmental fluctuations, thus their analysis provides an applied interest (Pimm et al. 1995; Root et al. 2003; Simons 2011; IPCC 2013). Indeed, ascertaining how organisms respond to environmental fluctuations is fundamental in biology (Chevin et al. 2010; Chown et al. 2010).

From the point of view of an organism, predictability involves a time scale (Chapter 1) and is related to organism's ability to anticipate and adjust to a future environmental condition (MacArthur and Levins 1964; Levins 1968). Any habitat shows some constancy and some variation in their features, and consistently the focus habitat feature can be decomposed in three components: constancy, predictable —periodic– fluctuations, and unpredictable fluctuations. The relative importance of these components is expected to produce diverging adaptive responses in organisms (Crozier et al. 2008). Thus, to quantify whether environmental fluctuations are prevalently predictable or unpredictable is highly relevant

to understanding evolutionary responses and to test ecological and evolutionary hypotheses (Simons 2011).

The characterization of fluctuations is a complex problem needing a robust methodology, since it is based on variances and needs assumptions regarding the relevant time-scales that can be associated with predictability (see Chapter 1). Last, one of the most important constraints is the need of long time-series data. Thus, (1) a methodology is required for the acquisition of sufficiently long time series of data, and then, (2) an appropriate metric has to be assumed and assessed, so that the habitats of interest can be distinguished.

The indices developed to estimate the degree of predictability from time-series data with stationary distribution decompose the time series by periodic and stochastic variation (Sabo and Post 2008), associating them with predictable and unpredictable fluctuations, respectively. For nominal data such as the presence/absence of water, Colwell (1974) proposed a predictability index based on information theory (Stearns 1981) that has been used especially in streams and rivers (Gallart et al. 2012). For continuous data, indices based on spectral analysis, like the Fourier transform and asymmetric eigenvector maps (Blanchet et al. 2008; Legendre and Gauthier 2014) have been widely used (Grossman and Sabo 2010). However, these

continuous-metric methods are very sensitive to gaps in the time series, so Colwell's method might be preferable despite the drawbacks of discretizing a quantitative variable.

The purpose of this chapter is using the time series from remote sensing data to assess a set of models for environmental predictability estimation. Taking advantage of the free access to long-term, high-frequency data recording, the group of Mediterranean shallow ponds and lakes presented in Chapter 2 are considered as the case study. It is worth noting that the Mediterranean basin is considered a priority in the 'silver bullet' conservation strategy for being one of the 25 biodiversity hotspots in the world (Myers et al. 2000). Biological communities confined in non-permanent ponds are expected to be adapted and strongly reliant on patterns of pond inundation. Moreover, migratory flyways and nesting places of many waterfowls depend on the dynamics of inundation, particularly if birds are dependent on the ponds' food webs (e.g. Krapu 1974). Migratory patterns of waterfowls, lifecycles of many short-lived animals and rainy patterns occur at an annual time scale or shorter. Accordingly, ecological unpredictability in these systems can be conceived as departures from an average seasonal variation, i.e. the main source of environmental unpredictability is the inter-annual variation of the within-year variation. Additionally, the water-surface area (hereafter *A*; Wetzel and Likens 2013) is

considered to be an ecologically relevant factor in lentic water bodies, expected to positively correlate with species diversity according to the predictions of island biogeography theory (MacArthur and Wilson 1967). An increase of the pond's *A* and its correlated increase in average pond depth will result in an increase of the habitat size for aquatic organisms, thus affecting the 'colonization *vs.* extinction' balance and the habitat heterogeneity with positive effects on species persistence, complexity of the community and species diversity (Reche et al. 2005). Fluctuations in *A* have multiple effects on the organisms above and below the waterline, for aquatic flora and fauna —e.g. macrophytes (Turner et al. 2005), fish (Cott et al. 2008) and zooplankton (Mageed and Heikal 2006)–. These fluctuations can strongly affect habitat conditions through variations in the physicochemical parameters —e.g. temperature, light, nutrient or solute concentration (Wurtsbaugh 1992; White et al. 2008)–. Solute concentration is particularly relevant in saline ponds and lakes, which are a significant, geographically widespread part of the world's inland aquatic ecosystems (Williams 1981; Hammer 1986). Remote sensing studies in saline water bodies are mostly based on large lakes (Bryant and Rainey 2002; Scuderi et al. 2010; Adams and Sada 2014). Band ratios are typically used for the identification of water bodies (e.g. Xu 2006; Chao Rodríguez et al. 2014), but are rarely used in saline ponds. In contrast, infrared bands have been

previously used for *A* assessment in saline ponds but not accompanied by band ratio indices (Ormeci and Ekercin 2007; Groeneveld and Barz 2014).

This chapter is divided in two parts: (1) the estimation of *A* and assessment of its accuracy; and (2) the quantification of environmental predictability. In (1) 27 years of Landsat 5/7 scenes are used to quantify saline ponds' *A*. The salt crust formed in some of these ponds when water evaporates can make it difficult to distinguish between soil and water. This difficulty was addressed by means of a novel approach that combines the sequential use of water indices and the short infrared band (salt filter). In (2) the predictable and unpredictable components of the variation in *A* were extracted using two different approaches. First, Colwell's methodology was applied, taking into account different models based on how the environmental variation could be relevant from the point of view of the focus organism (see Chapter 1). Second, a novel approach that parallels Colwell's but that is based on a regression model was developed. The similarity and divergence from the different predictability indices were analysed in order to determine their sensitivity to the assumptions needed to quantify predictability. It is proposed which assumptions are more reliable depending on the type of organism. The novel approach developed here obtains consistent results with most of the models based on

Colwell's approach, having similar discriminating power among the studied ponds with minimal data manipulation.

Materials and methods

Satellite scene data

Path 199 and row 33 of Landsat Thematic Mapper (Landsat 5) and Enhanced Thematic Mapper Plus (Landsat 7) scenes were downloaded. The total number of scenes was 432. From these scenes, 314 were obtained from the European Space Agency (ESA) as an L1T product, which is radiometrically corrected and includes geometric correction (Northdrop 2014). The remaining 118 scenes were obtained from the United States Geological Survey (USGS) and were generated as a Climate Data Record (CDR) product from the software Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) (Masek et al. 2013). Landsat CDR is radiometrically, geometrically and atmospherically corrected, thus providing a surface reflectance product besides an accurate cloud mask developed by Zhu and Woodcock (2012). Scenes from ESA were processed with LEDAPS using atmospheric correction routines developed for the Terra MODIS instrument (Vermote et al. 1997) by homogenizing these scenes with respect to those from USGS. The resulting set of scenes corresponds to the period of 1984-2011, with a spatial

resolution of 30 m and a revisit time of 16 days for each satellite (Frazier and Page 2000).

Estimation of the water-surface area (A)

Water presence in a pixel was a two-condition assessment (2cA). The first condition of 2cA was addressed to differentiate potentially (see below) water-covered areas from soil. Modified normalized difference water index (MNDWI Xu 2006) was calculated using the reflectance from Landsat's band 2 (green, 0.52-0.60 µm) and Landsat's band 5 (middle infrared band, MIR, 1.55-1.75 µm) as follows:

$$\text{MNDWI} = \frac{\text{Green-MIR}}{\text{Green+MIR}}, \quad (1)$$

Following Xu (2006), the threshold value used was zero, and the pixels with positive MNDWI values were selected as potential water-covered areas.

Those potentially water-covered pixels were evaluated for a second condition. The salt crust left by water evaporation produced the appearance of pixels that were false positives for water-covered areas. To exclude them, a second filter (salt filter) was included using the approach by Ormeci and Ekercin (2007), based on the condition that band 4 (near infrared) reflectances lower than 0.4 are classified as non-salt. However, while these authors applied the condition directly to

land pixels, here it was applied after the first condition was assessed (Figure 3.1). These automatically processed results were compared with visual interpretations of the satellite scenes, aerial photographs (years 2006 and 2009; four bands; spatial resolution: 0.25-0.5 m) and qualitative field observations (presence/absence of water). Table A.1 contains information for the two latter sources of validation. Satellite data analyses were performed using ENVI/IDL (Exelis Visual Information Solutions, Boulder, Colorado).

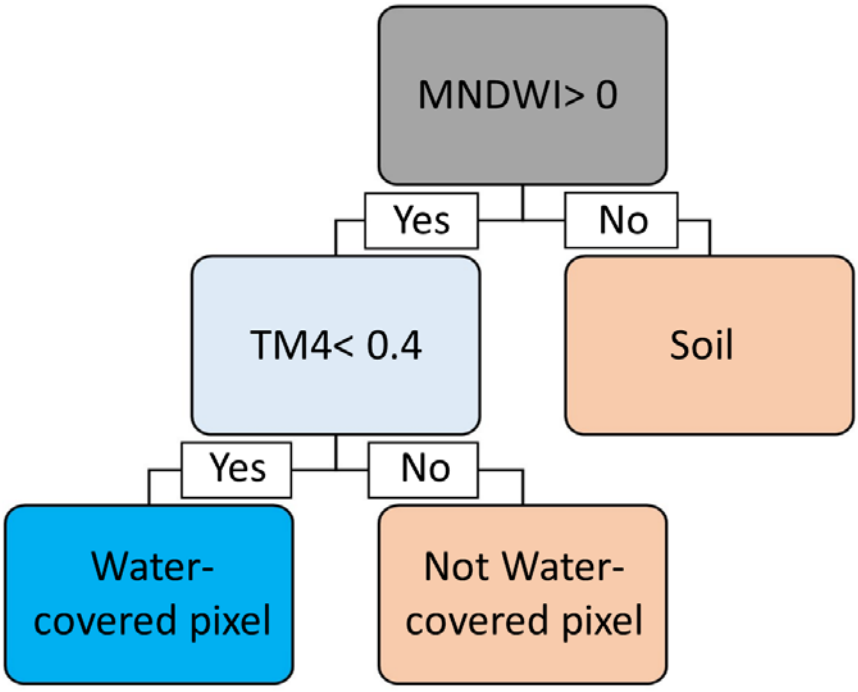


Figure 3.1. Water presence in a pixel: a two-condition assessment (2cA). First condition differentiates potentially water covered pixels from soil. Second condition differentiates salt-covered from water-covered pixels. MNDWI: modified normalized difference water index; TM4: Landsat Thematic Mapper 4.

Quantification of environmental predictability

Using each pond as an individual case, the degree of predictability of the variation in *A* was estimated. Seven data analysis models with different assumptions were evaluated. Note that here the term 'model' means a different way to implement quantitatively *a priori* concepts (i.e. assumptions) in order to calculate the predictability indices. Five models were developed in the present thesis according to Colwell's approach, thus classifying observations (*A*) into discrete categories. Two additional models used a continuous approach, and therefore raw *A* observations were retained as the focus variable; these models are novel contributions of this thesis. The seven models were applied to the within-year variation, and different years in the time series were treated as replicates. Mean water-surface area ($\overline{A}$) for each pond was computed by averaging each month's mean *A* (excluding cloud-covered observations). Hydroperiod was estimated as the annual average of the fraction of observations with $A > 0$ in each month (excluding cloud-covered observations).

Discrete models based on Colwell's approach

Colwell's predictability (*P*) index (Colwell 1974; Stearns 1981) is composed by the summation of two metrics: constancy (which measures the degree in which a pond remains in the same state) and contingency (which measures the

Table 3.1. Models for predictability estimation. Statistic: sample parameter used to define the states. Scaling: proportion between consecutive ranges defining states; s_i: i-th state (discrete); raw data: water-surface area (A).

Model	**Features**			
Type / ***Acronym***	**Data excluded**	**Statistic (Scaling)**	**Range**	**States definition**
Discrete				
COL_wd[a]	None	-	(0, 1)	s_1, if $A = 0$; s_2, if $A > 0$.
COL_ANa[b]	None	Mean	(0, 1)	s_1, if $A < (1\text{-}0.7) \cdot \overline{A}$; s_2, if $(1\text{-}0.7) \cdot \overline{A} \leq A \leq (1\text{+}0.7) \cdot \overline{A}$; s_3, if $(1\text{+}0.7) \cdot \overline{A} < A$.
COL_ANw[c]	Dried pond	Mean	(0, 1)	As in COL_ANa
COL_MAXlin[d]	None	Maximum (Linear)	(0, 1)	s_1, if $A < 1/3 \cdot MAX(A)$; s_2, if $1/3 \cdot MAX(A) \leq A \leq 2/3 \cdot MAX(A)$; s_3, if $2/3 \cdot MAX(A) < A$.
COL_MAXg[e]	None	Maximum (Geometric)	(0, 1)	s_1, if $A < 1/4 \cdot MAX(A)$; s_2, if $1/4 \cdot MAX(A) \leq A \leq 2/4 \cdot MAX(A)$; s_3, if $2/4 \cdot MAX(A) < A$.

Table 3.1 (continued)

Model	**Features**			
Type / *Acronym*	**Data excluded**	**Statistic (Scaling)**	**Range**	**States definition**
Continuous				
GAM_a[f]	None	-	(0, ∞)	Raw data
GAM_w[g]	Dry states excluded for mean computation	-	(0, ∞)	Raw data

a. Colwell water/dry model.
b. Colwell average neighbourhood with all data included model.
c. Colwell average neighbourhood with water presence model.
d. Colwell maximum value linear scaling model.
e. Colwell maximum value geometric scaling model.
f. GAM with mean calculated with all data included model.
g. GAM with mean calculated when water is detected model.

repeatability of the variation pattern), the residual being regarded as unpredictability. Colwell's approach applies to nominal variables, so that, if values are for a continuous variable, they are transformed to the so-called 'states'. Therefore *A* (continuous variable) was transformed into a nominal variable. The five discrete models used are described in Table 3.1.

The rationale for the number of states and thresholds was straightforward in some models, such as COL_wd, where two states (presence/absence of water) were considered. The number of states in the remaining discrete models was limited to three in order to (1) avoid excessive proliferation of states and (2) accumulate enough data in each state. The thresholds between states in COL_ANa and COL_ANw were defined under the assumption that most organisms have a rather broad range of tolerance around the mean value of *A*, at which they are expected to be best adapted (being vulnerable only to extreme values), although opportunistic extreme-conditions organisms may exist. In other words, they focus on separating extreme values from regular values. It is implicit that extreme high values and extreme low values have different effects, low values include absence of water, a condition that no organism depending on water could tolerate. Consequently, in these models the thresholds were $\overline{A}$ ± 70% $\overline{A}$, with a fairly wide intermediate range. In COL_MAXlin and COL_MAXg, organisms would have a narrow tolerance

range, confined in one of the proposed partitions of the maximum value. In these models, the thresholds were respectively linear and geometric divisions of the maximum value, assuming in the geometric scaling that variation in *A* around large values has lower effect than the same variation around small values. Additional criteria are explained in Table 3.1.

In Colwell's approach, time is also required to be a nominal variable. Accordingly, in the Colwell models used in this chapter, time steps were matched to calendar months. In those months with more than one *A* observation, the *A* value was obtained by averaging *A* observations with the lowest cloud cover. A number of cases were excluded because, due to cloud covering, its state could not be determined.

Continuous models: GAM predictability

This family of models is based on the dispersion of observations in the time series around a typical intra-annual curve of the normalized *A* (hereon *A'*; normalization: $A/\overline{A}$). Since a continuous approach was used, except when clouds occurred. The typical curve for a pond was fitted to *A'* (dependent variable) in relation to the day of the year. To avoid assuming any *a priori* general, very constrained shape for the curve (e.g. sine function), general additive models (GAM) with cubic splines as a smoothing function were

performed for identifying trends in the time-series data (Hastie and Tibshirani 1990). For this, the gam function ('MGCV' package, Wood 2000) in R statistical software v.4.1.1 (R Development Core Team 2009) was used.

Predictability was estimated from the closeness between the observed and expected values from GAM regression. The determination coefficient of the regression model regards constancy as lack of determination (prediction) from the independent variable (here, time) and therefore was not used. Hence, an index of predictability (P_{GAM}) was developed based on the dispersion of the data with respect to the fitted model:

$$P_{GAM} = \frac{1}{SD_{res}}, \quad (2)$$

where SD_{res} is the standard deviation of the residuals of the fitted model. This predictability index is related to a well-known statistic of regression models: the coefficient of variation reduced by regression, defined as the mean value of the dependent variable divided by SD_{res}. Note that the normalization of the independent variable (A') causes its mean values to approach 1, particularly when the values with $A = 0$ were not excluded.

Relationships among predictability estimations

Pearson's correlation coefficient was computed to evaluate the relationships between predictability estimates from each model using (1) the 20 ponds studied, and (2) 18 ponds, excluding the two most ephemeral ponds (HTU and HYE; more than 90% of time series with $A = 0$) because they were expected to pose problems in continuous metrics of predictability (see Sabo and Post 2008). Finally, the standard unweighted pair group method with arithmetic mean (UPGMA) cluster analysis was computed with the 'Pvclust' package in R (Suzuki and Shimodaira 2006) using a dissimilarity matrix based on the chord distance ($\sqrt{(2-2 \cdot r)}$), with r being Pearson's correlation coefficient (Greenacre and Primicerio 2013) and resampling the data 10,000 times.

Results

Estimation of water-surface area (A): assessment

In order to assess the accuracy of A estimations, they were compared with (1) visual interpretations of the satellite scenes, (2) aerial image inspection, and (3) qualitative field data (presence/absence of water in the pond). Visually estimated water-covered area from raw scenes were inspected and compared to A estimations (i.e. produced by automatic processing after applying 2cA), and both estimations yielded

consistent results. Moreover, a total of 33 estimations of *A* and estimations of water-covered area based on aerial scenes taken close in time (maximum time for matching values corresponding to the two estimations was 18 days) were compared. Both estimations were found to be strongly correlated (R^2= 0.98, *P*-value< 0.001; Figure A.1; Table A.1); and no statistically significant differences were found between them (paired *t*-test; *t*= 1.3, *d.f.*= 31, *P*-value= 0.2).

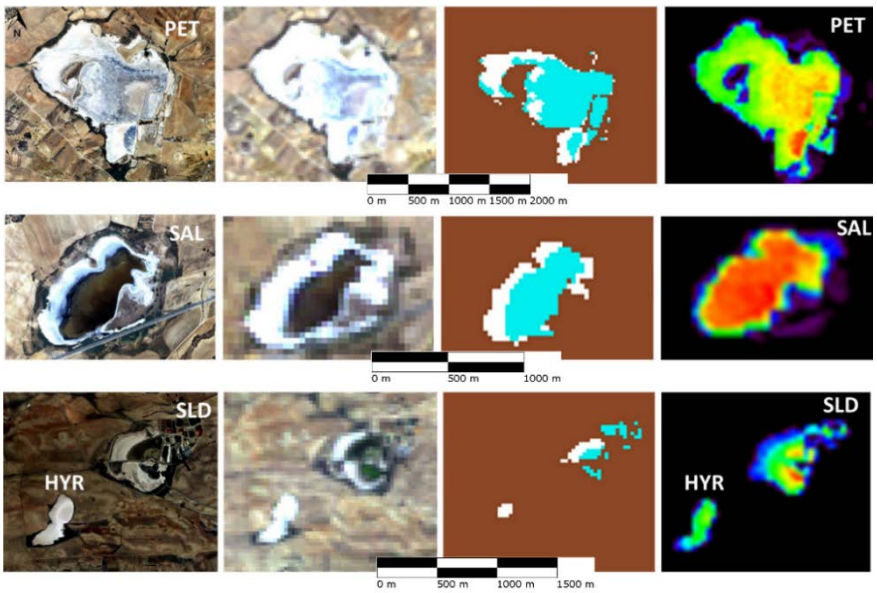


Figure 3.2. Raw and processed data for PET, SAL, HYR and SLD ponds; Chapter 2). First column: aerial scenes. Second column: Landsat 5 satellite raw scenes. Third column: automatically-processed satellite scenes (white: MNDWI> 0, TM4> 0.4; blue: MNDWI> 0, TM4< 0.4). Fourth column: integration over the time series of the *A* as estimated from 2cA applied to raw satellite scenes; grading from blue to red is the proportion of scenes with water covering (low to high proportion respectively).

Figure 3.2 shows instances of this comparison, stressing the role of applying the filter for salt crust involved in 2cA. Finally, *A* estimations were compared with direct qualitative field observations (presence/absence of water in the pond) from Lapesa (2004); García-Roger (2006) and Montero-Pau (personal communication), close in time (within the same month, Table A.1). This comparison showed consistent results of presence/absence of water.

The comparison between the automatically-processed results with and without the salt filter showed that this filter reduces *A* in 7.5% (overall mean). In some of the ponds, there were no pixels excluded by this condition (six ponds), but the reduction can achieve up to 100% of the pixels (Figure A.2). Therefore, a putative confusion between salt and water is not neutral in relation to the pond.

Estimation of water-surface area (A): historical data record

From the 27 years of monitoring, 8,640 data points (20 ponds x 432 raw scenes) were obtained. According to the presence of water inferred with 2cA, the maximum size of the ponds ranged from 5 to 2,513 pixels. Automatic application of 2cA to raw satellite scenes yielded 4,036 cloud-free *A* observations (average per pond= 201; range= 92-257). The range of *A* was 0-4,500 m^2 in the smallest pond and 17,000-2,263,000 m^2 in the largest one. Pearson's correlation coefficient (*r*) between *A* time series of pairwise ponds varied from -0.31 to 0.86. After a

principal component analysis, no consistent pattern was observed between *A* and pond size. Annual hydroperiod ranged from 0.07 to 1.00 (1.0: continuous presence of water; Table 3.2). Three ponds (PET, SAL and SLD) achieved maximum hydroperiod because these ponds never completely dried out in the data record.

Quantification of environmental predictability

Figure 3.3 illustrates how two models differing in the variable type (COL_Ana, discrete; GAM_a, continuous) estimate *A* predictability from the time series for two ponds (SAL and CAM), which were selected because they resulted to have very different predictabilities. Despite the reduction of information caused by the discrete classification or regression model, the indices captured the different fluctuation patterns observed in each pond's time series.

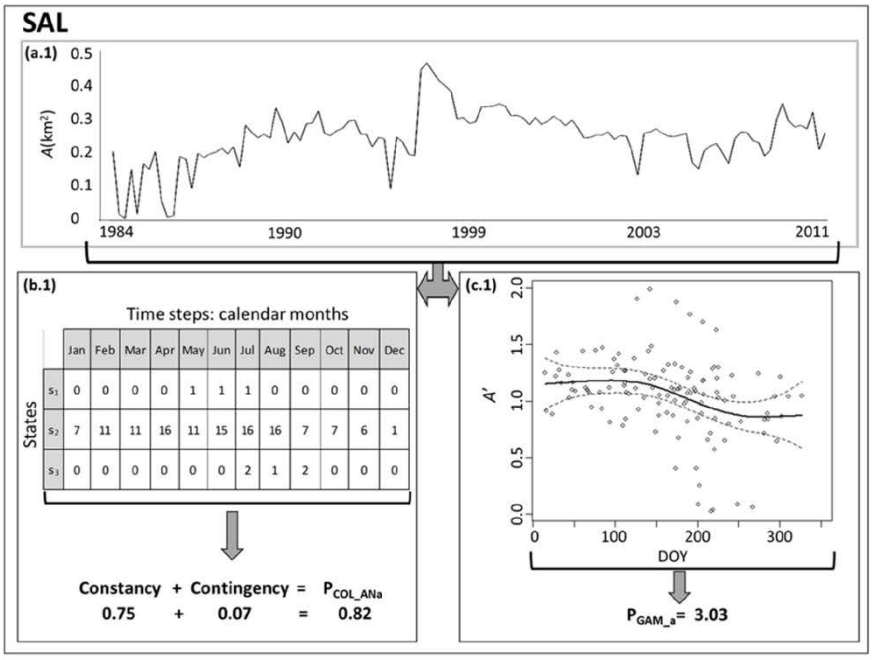
The predictability estimated for each pond and model is shown in Table 3.3. Maximum predictability (i.e. 1) is achieved for three ponds (PET, SAL and SLD) in the COL_wd model because these ponds never completely dried out in the data record and the COL_wd model only considers the presence/absence of water. The coefficient of variation for the predictability of the ponds, which can be associated to the discrimination power among ponds' predictability values, ranged from 0.32 to 0.72 across models.

Table 3.2. Mean water-surface area ($\bar{A}$) and hydroperiod of the ponds studied. Ponds ranked according to $\bar{A}$.

Pond name	$\bar{A}$ (m²)	*S.E.* (m²)	Hydroperiod
Pétrola	1190000	140000	1.00
Salobralejo	237000	18000	1.00
Ontalafia	209000	13000	0.99
Hoya Grande	141000	18000	0.69
El Saladar	111000	8000	1.00
Atalaya de los Ojicos	47000	3000	0.93
Horna	41000	7000	0.53
Hoya Rasa	40000	4000	0.87
Casa Villora	36000	6000	0.48
Hoya Redonda	34000	6000	0.30
Hoya del Norte	32000	6000	0.43
Hoya Chica	32000	4000	0.51
La Campana	29000	4000	0.63
Mojón Blanco	19000	1700	0.89
Hoya del Monte	15800	1900	0.51
Casa Villora2	5600	1000	0.26
Hoya de las Ánades	4900	900	0.22
Hoya Yerba	1060	230	0.23
Hoya Elvira	230	70	0.09
Hoya Turnera	130	50	0.07

The relationships among models are shown in Figure 3.4. Two clusters of models are identified using 0.90 as the distance threshold. Cluster A includes the models COL_wd, GAM_a, COL_ANa, GAM_w and COL_ANw (bootstrap= 85%). This cluster includes discrete and continuous models, thus continuous and discrete approaches can produce similar results depending on other assumptions. Comparison of the two hemi-matrixes in Figure 3.4 shows the increase in the correlation between the model GAM_a and the rest of the models when excluding the two most ephemeral ponds (lowest correlation: 0.27 *vs*. 0.72).

Figure 3.3. Instances outlining the data analysis procedure to compute predictability estimates of two ponds (SAL and CAM) using two models: COL_ANa (discrete) and GAM_a (continuous). ***(a.1,2):*** 27-year time series of water-surface area (*A*) obtained by Landsat 5/7 scenes after applying 2cA. ***(b.1,2):*** Counts for the two-way table (time steps *vs*. pond state) for COL_ANa model; constancy, contingency and predictability indices. ***(c.1,2):*** Scatter plot of the normalized water-surface area ($A' = A/\overline{A}$) and day of the year (DOY) from the time series; regression curves based on GAM (solid line), 95% CIs (dashed lines), and predictability index. Notice that predictability indices ranges differently depending on the model.


SAL
(a.1)
A(km²)
0.5
0.4
0.3
0.2
0.1
0
1984
1990
1999
2003
2011
(b.1)
Time steps: calendar months
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
States
s1 0 0 0 0 1 1 1 0 0 0 0 0
s2 7 11 11 16 11 15 16 16 7 7 6 1
s3 0 0 0 0 0 0 2 1 2 0 0 0
Constancy + Contingency = P_COL_ANa
0.75 + 0.07 = 0.82
(c.1)
A'
2.0
1.5
1.0
0.5
0.0
0
100
200
300
DOY
P_GAM_a = 3.03

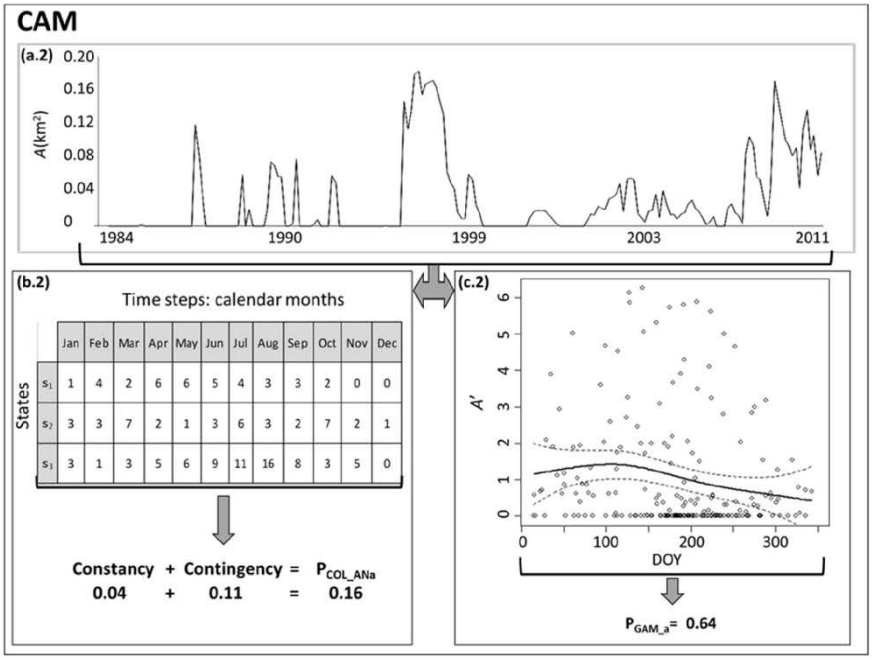

CAM
(a.2)
A(km²)
0.20
0.16
0.12
0.08
0.04
0
1984
1990
1999
2003
2011
(b.2)
Time steps: calendar months
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
States
s1 1 4 2 6 6 5 4 3 3 2 0 0
s2 3 3 7 2 1 3 6 3 2 7 2 1
s3 3 1 3 5 6 9 11 16 8 3 5 0
Constancy + Contingency = P_COL_ANa
0.04 + 0.11 = 0.16
(c.2)
A'
6
5
4
3
2
1
0
0
100
200
300
DOY
P_GAM_a = 0.64

Table 3.3. Predictability estimates and statistics for each pond and model combination. Ponds ranked by $\bar{A}$.

	Model						
Pond	**COL_wd**	**GAM_a**	**COL_ANa**	**GAM_w**	**COL_ANw**	**COL_MAXg**	**COL_MAXlin**
PET	1.00	2.23	0.66	2.28	0.64	0.48	0.26
SAL	1.00	3.03	0.82	3.07	0.86	0.39	0.55
ONT	0.97	2.04	0.62	2.08	0.67	0.17	0.26
HYG	0.32	0.74	0.22	1.14	0.12	0.41	0.42
SLD	1.00	2.21	0.80	2.26	0.77	0.34	0.21
ATA	0.75	2.27	0.67	2.43	0.74	0.48	0.32
HOR	0.23	1.02	0.29	1.91	0.56	0.34	0.34
HYR	0.66	2.07	0.66	2.35	0.72	0.48	0.30
CVI	0.13	0.73	0.31	1.62	0.51	0.34	0.44
HRE	0.17	0.54	0.39	1.90	0.52	0.45	0.53
HYN	0.18	0.58	0.35	1.46	0.35	0.53	0.67
HYC	0.12	0.83	0.21	1.61	0.45	0.23	0.33
CAM	0.11	0.64	0.16	1.02	0.20	0.41	0.52
BLA	0.63	1.94	0.55	2.13	0.71	0.37	0.24

Table 3.3. (continued)

Pond	Model						
	COL_wd	GAM_a	COL_ANa	GAM_w	COL_ANw	COL_MAXg	COL_MAXlin
HMT	0.19	0.72	0.28	1.58	0.35	0.39	0.47
CVI2	0.22	0.54	0.46	2.00	0.65	0.53	0.54
HAY	0.25	0.41	0.48	1.80	0.26	0.57	0.60
HYB	0.34	0.42	0.50	1.88	0.56	0.72	0.81
HYE	0.63	0.33	0.77	3.38	0.88	0.82	0.81
HTU	0.70	0.28	0.81	3.63	0.83	0.88	0.89
Mean	0.48	1.18	0.50	2.08	0.57	0.47	0.48
C.V.[a]	0.70	0.72	0.43	0.32	0.39	0.48	0.43

a. Coefficient of variation

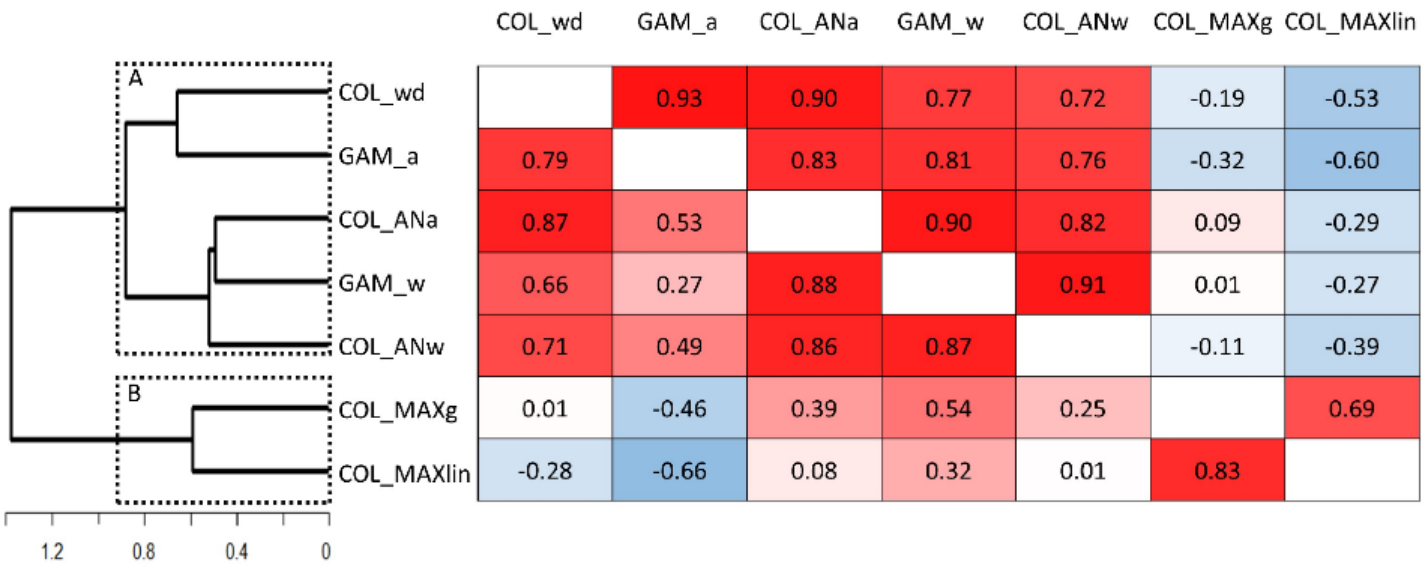


Figure 3.4. Relationships between models from the estimated predictabilities. On the left-side, dendrogram obtained by UPGMA cluster analysis based on the chord distance between predictability estimations. Using 0.90 as the distance threshold. The two clusters of models (A, and B; enclosed by dashed rectangles) are defined by distance threshold of 0.9. On the right-side, the matrix shows correlations using Pearson's coefficient excluding the two most ephemeral ponds (upper hemi-matrix, *n*= 18) and not excluding them (lower hemi-matrix, *n*= 20). Colour code uses a blue-red scale, from negative to positive.

Cluster B includes COL_MAXg and COL_MAXlin (bootstrap= 99%), the two discrete models where the states were defined with respect to the maximum observed *A*. Therefore, according to the results, linear and geometric scaling seems secondary for predictability estimations. Interestingly, some of the models in cluster A were negatively correlated with the models in cluster B. This diverging result seems to be associated to a different effect on the predictability index caused by those ponds that keep some water cover (*A*> 0) frequently. To show this, in Figure 3.5 predictability estimations are plotted against the proportion of observations in the state s_1 (Table 3.1; the state including *A*= 0) for a representative model from each cluster (COL_ANa and COL_MAXlin).

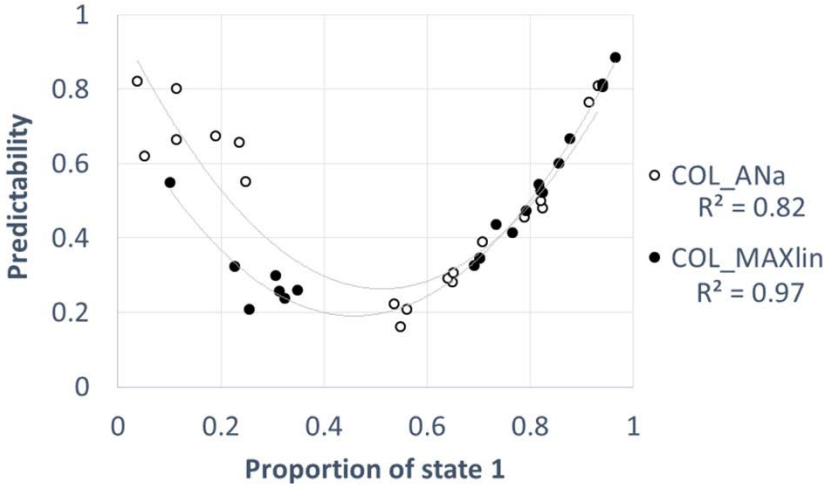


Figure 3.5. Relationship between predictability and the proportion of observations in the state s_1; the state that includes *A*= 0 for a model of cluster A (COL_ANa) and a model of cluster B (COL_MAXlin). Solid lines are the quadratic least-square fitting (R^2 is the determination coefficient associated to the corresponding fitting). Note that predictably dry ponds are at upper right part of the plot.

Discussion

Characterizing environmental fluctuations is a central issue in ecology and environmental sciences requiring long-term time series (Sabo and Post 2008). Time-series analysis often implies discarding a high number of observations; in the present case, 27 years of observations were reduced by 53% due to cloud filtering. Not surprisingly, obtaining long time series is regarded as costly, and they are consequently scarce (May and McLean 2007). In this context, remote sensing has been proven to be a timely, reliable, global and cost-efficient tool for analysing a given environmental variable over a long time series. Therefore, it is worthy developing procedures in order to solve the problems that can be associated to remote sensing data, such as confusing factors (e.g. salt *vs.* water) and non-regular recording frequency.

In the study system, the water-surface area (*A*) showed fluctuation patterns differing among ponds, what can be relevant to the fauna observed in the region (see below). The conditions for water detection applied here combine (1) the robustness of band ratios (MNDWI index> 0; Xu 2006) and (2) a refinement needed for saline ponds (near infrared band reflectance< 0.4, based on Ormeci and Ekercin 2007). The consistency between the satellite-based measurements and measurements based on more direct observations was found to be remarkable, and makes the automatic procedure

developed here reliable. Here it is shown that satellite scenes, after a convenient process, provided enough resolution to detect *A* variation in saline ponds ranging from 0.00013 to 1.19 km^2. Quantification of small ponds is most important when studying ecology in arid regions.

The final goal of this chapter was to implement and compare procedures oriented to quantify the degree of predictability of the focus habitat. A first challenge was to assess the effect of the measurement scale on the focus variable. Many studies need to be based on a nominal scale or just prefer this approach, while others use a quantitative scale. In the present study, the cluster analysis merged predictability estimates from discrete (nominal) and continuous (quantitative) models, and therefore, these approaches can yield similar predictability results, other assumptions having higher impact on predictability. Fourier analysis and AEMs (Blanchet et al. 2008; Legendre and Gauthier 2014) address quantitative data in fluctuation analysis but are not able to deal with aperiodic observations, so that additional assumptions, which may critically affect the analysis, are needed if many gaps occur in the periodic sampling scheme. Unfortunately, this case is not uncommon (Horton and Kleinman 2007), and it also affected sampling in the present study. The continuous (GAM) models used here —needed to assess the effect of the measurement scale– addressed this

problem. However, unlike Fourier and AEMs, GAM ignores the correlation between consecutive observations because the days of the year in different years are merged. This also occurs in models based on Colwell's approach, implying a shared principle in the continuous and discrete models. Developing time-series analysis able to address rather aperiodic observations would be very interesting. In the meantime Colwell approach, even involving a discretization of a continuous variable, seems to be reliable if compared to the regression methods explored here, and has the advantage of keeping the analysis simple (Sabo and Post 2008; Horton and Kleinman 2007; Kennard et al. 2009). Nevertheless, discretization criteria need more attention, as suggested by the obtained results, as will be discussed below.

A second challenge is to determine what kind of variation in data is the relevant one to estimate unpredictability. For instance, extreme values could be much more —or much less– important than what a linear approach would account for, or unpredictability estimation could be inflated by considering meaningless values. This problem translates in the scale transformation (e.g. normalization) needed to be applied to the focus variable and cannot be thoroughly analysed without considering the organisms for which unpredictability is evaluated. In previous stream studies using Fourier transform, the presence of many zeros in the time series caused the non-

zero observations to produce an extreme residual (Sabo and Post 2008). Here, different models that consider or not zeros in several ways were explored, and similar results were obtained in the majority of cases. However, one of the continuous models (GAM_a) differed from the remaining models when estimating predictability of the two most ephemeral ponds (HTU and HYE; $A = 0$ in more than 90% of observations), which likely caused the low correlation between GAM_a and some of the models in cluster A. Most of the models assigned a high degree of predictability to these two ponds, while GAM_a assigned them a low degree of predictability. Predictability evaluation by GAM_a is based on the coefficient of variation of A and, for these two ponds, it tended to be large —thus yielding low predictability– because the mean is very low. This effect on the coefficient of variation has been reported in statistical studies (Gulhar et al. 2012; Fay et al. 2013). In other words, GAM_a —and more generally, considering zeros to estimate the mean A in a continuous approach– could inflate the environmental unpredictability. This effect is supported by direct inspection of the data. Actually, HTU and HYE ponds rather than unpredictable systems seem to be ephemeral ponds, as most of the time they are dry.

As stressed long time ago (MacArthur and Levins 1964; Levins 1968), predictability depends on the point of view of the focus organism, and confusion is possible between what human

researchers and what other organisms can predict (Chapter 1). Therefore, no predictability metric can pretend to be an absolute quantification of the degree of predictability of a specific environment, but instead, time scales of periodic variations and other biological factors must be taken into account when referring to any specific predictability estimation (Southwood 1977). Models designed under different biological assumptions would be expected to produce different predictability results, which is the case when comparing models classified here in different clusters. An inspection of the model assumptions included in cluster A suggests that predictability in those models would be the one perceived by a generalist, eurioic, small-sized organism; holding the opposite for models in cluster B. For instance, in model COL_wd (cluster A), the presence of water —either with high *A* or low *A*– would be enough condition to make the environment suitable; in COL_ANa and COL_ANw, a wide variation around mean *A* does not result in a state change. Thus, taking into account the case of animals, the scaling in cluster A models could fit in the case of aquatic invertebrates (cladocerans, rotifers, copepods, insect larvae, etc.; sized less than ca. 1 cm) that can achieve high population sizes in small volumes and tolerate broad salinity ranges. This *a priori* assessment is confirmed after dissecting the outputs from the models. As shown by Figure 3.5, both cluster A and cluster B models did not confound frequently dry ponds with

unpredictable ponds. Both, frequently dry ponds and frequently flooded ponds achieved high predictability values. Thus, (a) to keep frequently some water and (b) to remain frequently dry or with low water cover can be regarded as predictable conditions. However, in relation to cluster A models, cluster B models seem to assign lower predictability to the environment of ponds keeping some water frequently, and this is interpretable as being sensitive to unpredictable variation in the amount of water when water occurs. This sensitivity is appropriate if the focus organism are affected by the *A* extension of the water cover, not being enough the presence of some water. This is likely the case for large animals that typically need a big exploitation area and are close to the top of the trophic chains (e.g. fish). The ponds studied here would offer low opportunities for these animals. The analysis performed would be detecting lack of such predictable environments. Consistently, the animals typically found in the study area are very much restricted to small invertebrates (Cirujano et al. 1988; Lapesa 2004), able to display diapause stages. To the best of found knowledge, no large animals have been reported in the study system with the exception of migratory birds (López Donate et al. 2004). For migratory birds, the seasonal home range would be the whole pond system rather than a single pond because they can move easily from one pond to another. Moreover, since results showed that the fluctuation pattern in *A* is poorly correlated among the

studied ponds, such behaviour allows them to experience the region as much more predictable than any single pond.

When characterizing a natural system, ecologists face the following two problems: (1) needing to gather long-term data and (2) handling that valuable and tough-acquiring data. In this chapter it is shown that remote sensing data has become more accessible and opens an opportunity for ecologists to obtain long time series needed to calculate predictability metrics. Here, the analysis at this step is restricted to a specific variable (*A*) and type of habitat, but this type is important in many geographic regions. Once this long-term data series is acquired, a model has to be assumed to produce a predictability metric. Interestingly, here GAM fitting is effectively introduced as an alternative approach for measuring predictability. The developed methodology extracted meaningful information about the degree of predictability of a set of Mediterranean ponds. Remarkably, some model assumptions have been shown to have negligible effects, while others can be associated with the species assemblages for which predictability needs to be assessed. The methodology described here is applicable to a wide variety of study systems and will be valuable for quantifying and characterizing predictability, which is essential considering the predicted scenario of upcoming global increases in the unpredictability of environmental

fluctuations (Pimm et al. 1995; IPCC 2013). Worth noting, the recently launched satellite Sentinel 2A in 2015 and the upcoming launch of Sentinel 2B, working together with Landsat 8, will provide images every 3-5 days. This improvement in the temporal resolution of remote sensing data will enable an improvement of the area estimations and a better study of the variability.

4

Adaptive responses of rotifer populations to environmental unpredictability

Summary

Here, rotifers are used as model organism to investigate adaptation to environmental fluctuations, a fundamental question in evolutionary biology. In cyclically parthenogenetic rotifers, clonal proliferation occurs along with occasional bouts of sex. These bouts contribute to the production of diapausing eggs, which allows survival between growing seasons. In this chapter, two diapause-related traits are studied using rotifer clones from nine Brachionus plicatilis *natural populations that vary in the degree of environmental unpredictability (Chapter 3). The tested hypotheses are that the level of environmental unpredictability is directly related to the propensity for sex (a proxy of the timing of diapausing egg production) and inversely related to the hatching fraction of diapausing eggs (the lower, the longer the diapause duration). Significant levels of genetic variation within populations for both traits were found. As predicted, a positive correlation between pond unpredictability and the propensity for sex was found. This correlation suggests a conservative bet-hedging strategy that provides protection against growing seasons unexpectedly short. In contrast, the hatching fraction of diapausing eggs was not related to the level of environmental predictability. Our results highlight the*

ability of rotifer populations to locally adapt to time-varying environments.

Introduction

Environmental fluctuations and their degrees of predictability are evolutionarily relevant since they are expected to produce diverging adaptive responses in organisms (Chapter 1). As showed in Chapter 3, environmental fluctuations can be decomposed into predictable and unpredictable components. These two components are expected to act as very different selection pressures on shaping life-history traits. Natural populations confined in shallow, non-permanent ponds are expected to be adapted to and strongly driven by inundation patterns (Angeler et al. 2000; James et al. 2008; Florencio et al. 2016). This is likely the case for rotifer populations that inhabit Mediterranean water bodies. These populations can be conceived as a collection of annual clones that cross sexually to produce diapausing eggs, some of which will hatch in the next growing season (Carmona et al. 2009; Gómez and Carvalho 2000). The number of diapausing eggs produced is a major component of clonal fitness because these eggs are the only way to survive unsuitable water column conditions between growing seasons (Serra and King 1999). Since the length of the growing season affects diapausing egg production, among-year fluctuations in the duration of the growing season can provide insight into how environmental unpredictability affects fitness. For instance, an unexpected critically short growing season may cause the density threshold for sex initiation (see Chapter

2) to be unmet. Failure of a clone to reproduce sexually means no production of diapausing eggs and, consequently, zero fitness. The end of the growing season in rotifer populations can arise not only due to abiotic factors (e.g. drought, extreme salinity, extreme temperature) but also due to biotic factors (e.g. occurrence of competitors and predators). Randomness in these ecological factors will determine the variance of the growing season and hence the uncertainty the population will face. Such unpredictable fluctuations are expected to be powerful drivers of life-history traits and might be increasingly important in periods of climate change. Several studies have reported high levels of genetic variation in diapause-related traits in rotifers (Carmona et al. 2009; Gilbert and Diéguez 2010; Walsh et al. 2014; Gabaldón and Carmona 2015). A few other studies have found signatures of local adaptation in these traits (Campillo et al. 2010; Schröder 2005). Nevertheless, the relationships between life-history traits related to diapause and the degree of environmental unpredictability in natural populations of aquatic organisms remain largely unknown, especially because quantifying unpredictability is a complex issue (Sabo and Post 2008; Chapter 3). This difficulty can be partially overcome in the laboratory by simulating different patterns of predictability. In a recent study using an experimental evolution approach, laboratory rotifer populations showed a rapid adaptive response in diapause-related traits under two

contrasting selective regimes of environmental predictability (Tarazona et al. 2017). These selection experiments are important because they show that traits evolve as expected in relation to environmental unpredictability after controlling for other factors in laboratory conditions. However, they do not provide unambiguous evidence about whether such an evolution is possible —and might be occurring– in the wild, where levels of unpredictability might be different to the experimental ones, and adaptation to unpredictability might be traded off by adaptation to other concomitant conditions or counterbalanced by non-selective evolutionary forces, namely, migration.

There are two key life-history traits in the life cycle of monogonot rotifers (see Chapter 2) that are associated with the entrance and exit from diapause. One is the propensity for sex, which is inversely related to the density threshold for sex initiation (Aparici et al. 2001; Carmona et al. 2009) and is a proxy of the timing of diapausing egg production (García-Roger et al. 2016). Other is the diapausing egg hatching fraction, which is inversely related to diapause duration. Both traits have been proposed as instances of bet-hedging strategies that might interact to reduce the risks associated with environmental unpredictability (Spencer et al. 2001).

The propensity for sex has been proposed as a case of conservative bet hedging in rotifer populations that inhabit unpredictable habitats (García-Roger et al. 2016). When there is uncertainty regarding the onset of unsuitable periods, a low-risk strategy might be to produce diapausing eggs as soon as possible to avoid an unexpectedly short growing season (Serra and King 1999). However, if the growing season is not short, early investment in sex and diapause will reduce the rate of clonal proliferation, thereby resulting in lower fitness (Carmona et al. 2009; Serra et al. 2004). These considerations lead to the prediction that the propensity for sex will increase with increasing unpredictability. However, in testing this prediction, a confounding factor should be considered: if the growing seasons are predictably short, a high propensity for sex is also expected to evolve.

In contrast, diapausing egg hatching fraction has been proposed in several theoretical studies as a form of diversified bet hedging. As far as rotifers cannot predict whether a particular growing season will be sufficiently long to complete the life cycle and ensure diapausing egg production, intermediate hatching rates are expected in habitats with both long growing seasons (where, ideally, all of the eggs should hatch; i.e. good seasons) and short growing seasons (where, ideally, no egg should hatch; i.e. bad seasons). According to bet-hedging theory, the optimal hatching

fraction should equal the frequency of good seasons (Cohen 1966) (reviewed in García-Roger et al. 2014). For example, in this 'good *vs.* bad growing season' scenario, a hatching fraction around 0.5 (i.e. intermediate) would be expected in a completely unpredictable habitat, since the frequency of good seasons would be somewhere around 0.5 (Cohen 1966).

In this chapter, it is studied whether natural rotifer populations locally adapt to the degree of environmental unpredictability by evolving diverging strategies for both the propensity for sexual reproduction and the fraction of diapausing egg hatching. A total of 270 clones from nine populations of *Brachionus plicatilis* from habitats in which environmental unpredictability was quantified (Chapter 3) are analysed. While most studies on bet hedging typically focus on single traits (Childs et al. 2010), the present study addresses whether bet hedging has evolved in either or both diapause-related traits in response to a natural gradient of environmental predictability. Moreover, it tests for the predictions from the well-established theory (Cohen 1966; Starrfelt and Kokko 2012) in a field where empirical evidence is scarce (Simons 2011). Specifically the predictions tested here are that environmental unpredictability will select for (1) a high propensity for sex (early diapausing egg production) and (2) intermediate diapausing egg hatching fractions.

Material and methods

Study populations and clone establishment and maintenance

To obtain a collection of experimental clones directly from the field, diapausing egg banks of *B. plicatilis* were sampled in nine of the saline ponds described in Chapter 2 (Figure 4.1). Both, the degree of environmental predictability and the hydroperiod length at these sites were previously quantified in Chapter 3 (Table 4.1). Here, the predictability metric derived from Colwell's approach (Colwell 1974) and based on the presence/absence of water (COL_wd model; Chapter 3) was used. This metric has the advantage of being simple and parameter-free, and was proposed as appropriate for aquatic invertebrates in Chapter 3.

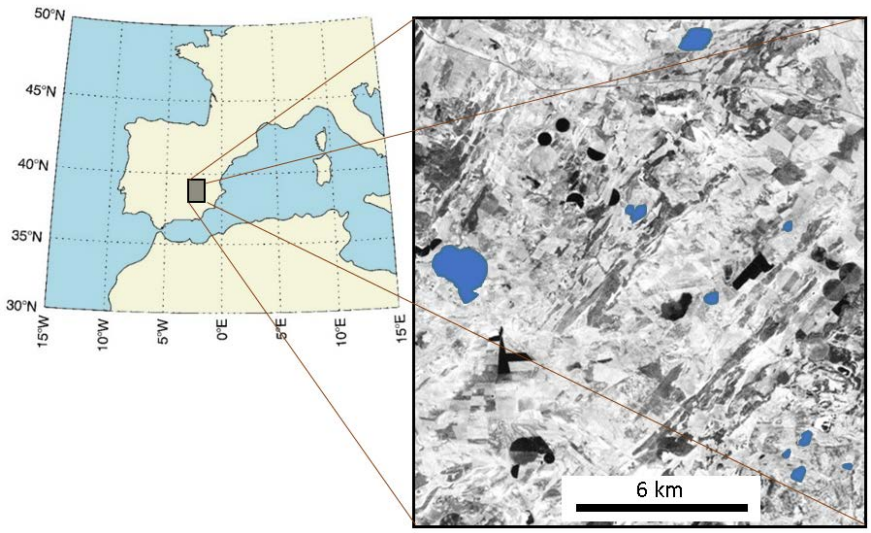


Figure 4.1. Location of the study region (38°55.4' to 38°41.803'N and 1°47.32' to 1°24.26'W), with the studied ponds highlighted in blue.

Table 4.1. Features of the studied ponds (Chapter 3) and nuclear genetic diversity of the populations (expected heterozygosity H_e, from Montero-Pau 2016). Broad-sense heritability H^2 of propensity for sex, as estimated in this chapter, is shown for the studied populations.

Pond / Population	Hydroperiod length[a]	Estimated environmental predictability[b]	H^2	H_e
PET	1.00	1.00	0.11	0.44
SAL	1.00	1.00	0.36	0.27
ATA	0.93	0.75	0.49	0.25
HYR	0.87	0.66	0.35	0.38
HYC	0.51	0.12	0.58	-
CAM	0.63	0.11	0.30	0.41
HMT	0.51	0.19	0.16	0.41
HYB	0.23	0.34	0.30	0.41
HTU	0.07	0.70	0.68	0.26

a. From Chapter 3, annual average of the fraction of observations with $A > 0$ in each month (excluding cloud-covered observations).

b. From Chapter 3, as calculated by COL_wd model.

A sample from the uppermost 10 cm of sediment at each of the nine ponds was obtained with a Van Veen grab (Eijkelkamp Agrisearch Equipment, Giesbeck, The Netherlands) in September 2013. Diapausing eggs were isolated from the sediment using a sugar flotation technique (Gómez and Carvalho 2000), and the diapausing eggs that looked healthy were transferred individually to wells in 96-

multiwell plates (NuncTM, Nalge Nunc Int., Roskilde, Denmark). The deterioration state (healthy *vs*. deteriorated) of the diapausing eggs was determined attending to the percentage of embryo integrity and visual aspect (García-Roger et al. 2005). Eggs were induced to hatch under standard hatching conditions: 25 °C, 6 g L^{-1} (artificial saline water, Instant Ocean; Aquarium Systems, Inc., Mentor, OH, USA), and constant illumination (PAR: 35 µmol photons m^{-2} s^{-1}) (details in García-Roger et al. (2005, 2006); Martínez-Ruiz and García-Roger (2014)). Hatchlings were monitored every 24 hours for a maximum of three weeks. Clonal lines were established by asexual proliferation of the resulting neonate females. Inadvertent selection in favour of clonal lines with high hatchability is unlikely. At the sediment depth sampled — which integrates several years (García-Roger et al. 2006)– eggs induced to hatch in the laboratory are a mixture of different cohorts, so that, if variation in timing of hatching occurs, the hatched eggs should include early hatchers and late hatchers (i.e. those eggs that remained in diapause when cues inducing hatching occurred in their habitats). Thirty clones from each field population were founded and maintained in 15 mL stock cultures at 12 g L^{-1} salinity and 20 °C. Every week, half of each clonal culture volume was renewed with fresh medium. This medium was f/2-enriched saline water (Guillard and Ryther 1962) in which the microalgae *Tetraselmis suecica* (Microalgae Culture

Collection of ICMAN-CSIC, Spain) had been grown as rotifer food. Unless otherwise indicated, pre-experimental and experimental rotifer culture media and conditions were the same as for the stock cultures (hereafter, 'standard conditions'). Clonal lines were identified to the species level by genetic analysis of cytochrome c oxidase subunit I (COI) based on PCR-RFLP (Campillo et al. 2005) since *B. plicatilis* belongs to a cryptic species complex. Data on the unbiased nuclear genetic diversity for these populations are available from Montero-Pau et al. (2016) (range 0.25-0.44, see Table 4.1). Rotifer genetic diversity and pond unpredictability were not significantly correlated (R^2= 0.23; *P*-value= 0.22).

Characterization of life-history traits: propensity for sex

Genetic variation in the propensity for sex was studied by conducting 810 bioassays (9 populations × 30 clones × 3 replicates) following the procedure in Carmona et al. (2009) with minor modifications as described below.

Pre-experiment. To control for maternal effects and avoid the early induction of sexual reproduction, each rotifer clone was pre-cultured at low density under standard conditions in darkness over three generations (Stelzer and Snell 2006). To accomplish this, three asexual females that each carried two asexual eggs were individually sampled from each stock clonal culture and transferred to Petri dishes with 40 mL of culture medium (initial microalgae concentration: 250,000

cells mL^{-1}). After 24 hours, daughters had been produced. Next, all of the females were removed except for one neonate female (F1) in each dish. Each of these kept females produced daughters after 48 hours. Next, a single neonate female from the second generation (F2) of each replicate was transferred individually into a new Petri dish containing 40 mL of fresh culture medium. This procedure was repeated to obtain F3 neonate females and begin the bioassay. Through this approach, the three experimental replicates were independent since the pre-experimental culture.

Bioassay. Each F3 neonate female was transferred individually to a Petri dish containing 15 mL of fresh culture medium (initial microalgae concentration: 500,000 cells mL^{-1}). The F3 females were allowed to reproduce and proliferate and were monitored every 12 hours until the first male was observed (i.e. the initiation of sexual reproduction). Then, the culture was fixed with Lugol's solution (final concentration 4%), and the population density was recorded as an inverse measure of the propensity for sex (Aparici et al. 2001). Additionally, the time for sex initiation was recorded.

Characterization of life-history traits: diapausing egg hatching fraction

Estimation of the diapausing egg hatching fraction was performed by randomly selecting from each pond a subset of ten clones that were used in the propensity-for-sex

experiment. In total, 8,640 bioassays (96 diapausing eggs × 10 clones × 9 populations) were performed.

Diapausing egg production. To estimate hatching fraction, a high number of diapausing eggs were produced under laboratory conditions for each clone by intraclonal sexual reproduction within a narrow time window (4 days). Twenty ovigerous, asexual neonate females of each stock clone were transferred into a Petri dish containing 40 mL of culture medium (initial microalgae concentration: 500,000 cells mL^{-1}), and parthenogenetic proliferation and sexual reproduction was allowed. Microalgae density was maintained at over 250,000 cells mL^{-1} by adding highly concentrated (centrifuged) microalgae daily. Cultures were inspected every 24 hours until the first mature diapausing egg was observed in any clone. Then, all of the cultures were maintained for 4 additional days. On the fourth day, mature diapausing eggs that looked healthy (García-Roger et al. 2005) were collected and cleaned with 6 g L^{-1} saline water. Most of the clones (83.3%) had produced 96 eggs by this stage; however, for the remaining clones, 8 additional days, under the same production conditions (collecting diapausing eggs every 4 days) were required to obtain 96 eggs. Immediately following collection, the eggs were individually transferred to wells in 96-multiwell plates and incubated under standard hatching conditions. As shown in some recent

studies in *Brachionus* (Becks and Agrawal 2012; Scheuerl and Stelzer 2013; Martínez-Ruiz and García-Roger 2014), a long diapause period after the production of diapausing eggs before hatching is not necessarily required.

Bioassay. Diapausing eggs were inspected every 24 hours over a 28-day period for the presence of hatchlings and deterioration. The unhatched fraction of diapausing eggs that looked healthy after that period (García-Roger et al. 2005) were dried out and held in darkness at 4 °C for 28 days. Afterwards, diapausing eggs were induced to hatch under the same hatching conditions described above. This was made to test their viability. Diapausing egg hatching fraction was calculated as the number of hatched eggs in the first incubation period out of those that looked healthy. The fraction of deteriorated diapausing eggs (i.e. excluded diapausing eggs) ranged 0.24-0.53 across populations.

Data analysis

Generalized linear mixed-effect models (GLMMs) were used to test for differences among populations in both life-history traits. The propensity for sex was analysed using a Poisson distribution of errors and the log link function. For the diapausing egg hatching fraction, a binomial distribution of errors and the logit function were used. For the analysis of the life-history traits, hydroperiod length and predictability were included as fixed-effect continuous predictors (factors),

whereas population (9 levels) and clone nested within population (*n*= 30 and 10 per population for propensity for sex and hatching fraction respectively) were considered as random-effects factors. In both analyses, maximum likelihood (ML) ratio tests were used to determine the structure of fixed effects by alternatively dropping hydroperiod length or predictability against a full model including all effects. Then, the significance of random effects was tested by means of restricted maximum likelihood (REML) ratio tests. GLMMs were performed in R v.4.1.1 (R Development Core Team 2009) by using the glmer function of the lme4 package (Bates et al. 2015).

Within-population genetic variation in the propensity for sex was measured using broad-sense heritability (H^2, i.e. the ratio of the among-clone variance to the total within- and among-clone variance), which is the relevant measure for selection during clonal proliferation (Lynch and Walsh 1998). H^2 was estimated according to Pfrender and Lynch (2000). H^2 for the hatching fraction was not calculated because variance components cannot be reliably estimated under a binomial distribution of errors (de Villemereuil et al. 2016).

Results

Propensity for sexual reproduction

Clones of *B. plicatilis* showed wide variation in the density and the time for sex initiation (Figure 4.2), clone mean ranges being of 0.2-31.4 females mL^{-1} and 3.5-14 days, respectively. These two measurements are highly correlated (overall Pearson correlation coefficient= 0.754, *P*-value< 0.0001). Propensity for sex, as inversely measured by density for sex initiation, differed significantly among populations and among clones within populations (Table 4.2). Broad-sense heritability values were significant in most populations (seven out of nine; Table 4.1) and were not significantly correlated with unpredictability or hydroperiod length.

The propensity for sex was significantly affected by the degree of predictability of each pond after controlling for hydroperiod length (Table 4.2, Figure 4.2).

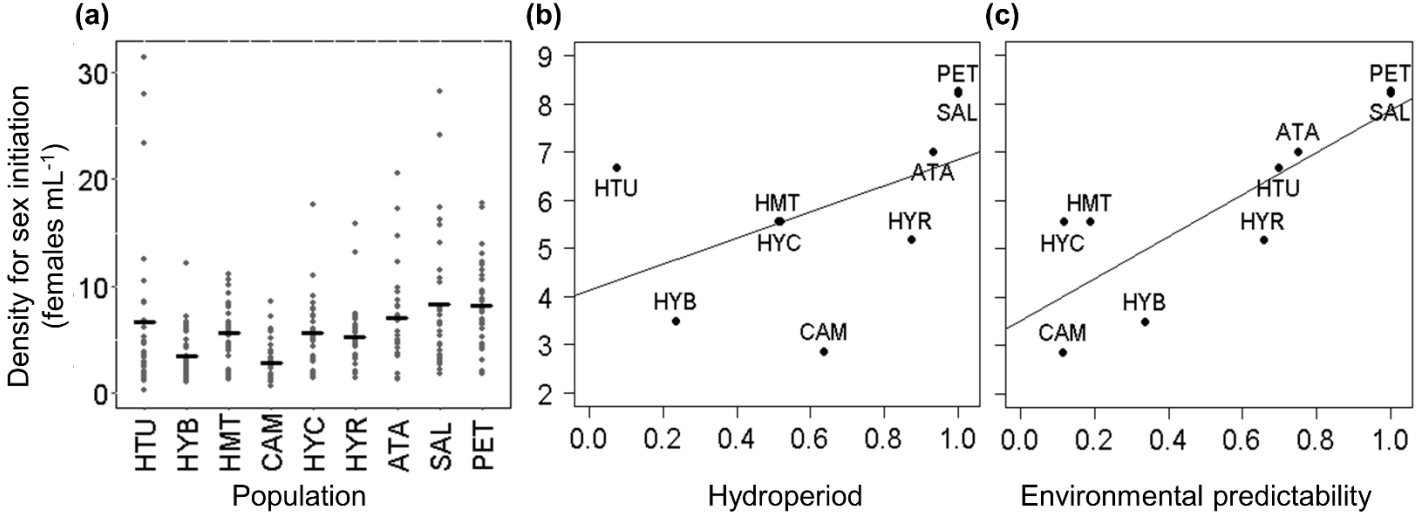


Figure 4.2. Propensity for sex in each pond as inversely measured by the population density for sex initiation. ***(a):*** Clonal means (dots) and population means (horizontal bars) as estimated for the nine populations. ***(b):*** Relationship between the estimated population means and the pond hydroperiod (fraction of the year flooded). ***(c):*** Relationship between the estimated population means and the estimated environmental predictability in the ponds. The solid line is the least-squares linear regression fitting.

Table 4.2. Summary of the generalized linear mixed effects model (GLMM) results on the propensity for sex and the diapausing egg hatching fraction with one degree of freedom.

	Propensity for sex		Diapausing egg hatching fraction	
Effect	X^2	***P*-value**	X^2	***P*-value**
Predictability	4.08	0.043*	0.43	0.511
Hydroperiod length	1.25	0.262	1.58	0.208
Population	9.40	< 0.001*	3.57	0.058
Clone (Population)	75.50	< 0.001*	1140.90	< 0.001*

* *P*-value < 0.05

Diapausing egg hatching fraction

Diapausing egg hatching fraction was neither significantly affected by predictability nor hydroperiod length (Table 4.2). Hatching fraction differed significantly among clones but not among populations (Table 4.2). Noteworthy, high predictability is not associated to high dispersion in hatching fraction values among-ponds, a hypothesized pattern expected if predictable habitats include both predictably adverse (low hatching fraction) and predictably favourable

(high hatching fractions). The clones of *B. plicatilis* showed wide variation in diapausing egg hatching fraction, which ranged from 0 to 100% among clones (Figure 4.3).

Differences among clones within populations indicate that heritable variation exists, although H^2 was not estimated.

Discussion

Diapause-related traits are proposed to be important fitness components, and they are expected to be subject to strong selection (Levins 1968). Accordingly, within-species adaptive divergence in the response to local conditions is expected where gene flow and genetic drift do not counterbalance selection. Here, it is demonstrated that the degree of environmental predictability strongly correlates with the propensity for sex in rotifer populations, with rotifer populations that inhabit more unpredictable environments being more prone to reproduce sexually. Patterns of differentiation among populations in diapause-related traits have been found in other studies of cyclical parthenogens, including cladocerans (Tessier and Cáceres 2004; Roulin et al. 2013, 2015), aphids (Simon et al. 1999; Dedryver et al. 2001) and rotifers (Schröder 2005; Campillo et al. 2009). Thus, our research extends previous studies showing that populations

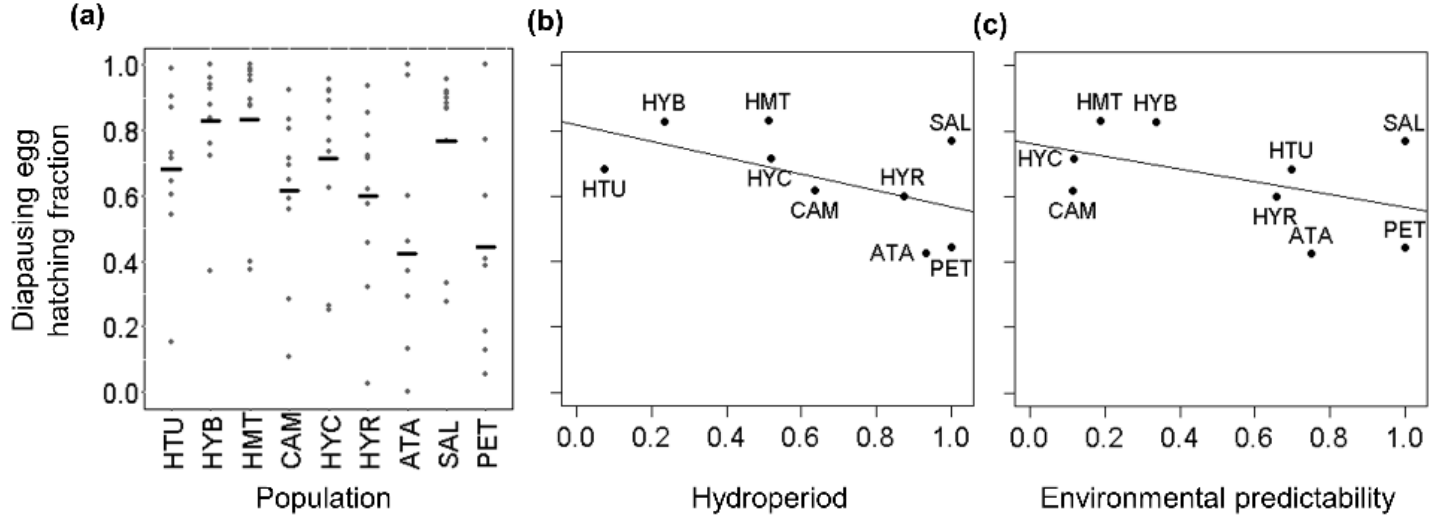


Figure 4.3. Diapausing egg hatching fraction in each pond. ***(a):*** Observed clonal hatching fractions (dots) and population means (horizontal bars) as estimated for the nine populations. ***(b):*** Relationship between the estimated population means and pond hydroperiod (fraction of the year flooded). The solid line indicates the least-squares regression fitting. ***(c):*** Relationship between the estimated population means and pond environmental predictability. The solid line indicates the least-squares regression fitting.

are locally adapted to the environment in the timing of diapausing egg production. However, these studies did not investigate whether these patterns are correlated with the environmental unpredictability of the habitats of the organisms. This study provides this correlational evidence across a well-established gradient of environmental unpredictability occurring in the wild, pointing to local adaptation in a small geographical range (240 km^2).

The results from this chapter are consistent with the theoretical prediction that environmental unpredictability selects for early timing of the production of diapausing stages (Hairston and Munns 1984; Taylor and Spalding 1989; Serra and King 1999), as propensity for sex is a major factor influencing diapausing egg production in cyclically parthenogenetic rotifers. As rotifer response was measured in many clones under laboratory conditions, our experimental design allows us to conclude that the differential response among populations is genetically based and is thus shaped by evolutionary forces, likely natural selection. The apparent suboptimality of early sex may be explained as a conservative bet-hedging strategy that has evolved in response to environmental unpredictability. This low-risk strategy protects against unexpectedly short growing seasons (e.g. due to drought or to the occurrence of predators or competitors) by ensuring that some diapausing egg production occurs

despite the cost to current growth rate in long growing seasons. Delay of sex would be fatal if the habitat becomes unsuitable when reproduction is still exclusively asexual or when sexual reproduction is incomplete and no diapausing eggs have yet been produced. Accordingly, by reducing the variance in diapausing egg production across growing seasons, a higher propensity for sex would overcome the disadvantage of producing a lower average yield of diapausing eggs. A high propensity for sex could also be expected to evolve as a response to predictable short hydroperiods, as found by Smith and Snell (2012). However, this latter explanation appears unlikely in our case because our analysis revealed a highly significant correlation between predictability and propensity for sex when the effect of hydroperiod length was controlled for. Nevertheless, unpredictability should select for early sex because short hydroperiods may occur, and these events have an overwhelming effect on shaping optimal investment in sex and diapause.

A previous experimental evolution study using the same model organism showed that a selective regime simulating environmental unpredictability rapidly selected for early sex in multiclonal, highly-diverse populations created in the laboratory (Tarazona et al. 2017). Now, in this chapter, the degree of environmental unpredictability has been identified

as an important contributor to explain the existing patterns of timing of sex in nature. Unpredictability works as an effective selective factor in the studied rotifer natural populations. Their evolutionary effects are not traded off by adaptation to unknown selective factors that might act in the wild, and not counter-balanced by other evolutionary forces as migration or genetic drift. Of most importance when comparing to the results in Tarazona et al. (2017), our results indicate that the studied natural populations harboured enough genetic diversity to fuel adaptation to unpredictability.

The significant levels of within-population genetic variation in the propensity for sex found here may be the combined consequence of fluctuating selection and the buffering effect on genetic variation that is provided by the diapausing egg banks (De Meester et al. 2002). Adaptive tracking may be acting simultaneously with bet hedging in these fluctuating environments. Nevertheless, if the rate of change in the environment is high, the mean phenotype selected through adaptive tracking may lag behind the optimum, and bet-hedging strategies may become more important (Lande and Shannon 1996; Simons 2011; Tufto 2015). However, adaptive tracking of the propensity for sex appears to be less likely in the studied populations since no relation between unpredictability and the heritability of this trait was found.

Intermediate hatching fractions (mean population range= 44-88%) as well as among-population genetic differentiation in relation to this trait were found in all populations. However, in contrast to expectations, a relationship of this trait with environmental predictability was not found, the same being held to the hydroperiod length. The lack of any expectable pattern in Figure 4.3 suggests that low statistical power is not the explanation for the negative result founded. Notice that according to theory (García-Roger et al. 2014), those populations inhabiting ponds with predictable but short growing seasons (predictably adverse habitats; HYB and HTU) could be expected to present low hatching fractions; however, this was not observed in the studied ponds. Several factors might explain the observed lack of association between unpredictability and the hatching fraction of diapausing eggs. First, since clones were produced by intra-clonal crosses, inbreeding depression could be affecting the hatching fractions (Tortajada et al. 2009). Inbreeding depression caused by intra-clonal crosses forced in the laboratory is expected in populations where clones normally do not inbred in nature; that is, in genetically diverse populations. The studied populations embrace a range of genetic diversity. However, such diversity does not correlate with environmental unpredictability. Thus, despite inbreeding depression could be causing some noise in the results, it seems very unlikely that it could counterbalance the hypothesized

effect of unpredictability on hatching fractions. Second, theoretical work has suggested that organisms that hedge bets successfully with one strategy do not need to bet hedge to the same extent with another (Spencer et al. 2001). This phenomenon might occur here, a high propensity for sex being sufficient to avoid the risk of zero-fitness growing seasons, particularly if any population has intermediate hatching fractions. Although further evidence might be needed to discard an effect of unpredictability on hatching fraction in the wild, the findings in this chapter highlight the importance of studying multiple traits involved in the same strategy to get a sensible test for bet-hedging (Childs et al. 2010). Third, a recent study in laboratory populations has addressed the effects of two selective regimes, predictable and unpredictable, on this trait using an experimental evolution approach (Tarazona et al. 2017). That study found a rapid adaptation of hatching fractions to experimental conditions. Accordingly, the hatching fraction might adaptively track rows of similar growing seasons, instead of being optimized to the predictability of the overall time-series of growing seasons.

The present chapter reveals that rotifers are able to locally diverge in diapause-related traits even within a small geographical range (240 km^2) despite their potential for widespread genetic exchange through the passive dispersal

of diapausing eggs (De Meester et al. 2002). This finding is in agreement with the genetic differentiation observed in neutral and ecologically relevant traits among populations of cyclical parthenogenetic zooplankters at local scales (De Meester 1996; Gómez et al. 2000; 2002a; De Meester et al. 2002; Campillo et al. 2009). These populations are likely to adapt within short time spans (De Meester 2004), with their huge local abundances diluting the effect of immigrants. Several studies have indeed shown rapid evolution in response to environmental changes in natural populations of cladocerans (e.g. Hairston et al. 1999; Cousyn et al. 2001; Hairston et al. 2001).

This chapter supports the expectation that wild populations of *B. plicatilis* can develop evolutionary responses to face environmental unpredictability. Given that scenarios of increased environmental variability are expected to occur in the near future (IPCC 2013), the persistence of rotifer natural populations under these circumstances may depend on the evolution of bet hedging in key life-history traits (Childs et al. 2010; Simons 2011; Gremer and Venable 2014). Therefore, a comprehensive understanding of the role of bet-hedging strategies is necessary for predicting population responses to environmental change (Lawson et al. 2015). The present study contributes to this understanding by relating two potentially bet-hedging life-history traits with a quantitative measure of

environmental unpredictability, with the result that one of them is actually related. This makes this contribution particularly relevant in a field of study —bet-hedging strategies– with strong theoretical development (Cohen 1966; Starrfelt and Kokko 2012) but where empirical evidence is still scarce, especially in natural populations (Simons 2011).

5

The genomic basis of local adaptation in wild populations of the rotifer *Brachionus plicatilis*

Summary

Unravelling the genetic mechanisms underlying phenotypic variation is needed to thoroughly understand the adaptive response of organisms to their environments. As seen in Chapter 4, life-history traits —namely, investment in sexually produced diapausing eggs– of the rotifer Brachionus plicatilis *show diverging local adaptation in relation to environmental predictability. However, little is known of the genetic basis of this diverging adaptation. Here, an integrative approach was used, combining environmental (Chapter 3), phenotypic (Chapter 4) and genotypic data. As a reference, 11.7 Gb of Illumina data were generated to assemble and annotate a draft genome for B. plicatilis. Then, polymorphism was studied using genotyping by sequencing (GBS) data obtained at a genomic scale for the set of 270 clones of the rotifer* B. plicatilis *from the nine populations in eastern Spain studied in Chapter 3. As a result, 4543 high quality SNPs were discovered and genotyped in a reduced sized —but dense in gene regions– genome. More than 90 SNPs showed* Fst *values indicating that they are putatively under selection, and signatures of diverging and balancing/purifying selection were found. More than 160 SNPs were correlated to a set of environmental or phenotypic parameters revealing signatures of*

local adaptation to environmental predictability. Remarkably, most of these SNPs were found to be located within genes annotated in the genome, so that putative functions are associated to those.

Introduction

Mapping genotype on phenotype and phenotype on environment is of fundamental importance for understanding how natural populations adapt to their local environments (Storz et al. 2005; Stinchcombe and Hoekstra 2008). As described in Chapter 1, adaptive tracking, phenotypic plasticity and bet hedging are different ways by which organisms evolutionarily respond to unpredictable variation (Philippi and Seger 1989; Reed et al. 2010; Simons 2011; Tufto 2015). Recently, some studies have identified molecular mechanisms underlying some of these evolutionary strategies —mainly to phenotypic plasticity– (see for example in *Daphnia sp.* Yampolski et al. 2014 and Roulin et al. 2016). However, the genomic signatures of the adaptive evolutionary responses to local unpredictability, especially those related to bet-hedging strategies, are little known.

In order to advance in understanding the genetic basis of diverging adaptation to environmental predictability, a system of natural populations is required so that they (1) inhabit a wide, quantifiable range of environmental predictability, and (2) their phenotypic analysis show quantitative genetic variation correlated to that environmental predictability. As shown in Chapter 3, the Mediterranean ponds studied there have a wide range of environmental unpredictability. Moreover, findings in Chapter

4 show that two key life-history traits —the propensity for sex and the diapausing egg hatching fraction– associated to the entrance to and exit from diapause in the rotifer *Brachionus plicatilis* have significant levels of genetic variation within populations. Of most importance, the propensity for sex was positively correlated to pond unpredictability, this strongly supporting the conclusion that environmental unpredictability is a selective factor shaping rotifer populations. Thereby, these results highlight that this study system —*B. plicatilis* rotifers in Mediterranean temporary water bodies– offers an opportunity to study the genetic basis of local adaptation to environmental unpredictability.

In analysing the association between environmental unpredictability and genomic variation, some caveats are needed in order to minimize the risk of spurious associations. Not being independent, an analysis of environmental unpredictability requires including the annual hydroperiod length, paralleling the approach performed in Chapter 4.

The study of the genetic basis of adaptation in non-model organisms has recently been fuelled by the rapid development of genomic technologies —especially next generation sequencing (Van Straalen and Roelofs 2006; Roff 2007; Stapley et al. 2010)–. One of these technologies, genotyping by sequencing (GBS, Elshire et al. 2011; Chapter

2) is a cost-effective technique to screen genome-wide patterns of diversity by using restriction enzymes to reduce the genome complexity. It allows the genotyping of a high number of individuals and the discovery of typically large numbers of single nucleotide polymorphisms (SNPs). By using these SNPs, relevant information about the genetic population structure of the natural populations can be obtained, and also, genome locations with signatures of selection can be identified. There are several methods that aim to detect loci under selection. A first group of methods are based on F_{ST} values (Beaumont and Balding 2004), and use the statistical distribution of F_{ST} values at a large number of polymorphic loci in order to detect outliers; that is, those that present significatively higher —i.e. divergent selection– or lower —i.e. balancing/purifying selection– differentiation among populations than expected under a model of neutral evolution (Storz 2005). These methods, however, have been described to have low power, especially when a high differentiation in neutral markers exists (De Kovel 2006; Butlin 2010), as reported in *B. plicatilis* populations (Campillo et al. 2009). Recently, another group of outlier detection methods has been developed. It is based on correlating allele frequencies to a set of environmental parameters (i.e. genome-wide association methods, GWAS; Coop et al. 2010). These correlation-based models have been reported to have higher power than the differentiation-based models, but also

to show higher false positive rates (Savolainen et al. 2013), which is the main drawback of the outlier analysis approaches currently available (Storz 2005). This approach can be extended straightforwardly to phenotypic parameters, besides or instead of environmental ones. It is worth noting that all of these methods identify markers that are not necessarily under direct selection, but instead some of them can be neutral markers 'hitchhiking' with the true locus (or loci) due to genetic linkage (Smith and Haigh 1974; Kaplan et al. 1989). Nevertheless, these techniques can be used to shed light into putative genes under selection, as those involved in adaptive population divergence.

In this chapter the genome-wide patterns of genetic variability and the differentiation in response to environmental predictability are studied. GBS is used to obtain genome data from the same 270 *B. plicatilis* clones previously characterised for life-history traits (Chapter 4). These clones were isolated from nine populations in eastern Spain with a wide range of environmental predictability (Chapter 3). Since there is no *B. plicatilis* genome available to date, whole genome sequencing data (Chapter 2) are obtained in order to assemble the first draft genome for this species —including structural and functional annotation–. This genome is used here as the reference genome to facilitate genotyping and discovery of genes under selection. More than 4500 SNPs —

discovered and genotyped here– are used to (1) explore the genetic population structure of these populations and (2) identify putative loci under selection. In (2) the association to life-history traits and environmental parameterss is investigated. Finally, the genes associated to these loci are identified, making an effort to unravel the genomic signatures of the evolutionary responses leading to local adaptation with respect to environmental predictability.

Materials and methods

Brachionus plicatilis *genome*

A clone from Hoya Rasa pond (Table 2.1; RAS1 clone, from those studied in Chapter 4) was allowed to grow at 12 g L^{-1} salinity, 20 ºC and *Tetraselmis suecica* as food for 15 days, starting from a low-density stock culture. During the first 5 days, rotifers were kept in a flask containing 200 mL of culture medium and, after that, they were transferred to eight bottles containing 1.5 L of fresh culture medium. On day 10, when population size was larger than 10,000 individuals in each bottle, rotifers were filtered out through a 30-µm Nytal mesh sieve and released on saline water 12 g L^{-1} for 24 hours in order to purge their digestive tracts and, thus, minimize contamination with microalgae DNA. Finally, the rotifers were concentrated by filtering again with a new 30-µm Nytal mesh sieve, and the genomic DNA extracted with DNeasy Blood

and Tissue kit (Qiagen, Valencia, CA, USA) following the manufacturer's protocol. DNA quality was assessed by visual inspection on 1% agarose gel electrophoresis and the quantification of DNA was performed with a Qubit 2.0 Fluorometer (Life Technologies). For library construction, DNA was sheared through sonication down to ca. 550 bp by using a Diagenode Bioruptor. DNA libraries were prepared with the sonicated DNA using the NEBNext Ultra DNA library Prep kit for Illumina combined with NEBNext Singleplex Adaptors, without including PCR enrichment stage. The library was quantified using Qubit and assessed by qPCR using the Illumina Library Quantification Kit (NEBNext Library Quant Kit) and a StepONE plus Real time PCR system (Applied Biosystems). The template DNA was denatured according to the protocol described in the Illumina system guide and loaded at 20 pM concentration. To improve sequencing quality, 1% PhiX control was spiked-in. The library was sequenced on a single Illumina MiSeq sequencer run with v3 chemistry (2x250 bp paired-end read lengths). Library preparation and sequencing was carried out at the EvoHull Genomics Lab at the University of Hull (Hull, United Kingdon).

The sequenced reads were trimmed using Trimmomatic (MacManes 2014) to remove (1) adapters added for PCR, (2) low quality leading or tailing regions (quality measured as Phred score; Phred score> 30) or unidentified nucleotides, (3)

reads with average Phred score< 100, and (4) regions (5-base sliding window) with Phred score< 20. Thereafter, reads were corrected using a bloom filter-based error correction tool (BLESS v0.16; Heo et al. 2014) and user developed scripts in order to correct sequencing errors. The paired reads were then merged using Userach (Edgar 2010). Pre-assembly assessment of k-mer distributions were obtained by KMC (Deorowicz et al. 2015) and profiling by GenomeScope (Vurture et al. 2017). *De novo* genome assemblies were constructed with several assemblers using default parameters (Celera, SPAdes, DISCOVAR, and Platanus; Table 5.1) and discarding in the final assembly contigs shorter than 500 bp. Assembly quality was evaluated using (1) Blobtools (Kumar et al. 2013) to assess contamination; (2) CEGMA (core eukaryotic genes mapping approach; Parra et al. 2009), a tool that assesses genome completeness based on the presence of complete or partial sequences of a set of core eukaryotic genes; and (3) Assemblathon scripts (Bradnam et al. 2013) to obtain length-based statistics, such as N50 (i.e the sequence length at which 50% of all bases in the genome are in sequences of this length or shorter). Platanus assembly was chosen as the reference assembly (see Results and Table 5.1) and downstream analyses were performed on this assembly.

Table 5.1. Assembly features resulting from the different genome assemblers assayed (see text for details).

Assembler	Contamination (Blobtools)	*N50* (Kb)	CEGMA complete (%)	CEGMA partial (%)	Reference
Celera	-	18.7	62.9	66.5	Myers et al. (2000)
DISCOVAR	+	14.0	89.1	97.2	Weisenfeld et al. (2014)
Platanus	-	20.4	88.7	96.5	Kajitani et al. (2014)
SPAdes	+	15.4	91.9	95.6	Bankevich et al. (2012)

In order to perform the structural annotation of the assembly, GeneMark (Lomsadze et al. 2005) was initially used as an *ab initio* gene prediction software. After that, GeneMark predictions were provided to run MAKER2 v. 2.31.8 (Cantarel et al. 2008), which is an annotation pipeline that identifies repetitive elements, aligns expressed sequence tags (ESTs), and uses protein homology evidence to generate further gene predictions. MAKER2 was run twice. The first MAKER2 round was performed using SNAP (Korf 2004) and RepeatMasker v.3.0. (Smit et al. 1996). In order to create evidence-based gene model predictions, MAKER2 was provided in this first run with: (1) the general feature file obtained from CEGMA; (2) EST from *B. plicatilis* (NCBI database, 10-07-2016); (3) available transcriptomes from congeneric species (Table A.2); (4) Proteins from *Adineta*

vaga (a high-quality bdelloid rotifer genome; Flot et al. 2013); and (5) the Swiss-Prot database. The annotation output from MAKER2 first run was converted into a training set to run Augustus. Finally, the Augustus and MAKER2 (1st pass) outputs were provided to perform a second MAKER2 pass in order to obtain the final set of gene predictions.

The functional annotation of the final set of predicted genes was performed by identifying protein domains with InterProScan (Jones et al. 2014), and by performing Blastp searches (e-value< 10^{-5}; Altschul 1990) in the non-redundant metazoan subset of the NCBI's protein database. Finally, Blast2GO (Conesa et al. 2005) was run to retrieve Gene Ontology (GO) terms and annotation.

Brachionus plicatilis *genotyping by sequencing*

The 270 *B. plicatilis* clones established directly from diapausing egg banks in Chapter 4, from nine saline inland ponds located at the east of the Iberian Peninsula, were genotyped. For each clone, DNA extraction was performed with JETFLEX genomic DNA purification kit (GENOMED, Löhne, Germany), following manufacturer's protocol. After quality assessment and quantification (see above), samples were submitted to Cornell University Genomic Diversity Facility at the Institute of Biotechnology (Ithaca, NY, US), where libraries were constructed and sequenced.

GBS libraries were constructed and sequenced on an Illumina HiSeq 2000/2500 (100 bp, single-end) according to Elshire et al. (2011), using the restriction enzyme ApeK1 (GC[A-T]GC) for digestion. A library with unique barcodes for each clone, plus blank samples, was created.

SNPs were called from the raw DNA sequences using the GBS pipeline as implemented in TASSEL-GBS v2 (TASSEL 5) (Glaubitz et al. 2014), in which all reads were trimmed to the same length (64 bp) and identical reads were collapsed into tags (i.e reads putatively representing the same genome locations). These tags were then aligned against the reference genome using the Burrows-Wheeler alignment tool (BWA; Li and Durbin 2010) and SNPs were called from aligned tags. Pipeline default parameters were used except from two cases, in which more conservative values than the default were used: (1) at the minimum length of aligned base pair to store the sequence alignment/map (SAM) entry (SAMToGBSdbPlugin plugin, option aLen being 30 instead of default 0); and (2) at the minimum locus coverage (i.e. proportion of samples with a genotype; DiscoverySNPCallerPluginV2 plugin, option mnLCov being 0.8 instead of default 0.1).

The set of SNPs were quality filtered using custom developed scripts and VCFtools (Danecek 2011). At each genotype from

each sample, in order to obtain confidently assigned genotypes, those genotypes supported by less than six reads were excluded (i.e. considered as not genotyped). At the SNP level, a set of filters were applied in order to keep just those SNPs highly informative, of high quality and that were not over-merged repetitive regions. First, SNPs required at least 50% of the clones genotyped in each population. Second, minor allele frequency (MAF) had to be higher than 1%. Third, only two alleles were present. Fourth, the average read depth among individual had to be lower than 150 reads. Fifth, less than 60% of individuals had to be heterozygotes in each SNP.

To estimate genome-wide genetic variation and differentiation: (1) principal component analysis (PCA), a model-free multivariate ordination method, was implemented in the adegenet package (Jombart 2008) in R v4.1.1 (R Core Team 2009). Using the package hierfstat (Goudet 2005), (2) genetic divergence between populations was estimated by the fixation index (F_{ST}) for each pairwise comparison, and (3) within-population genetic diversity was measured as the average among loci expected heterozygosity (H_e), observed heterozygosity (H_o) and inbreeding coefficient (F_{IS}). Mantel test were used to test for isolation-by-distance patterns in IBDWS v. 3.23 (Jensen et al. 2005). In order to perform Mantel test, the geographic distance between ponds was calculated using Google Earth

(2008) and pairwise F_{ST} was used as genetic distance. The relationship between the H_e and the log-transformed pond size (Table 3.2) was tested using linear correlation analysis. File format conversions were performed using PGDSpider (Lischer and Excoffier 2012) and Plink (Purcell 2007, URL: http://pngu.mgh.harvard.edu/purcell/plink/).

To identify SNPs candidates to be the result of selection, two different approaches were used. Firstly, BayeScan 2.1 (Foll et al. 2008) was used to estimate the posterior probability that a given SNP is affected by selection. Briefly, prior odds of 10 were used for identifying the top candidates of the SNPs with a false discovery rate (FDR) of 0.05. Secondly, BAYENV (Coop et al. 2010) was used to test for correlations between (1) genetic differentiation and (2) phenotypic and environmental parameters. BAYENV uses a Bayesian approach that takes into account the population structure —from covariance matrices– when testing for correlations between environmental and genetic differentiation (Coop et al. 2010). To do so, 10 input files with 2000 randomly selected SNPs were created for the construction of each covariance matrix. The correlation with genetic variation was tested for two pond environmental parameters (environmental predictability and hydroperiod length), and two phenotypic parameters (life-history traits: propensity for sex and hatching fraction). Data on environmental predictability and hydroperiod length of

the ponds are from Chapter 3 (Tables 3.2 and 3.3 respectively), data on the life-history traits from the *B. plicatilis* clones are from Chapter 4. Each environmental or phenotypic parameter was standardized to mean equal to 0 and standard deviation equal to 1. Ten independent runs with different random seeds were run and the mean rank and standard deviation of each SNP between runs was computed. Based on their Bayesian factors (BF), a transformed rank statistic (scaled between 0 and 1; corresponding to the lowest and highest BF, respectively) was calculated and averaged between runs. SNPs were considered as potentially under selection when their mean rank was over 0.99.

In order to identify the genes associated to the putative SNPs under selection —as identified by BayeScan or BAYENV–, BEDtools (Quinlan and Hall 2010) was used to find genes in a flanking region of 0, 2.5 or 5 Kb upstream and downstream from the focus SNP. Due to the compact genome of *B. plicatilis* (Table 5.2), a high number of putative genes were found (Table A.3). Thus, additional analyses were performed using only those genes including a SNP in its coding region (i.e. 0 Kb). A GO enrichment analysis was conducted to test if certain gene ontologies were over- or under- represented in the lists of genes putatively under selection with respect to the genome using Blast2GO (Conesa et al. 2005).

Table 5.2. Summary for *Brachionus plicatilis* draft genome assembly and gene annotation

Number of scaffolds	14326
Total length (Mb)	108.5
Longest scaffold	169985
Scaffold %GC	26.4
Number of predicted genes	54725
Genes with functional annotation	16674

Results

Draft genome features

A total of 11.7 Gb of raw genomic sequence data were obtained. Based on k-mer statistics (k-mer length= 21; Table A.4), *B. plicatilis* RAS-1 genome was estimated to be 115.77 Mb with a heterozygosity of 0.65%. Overall, the best assembly features were found in the Platanus assembly (Table 5.1; highest N50, no evidence of contamination and high proportion in CEGMA completeness report). Therefore downstream analyses were performed using this assembly as reference. The assembly included 108.5 Mb of genomic sequence scaffolds with 50% of all bases in scaffolds longer

than 20.4 Kb (N50) and a maximum scaffold length of 169.9 Kb (Table 5.2). Genome completeness analysis by CEGMA found 88.7 and 96.4% of complete and partial sequences respectively of the set of core eukaryotic genes included in the analysis (Table 5.1). Finally, Blobtools did not find evidence of contamination based on G+C content and contig coverage (Table 5.1; Figure A.3).

The structural gene annotation yielded a predicted set of 54,725 gene models, from which functions were tentatively assigned to 30% of them.

GBS raw data, SNP calling and filtering

A total of 52 Gb of raw sequence data were obtained for the 270 clones from the nine field populations. Out of the 630,460,573 raw reads, 379,536,357 quality-filtered barcoded reads were obtained. The SNP calling pipeline yielded 12,856 SNPs, resulting in 4543 SNPs after filtering. Hardy-Weinberg proportions for each SNP and population showed that most SNPs do not differ from Hardy-Weinberg expectations (cases with *P*-value> 0.05= 95.6%; after Bonferroni correction 99.9%).

Genome-wide population differentiation

The overall expected heterozygosity (averaged over loci) was 0.17 and the within population expected heterozygosity ranged from 0.15 to 0.23 (Table 5.3). Heterozygosity correlated significantly to the log-transformed pond size (R^2= 0.40; t= 2.15,

d.f.= 7, *P*-value= 0.034). The overall F_{ST} (averaged over loci) was 0.18. Pairwise F_{ST} averaged over loci (Table 5.4) ranged from 0.07 (HYB and HYC) to 0.18 (SAL or HTU and HMT). PCA (Figure 5.1) clustered clones of the same populations with some overlap between populations. There was no significant pattern of isolation by distance (*r*= 0.31, *P*-value= 0.07, Mantel's test).

Table 5.3. Population values of heterozygosity (*Ho*), expected heterozygosity (*He*) and inbreeding index (F_{IS}) averaged over SNPs.

Population	H_o	H_e	F_{IS}
PET	0.23	0.23	-0.014
SAL	0.17	0.17	0.009
ATA	0.17	0.17	0.025
HYR	0.15	0.15	0.000
HYC	0.20	0.19	-0.017
CAM	0.17	0.18	0.045
HMT	0.15	0.15	-0.016
HYB	0.17	0.17	0.009
HTU	0.15	0.15	0.012

Table 5.4. Pairwise F_{ST} values averaged over SNPs.

	ATA	CAM	HYC	HMT	PET	HYR	SAL	HTU
CAM	0.13							
HYC	0.09	0.08						
HMT	0.16	0.15	0.10					
PET	0.13	0.07	0.08	0.13				
HYR	0.11	0.15	0.10	0.17	0.13			
SAL	0.13	0.12	0.10	0.18	0.12	0.15		
HTU	0.14	0.12	0.11	0.18	0.13	0.15	0.15	
HYB	0.09	0.14	0.07	0.14	0.13	0.08	0.13	0.12

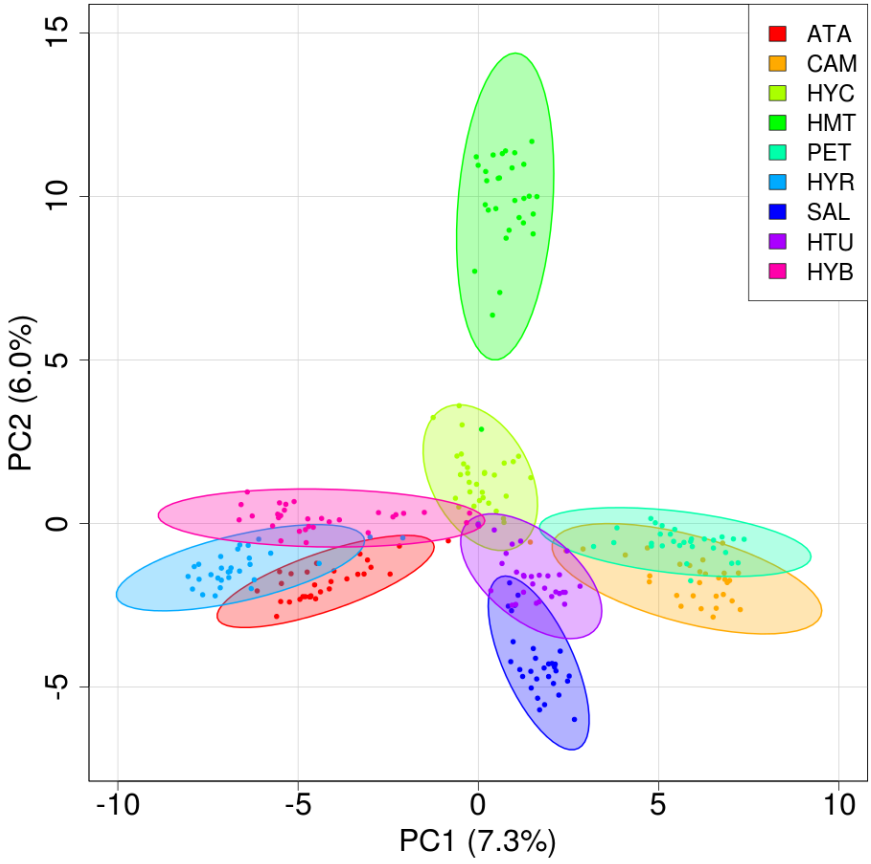


Figure 5.1. Principal component analysis based on 4,543 genome-wide SNPs for 30 clones (dots) from each of the nine *B. plicatilis* populations (acronyms on the upper-right side). Ellipsoids are the 95% confidence interval for each population. Percentage of variance explained by each principal component is shown between parentheses in the axis label.

Putative SNPs under selection

Using the BayeScan approach, 93 SNPs were identified as candidates to be under selection. From them, 12 are putatively under diversifying selection and 81 under balancing/purifying selection (Figure 5.2).

Using the BAYENV approach, significant correlation with environmental and phenotypic parameters was found for a number of SNPs ranging from 38 to 44 (average= 40.7 SNPs; Figures 5.3 and 5.4). Six (50%) of the SNPs detected by

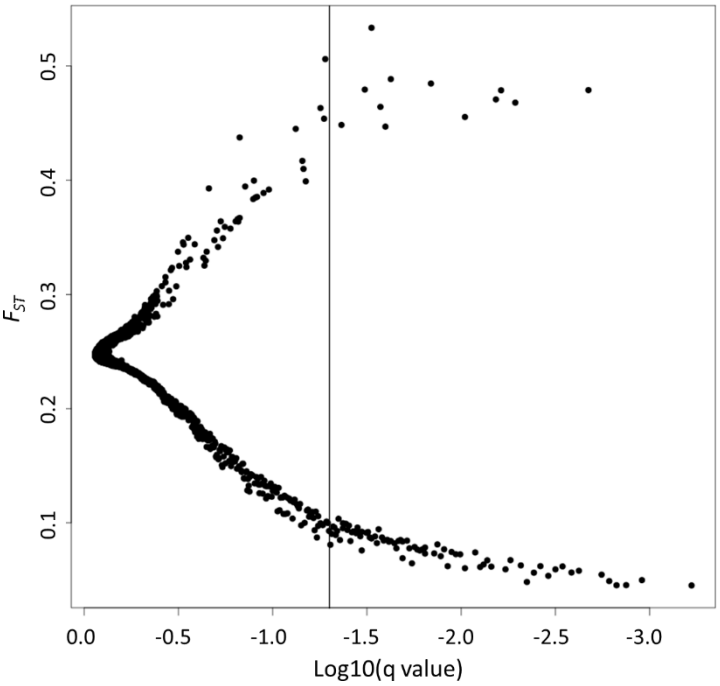


Figure 5.2. Relationship between F_{ST} and log10(*q* value) based on 4,543 genome-wide SNPs in nine rotifer field populations according to BayeScan. Adopting a *P-value*< 0.05 criterion, SNPs on the right side of the vertical line are putatively under selection.

BayeScan as being under diversifying selection were also found to be correlated to one of the parameters tested in BAYENV (three to hatching fraction, two to propensity for sex and one to environmental predictability; none of them being found in more than one parameter).

Genes associated to putative SNPs under selection

Most of the SNPs detected to be putatively under selection were located within genes (Table A.3). Out of the 12 SNPs that BayeScan detected under diversifying selection, 11 (91.7%) were located within 12 genes; the one that was not found within any gene was located at 1783 bp from the closest gene. Out of the 81 outlier SNPs under balancing/purifying selection, 43 (53.1 %) were located within 39 genes. Regarding BAYENV, the number of genes found in each parameter tested ranged from 28 to 43.

The 162 genes including SNPs identified under selection by BayeScan or BAYENV (Table A.3) had a broad range of gene ontology (GO; Table A.5). According to the enrichment analysis, the list of genes associated to propensity for sex was found to be enriched in histone acetyltransferase activity (GO: 0004402). For the rest of genes (associated to hatching fraction, environmental predictability, and hydroperiod length) no significant gene enrichment was found.

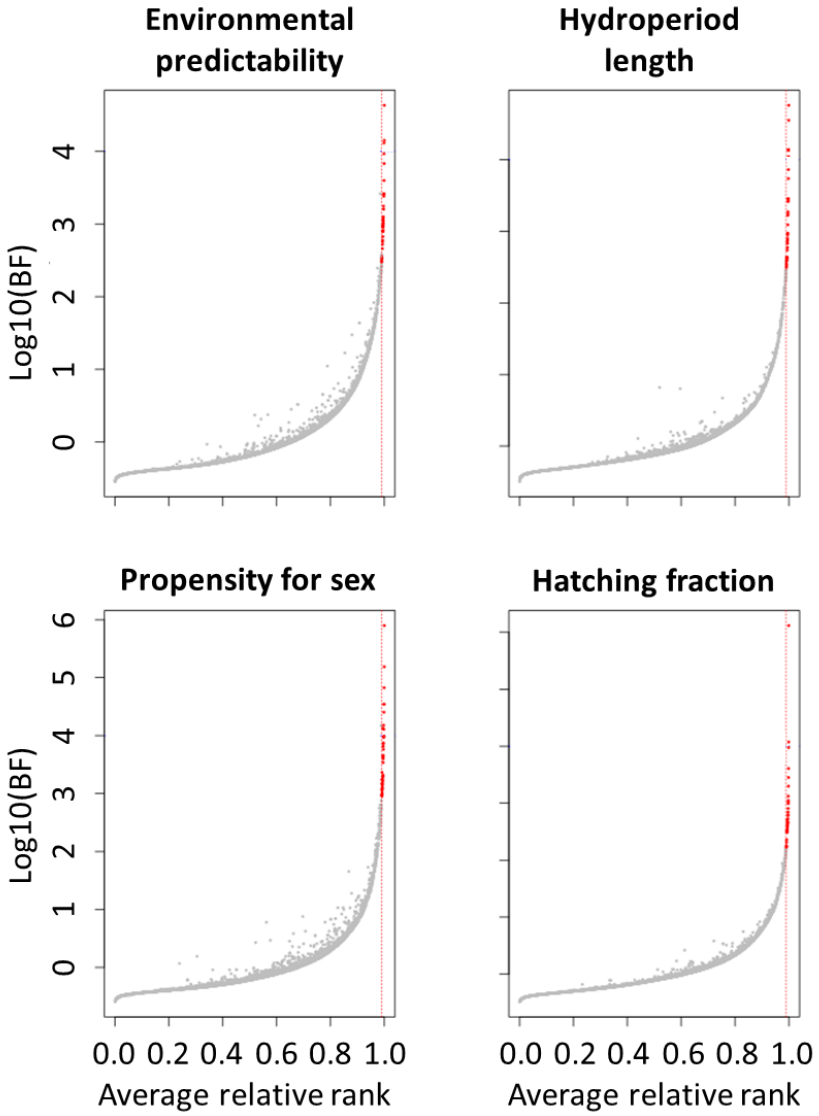


Figure 5.3. Bayesian factors (BF) and average relative rank for the correlation between SNPs (dots) and two environmental (upper row panels) and two phenotypic (lower row panels) parameters. Results are based on ten replicate runs of BAYENV. Vertical red line shows the average relative threshold (Average relative rank > 0.99) to consider a SNP to be outlier (red dots).

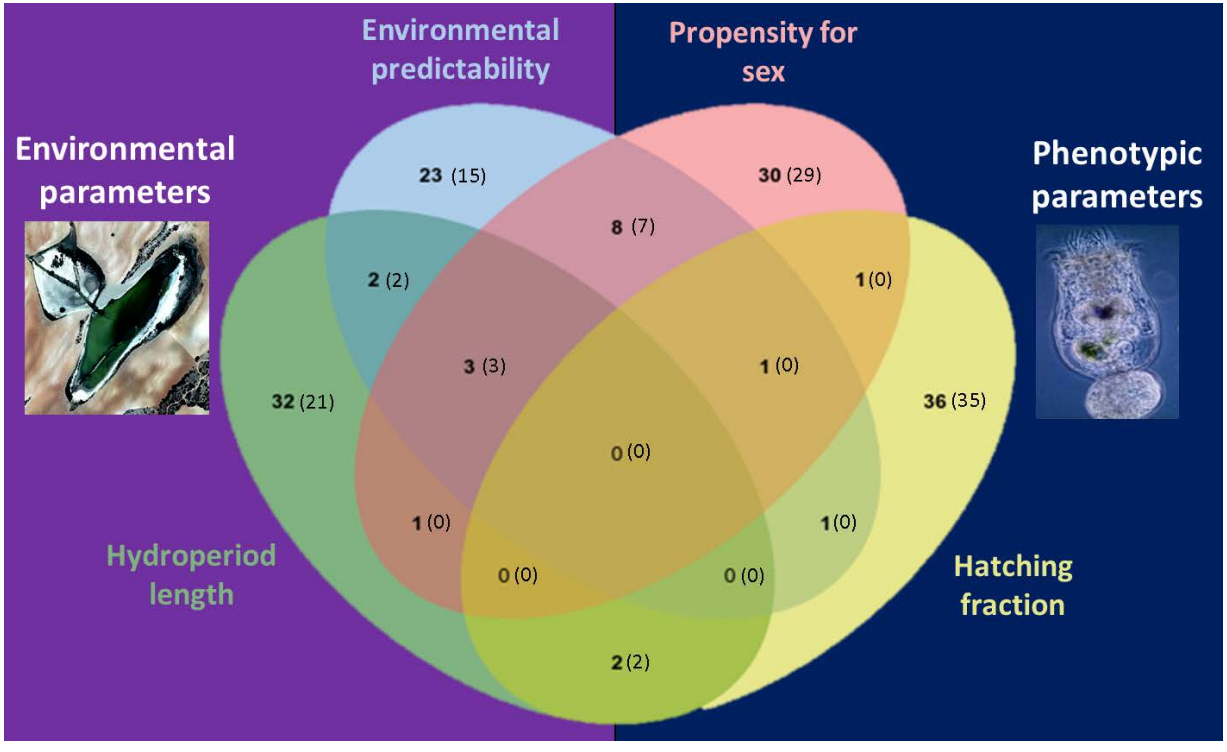


Figure 5.4. Venn diagram showing the number of SNPs putatively under selection that are related to each parameter according to BAYENV. The number of genes with at least a SNP in its coding regions is shown between parentheses.

Discussion

The integrated approach used here, combining environmental, phenotypic and genomic data, confirms previous evidence for selection in response to environmental predictability in rotifer populations (Chapter 4). This is a remarkable finding in populations located within a small geographical area (240 km^2). Six SNPs showed higher F_{ST} values than expected by chance and are correlated with phenotypic or environmental parameters. Therefore, they are very strong candidates for diversifying selection causing local adaptation. Additionally, 38 SNPs are correlated to environmental predictability, what points out that it is a relevant selective factor shaping rotifer populations. To our knowledge, this is the first genome-wide study aiming to unravel the selection footprints of environmental predictability.

The genome-wide patterns of genetic variation did not reveal a geographically-based population structure. In previous studies isolation-by-distance patterns have been reported in this species at a larger regional scale (Gómez et al. 2002a; Campillo 2011), but —consistently to what was found in the same geographical area by Montero-Pau et al. (2016)– no clear evidence was observed here. This is not unexpected because of the proximity of the studied populations. The

overall population differentiation (F_{ST}= 0.18) was slightly lower than the estimation obtained by Montero-Pau et al. (2016) with microsatellite markers (F_{ST}= 0.25) for the same geographical area. Consistently to what Montero-Pau et al. (2016) observed using mitochondrial data, a significant relation between genetic diversity (measured as estimated heterozygosity) and pond size was found, suggesting that pond size could be a reliable proxy for effective population size in *B. plicatilis* populations.

The results found here confirm that rotifers adapt to local conditions, as shown previously at medium (Campillo 2010; 29,000 km^2) and small (Chapter 4) regional scales, at least in a highly heterogeneous region as it is the case of the Mediterranean one. On one hand, it has been increasingly recognized that continental zooplankton populations are highly differentiated in neutral genetic markers, with strong phylogeographic structure (Campillo et al. 2011). On the other, studies had found that rotifer populations harbour a high degree of genetic variation, frequently associated to diapause traits (e.g. Carmona et al. 2009; Campillo et al. 2011; Gabaldón and Carmona 2015), and that genetic divergence among populations is correlated with environmental features (Campillo et al. 2011; Chapter 4), thus providing evidence for local adaptation. However, in these studies the genetic basis of local adaptation remained unknown. In this chapter,

several candidate genes underlying *B. plicatilis* local adaptation have been detected. Several studies have recently addressed the genomic signatures of local adaptation in different species of aquatic organisms. For example, Orsini et al. (2012) related several genes of *Daphnia magna* to a set of natural anthropogenic stressors using an experimental evolution approach, and found repeatable patterns of local adaptation. Another example is the studies of natural populations of sticklebacks (e.g. Guo et al. 2016; Hohenlohe et al. 2010), where several genomic regions and genes of relevance to local adaptation have been detected. The results presented here are in the line of the findings in these studies, and show how genome-wide methods can be applied to non-model species in order to understand the genomic basis of local adaptation, which is still poorly understood (Savolainen et al. 2013).

Overall, the correlation analysis (BAYENV) provided 163 SNPs; a much larger number of SNPs putatively under selection than the 12 SNPs (93 SNPs counting those putatively under balancing/purifying selection) provided by the F_{ST} method (BayeScan). This is not surprising given that the correlation methods are expected to have higher power for the type of data available here. First, correlation-based analysis is provided with additional information for each pond/population (environmental and phenotypic

parameters) which is not considered in the F_{ST} methods. This results in a more accurate hypothesis (genetic differentiation in relation to a parameter vs. genetic differentiation). Second, F_{ST} methods have low power when a high differentiation in neutral markers exists (De Kovel 2006; Butlin 2010). According to several authors (Gómez et al. 2002a; Montero-Pau et al. 2016; Campillo et al. 2009) and the results presented here (F_{ST}= 0.18), this is the case in *B. plicatilis* populations.

Half of the SNPs found to be putatively under diverging selection by BayeScan were also detected in BAYENV, so that they are very strong candidates to be under diverging selection. These SNPs are correlated to hatching fraction, propensity for sex and environmental predictability, so that this conservative restriction highlights the importance of these parameters in driving local rotifer adaptation. However, no SNP was found to be under selection by BayeScan and correlated simultaneously to environmental predictability and to propensity for sex. This is not surprising given the low statistical power of BayeScan already commented above.

According to the BAYENV, 12 SNPs (10 genes) were correlated simultaneously to propensity for sex and the degree of environmental predictability (Figure 5.4), the highest overlap found. A 31.6% of the SNPs correlated to the degree of predictability are also correlated to propensity for sex. This

provides a genetic base to the tight relation between these two parameters described in Chapter 4. In addition the weak relationship between the degree of predictability and the hatching fraction observed in Chapter 4 is consistent with the results at genomic level. Only two SNPs (no genes) are correlated to both parameters. This implies that only 5.3% of the SNPs associated to the degree of predictability are also associated to hatching fraction. Finally, hydroperiod is also confirmed to have low influence on propensity for sex once the influence of environmental predictability is discounted.

A high number of SNPs are correlated to environmental parameters (i.e. to hydroperiod length or to environmental predictability) and not to the phenotypic parameters (i.e. propensity for sex or hatching fraction), suggesting that the former acted as selective pressures on traits not considered here. Conversely, a high number of SNPs are correlated to phenotypic parameters and not to the environmental parameters. Obviously, not all the genetic variability associated to this fitness components is responding to selection by hydroperiod length and environmental predictability. Other selective pressures not considered here —such as the trophic status of the pond or salinity– could be acting. Nonetheless, by genetic tracking, genes causing divergence in life-history traits might be responding to short-term fluctuations that are not coupled among localities.

Genetic differentiation in *Daphnia* populations caused by this scenario has been reported (Lynch 1987). Interestingly, fluctuations of water-surface area —an important feature and a proxy for other features– are poorly correlated over the ponds studied (Chapter 3).

According to BAYENV, most of the SNPs correlated to any of the parameters studied here are located within genes, which showed a range of gene ontologies, so that reliable inferences are difficult. However, a specific insight on gene functionality was achieved for the genes associated to the propensity for sex parameter, because the set shows an enrichment of genes that were putatively associated to gene regulation processes (histones). Interestingly, both phenotypic plasticity and bet hedging have been proposed to be regulated by epigenetic mechanisms (Duncan et al. 2014; O'Dea et al. 2016). Undoubtedly, additional information on rotifer genome would refine the functionality of genes that are candidates to be targets of selection by the degree of predictability.

The genome of *B. plicatilis* assembled and annotated here showed several features worthy to stress. First, the estimated (haploid) size is 108.5 Mb, what is regarded as small. This estimation is similar to the ones in Stelzer et al. (2011) using flow-cytometry with different *B. plicatilis* strains (mean haploid

genome size= 118.9 Mb, range= 111.7-128.8 Mb). Second, the k-mer analysis revealed that the heterozygosity of the *B. plicatilis* genome is relatively high (0.65%). Likely, this made Platanus assembler to obtain better results than other assemblers, since it has been described to have a good performance when assembling genomes with heterozygosity higher than 0.5% (Kajitani et al. 2014). Moreover, genomic diversity favoured to detect a large number of high quality SNPs (4,543). Third, the genome is highly compact, with a high number of gene models (54,725). This compactness produced a high number of genes physically associated to the SNPs putatively under selection. Therefore the analysis had to be restricted to those genes where these SNPs were in their coding regions. This approach can potentially produce a loss of information, because inter-genic space changes can affect gene expression, for example by a modification of regulatory regions. Unfortunately, there is no information about the linkage disequilibrium in the *B. plicatilis* genome. This information is meaningful when defining the window size to search upstream or downstream of a SNP. Finally, the quality of the drafted genome assembly is high enough to achieve the objectives of this thesis. However, being *B. plicatilis* a common species in rotifer studies, it is desirable to improve the quality of the genome assembly, for instance, by the inclusion of both more sequencing data (such as mate-pairs or PacBio platform sequencing) and gene mapping information. For

instance, with an improved genome assembly, additional tests on the role of selection on the observed SNPs are possible. For the SNPs located in genes, it could be determined whether variation is synonymous or non-synonymous (Smith and Haigh 1974; Kaplan et al. 1989).

According to BayeScan, a remarkably high proportion of SNPs have signatures of balancing or purifying selection (F_{ST} lower than expected by chance). Not being included in the topic of this thesis, additional analysis of this data remains for future work. This work should be addressed to discriminate between balancing and purifying selection for the SNPs with significantly low F_{ST}, and to propose the selective mechanisms compatible with the observed patterns (e.g. heterozygote advantage, negative frequency-dependent selection, spatial or temporal habitat heterogeneity, genetic slippage, conflicting optimal fitness between life-cycle stages, weak negative selection; Fisher 1930; Haldane and Jayakar 1963; Clarke 1979; Lynch and Deng 1994; Patten and Haig 2009).

Admittedly, there is pending work to get a deeper understanding on how rotifer genome maps to adaptation to environmental factors though life-history traits. Most of this work depends on a research effort that is beyond what can be addressed in this thesis. Nevertheless, the results reported here are a first but important step in applying ecological

genomics to rotifers. This made possible to show relationship between genomic variation to diapause-related traits and environmental predictability, pointing out how efficiently local rotifer populations can adapt.

6

Final remarks and conclusions

An essential question in ecology is to understand how environments and organisms match (e.g. McIntosh 1986; Begon et al. 2006). Not surprisingly, this question crosses trough a bunch of scientific fields as geography (e.g. migration barriers), physiology (e.g. metabolic response to environmental factors), genetics (e.g. ecotypes), demography (e.g. population age structure), behavioural biology (e.g. habitat choices), etc. In investigating this central question evolutionary ecology focuses on how the living beings are shaped by selective pressures; that is, on adaptation (Collins 1986).

Investigating to what extent natural populations have diverged in response to local environmental conditions —i.e. local adaptation– is of great interest to evolutionary ecology (Kawecki and Ebert 2004; Fox and Wolf 2006; Blanquart et al. 2013). The emergence of locally adapted genotypes can produce the expansion of home ranges (Kirkpatrick and

Barton 1997). Moreover, adaptation to local environment also affects the ecological and evolutionary dynamics of biotic interactions (e.g. Kaltz and Shykoff 1998; Gandon and Michalakis 2002). In addition, the population differentiation associated to and reinforced by local adaptation processes is recognized by many theoretical models to have a crucial role in early speciation processes (Schluter 2001; Turelli et al. 2001; Via 2001). However, local adaptation is not the unique evolutionary output because other selective forces such as migration —especially at small geographical scales– or drift can counterbalance the effect of local selective pressures (Campillo et al. 2010; Roulin et al. 2015). Therefore, empirical studies are needed in order to know whether local adaptation does occur. Of particular interest is to determine if and how organism with huge dispersal capability adapt locally. In this thesis we provide evidence for local adaptation in populations of the rotifer *Brachionus plicatilis* at a small geographical scale (240 m^2). These adaptation patterns were associated to the degree of predictability, what highlights the ability of rotifer populations to locally adapt to time-varying environments. Strikingly, this adaptive response involves diapause-related life-history traits (Chapter 4). Adaptation to environmental fluctuations could be conceived as a difficult one, as selective pressure has to integrate rows of values for the relevant environmental factors. Chapters 4 and 5, by identifying diapause patterns as the target of selection, stress

its important consequences on the dynamics and persistence of rotifer populations. Several studies had already reported the high degree of genetic differentiation of rotifer populations in the eastern Iberian Peninsula (Gómez et al. 2002a; Campillo et al. 2009), and local adaptation signatures had been observed (Schröeder 2005; Campillo et al. 2010). However, before this thesis, the relationships between life-history traits related to diapause and the degree of environmental unpredictability in natural populations of aquatic organisms remained largely unknown.

At first glance, adaptation to local conditions is conceived as adaptation to the local average values of the environmental factors. Consequently, during the first half of the 20th century ecological populations were frequently conveniently conceived as inhabiting a homogeneous environment (McIntosh 1991). This conception does not mean that population biologists ignored that the environment fluctuates. Contrarily, the observation of the plant or animal life cycles and their association to seasonal changes (i.e. phenology) has been of great interest —especially for botanists– since the mid-18th century, when Linnaeus compiled annual data on leaf opening, flowering, fruiting, and leaf fall in '*Philosophia Botanica'* (1751). Nevertheless, frequently these complexities were dealt with by using verbal descriptions or quantitate surveys, rather than metrics able to test models. Indeed,

fluctuations were gradually incorporated to mathematical models in the second half of the 20th century (see Chapter 1; McIntosh 1991; Wiens 2000). In contrast, the assumption of environmental homogeneity affected mainly to empirical ecologists as a consequence of the methodological difficulties to deal with fluctuations, particularly when they are not periodic. Firstly, long time series are needed. As gathering this type of data is costly, these long time series are consequently scarce (May et al. 2007). Fortunately, remote sensing data can be especially helpful to expand temporally and spatially the knowledge of the focus environments (Chapter 1). In this thesis, remote sensing data —Landsat 5/7 scenes– have been found to be a timely, reliable, global and cost-efficient tool for analysing a given environmental variable over a long time series (Chapter 3). Remarkably, there is a promising perspective for remote sensing data in the next years, since new satellites have been launched that provide increased temporal and spatial resolutions (e.g. Landsat 8; Sentinel 2A and 2B). Once a long-time series is acquired, to quantify the relative importance of predictable and unpredictable components is a challenging task, since it implies modelling environmental variances; that is, modelling environment beyond the mean (Chapter 1), in contrast to the traditional and most prevalent statistical characterization used to think about (Kneib 2013). To make things more complex, unpredictability involves a time scale and is related

to the organism's ability to anticipate and adjust to a future environmental condition, as it was long time ago recognized by Robert MacArthur and Richard Levins (MacArthur and Levins 1964; Levins 1968; see Chapter 1). As a result, assumptions regarding this matter have to be made in order to quantify unpredictability. In Chapter 3, procedures oriented to quantify the degree of predictability of the habitat taking into account the point of view of the focus organism have been implemented and compared. Some model assumptions have been shown to have negligible effects, while others can be associated with the species assemblages for which predictability needs to be assessed. Discrepancy between the outputs of these models calls for caution and a careful assessment of the assumptions in relation of the species of interest. A cluster of models to quantify predictability was proposed as appropriate for aquatic invertebrates, and one of them was used to test correlations to life-history traits related to diapause in rotifers (Chapter 4). However, although procedures oriented to perform an educated guess about the point of view of the focus organism are provided in the present thesis, removing subjectivism in choosing relevant scales is still a pending task in the quantification of environmental unpredictability. This requires further study, where intense conceptual and philosophical analysis about the notion of unpredictability might be required.

From a first-order environmental characterization (e.g. expectation of factor values) to a second-order one (e.g. expectation based on squared values), the focus moves from environment average to environment variance. This has a number of relevant consequences since environmental fluctuations affect the fitness, which is no longer constant but varies through time; fitness in the following environmental state might be lower than before (Sæther and Engen 2015). In a fluctuating environment, several evolutionary responses are possible, each one implying different costs and needing different circumstances. If genetic change tracks the environment in a continuous change, then the mean phenotype lags behind the optimum. Although, phenotypic plasticity can be an alternative, it incurs in the inherent costs of both having large genetic equipment and passing through transient maladaptation if the adjustment to the environmental condition is consequent, rather than anticipatory. In these situations risk-avoiding strategies can evolve and be important to reduce fitness variance (Philippi and Seger 1989), increasing the geometric mean at the cost —again– of a decrease in the arithmetic mean (Gillespie 1974). Although bet hedging is rather new to the field of evolutionary ecology, its roots trace back to Bernoulli (1738) (reviewed in Stearns 2000). The current growing influence of bet hedging in this field is patent, but although its theoretical basis is well developed (Cohen 1966; Starrfelt and Kokko

2012), empirical evidence is still scarce (Childs et al. 2010; Simons 2011). Several causes can explain this situation. Firstly, as stressed above, accurate measurements for habitat predictability are challenging. Secondly, adversity and predictability are often confounded and can produce similar adaptive responses (García-Roger 2016). Thirdly, bet-hedger traits are sub-optimal in the short term and adaptive on the long term (i.e. the geometric fitness increases; Gillespie 1974), what makes them difficult to be recognized (Simons 2011). Finally, most studies on bet hedging typically focus on single traits (Childs et al. 2010), but bet hedging could work on one or more traits belonging to a set of traits having similar effects on buffering environmental variance (García-Roger et al. 2016); that is, if several traits have the same risk-avoidance role not all of them might evolve. Hence, studies focusing in just one trait can result in negative results, and in an underestimation of the importance of bet hedging in nature. This thesis overcomes some of these difficulties and contributes to the empirical evidence of bet hedging by relating two potentially bet-hedging life-history traits (propensity for sex and diapausing egg hatching fraction; Chapter 4) with a quantitative measure of environmental unpredictability (Chapter 3). Interestingly, we found that propensity for sex showed the expected correlation, while the hatching fraction, although showing intermediate values, was not related to the degree of environmental unpredictability

Evolutionary ecology was developed in the 1980s in the framework of a research program that tried to bring together population genetics, ecology and evolution (McIntosh 1986). Admittedly, this unifying program contributed to our understanding on how organisms adapt to their environments, and the population dynamics consequences of this adaptation. The *r-K* scheme (Pianka 1970) is just an instance. Nevertheless, many ecologists adopted the so-called 'adaptationist program'. This approach attempts to describe all aspects of living organisms as optimal adaptive solutions to problems set by both the environment and the biology of the species (Lewontin 1979). The adaptationist program focus on phenotypes; and neglects the complexities of the genetic dynamics as affected by epistasis, over-dominance, etc. Even being controversial (e.g. Gould and Lewontin 1979, Mayr 1983), it has been able to make qualitative predictions for the phenotype that is expectable in a given ecological scenario. Behavioural ecology is a robust field that credits to this program. Nowadays, unification of ecology and evolution runs on other courses. One of them is the recognition that ecological dynamics and evolutionary dynamics overlap their time scales (e.g. special issue 364 of the *Philosophical transactions of the Royal Society B* or the volume 21, issue 3 of *Functional Ecology*). This brings, for instance, to assume that, on one side, parameters in the ecological models are under evolutionary change (i.e.

varying with time) and, on the other, that evolutionary equations (e.g. Price equation, Price 1970; 1972; Frank 1997) have to incorporate ecological features (Ellner et al. 2011; Kerr and Godfrey-Smith 2008; Coulson et al. 2010; Collins and Gardner 2009; Fox and Harpole 2008). Other of them is that, in the '-omics' era, it becomes approachable to map genotypes into phenotypes and phenotypes into environments, even using non-model organism. Rotifers and its adaptation to unpredictability provide good instances of these two issues (Tarazona et al. 2017; Chapters 4 and 5). The use of next-generation sequencing (NGS) methods shows how advance in knowledge is promoted from technical advance. This thesis has been benefited also from NGS, which allowed obtaining a large amount of genomic data from a non-model species (Chapter 5). Here it is shown that the genetic basis of adaptation to different environments can be directly disentangled at the genome level, without any prerequisites about the selectively advantageous genes or traits. State-of-the-art bioinformatics use background variation in the genome as a null hypothesis for detecting genome regions under direct and indirect selection (Storz 2005; Savolainen et al. 2013; Coop et al. 2010). Nevertheless, bioinformatic analysis is empowered if genomic information is combined with both environmental data (Chapter 3) and life-history trait data (Chapter 4). This enables to identify genes putatively associated to local adaptation processes. This is

relevant since the effect of local adaptation at the genomic level remains poorly understood (Savolainen et al. 2013). Fortunately, the studies on the genomic basis of local adaptation in aquatic species such as *Daphnia magna* (Orsini et al. 2012) or sticklebacks (e.g. Guo and Merila 2016; Hohenlohe et al. 2010) are becoming more frequent and contributing to expand our knowledge of genomics of local adaptation to non-model organisms. However, gene functional annotations based on experimental data for non-model organism in databases are still scarce (Baric et al. 216). Thus, genes from NGS data are mostly annotated based on their similarity to genes described in model organisms. This leads to a low fraction of genes functionally annotated —in *B. plicatilis* 30% (Chapter 5)– and to vague functional annotations (Orsini et al. 2012; Tagu et al. 2014). Functional annotation is essential since it is what connects nucleotide sequences to the biology of organisms (Stein 2001). Despite the scientific community is making an effort to improve the functional information available in the databases, it is still lagging behind the great amount of data obtained through the new NGS approaches (Tagu et al. 2014; Baric et al. 2016). To this regard, it is worth noting that biology is to a large extent a science of diversity. In order to achieve broader conclusions beyond a few non-representative model organisms, ecological genomics needs to be expanded towards non-model organisms.

The appearance of NGS technologies is providing a large amount of data and generating enormous databases (Tagu et al. 2014). This has increased the interest on explorative surveys (Straalen and Roelofs 2006). Although most of this thesis is hypothesis-driven, some of the findings in Chapter 5 are instances of discovery-driven research. Descriptive studies are frequently the starting point of new sciences —such as ecological genomics– and, although this is frequently associated to pejorative connotations and has several limitations (Casadevall and Fang 2008), an inductive phase is essential to the generation of new hypothesis (Straalen and Roelofs 2006). In Chapter 5 our hypothesis and predictions were to found SNPs with high F_{ST} values (i.e. highly differentiated between populations). However, the exploration of the data revealed genes having SNPs with low F_{ST} that were found unintentionally. The role of these SNPs in rotifer adaptation is one of the questions that remains open.

Our study on rotifers stresses that the persistence of natural populations may depend on the evolution of bet hedging in key life-history traits (Childs et al. 2010; Simons 2011; Gremer and Venable 2014). This type of conclusions is becoming of major applied relevance under the predicted scenarios of increased environmental variability in the near future (IPCC 2013). Thus, a comprehensive understanding of the role of

bet-hedging strategies is necessary for predicting population responses to environmental change (Lawson et al. 2015).

Conclusions

The main conclusions derived from this thesis are enumerated below:

1. The scenes from the satellites Landsat 5 and 7 —after applying a combination of procedures to discriminate water from background– provided a long time series (27 years) of the variation in water-surface area of a group of twenty Mediterranean saline water bodies.

2. Following the conception that predictability depends on the point of view of the focus organism and using Colwell metrics, different models for predictability estimation were developed here. Furthermore, GAM fitting was developed in this thesis as an alternative continuous approach for measuring predictability. All these models were assessed by considering how the variation in water-surface area could be relevant for the focus organism.

3. The application of the predictability metrics allowed quantifying predictability in a group of Mediterranean ponds. These ponds showed a wide range of predictability.

4. *A posteriori* classification of the models for predictability estimation showed that some assumptions had negligible effects, while others can be associated with the species assemblages for which predictability needs to be assessed.

5. *Brachionus plicatilis* populations inhabiting a set of nine Mediterranean saline ponds showed significant levels of within-population genetic variation for propensity for sex and for hatching fraction of diapausing eggs.

6. The propensity for sex in rotifer populations, and hence the early investment in diapause, decreased with environmental predictability, while the relationship of that trait with hydroperiod length was relatively weak. This suggests a conservative, bet-hedging strategy that provides protection against unexpectedly short growing seasons.

7. Diapausing egg hatching fractions had intermediate values (from 44 to 88%) in all the studied populations, but hatching fractions were neither related to the level of environmental predictability nor to hydroperiod length.

8. Rotifer populations are able to locally diverge in diapause-related traits within a small geographical range (240 km^2) despite their potential for widespread genetic exchange through the passive dispersal of diapausing eggs.

9. The *B. plicatilis* genome was assembled in this thesis. Its structural annotation yielded 54,725 predicted genes. Functions were tentatively assigned to 30% of them

10. Genotyping by sequencing (GBS) and the subsequent bioinformatics analyses provided a large number (4,543) of high quality single nucleotide polymorphisms (SNPs).

11. A number of SNPs —most of them located within genes– showed higher between-population differentiation than expected by chance and were correlated with life-history traits and environmental factors, so that they are candidates for diversifying selection for local adaptation.

12. Unexpectedly, a large set of SNPs, more than half of them located within genes were found to present signatures of balancing/purifying selection in *B. plicatilis*. This finding requires further research.

13. A number of genes were identified as strong candidates to be part of the genomic basis of local adaptation to fluctuating environments. These genes constitute a database for future studies.

Literature cited

Adams, K.D. & Sada, D.W. (2014). Surface water hydrology and geomorphic characterization of a playa lake system: implications for monitoring the effects of climate change. *Journal of Hydrology*, **510**, 92–102.

Alcántara-Rodríguez, J., Ciros-Pérez, J., Ortega-Mayagoitia, E., Serrania-Soto, E. & Piedra-Ibarra, E. (2012). Local adaptation in populations of a *Brachionus* group *plicatilis* cryptic species inhabiting three deep crater lakes in Central Mexico. *Freshwater Biology,* **57**, 725–740.

Alonso, M. (1998). Las lagunas de la España peninsular. *Limnetica,* **15**, 1–176.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403–410.

Angeler, D.G., Álvarez-Cobelas, M., Rojo, C. & Sánchez-Carrillo, S. (2000). The significance of water inputs to plankton biomass and trophic relationships in a semi-arid freshwater wetland (central Spain). *Journal of Plankton Research*, **22**, 2075–2093.

Antón-Pardo, M. & Armengol, X. (2010). Zooplankton community from restored peridunal ponds in the Mediterranean region (L'Albufera Natural Park, Valencia, Spain). *Limnetica*, **29**, 133–144.

Aparici, E., Carmona, M.J. & Serra, M. (2001). Variability for mixis initiation in *Brachionus plicatilis*. *Hydrobiologia,* **446/447**, 45–50.

Aparici, E., Carmona, M. J., & Serra, M. (2002). Evidence for an even sex allocation in haplodiploid cyclical parthenogens. *Journal of Evolutionary Biology*, **15**, 65–73.

Armengol, X., Boronat, L., Camacho, A. & Wurtsbaugh, W.A. (2001). Grazing by a dominant rotifer *Conochilus unicornis* Rousselet in a mountain lake: in situ measurements with synthetic microspheres. *Hydrobiologia*, **446/447,** 107–114.

Asner, G.P. (2009). Tropical forest carbon assessment: integrating satellite and airborne mapping approaches. *Environmental Research Letters*, **4**, 34009.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S. *et al.* (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, **19**, 455–477.

Baric, R.S., Crosson, S., Damania, B., Miller, S.I. & Rubin, E.J. (2016). Next-generation high-throughput functional annotation of microbial genomes. *mBio*, **7**, e01245-16.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**, 1–48.

Beaumont, M.A. & Balding, D.J. (2004). Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13**, 969–980.

Becks, L. & Agrawal, A.F. (2012). The evolution of sex is favoured during adaptation to new environments. *PLoS Biology*, **10,** e1001317.

Begon, M., Harper, J. & Townsend, C. (2006). *Ecology: from individuals to ecosystems*, 4th edn. Blackwell Publishing Ltd.

Bernoulli, D. (1738). Specimen theoriae novae de mensura sortis. *Classe Mathematica*.

Berryman, A.A. (1999). *Principles of population dynamics and their application* (ed. Stanley Thornes) Cheltenham, UK.

Blanchet, F.G., Legendre, P. & Borcard, D. (2008). Modelling directional spatial processes in ecological data. *Ecological Modelling*, **215**, 325–336.

Blanquart, F., Kaltz, O., Nuismer, S. & Gandon, S. (2013). A practical guide to measuring local adaptation. *Ecology letters*, **16**, 1195–1205.

Botero, C.A., Weissing, F.J., Wright, J. & Rubenstein, D.R. (2015). Evolutionary tipping points in the capacity to adapt to environmental change. *Proceedings of the National Academy of Sciences*, **112**, 184–189.

Bradnam, K.R., Fass, J.N., Alexandrov, A., *et al.* (2013). Assemblathon 2: evaluating de novo methods of genome assembly in three vertebrate species. *GigaScience*, **2**, 10.

Brendonck, L. & De Meester, L. (2003). Egg banks in freshwater zooplankton: Evolutionary and ecological archives in the sediment. *Hydrobiologia*, **491**, 65–84.

Brown, J.H. (1997). An ecological perspective on the challenge of complexity. *EcoEssay*, **1**. National Center for Ecological Analysis and Synthesis, www.nceas,ucsb.edu.

Brown, J.S. & Venable, D.L. (1986). Evolutionary ecology of seed-bank annuals in temporally varying environments. *The American Naturalist*, **127**, 31.

Bryant, R.G. & Rainey, M.P. (2002). Investigation of flood inundation on playas within the Zone of Chotts, using a time-series of AVHRR. *Remote Sensing of Environment*, **82**, 360–375.

Butlin, R.K. (2010). Population genomics and speciation. *Genetica*, **138**, 409–418.

Campbell, J. (2002). *Introduction to remote sensing*, 3rd edn. Guilford Press, New York.

Campillo, S. (2009). *The dispersal-gene flow paradox in zooplankton: the case of rotifers dwelling in temporary rotifers*. Tesis doctoral, Universitat de València, Valencia.

Campillo, S., García-Roger, E.M., Carmona, M.J., Gómez, A. & Serra, M. (2009). Selection on life-history traits and genetic population divergence in rotifers. *Journal of Evolutionary Biology*, **22**, 2542–2553.

Campillo, S., García-Roger, E.M., Carmona, M.J. & Serra, M. (2010). Local adaptation in rotifer populations. *Evolutionary Ecology*, **25**, 933–947.

Campillo, S., García-Roger, E.M., Martínez-Torres, D. & Serra, M. (2005). Morphological stasis of two species belonging to the L-morphotype in the *Brachionus plicatilis* species complex. *Hydrobiologia*, **546**, 181–187.

Campillo, S., Serra, M., Carmona, M.J. & Gómez, A. (2011). Widespread secondary contact and new glacial refugia in the

halophilic rotifer *Brachionus plicatilis* in the Iberian Peninsula. *PLoS ONE*, **6**.

Cantarel, B.L., Korf, I., Robb, S.M., Parra, G., Ross, E., Moore, B., *et al.* (2008). MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research*, **18**, 188–196.

Carmona, M.J., Dimas-Flores, N., García-Roger, E.M. & Serra, M. (2009). Selection of low investment in sex in a cyclically parthenogenetic rotifer. *Journal of Evolutionary Biology*, **22**, 1975–1983.

Carmona, M.J., Dimas-Flores, N., Montero-Pau, J. & Serra, M. (2011). Effect of experimental methodology on estimation of density at sex initiation in cyclically parthenogenetic rotifers. *Hydrobiologia*, **662**, 131–139.

Carmona, M.J., Gómez, A. & Serra, M. (1995). Mictic patterns of the rotifer *Brachionus plicatilis* Müller in small ponds. *Hydrobiologia*, **313/314**, 365–371.

Carmona, M.J., Serra, M. & Miracle, M.R. (1993). Relationships between mixis in *Brachionus plica*tilis and preconditioning of culture medium by crowding. *Hydrobiologia*, **255**/**256**, 145–152.

Carrillo, P., Cruz-Pizarro, L. & Morales, R. (1987). Cambios estacionales en las comunidades de fitoplancton y de zooplancton de la Albufera de Adra. *Limnetica*, **3**, 243–254.

Casadevall, A. & Fang, F. (2008). Descriptive science. *Infection and Immunity*, **76**, 3835–3836.

Chao Rodríguez, Y., el Anjoumi, A., Domínguez Gómez, J.A., Rodríguez Pérez, D. & Rico, E. (2014). Using Landsat image time series to study a small water body in Northern Spain. *Environmental monitoring and assessment*, **186**, 3511–22.

Chevin, L.-M., Lande, R. & Mace, G.M. (2010). Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLoS Biology*, **8**, e1000357.

Childs, D.Z., Metcalf, C.J.E. & Rees, M. (2010). Evolutionary bet-hedging in the real world: empirical evidence and challenges revealed by plants. *Proceedings of the Royal Society of London B: Biological Sciences*, **277**, 3055–3064.

Chown, S., Hoffmann, A., Kristensen, T., Angilletta, M.J., Stenseth, N.C. & Pertoldi, C. (2010). Adapting to climate change: a perspective from evolutionary physiology. *Climate Research*, **43**, 3–15.

Ciros-Pérez, J., Carmona, M.J., Lapesa, S., & Serra, M. (2004). Predation as a factor mediating resource competition among rotifer sibling species. *Limnology and Oceanography*, **49**, 40–50.

Ciros-Pérez, J., Gómez, A. & Serra, M. (2001). On the taxonomy of three sympatric sibling species of the *Brachionus plicatilis* (Rotifera) complex from Spain, with the description of *B. ibericus* n. sp. *Journal of Plankton Research*, **23**, 1311–1328.

Cirujano, S. (1981). Las lagunas manchegas y su vegetación. II. *Anales del Jardín Botánico de Madrid*, **38**, 187–232.

Cirujano, S., Montes, C. & García, L. (1988). Los humedales de la provincia de Albacete: una panorámica general. *Al-Basit: Revista de estudios albacetenses*, **24**, 77–95.

Clarke, B.C. (1979). The evolution of genetic diversity. *Proceedings of the Royal Society of London B: Biological Sciences*, **205**, 453–474.

Cohen, D. (1966). Optimizing reproduction in a randomly varying environment. *Journal of Theoretical Biology*, **12**, 119–129.

Collins, J. (1986). Evolutionary ecology and the use of natural selection in ecological theory. *Journal of the History of Biology*, **19**, 257–288.

Collins, S. & Gardner, A. (2009). Integrating physiological, ecological and evolutionary change: a Price equation approach. *Ecology Letters*, **12**, 744–757.

Colwell, R.K. (1974). Predictability, constancy, and contingency of periodic phenomena. *Ecology*, **55**, 1148–1153.

Colwell, R.N., Ųaby, F.T., Simonett, D.S., Estes, J.E. & Thorley, G.A. (1983). *Manual of remote sensing. Volume 2. Interpretation and applications,* 2nd edn. American Society of Photogrammetry, Falls Church, Virginia, USA.

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M. & Robles, M. (2005). Blast2GO: A universal annotation and

visualization tool in functional genomics research. Application note. *Bioinformatics*, **21**, 3674–3676.

Coop, G., Witonsky, D., Di Rienzo, A. & Pritchard, J.K. (2010). Using environmental correlations to identify loci underlying local adaptation. *Genetics*, **185**, 1411–1423.

Cott, P.A., Sibley, P.K., Somers, W.M., Lilly, M.R. & Gordon, A.M. (2008). A review of water level fluctuations on aquatic biota with an emphasis on fishes in ice-covered lakes. *Journal of the American Water Resources Association*, **44**, 343–359.

Coulson, T., Tuljapurkar, S. & Childs, D.Z. (2010). Using evolutionary demography to link life history theory, quantitative genetics and population ecology. *Journal of Animal Ecology*, **79**, 1226–1240.

Cousyn, C., De Meester, L., Colbourne, J.K., Brendonck, L., Verschuren, D. & Volckaert, F. (2001). Rapid, local adaptation of zooplankton behavior to changes in predation pressure in the absence of neutral genetic changes. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 6256–60.

Crozier, L.G., Hendry, A.P., Lawson, P.W., Quinn, T.P., Mantua, N.J., Battin, J., *et al.* (2008). Potential responses to climate change in organisms with complex life histories: evolution and plasticity in Pacific salmon. *Evolutionary Applications*, **1**, 252–270.

Dahms, H.U., Hagiwara, A. & Lee, J.S. (2011). Ecotoxicology, ecophysiology, and mechanistic studies with rotifers. *Aquatic Toxicology*, **101**, 1–12.

Danecek, P., Auton, A., Abecasis, G., Cornelis, A.A., Banks, E., DePriesto, M.A., *et al.* (2011). The variant call format and VCFtools. *Bioinformatics*, **27**, 2156–2158.

Danks, H. V. (1987). Insect dormancy: an ecological perspective (Vol. 1, pp. 439-439). Biological Survey of Canada (Terrestrial Artropods).

Dedryver, C.A., Hullé, M., Le Gallic, J.F., Caillaud, M.C. & Simon, J.C. (2001). Coexistence in space and time of sexual and asexual populations of the cereal aphid *Sitobion avenae*. *Oecologia*, **128**, 379–388.

Deorowicz, S., Kokot, M., Grabowski, S. & Debudaj-Grabysz, A. (2015). KMC 2: fast and resource-frugal k-mer counting. *Bioinformatics*, **31**, 1569–1576.

Duncan, E.J., Gluckman, P.D. & Dearden, P.K. (2014). Epigenetics, plasticity, and evolution: How do we link epigenetic change to phenotype? *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, **322**, 208–220.

Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, **26**, 2460–2461.

Ellner, S.P. (1997). You bet your life: life-history strategies in fluctuating environments. *Case studies in mathematical modeling: ecology, physiology and cell biology* (eds. H. Othmer, F.R. Adler, M. Lewis & J.C. Dallon), pp. 3–24. Prentice Halle, New Jersey.

Ellner, S.P., Geber, M.A. & Hairston, N.G. (2011). Does rapid evolution matter? Measuring the rate of contemporary evolution and its impacts on ecological dynamics. *Ecology Letters*, **14**, 603–614.

Ellner, S.P., Hairston, N.G. & Babaï, D. (1998). Long-term diapause and spreading of risk across the life cycle. *Archiv für Hydrobiologie*, **52**, 297–312.

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S. *et al.* (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, **6**, e19379.

Fay, D.S. (2013). *A biologist's guide to statistical thinking and analysis* (ed. The *C. elegans* Research Community). WormBook. https://www.ncbi.nlm.nih.gov/books/NBK153593/

Fisher, R.A. (1999). *The genetical theory of natural selection* (ed. J. H. Bennett,). *O*xford University Press, UK.

Flatt, T. & Heyland, A. (2011). *Mechanisms of life history evolution: the genetics and physiology of life history traits and trade-offs* (eds.T. Flatt and A. Heyland). Oxford University Press, Oxford and New York.

Florencio, M., Díaz-Paniagua, C. & Serrano, L. (2016). Relationships between hydroperiod length, and seasonal and spatial patterns of beta-diversity of the microcrustacean assemblages in Mediterranean ponds. *Hydrobiologia*, **774**, 109–121.

Flot, J.-F., Hespeels, B., Li, X., Noel, B., Arkhipova, I., Danchin, E.G.J., *et al.* (2013). Genomic evidence for ameiotic evolution in the bdelloid rotifer *Adineta vaga*. *Nature*, **500**, 453–457.

Foll, M. & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, **180**, 977–993.

Fontaneto, D., I. Giordani, G. Melone, & M. Serra, 2007. Disentangling the morphological stasis in two rotifer species of the *Brachionus plicatilis* species complex. *Hydrobiologia,* 583: 297–307.

Fontaneto, D. & Jondelius, U. (2011). Broad taxonomic sampling of mitochondrial cytochrome c oxidase subunit I does not solve the relationships between Rotifera and Acanthocephala. *Zoologischer Anzeiger - A Journal of Comparative Zoology*, **250**, 80–85.

Fontaneto D., Tang C.Q., Obertegger U., Leasi F. & Barraclough T.G., (2012). Different diversification rates between sexual and asexual organisms. *Evolutionary Biology*, **39,** 262–270.

Fox, J.W. & Harpole, W.S. (2008). Revealing how species loss affects ecosystem function: the trait-based Price equation partition. *Ecology*, **89**, 269–279.

Fox, C. & Wolf, J. (2006). *Evolutionary genetics: concepts and case studies* (*eds.* Charles W. Fox and Jason B. Wolf*)* Oxford University Press, New York.

Frantz, A., Plantegenest, M. & Simon, J.-C. (2006). Temporal habitat variability and the maintenance of sex in host populations of the pea aphid. *Proceedings of the Royal Society of London B: Biological Sciences*, **273**, 2887–2891.

Frazier, P.S. & Page, K.J. (2000). Water body detection and delineation with Landsat TM data. *Photogrammetric Engineering Remote Sensing*, **66**, 1461–1467.

Frisch, D., Green, A.J. y Figuerola, J. 2007. High dispersal capacity of a broad spectrum of aquatic invertebrates via waterbirds. *Aquatic Sciences,* **69**, 568–574.

Gabaldón, C. & Carmona, M.J. (2015). Allocation patterns in modes of reproduction in two facultatively sexual cryptic rotifer species. *Journal of Plankton Research*, **37**, 429–440.

Gabaldón, C., Fontaneto, D., Carmona, M.J., Montero-Pau, J. & Serra, M. (2016). Ecological differentiation in cryptic rotifer species: what we can learn from the *Brachionus plicatilis* complex. *Hydrobiologia, in press*.

Gabaldón, C., M. Serra, M. J. Carmona, & J. Montero-Pau, 2015a. Life-history traits, abiotic environment and coexistence: The case of two cryptic rotifer species. *Journal of Experimental Marine Biology and Ecology,* **465,** 142–152

Gabaldón, C., Montero-Pau J., Carmona M. J., & Serra M., 2015b. Life-history variation, environmental fluctuations and competition in ecologically similar species: modeling the case of rotifers. *Journal of Plankton Research,* **37,** 953–965

Gabaldón, C., Montero-Pau, J., Serra, M. & Carmona, M. (2013). Morphological similarity and ecological overlap in two rotifer species. *PLoS ONE,* **8**, e57087.

Gallart, F., Prat, N., García-Roger, E.M., *et al.* (2012). A novel approach to analysing the regimes of temporary streams in relation to their controls on the composition and structure of aquatic biota. *Hydrology and Earth System Sciences*, **16**, 3165–3182.

Gandon, S. & Michalakis, Y. (2002). Local adaptation, evolutionary potential and host–parasite coevolution: interactions between migration, mutation, population size and generation time. *Journal of Evolutionary Biology*, **15**, 451–462.

García-Ferrer, I., Camacho, A., Armengol, X. & Miracle, M. (2003). Seasonal and spatial heterogeneity in the water chemistry of two sewage-affected saline shallow lakes from central Spain. *Hydrobiologia,* **506/509**, 101–110.

García-Roger, E. (2006). *Análisis demográfico de bancos de huevos diapáusicos de rotíferos.* Tesis doctoral, Universitat de València, Valencia.

García-Roger, E.M., Carmona, M.J. & Serra, M. (2005). Deterioration patterns in diapausing egg banks of *Brachionus* (Müller, 1786) rotifer species. *Journal of Experimental Marine Biology and Ecology*, **314**, 149–161.

García-Roger, E., Carmona, M.J. & Serra, M. (2006). Patterns in rotifer diapausing egg banks: density and viability. *Journal of Experimental Marine Biology and Ecology*, **336**, 198–210.

García-Roger, E.M., Carmona, M.J. & Serra, M. (2016). Modes, mechanisms and evidence of bet hedging in rotifer diapause traits. *Hydrobiologia, in press.*

García-Roger, E.M., Serra, M. & Carmona, M.J. (2014). Bet-hedging in diapausing egg hatching of temporary rotifer populations - A review of models and new insights. *International Review of Hydrobiology*, **99**, 96–106.

García, C., García-Ruiz, R., Rendón, M. & Niell, F. (1997). Hydrological cycle and interannual variability of the aquatic community in a temporary saline lake (Fuente de Piedra, Southern Spain). *Hydrobiologia,* **345**, 131–141.

Gilbert, J.J. (1974). Dormancy in Rotifers. *Transactions of the American Microscopical Society*, **93**, 490–513.

Gilbert, J.J. & Diéguez, M.C. (2010). Low crowding threshold for induction of sexual reproduction and diapause in a Patagonian rotifer. *Freshwater Biology*, **55**, 1705–1718.

Gillespie, J.H. (1974). Natural selection for within-generation variance in offspring number. *Genetics*, **76**, 601–606.

Ginzburg, L.R. (1987). The theory of population dynamics: back to first principles, (eds. Teramoto, Ei and Yamaguti Masaya), pp. 70–79. Springer Berlin Heidelberg.

Glaubitz, J.C., Casstevens, T.M., Lu, F., Harriman, J., Elshiere, R.J., Sun, Q. *et al.* (2014). TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PLoS ONE*, **9**, e90346

Gómez, A. (2005). Molecular ecology of rotifers: from population differentiation to speciation. *Hydrobiologia*, **546**, 83–99.

Gómez-Alday, J.J., Carrey, R., Valiente, N., Otero, N., Soler, A., Ayora, C., *et al.* (2014). Denitrification in a hypersaline lake-aquifer system (Pétrola Basin, Central Spain): the role of recent organic matter and Cretaceous organic rich sediments. *The Science of the total environment*, **497**/**498**, 594–606.

Gómez, A., Carmona, M.J. & Serra, M. (1997). Ecological factors affecting gene flow in the *Brachionus plicatilis* complex (Rotifera). *Oecologia*, **111**, 350–356.

Gómez, A. & Carvalho, G.R. (2000). Sex, parthenogenesis and genetic structure of rotifers: microsatellite analysis of

contemporary and resting egg bank populations. *Molecular Ecology*, **9**, 203–214.

Gómez, A., Carvalho, G. & Lunt, D. (2000). Phylogeography and regional endemism of a passively dispersing zooplankter: mitochondrial DNA variation in rotifer resting egg banks. *Proceedings of the Royal Society of London B: Biological Sciences,* **267**, 2189–97.

Gómez, A., Montero-Pau, J. & Lunt, D. (2007). Persistent genetic signatures of colonization in *Brachionus manjavacas* rotifers in the Iberian Peninsula. *Molecular Ecology,* **16**, 3228–3240.

Gómez, Á. & Serra, M. (1995). Behavioral reproductive isolation among sympatric strains of *Brachionus plicatilis* Müller 1786: insights into the status of this taxonomic species. *Hydrobiologia*, **313**/**314**, 111–119.

Gómez, A., Adcock, G.J., Lunt, D.H. & Carvalho, G.R. (2002a). The interplay between colonization history and gene flow in passively dispersing zooplankton: microsatellite analysis of rotifer resting egg banks. *Journal of Evolutionary Biology*, **15**, 158–171.

Gómez, A., Serra, M., Carvalho, G.R. & Lunt, D.H. (2002b). Speciation in ancient cryptic species complexes: evidence from the molecular phylogeny of *Brachionus plicatilis* (Rotifera). *Evolution*, **56**, 1431–1444.

Gómez, A. & Snell, T. (1996). Sibling species and cryptic speciation in the *Brachionus plicatilis* species complex (Rotifera). *Journal of Evolutionary Biology*, **9**, 953–964.

Gómez, A., Temprano, M. & Serra, M. (1995). Ecological genetics of a cyclical parthenogen in temporary habitats. *Journal of Evolutionary Biology*, **8**, 601–622.

Goudet, J. (2005). HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, **5**, 184–186.

Gould, S.J. & Lewontin, R.C. (1979). The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London B: Biological Sciences*, **205**, 581–598.

Greenacre, M. & Primicerio, R. (2013). *Multivariate analysis of ecological data*, 1st edn. Rubes Editorial, Bilbao, Spain.

Gremer, J.R. & Venable, D.L. (2014). Bet hedging in desert winter annual plants: optimal germination strategies in a variable environment. *Ecology Letters*, **17**, 380–387.

Groeneveld, D.P. & Barz, D.D. (2014). Dixie Valley, Nevada playa bathymetry constructed from Landsat TM data. *Journal of Hydrology*, **512**, 435–441.

Grossman, G. & Sabo, J. (2010). Incorporating environmental variation into models of community stability: examples from stream fish. *American Fisheries Society Symposium*, **73**, 407–426.

Guillard, R.R.L. & Ryther, J.H. (1962). Studies of marine planktonic diatoms: I. *Cyclotella nana* Hustedt, and *Detonula confervacea* (Cleve) Gran. *Canadian Journal of Microbiology*, **8**, 229–239.

Gulhar, M., Kibria, G., Albatineh, A.N. & Ahmed, N.U. (2012). A comparison of some confidence intervals for estimating the population coefficient of variation: A simulation study. *SORT: statistics and operations research transactions*, **36**, 45–68.

Guo, B., Li, Z. & Merilä, J. (2016). Population genomic evidence for adaptive differentiation in the Baltic Sea herring. *Molecular Ecology,* **25**, 2833–2852.

Hagiwara, A., Balompapueng, M.D., Munuswamy, N. & Hirayama, K. (1997). Mass production and preservation of the resting eggs of the euryhaline rotifer *Brachionus plicatilis* and *B. rotundiformis*. *Aquaculture*, **155**, 223–230.

Hagiwara, A. & Hino, A. (1989). Effect of incubation and preservation on resting egg hatching and mixis in the derived clones of the rotifer *Brachionus plicatilis*. *Hydrobiologia*, **186**/**187**, 415–421.

Hairston, N. (1998). Time travelers: What's timely in diapause research? *Archiv für Hydrobiologie. Advances in Limnology*, **52**, 1–15.

Hairston, N.G., Van Brunt, R.A., Kearns, C.M. & Engstrom, D.R. (1995). Age and survivorship of diapausing eggs in a sediment egg bank. *Ecology*, **76**, 1706–1711.

Hairston, N.G., Holtmeier, C.L., Lampert, W., Weider, L.J., Post, D.M., Fischer, J.M., *et al.* (2001). Natural selection for grazer resistance to toxic cyanobacteria: evolution of phenotypic plasticity? *Evolution*, **55**, 2203–2214.

Hairston Jr., N.G., Lampert, W., Cáceres, C.E., Holtmeier, C.L., Weider, L.J., Gaedke, U. *et al.* (1999). Lake ecosystems: rapid evolution revealed by dormant eggs. *Nature*, **401**, 446.

Hairston Jr., N.G. & Munns Jr., W.R. (1984). The timing of copepod diapause as an evolutionarily stable strategy. *The American Naturalist*, **123**, 733–751.

Haldane, J.B.S. & Jayakar, S.D. Polymorphism due to selection depending on the composition of a population. *Journal of Genetics*, **58**, 318–323.

Hammer, U.T. (1986). Saline lake ecosystems of the world, 1st edn. Dr W. Junk Publishers, Springer Netherlands, Dordrecht, The Netherlands.

Hand, S.C. & Podrabsky, J.E. (2000). Bioenergetics of diapause and quiescence in aquatic animals. *Thermochimica Acta*, **349**, 31–42.

Hanson, S.J., Stelzer, C.-P., Welch, D.B., Logsdon Jr. J. (2013). Comparative transcriptome analysis of obligately asexual and cyclically sexual rotifers reveals genes with putative functions in sexual reproduction, dormancy, and asexual egg production. *BMC Genomics*, **14**, 412.

Hastie, T. & Tibshirani, R. (1990). *Generalized additive models*, 1st edn. Chapman and Hall/CRC.

Heo, Y., Wu, X., Chen, D., Ma, J. & Hwu, W. (2014). BLESS: bloom filter-based error correction solution for high-throughput sequencing reads. *Bioinformatics,* **30**, 1354–1362.

Hickman, C., Roberts, L. & Larson, A. (1997). *Zoología. Principios integrales.* McGraw-Hill Interamericana, Madrid, Spain.

Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A., & Cresko, W.A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics*, **6**, e1000862.

Horton, N.J. & Kleinman, K.P. (2007). Much Ado About Nothing. *The American Statistician*, **61**, 79–90.

Hutchinson, GE. (1978). *An introduction to population ecology,* Yale University Press, New Haven.

IPCC. (2013). Climate change 2013: the physical science basis. *Contribution of Working Group I to the Fifth Assessment Report of*

the Intergovernmental Panel on Climate Change (eds. T.F. Stocker, D. Qin, G.-K. Plattner, *et al.*). Cambridge University Press, UK.

James, C.S., Thoms, M.C. & Quinn, G.P. (2008). Zooplankton dynamics from inundation to drying in a complex ephemeral floodplain-wetland. *Aquatic Sciences*, **70**, 259–271.

Jensen, J.L., Bohonak, A.J. & Kelley, S.T. (2005). Isolation by distance, web service. *BMC Genetics*, **6**, 13.

Jombart, T. (2008). Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics,* **24**, 1403–1405.

Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., *et al.* (2014). InterProScan 5: Genome-scale protein function classification. *Bioinformatics*, **30**, 1236–1240.

de la Peña J.A. & Marfil, R. (1986). La sedimentación salina actual en las lagunas de La Mancha: una síntesis. *Cuadernos de Geologia Ibérica*, **10**, 235–270.

Kajitani, R., Toshimoto, K., Noguchi, H., Toyoda, A., Ogura, Y., Okuno, M. *et al.* (2014). Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. *Genome Research*, **24**, 1384–1395.

Kaltz, O. & Shykoff, J. (1998). Local adaptation in host–parasite systems. *Heredity,* **81**, 361–370.

Kaplan, N.L., Hudson, R.R. & Langley, C.H. (1989). The "hitchhiking effect" revisited. *Genetics*, **123**, 887-889.

Kawecki, T. & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology letters,* **7**, 1225–1241.

Kennard, M.J., Mackay, S.J., Pusey, B.J., Olden, J.D. & Marsh, N. (2009). Quantifying uncertainty in estimation of hydrologic metrics for ecohydrological studies. *River Research and Applications*, **22**, 137–156.

Kerr, B. & Godfrey-Smith, P. (2009). Generalization of the Price equation for evolutionary change. *Evolution*, **63**, 531–536.

Kirkpatrick, M. & Barton, N. (1997). Evolution of a species' range. *The American Naturalist,* **150**, 1–23.

Korf, I. (2004). Gene finding in novel genomes. *BMC bioinformatics*, **5**, 59.

Koste, W. (1978). Rotatoria die rädertiere Mitteleuropas begründet von Max Voigt–Monogononta. 2. *Auflage neubearbeitet von Walter Koste.* Gebrüder Borntraeger, Berlin, Stuttgart.

Kostopoulou, V., Carmona, M.J. & Divanach, P. (2012). The rotifer *Brachionus plicatilis*: an emerging bio-tool for numerous applications. *Journal of Biological Research- Thessaloniki*, **17**, 97–112.

Kotani, T., Ozaki, M., Matsuoka, K., Snell, T.W. & Hagiwara, A. (2001). Reproductive isolation among geographically and temporally isolated marine *Brachionus* strains. *Hydrobiologia*, **446/447**, 283–290.

De Kovel, C.G.F. (2006). The power of allele frequency comparisons to detect the footprint of selection in natural and experimental situations. *Genetics Selection Evolution*, **38**, 3–23.

Krapu, G. (1974). Feeding ecology of pintail hens during reproduction. *The Auk*, **91**, 278–290.

Kumar, S., Jones, M., Koutsovoulos, G., Clarke, M. & Blaxter, M. (2013). Blobology: exploring raw genome data for contaminants, symbionts and parasites using taxon-annotated GC-coverage plots. *Frontiers in Genetics*, **4**, 237.

Lande, R. & Shannon, S. (1996). The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution*, **50**, 434–437.

Lapesa, S. (2004). *Efecto de la depredación por invertebrados sobre poblaciones simpátricas de rotíferos*. Tesis doctoral, Universitat de València, Valencia.

Lapesa S, Snell TW, Fields DM, Serra M (2004) Selective feeding of *Arctodiaptomus salinus* (Copepoda, Calanoida) on co-occurring sibling rotifer species. *Freshwater Bioogy* **49,** 1053–1061.

Lawson, C.R., Vindenes, Y., Bailey, L. & van de Pol, M. (2015). Environmental variation and population responses to global change. *Ecology Letters*, **18**, 724–736.

Lee, B.Y., Kim, H.S., Hwang, D.S., Won, E.J., Choi, I.Y., Park, H.G., *et al.* (2015). Whole transcriptome analysis of the monogonont rotifer

Brachionus koreanus provides molecular resources for developing biomarkers of carbohydrate metabolism. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, **14**, 33–41.

Legendre, P. & Gauthier, O. (2014). Statistical methods for temporal and space-time analysis of community composition data. *Proceedings of the Royal Society of London B: Biological Sciences*, **281**, 20132728.

Levins, R. (1968). *Evolution in changing environments: some theoretical explorations*, 1st edn (ed. Robert H. MacArthur). Princeton University Press, Princeton, New Yersey, USA.

Lewontin, R. (1979). Sociobiology as an adaptationist program. *Systems Research and Behavioral Science,* **24**, 5–14.

Li, H. & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, **26**, 589–595.

Linnaeus, C. Von (1751). *Philosophia botanica*. Stockholm: G. Kiesewetter.

Lischer, H.E.L. & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, **28**, 298–299.

Lomsadze, A., Ter-Hovhannisyan, V., Chernoff, Y.O. & Borodovsky, M. (2005). Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Research*, **33**, 6494–6506.

López Donate, J.A., Montesinos Ibáñez, J.G., López Cano, J.A. & Martínez Núñez, J.C. (2004). Estudio descriptivo del sector endorreico-salino de Pétrola, Corral Rubio y La Higuera (Albacete). *Jornadas sobre el Medio Natural Albacetense*, 357–370.

Lowe, C., Kemp, S., Bates, A. & Montagnes, D. (2005). Evidence that the rotifer *Brachionus plicatilis* is not an osmoconformer. *Marine Biology,* **146**, 923–929.

Lubzens, E., Tandler, A. & Minkoff, G. (1989). Rotifers as food in aquaculture. *Hydrobiologia*, **186**/**187**, 387–400.

Lubzens, E., Zmora, O. & Barr, Y. (2001). Biotechnology and aquaculture of rotifers. *Hydrobiology*, **446**, 337–353.

Lynch, M. (1987). The consequences of fluctuating selection for isozyme polymorphisms in *Daphnia*. *Genetics,* **115**, 657–669.

Lynch, M. & Deng, H.-W. (1994). Genetic slippage in response to sex. *The American Naturalist*, **144**, 242–261.

Lynch, M. & Lande, R. (1993). Evolution and extinction in response to environmental change. *Biotic Interactions and Global Change* (eds. P. Kareiva, J. Kingsolver and R. Huey), pp. 234–250. Sinauer Assocs., Inc. Sunderland, MA.

Lynch, M. & Walsh, B. (1998). Genetics and analysis of quantitative traits. Sinauer Assocs., Inc. Sunderland, MA.

MacArthur, R.H. & Levins, R. (1964). Competition, habitat selection, and character displacement in a patchy environment. *Proceedings of the National Academy of Sciences of the United States of America*, **51**, 1207–1210.

MacArthur, R.H. & Wilson, E.O. (1967). *The theory of island biogeography*, 1st edn. Princeton University Press, Princeton, New Yersey.

Mageed, A.A.A. & Heikal, M.T. (2006). Factors affecting seasonal patterns in epilimnion zooplankton community in one of the largest man-made lakes in Africa (Lake Nasser, Egypt). *Limnologica*, **36**, 91–97.

Malthus, T.R. (1798). *An essay on the principle of population* (ed. J. Johnson). London, UK.

Marcus, N.H., Lutz, R., Burnett, W. & Cable, P. (1994). Age, viability, and vertical distribution of zooplankton resting eggs from an anoxic basin: Evidence of an egg bank. *Limnology and Oceanography*, **39**, 154–158.

Martínez-Ruiz, C. & García-Roger, E.M. (2014). Being first increases the probability of long diapause in rotifer resting eggs. *Hydrobiologia*, **745**, 111–121.

Masek, J., Vermote, E., Saleous, N., *et al.* (2013). LEDAPS calibration, reflectance, atmospheric correction preprocessing code, version 2. *Model product*. ORNL DAAC, Oak Ridge, Tennessee, USA.

May, R.M. (1974). *Stability and complexity in model ecosystems*, 2nd edn. Princeton University Press, Princeton and Oxford.

May, R. & McLean, A. (2007). T*heoretical ecology: Principles and applictions*, 1st edn. Oxford University Press, Oxford, UK.

Maynard Smith, J. (1978). *The evolution of sex*. Cambridge University Press, Cambridge, UK.

Mayr, E. (1983). How to carry out the adaptationist program? *The American Naturalist*, **121**, 324–334.

McIntosh, R.P. (1991). *Concept and terminology of homogeneity and heterogeneity in ecology* (eds. J. Kolasa *et al*.,). pp. 24–46. Springer-Verlag New York.

De Meester, L. (2004). Evolutionary and ecological genetics of cyclical parthenogens. *Evolution: from molecules to ecosystems* (eds. A. Moya and E. Font), pp. 122–134. Oxford University Press Inc., New York.

De Meester, L. (1996). Local genetic differentiation and adaptation in freshwater zooplankton populations: Patterns and processes. *Ecoscience*, **3**, 385–399.

De Meester, L., Gómez, A., Okamura, B. & Schwenk, K. (2002). The monopolization hypothesis and the dispersal-gene flow paradox in aquatic organisms. *Acta Oecologica*, **23**, 121–135.

Mills, S., Alcántara-Rodríguez, J.A., Ciros-Pérez, J., Gómez, A., Hagiwara, A., *et al.* (2016). Fifteen species in one: deciphering the *Brachionus plicatilis* species complex (Rotifera, Monogononta) through DNA taxonomy. *Hydrobiologia, in press*.

Montero-Pau, J. & Serra, M. (2011). Life-cycle switching and coexistence of species with no niche differentiation. *PLoS ONE*, **6**, e20314.

Montero-Pau, J. 2012. *Ecological and evolutionary impact of diapause on zooplankton*. Universitat de València.

Montero-Pau, J., Serra, M. & Gómez, A. (2016). Diapausing egg banks, lake size, and genetic diversity in the rotifer *Brachionus plicatilis* Müller (Rotifera, Monogononta). *Hydrobiologia*, *in press*.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A. & Kent, J. (2000a). Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.

Myers, E.W., Sutton, G.G., Delcher, A.L., Dew, I.M., Fasulo, D.P., *et al.* (2000b). A whole-genome assembly of *Drosophila*. *Science*, **287**, 2196–2204.

O'Dea, R., Noble, D. & Johnson, S. (2016). The role of non-genetic inheritance in evolutionary rescue: epigenetic buffering, heritable bet hedging and epigenetic traps. Environmental *Epigenetics,* **2**, dvv014.

Ormeci, C. & Ekercin, S. (2007). An assessment of water reserve changes in Salt Lake, Turkey, through multi-temporal Landsat imagery and real-time ground surveys. *Hydrological Processes*, **21**, 1424–1435.

Orsini, L., Spanier, K.I. & De Meester, L. (2012). Genomic signature of natural and anthropogenic stress in wild populations of the waterflea *Daphnia magna*: validation in space, time and experimental evolution. *Molecular Ecology*, **21**, 2160–2175.

Ortells, R. (2002). *Diversidad genética y ecológica en especies crípticas de rotíferos: patrones y procesos*. Tesis doctoral, Universitat de València, Valencia.

Ortells, R., Gómez, A. & Serra, M. (2003). Coexistence of cryptic rotifer species: ecological and genetic characterisation of *Brachionus plicatilis*. *Freshwater Biology*, **48**, 2194–2202.

Ortells, R., Gómez, A. & Serra, M. (2006). Effects of duration of the planktonic phase on rotifer genetic diversity. *Archiv für Hydrobiologie*, **167**, 203–216.

Ortells, R., Snell, T., Gómez, A. & Serra, M. (2000). Patterns of genetic differentiation in resting egg banks of a rotifer species complex in Spain. *Archiv für Hydrobiologie,* **149**, 529–551.

Parra, G., Bradnam, K., Ning, Z., Keane, T. & Korf, I. (2009). Assessing the gene space in draft genomes. *Nucleic Acids Research*, **37**, 289–297.

Patten, M.M. & Haig, D. (2009). Maintenance or loss of genetic variation under sexual and parental antagonism at a sex-linked locus. *Evolution*, **63**, 2888–2895.

Peck, J.R., Jonathan, Y. & Barreau, G. (1999). The maintenance of sexual reproduction in a structured population. *Proceedings of the Royal Society of London B: Biological Sciences*, **266**, 1857–1863.

Pettorelli, N., Laurance, W.F., O'Brien, T.G., Wegmann, M., Nagendra, H. & Turner, W. (2014). Satellite remote sensing for applied ecologists: opportunities and challenges. *Journal of Applied Ecology*, **51**, 839–848.

Pfrender, M.E. & Lynch, M. (2000). Quantitative genetic variation in *Daphnia*: temporal changes in genetic architecture. *Evolution*, **54**, 1502–1509.

Philippi, T. & Seger, J. (1989). Hedging one's evolutionary bets, revisited. *Trends in Ecology and Evolution*, **4**, 41–44.

Pianka, E. (1970). On *r*-and *K*-selection. *The American Naturalist,* **104**, 592-597.

Pimm, S.L., Russell, G.J., Gittleman, J.L. & Brooks, T.M. (1995). The Future of Biodiversity. *Science*, **269**, 347–350.

Pound, G., Doncaster, C. & Cox, S. (2002). A Lotka–Volterra model of coexistence between a sexual population and multiple asexual clones. *Journal of theoretical biology,* **217**, 535–545.

Pourriot, R. & Snell, T.W. (1983). Resting eggs in rotifers. *Hydrobiologia*, **104**, 213–224.

Price, G.R. (1972). Extension of covariance selection mathematics. *Annals of Human Genetics*, **35**, 485–490.

Price, G.R. (1970). Selection and covariance. *Nature*, **227**, 520–521.

Purcell, S., Neale, B., Todd-Brown, K., *et al.* (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, **81**, 559–575.

Quinlan, A.R. & Hall, I.M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.

R Development Core Team, R. (2009). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.

Reche, I., Pulido-Villena, E., Morales-Baquero, R. & Casamayor, E.O. (2005). Does ecosystem size determine aquatic bacterial richness? *Ecology*, **86**, 1715–1722.

Reed, T.E., Waples, R.S., Schindler, D.E., Hard, J.J. & Kinnison, M.T. (2010). Phenotypic plasticity and population viability: the

importance of environmental predictability. *Proceedings of the Royal Society of London B: Biological Sciences*, **277**, 3391–400.

Rees, M., Childs, D.Z., Metcalf, J.C., Rose, K.E., Sheppard, A.W. & Grubb, P.J. (2006). Seed dormancy and delayed flowering in monocarpic plants: selective interactions in a stochastic environment. *The American Naturalist*, **168**, E53–E71.

Ricci, C., Serra, M. & Snell, T.W. (2000). Small, beautiful and sexy: what rotifers tell us about ecology and evolution. *Trends in Ecology & Evolution*, **15**, 220–221.

Roach, J.K., Griffith, B. & Verbyla, D. (2013). Landscape influences on climate-related lake shrinkage at high latitudes. *Global Change Biology*, **19**, 2276–2284.

Roff, D. a. (2007). Contributions of genomics to life-history theory. *Nature reviews. Genetics*, **8**, 116–125.

Root, T.L., Price, J.T., Hall, K.R., *et al.* (2003). Fingerprints of global warming on wild animals and plants. *Nature*, **421**, 57–60.

Roulin, A.C., Bourgeois, Y., Stiefel, U., Walser, J.-C. & Ebert, D. (2016). A photoreceptor contributes to the natural variation of diapause induction in *Daphnia magna*. *Molecular Biology and Evolution*, **33**, msw200.

Roulin, A.C., Mariadassou, M., Hall, M.D., Walser, J.-C, Haag, C.R. & Ebert, D. (2015). High genetic variation in resting-stage production in a metapopulation: Is there evidence for local adaptation? *Evolution*, **69**, 2747–2756.

Roulin, A.C., Routtu, J., Hall, M.D., Janicke, T., Colson, I., Haag, C.R. *et al.* (2013). Local adaptation of sex induction in a facultative sexual crustacean: insights from QTL mapping and natural populations of *Daphnia magna*. *Molecular Ecology*, **22**, 3567–3579.

Sabo, J.L. & Post, D.M. (2008). Quantifying periodic, stochastic, and catastrophic environmental variation. *Ecological Monographs*, **78**, 19–40.

Sæther, B.E. & Engen, S. (2015). The concept of fitness in fluctuating environments. *Trends in Ecology and Evolution*, **30**, 273–281.

Savolainen, O., Lascoux, M. & Merilä, J. (2013). Ecological genomics of local adaptation. *Nature reviews. Genetics*, **14**, 807–20.

Scheuerl, T. & Stelzer, C.P. (2013). Patterns and dynamics of rapid local adaptation and sex in varying habitat types in rotifers. *Ecology and Evolution*, **3**, 4253–4264.

Schluter, D. (2001). Ecology and the origin of species. *Trends in Ecology & Evolution,* **16**, 372–380.

Schowengerdt, R.A. (2007). *Remote sensing: models and methods for image processing*, 3th edn, pp. 37–38. Academic Press, Elsevier, London, UK.

Schröder, T. (2005). Diapause in monogonont rotifers. *Hydrobiologia*, **546**, 291–306.

Scuderi, L. A., Laudadio, C.K. & Fawcett, P.J. (2010). Monitoring playa lake inundation in the western United States: modern analogues to late-Holocene lake level change. *Quaternary Research*, **73**, 48–58.

Seger, J. & Brockman. (1987). What is bet-hedging? *Oxford Surveys in Evolutionary Biology*, **4**, 182–211.

Serra, M. & Carmona, M.J. (1993). Mixis strategies and resting egg production of rotifers living in temporally-varying habitats. *Hydrobiologia*, **255**/**256**, 117–126.

Serra, M., Gómez, A. & Carmona, M.J. (1998). Ecological genetics of *Brachionus* sympatric sibling species. *Hydrobiologia*, **387**, 373–384.

Serra, M. & King, C.E. (1999). Optimal rates of bisexual reproduction in cyclical parthenogens with density-dependent growth. *Journal of Evolutionary Biology*, **12**, 263–271.

Serra, M., Snell, T.W. & King, C.E. (2004). The timing of sex in cyclically parthenogenetic rotifers. *Evolution: from molecules to ecosystems.* (eds. A. Moya and E. Font), pp. 135–146. Oxford University Press, Oxford.

Sielaff, M., Schmidt, H., Struck, T.H., Rosenkranz, D., Mark Welch, D.B., Hankeln, T., *et al.* (2016). Phylogeny of Syndermata (syn. Rotifera): Mitochondrial gene order verifies epizoic Seisonidea as sister to endoparasitic Acanthocephala within monophyletic Hemirotifera. *Molecular Phylogenetics and Evolution*, **96**, 79–92.

Simon, J.C., Baumann, S., Sunnucks, P., Hebert, P.D., Pierre, J.S., Le Gallic, J.F., *et al.* (1999). Reproductive mode and population

genetic structure of the cereal aphid *Sitobion avenae* studied using phenotypic and microsatellite markers. *Molecular Ecology*, **8**, 531–545.

Simons, A.M. (2011). Modes of response to environmental change and the elusive empirical evidence for bet hedging. *Proceedings of the Royal Society of London B: Biological Sciences*, **278**, 1601–9.

Smit, A., Hubley, R. & Green, P. (1996). *RepeatMasker Open-3.0*, www.repeatmasker.org.

Smith, J.M. & Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genetical research*, **23**, 23–35.

Smith, H.A. & Snell, T.W. (2012). Rapid evolution of sex frequency and dormancy as hydroperiod adaptations. *Journal of Evolutionary Biology*, **25**, 2501–2510.

Snell, T.W. (2014). Rotifers as models for the biology of aging. *International Review of Hydrobiology*, **99**, 84–95.

Snell, T. & Boyer, E. (1988). Thresholds for mictic female production in the rotifer *Brachionus plicatilis* (Muller). *Journal of Experimental Marine Biology and Ecology,* **124**, 73–85.

Snell, T. & Carmona, M. (1995). Comparative toxicant sensitivity of sexual and asexual reproduction in the rotifer *Brachionus calyciflorus*. *Environmental Toxicology and Chemistry,* **14**, 415–420.

Snell, T.W. & Janssen, C.R. (1995). Rotifers in ecotoxicology: a review. *Hydrobiologia*, **313**/**314**, 231–247.

Snell, T. & Joaquim-Justo, C. (2007). Workshop on rotifers in ecotoxicology. *Hydrobiologia,* **593**, 227–232.

Snell, T., Johnston, R. & Gribble, K. (2015). Rotifers as experimental tools for investigating aging. *Invertebrate Reproduction & Development,* **59**, 5–10.

Snell, T.W., Kubanek, J., Carter, W., Payne, A.B., Kim, J., Hicks, M.K., *et al.* (2006). A protein signal triggers sexual reproduction in *Brachionus plicatilis* (Rotifera). *Marine Biology*, **149**, 763–773.

Southwood, T. (1977). Habitat, the templet for ecological strategies? *Journal of Animal Ecology*, **46**, 337–365.

Spencer, M., Colegrave, N. & Schwartz, S.S. (2001). Hatching fraction and timing of resting stage production in seasonal environments: effects of density dependence and uncertain season length. *Journal of Evolutionary Biology*, **14**, 357–367.

Stapley, J., Reger, J., Feulner, P.G.D., Smadja, C., Galindo, J., Ekblom, R., *et al.* (2010). Adaptation genomics: the next generation. *Trends in Ecology and Evolution*, **25**, 705–712.

Starkweather, P. L. (1987) Rotifera. T. J. Pandian and F. J. Vernberg (eds.), Animal energetics. Vol. 1, Protozoa through Insecta, pp. 159–183, Academic Press, Orlando.

Starrfelt, J. & Kokko, H. (2012). Bet-hedging-a triple trade-off between means, variances and correlations. *Biological Reviews*, **87**, 742–755.

Stearns, S. (2000). Daniel Bernoulli (1738): evolution and economics under risk. *Journal of Biosciences*, **25**, 221–228.

Stearns, S. (1981). On measuring fluctuating environments: predictability, constancy, and contingency. *Ecology*, **62**, 185–199.

Stein, L. (2001). Genome annotation: from sequence to biology. *Nature Reviews Genetics,* **2**, 493–503.

Stelzer, C.-P., Riss, S. & Stadler, P. (2011). Genome size evolution at the speciation level: the cryptic species complex *Brachionus plicatilis* (Rotifera). *BMC Evolutionary Biology*, **11**, 90.

Stelzer, C.P. & Snell, T.W. (2003). Induction of sexual reproduction in *Brachionus plicatilis* (Monogononta, Rotifera) by a density-dependent chemical cue. *Limnology and Oceanography*, **48**, 939–943.

Stelzer, C.-P. & Snell, T.W. (2006). Specificity of the crowding response in the *Brachionus plicatilis* species complex. *Limnology and Oceanography*, **51**, 125–130.

Stinchcombe, J.R. & Hoekstra, H.E. (2008). Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity*, **100**, 158–170.

Storz, J.F. (2005). Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, **14**, 671–688.

Straalen, N. Van & Roelofs, D. (2011). *An introduction to ecological genomics*. Oxford University Press.

Suatoni, E., Vicario, S., Rice, S., Snell, T.W. & Caccone, A. (2006). An analysis of species boundaries and biogeographic patterns in a cryptic species complex: the rotifer-*Brachionus plicatilis*. *Molecular Phylogenetics and Evolution*, **41**, 86–98.

Suzuki, R. & Shimodaira, H. (2006). Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics*, **22**, 1540–1542.

Tagu, D., Colbourne, J.K. & Nègre, N. (2014). Genomic data integration for ecological and evolutionary traits in non-model organisms. *BMC Genomics*, **15**, 490.

Tarazona, E., García-Roger, E.M. & Carmona, M.J. (2017). Experimental evolution of bet hedging in rotifer diapause traits as a response to environmental unpredictability. *Oikos*, *in press.*

Taylor, F. & Spalding, J.B. (1989). Timing of diapause in relation to temporally variable catastrophes. *Journal of Evolutionary Biology*, **2**, 285–297.

Tessier, A.J. & Cáceres, C.E. (2004). Differentiation in sex investment by clones and populations of *Daphnia*. *Ecology Letters*, **7**, 695–703.

Tortajada, A. & Carmona, M. (2010). Effects of population outcrossing on rotifer fitness. *BMC Evolutionary Biology,* **10**, 312.

Tortajada, A.M., Carmona, M.J. & Serra, M. (2009). Does haplodiploidy purge inbreeding depression in rotifer populations? *PLoS ONE*, **4**, e8195.

Tufto, J. (2015). Genetic evolution, plasticity, and bet-hedging as adaptive responses to temporally autocorrelated fluctuating selection: A quantitative genetic model. *Evolution*, **69**, 2034–2049.

Turchin, P. (2001). Does population ecology have general laws? *Oikos*, **94**, 17–26.

Turelli, M., Barton, N. & Coyne, J. (2001). Theory and speciation. *Trends in Ecology & Evolution,* **16**, 330–343.

Turner, M.A., Huebert, D.B., Findlay, D.L., *et al.* (2005). Divergent impacts of experimental lake-level drawdown on planktonic and

benthic plant communities in a boreal forest lake. *Canadian Journal of Fisheries and Aquatic Sciences*, **62**, 991–1003.

Vermote, E.F., Saleous, N., Justice, C.O., Kaufman, Y.J., Privette, J.L., Remer, L., *et al.* (1997). Atmospheric correction of visible to middle-infrared EOS-MODIS data over land surfaces: Background, operational algorithm and validation. *Journal of Geophysical Research: Atmospheres*, **102**, 17131–17141.

Via, S. (2001). Sympatric speciation in animals: the ugly duckling grows up. *Trends in Ecology & Evolution,* **16**, 381–390.

de Vicente, I., Amores, V. & Cruz-Pizarro, L. (2006). Instability of shallow lakes: A matter of the complexity of factors involved in sediment and water interaction? *Limnetica,* **25**, 253–270.

de Villemereuil, P., Schielzeth, H., Nakagawa, S. & Morrissey, M. (2016). General methods for evolutionary quantitative genetic inference from generalized mixed models. *Genetics*, **204**, 1281–1294.

Vurture, G.W., Sedlazeck, F.J., Nattestad, M., *et al.* (2017). GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics*, btx153.

Wallace, R.L., Nogrady, T., Snell, T.W. & Ricci, C. (2006). *Rotifera : biology, ecology and systematic*s. 2nd edn. Backhuys Publishers, Leiden. Wallace, R.L. & Smith, H.A. (2009). Rotifera. *Encyclopedia of inlands waters* (ed. G. Likens). Elsevier, Oxford, UK.

Wallace, R.L. & Snell, T.W. (1991). Rotifera. *Ecology and classification of North American freshwater invertebrates.* (eds. J. Thorpe and P. Covich). Academic Press, New York.

Wallace, R., Snell, T. & Smith, H. (2015). Rotifer: ecology and general biology. *Thorp and Covich's freshwater invertebrates.* 4th edn. (eds. J.H. Thorp and D.C. Rogers). Academic Press, Elsevier, UK and USA.

Walsh, E.J., Smith, H.A. & Wallace, R.L. (2014). Rotifers of temporary waters. *International Review of Hydrobiology*, **99**, 3–19.

Walz, N. (1997). Rotifer life history strategies and evolution in freshwater plankton communities. (eds B. Streit, T. Städler, and C. M. Lively), Evolutionary ecology of freshwater animals, pp. 119–149, Birkhäuser Verlag, Basel.

Weisenfeld, N.I., Yin, S., Sharpe, T., Lau, B., Hegarty, R., Holmes, L., *et al.* (2014). Comprehensive variation discovery in single human genomes. *Nature Genetics*, **46**, 1350–1355.

Welch, D.B.M. (2000). Evidence from a protein-coding gene that acanthocephalans are rotifers. *Invertebrate Biology*, **119**, 17–26.

Wetzel, R.G. & Likens, G. (2013). *Limnological analysis*, 3rd edn. Springer Science & Business Media.

Wey-Fabrizius, A.R., Herlyn, H., Rieger, B., Rosenkranz, D., Witek, A. & Mark Welch, D.B. (2014). Transcriptome data reveal Syndermatan relationships and suggest the evolution of endoparasitism in Acanthocephala via an epizoic stage. *PLoS ONE*, **9**, e88618.

White, M.S., Xenopoulos, M.A., Hogsden, K., Metcalfe, R.A. & Dillon, P.J. (2008). Natural lake level fluctuation and associated concordance with water quality and aquatic communities within small lakes of the Laurentian Great Lakes region. *Hydrobiologia*, **613**, 21–31.

Wiens, J.A. (2000). Ecological heterogeneity: an ontogeny of concepts and approaches. *The ecological consequences of environmental heterogeneity* (eds. M.J. Hutchings, E.A. John and A.J.A. Stewart), pp. 9–31. Blackwell Science Cambridge.

Williams, W.D. (1981). Inland salt lakes: an introduction. *Hydrobiologia*, **81**/**82**, 1–14.

Williams, P.D. & Hastings, A. (2011). Paradoxical persistence through mixed-system dynamics: towards a unified perspective of reversal behaviours in evolutionary ecology. *Proceedings of the Royal Society of London B: Biological Sciences*, **278**, 1281–1290.

Wood, S.N. (2000). Modelling and smoothing parameter estimation with multiple quadratic penalties. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **62**, 413–428.

Wurtsbaugh, W.A. (1992). Food-web modification by an invertebrate predator in the Great Salt Lake (USA). *Oecologia*, **89**, 168–175.

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, **27**, 3025–3033.

Yampolsky, L.Y., Zeng, E., Lopez, J., *et al.* (2014). Functional genomics of acclimation and adaptation in response to thermal stress in *Daphnia*. *BMC Genomics*, **15**, 859.

Zhu, Z. & Woodcock, C.E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, **118**, 83–94.

APPENDIX

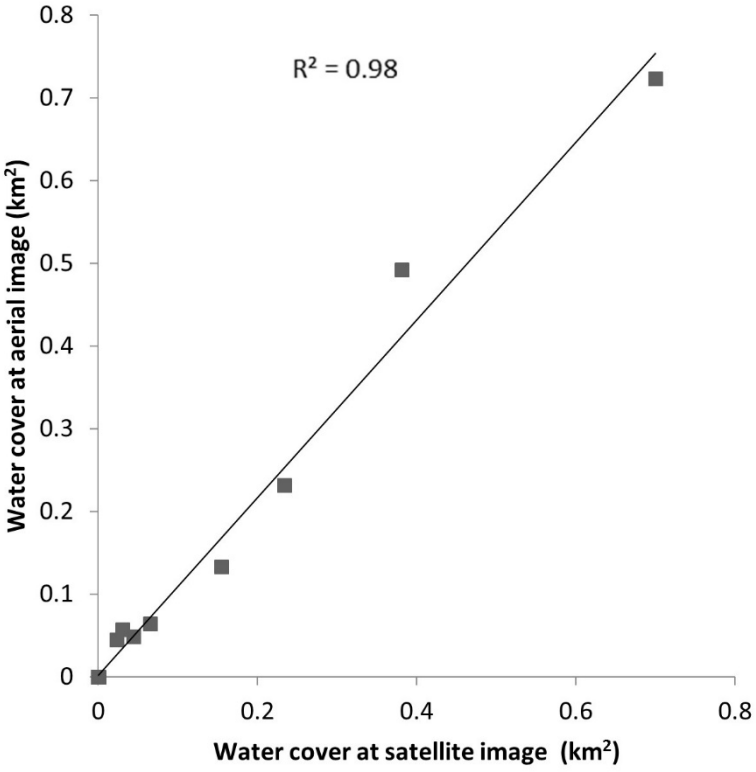


Figure A.1. Relationship between water-surface area (*A*) from direct inspection of aerial scenes and *A* after applying 2cA to raw satellite scenes (X axis). Dots are values for 19 out of 20 ponds (Y axis) recorded in years 2006 and 2009. Maximum time for matching values corresponding to the two estimations was 18 days. Linear least squares fitting and the corresponding determination coefficient are shown. *n*= 33; the pair 0-0 was found in 22 cases.

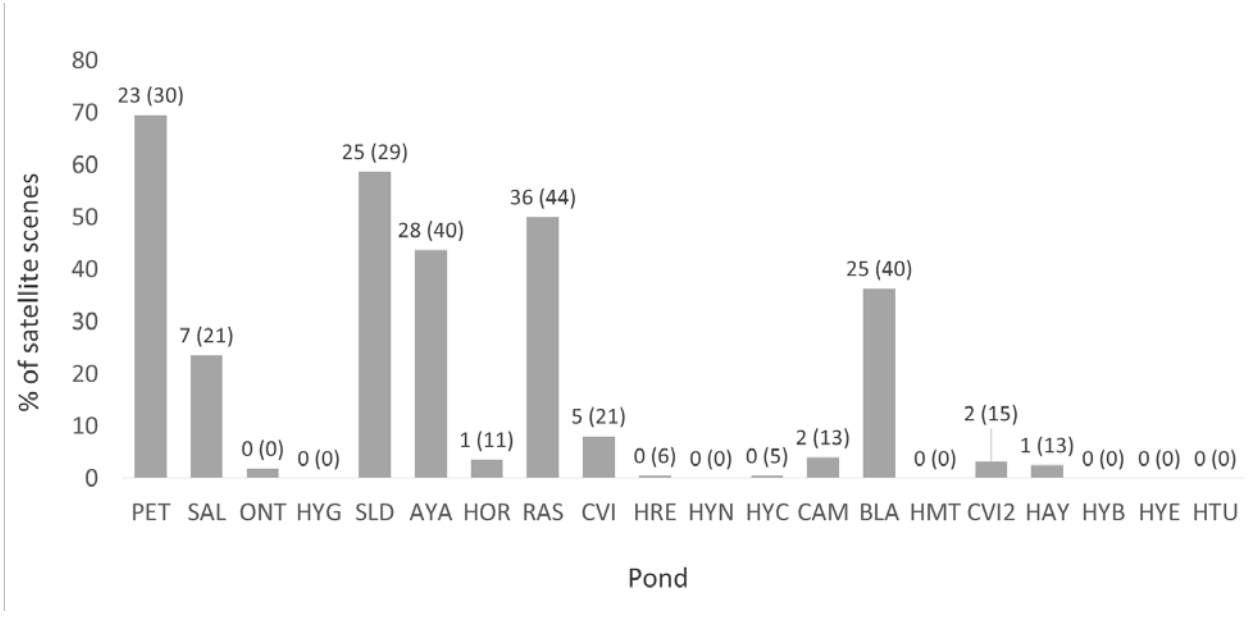


Figure A.2. Percentage of satellite scenes where salt-covered pixels (TM4< 0.4) were detected after retaining potentially water covered pixels (MNDWI> 0) at each pond. The mean and standard deviation (the latter between parentheses) of the percentage of reduction by the salt filter (TM4< 0.4)are shown above.

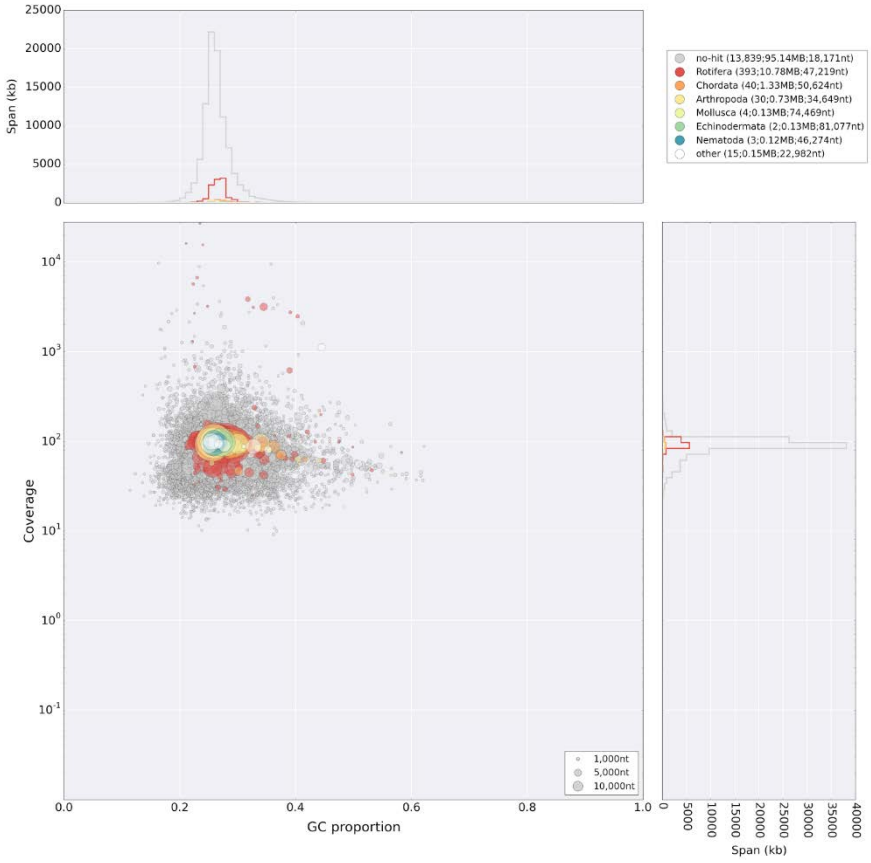


Figure A.3. Blobplot of the draft genome of *Brachionus plicatilis* based on G+C content and contig coverage.

Table A.1. Data used for validation of satellite water-surface area estimation (from aerial scenes) and presence/absence of water (from field observations).

Date	Source	Ponds
May 2000	Field	PET; SAL
June 2000	Field	PET; SAL
July 2000	Field	PET; SAL
August 2000	Field	PET; SAL
September 2000	Field	PET; SAL
October 2000	Field	PET; SAL
November 2000	Field	PET; SAL
December 2000	Field	PET
January 2001	Field	PET; SAL
March 2001	Field	PET; SAL
April 2001	Field	PET; SAL
July 2001	Field	PET; SAL; HYR
November 2001	Field	PET; SAL; HYR
July 2002	Field	PET
September 2004	Field	SAL
October 2004	Field	PET; SAL; HYR; HYM
November 2004	Field	PET
January 2005	Field	SAL
February 2005	Field	PET; SAL
March 2005	Field	PET; SAL
April 2005	Field	PET; SAL

Table A.1. (continued)

Date	Source	Ponds
May 2005	Field	PET; SAL
June 2005	Field	PET; HYR
August 2005	Field	PET
September 2005	Field	PET
February 2006	Field	PET
March 2006	Field	PET; SAL
14/07/2006 - 2/08/2006	Aerial	PET; SAL ; HYG; SLD; ATA; HOR; HYR; CVI; HRE; HYN; HYC; CAM; BLA; HMT; CVI2; HYA; HYB; HYE; HTU
16/07/2009 - 5/07/2009	Aerial	PET; SAL; HYG; SLD; ATA; HYR; HRE; HYN; CAM; BLA; HMT; HYB; HYE; HTU
July 2010	Field	PET; SAL

Table A.2. Transcriptomes used in the *Brachionus plicatilis* genome assembly from congeneric species.

Species	Accession number	Reference
B. manjavacas	GARS01000001-GARS01014244	Welch and Lea (2014)
B. calyciflorus	GACL00000000.1 and GACQ00000000.1	Hanson et al. (2013)
B. koreanus	GBXV00000000	Lee et al. (2015)

Table A.3. Number of genes detected to be physically associated to SNPs putatively under selection. Results for (1) three window sizes around each SNP and (2) two approaches to detect selection are reported. Overall number of genes is the sum of genes that appeared in at least one parameter or type of selection.

Approach	Type of selection/ Associated parameter	Window size		
		0 Kb	2.5 Kb	5 Kb
BayeScan	**Balancing**	39	144	199
	Diversifying	12	37	49
BAYENV	**Propensity**	41	118	174
	Hatching	40	105	136
	Predictability	28	81	127
	Hydroperiod	28	76	112
Overall		162	483	684

Table A.4. K-mer statistics for filtered and error-corrected raw data obtained from genome sequencing (k-mer length= 21)

Property	min	max
Heterozygosity	0.654 %	0.658%
Genome haploid length	115,775,183 bp	115,850,549 bp
Genome repeat length	37,904,922 bp	37,929,597 bp
Genome unique length	77,870,261 bp	77,920,952 bp
Model fit	90.069%	93.868 %
Read error rate	0.301%	0.301 %

Table A.5. Genes (162) having at least a SNP under selection within its coding region. The software employed to detect the SNP(s) is reported as well as the type of selection (H: high F_{ST}; L: low F_{ST}) and the parameter to which it is associated (pre.: predictability; hyd.: hydroperiod; pro.: propensity for sex; hat.: hatching). NA in gene description means that no functional annotation was obtained.

Gene name	Software (Parameter)	Description	Gene ontology number
augustus_masked-scaffold1187_len21429_cov63-processed-gene-0.1-mRNA-1	BayeScan (H); BAYENV (pre.)	---NA---	
augustus_masked-scaffold26640_len14398_cov39-processed-gene-0.0-mRNA-1	BayeScan (H)	E3 ubiquitin- ligase HUWE1 isoform X1	F:GO:0004842; F:GO:0008270; F:GO:0005515; P:GO:0016567
augustus_masked-scaffold29685_len4788_cov57-processed-gene-0.1-mRNA-1	BayeScan (H)	BTB POZ domain-containing 9	F:GO:0005515
augustus_masked-scaffold63942_len2227_cov40-processed-gene-0.0-mRNA-1	BayeScan (H)	liprin-alpha-1-like isoform X3	P:GO:0050808; C:GO:0045202
maker-scaffold14598_len41824_cov86-augustus-gene-0.32-mRNA-1	BayeScan (H); BAYENV (pro.)	opioid-binding cell adhesion molecule- partial	F:GO:0005515

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
maker-scaffold1996_len13316_cov70-augustus-gene-0.10-mRNA-1	BayeScan (H); BAYENV (hat.)	ubiquitin carboxyl-terminal hydrolase 15 isoform X1	P:GO:0016579; F:GO:0036459; P:GO:0006511
maker-scaffold22152_len11339_cov70-augustus-gene-0.14-mRNA-1	BayeScan (H); BAYENV (pro.)	zinc finger 1 isoform X2	F:GO:0003676; F:GO:0046872
maker-scaffold28017_len25204_cov83-augustus-gene-0.31-mRNA-1	BayeScan (H); BAYENV (hat.)	biotin-- ligase	P:GO:0006768; P:GO:0006464; F:GO:0004077
maker-scaffold37676_len28730_cov74-augustus-gene-0.32-mRNA-1	BayeScan (H)	---NA---	
snap_masked-scaffold29685_len4788_cov57-processed-gene-0.6-mRNA-1	BayeScan (H)	---NA---	
snap_masked-scaffold70385_len24210_cov75-processed-gene-0.14-mRNA-1	BayeScan (H)	---NA---	
snap_masked-scaffold7492_len27729_cov82-processed-gene-0.19-mRNA-1	BayeScan (H)	serine threonine- kinase minibrain isoform X1	F:GO:0005524; F:GO:0004672; P:GO:0006468

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
augustus_masked-scaffold103579_len11551_cov86-processed-gene-0.1-mRNA-1	BayeScan (L)	piggyBac transposable element-derived 3-like	
augustus_masked-scaffold10521_len6223_cov80-processed-gene-0.0-mRNA-1	BayeScan (L)	ankyrin repeat domain-containing partial	F:GO:0005515
augustus_masked-scaffold11529_len1245_cov71-processed-gene-0.0-mRNA-1	BayeScan (L)	Retrovirus-related Pol poly from transposon	F:GO:0003676; F:GO:0008270
augustus_masked-scaffold1174_len156180_cov80-processed-gene-0.14-mRNA-1	BayeScan (L)	chloride intracellular channel exc-4	F:GO:0016740
augustus_masked-scaffold118004_len671_cov123_single-processed-gene-0.0-mRNA-1	BayeScan (L)	---NA---	
augustus_masked-scaffold16172_len3098_cov131-processed-gene-0.1-mRNA-1	BayeScan (L)	---NA---	

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
augustus_masked-scaffold16380_len7967_cov69-processed-gene-0.3-mRNA-1	BayeScan (L)	---NA---	
augustus_masked-scaffold2345_len24629_cov66-processed-gene-0.2-mRNA-1	BayeScan (L); BAYENV (hyd.)	midasin	F:GO:0016887; C:GO:0005634; F:GO:0005524; P:GO:0000027
augustus_masked-scaffold24245_len1375_cov54-processed-gene-0.0-mRNA-1	BayeScan (L)	PREDICTED: uncharacterized protein LOC101238727	
augustus_masked-scaffold2438_len71187_cov80-processed-gene-0.1-mRNA-1	BayeScan (L)	---NA---	
augustus_masked-scaffold25654_len3753_cov62-processed-gene-0.0-mRNA-1	BayeScan (L)	photoreceptor-specific nuclear receptor-like	C:GO:0005634; F:GO:0003677; F:GO:0003707; P:GO:0006355; P:GO:0043401

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
augustus_masked-scaffold26462_len1294_cov60-processed-gene-0.0-mRNA-1	BayeScan (L)	---NA---	
augustus_masked-scaffold3012_len37027_cov63-processed-gene-0.6-mRNA-1	BayeScan (L)	---NA---	
augustus_masked-scaffold31186_len5218_cov49-processed-gene-0.0-mRNA-1	BayeScan (L)	---NA---	
augustus_masked-scaffold3260_len26479_cov81-processed-gene-0.0-mRNA-1	BayeScan (L)	---NA---	
augustus_masked-scaffold34844_len8684_cov79-processed-gene-0.0-mRNA-1	BayeScan (L)	zinc finger partial	F:GO:0003676; F:GO:0008270; F:GO:0004190; P:GO:0006508
augustus_masked-scaffold39332_len18654_cov78-processed-gene-0.0-mRNA-1	BayeScan (L)	---NA---	

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
augustus_masked-scaffold47287_len1611_cov32-processed-gene-0.0-mRNA-1	BayeScan (L)	RNA-directed DNA polymerase from mobile element jockey-like	F:GO:0003964; F:GO:0003723; F:GO:0016706; P:GO:0006278; P:GO:0055114; F:GO:0008168; P:GO:0032259
augustus_masked-scaffold65693_len4794_cov38-processed-gene-0.2-mRNA-1	BayeScan (L)	---NA---	
augustus_masked-scaffold77279_len1917_cov60_single-processed-gene-0.1-mRNA-1	BayeScan (L)	WD repeat and FYVE domain-containing 3 isoform X1	C:GO:0044428; C:GO:0012505; C:GO:0019898; F:GO:0005543; C:GO:0005776
augustus_masked-scaffold8880_len1179_cov45-processed-gene-0.0-mRNA-1	BayeScan (L)	zinc finger BED domain-containing 4-like	F:GO:0003676

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
genemark-scaffold521_len9866_cov60-processed-gene-0.6-mRNA-1	BayeScan (L)	histone acetyltransferase KAT6A	F:GO:0016747; F:GO:0046872; P:GO:0006355
maker-scaffold12277_len15072_cov77-augustus-gene-0.17-mRNA-1	BayeScan (L)	enhancer of mRNA-decapping 4	F:GO:0005515
maker-scaffold9159_len15155_cov69-augustus-gene-0.13-mRNA-1	BayeScan (L)	---NA---	
snap_masked-scaffold11529_len1245_cov71-processed-gene-0.2-mRNA-1	BayeScan (L)	---NA---	
snap_masked-scaffold14188_len19385_cov89-processed-gene-0.9-mRNA-1	BayeScan (L)	RNA-directed DNA polymerase from mobile element jockey-like	
snap_masked-scaffold25654_len3753_cov62-processed-gene-0.4-mRNA-1	BayeScan (L)	---NA---	

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
snap_masked-scaffold29981_len20176_cov62-processed-gene-0.3-mRNA-1	BayeScan (L)	---NA---	
snap_masked-scaffold31186_len5218_cov49-processed-gene-0.3-mRNA-1	BayeScan (L)	---NA---	
snap_masked-scaffold3260_len26479_cov81-processed-gene-0.19-mRNA-1	BayeScan (L)	---NA---	
snap_masked-scaffold34844_len8684_cov79-processed-gene-0.7-mRNA-1	BayeScan (L)	---NA---	
snap_masked-scaffold3615_len1478_cov37_single-processed-gene-0.2-mRNA-1	BayeScan (L)	chitin synthase partial	F:GO:0016758
snap_masked-scaffold413_len24703_cov83-processed-gene-0.16-mRNA-1	BayeScan (L)	---NA---	

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
snap_masked-scaffold47105_len1627_cov43-processed-gene-0.2-mRNA-1	BayeScan (L)	---NA---	
snap_masked-scaffold48206_len4282_cov83-processed-gene-0.8-mRNA-1	BayeScan (L)	---NA---	
snap_masked-scaffold521_len9866_cov60-processed-gene-0.11-mRNA-1	BayeScan (L)	---NA---	
snap_masked-scaffold56321_len989_cov47-processed-gene-0.1-mRNA-1	BayeScan (L)	---NA---	
snap_masked-scaffold77279_len1917_cov60_single-processed-gene-0.3-mRNA-1	BayeScan (L)	---NA---	
snap_masked-scaffold876_len10758_cov81-processed-gene-0.7-mRNA-1	BayeScan (L)	---NA---	
augustus_masked-scaffold10378_len9837_cov66-processed-gene-0.4-mRNA-1	BAYENV (pre.); BAYENV (pro.)	pre-mRNA-splicing factor ATP-dependent RNA helicase DHX16	F:GO:0003676; F:GO:0005524; F:GO:0008026

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
augustus_masked-scaffold14533_len21570_cov82-processed-gene-0.1-mRNA-1	BAYENV (pre.); BAYENV (pro.)	splicing factor 3B subunit 1 isoform X2	F:GO:0005488
augustus_masked-scaffold2001_len31722_cov84-processed-gene-0.4-mRNA-1	BAYENV (pre.); BAYENV (pro.)	regulation of nuclear pre-mRNA domain-containing 1B-like isoform X1	
augustus_masked-scaffold32254_len2672_cov45-processed-gene-0.0-mRNA-1	BAYENV (pre.)	beta- isoform X1	F:GO:0046983
augustus_masked-scaffold39615_len1017_cov43_single-processed-gene-0.0-mRNA-1	BAYENV (pre.); BAYENV (hyd.); BAYENV (pro.)	Furin-like protease isoforms 1 1-X partial	P:GO:0006508; F:GO:0004252
augustus_masked-scaffold4796_len24763_cov74-processed-gene-0.2-mRNA-1	BAYENV (pre.)	probable ATP-dependent RNA helicase DDX27	F:GO:0003676; F:GO:0005524

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
augustus_masked-scaffold90700_len860_cov33_single-processed-gene-0.0-mRNA-1	BAYENV (pre.); BAYENV (pro.)	splicing factor 3B subunit partial	F:GO:0005488
maker-scaffold10677_len62034_cov82-augustus-gene-0.46-mRNA-1	BAYENV (pre.); BAYENV (pro.)	ATP-binding cassette sub-family F member 3	F:GO:0016887; F:GO:0005524
maker-scaffold117958_len6680_cov80_single-augustus-gene-0.4-mRNA-1	BAYENV (pre.)	Triple functional domain	F:GO:0005515
maker-scaffold12216_len29004_cov73-augustus-gene-0.33-mRNA-1	BAYENV (pre.)	aristaless-like homeobox partial	P:GO:0006355; F:GO:0043565
maker-scaffold13979_len32988_cov73-snap-gene-0.20-mRNA-1	BAYENV (pre.); BAYENV (pro.)	titin isoform X5	F:GO:0005515
maker-scaffold147530_len9178_cov83_single-augustus-gene-0.10-mRNA-1	BAYENV (pre.)	RNA exonuclease 1 homolog	F:GO:0003676
maker-scaffold15317_len39562_cov81-augustus-gene-0.36-mRNA-1	BAYENV (pre.)	probable ATP-dependent RNA helicase DDX5 isoform X2	F:GO:0003676; F:GO:0005524

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
maker-scaffold15317_len39562_cov81-augustus-gene-0.36-mRNA-2	BAYENV (pre.)	probable ATP-dependent RNA helicase DDX5 isoform X2	F:GO:0003676; F:GO:0005524
maker-scaffold16727_len43249_cov76-snap-gene-0.49-mRNA-1	BAYENV (pre.); BAYENV (hyd.)	tubulin polyglutamylase ttll6-like isoform X1	F:GO:0005524; P:GO:0006464
maker-scaffold27749_len52370_cov82-snap-gene-0.68-mRNA-1	BAYENV (pre.)	Wee1 kinase	F:GO:0000287; C:GO:0005634; F:GO:0005524; F:GO:0004715; P:GO:0007067; P:GO:0006468

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
maker-scaffold2803_len16312_cov62-snap-gene-0.14-mRNA-1	BAYENV (pre.); BAYENV (hyd.); BAYENV (pro.)	CREB binding	F:GO:0004402; C:GO:0005667; F:GO:0008270; P:GO:0042967; F:GO:0003712; F:GO:0005515; P:GO:0006355; P:GO:0016573; P:GO:0006355; C:GO:0000123
maker-scaffold30593_len90354_cov83-augustus-gene-0.80-mRNA-1	BAYENV (pre.)	serine threonine- kinase NLK	F:GO:0005524; F:GO:0004707; P:GO:0007178; P:GO:0009069; P:GO:0006468
maker-scaffold3316_len54093_cov76-augustus-gene-0.52-mRNA-1	BAYENV (pre.)	ankyrin repeat domain-containing	F:GO:0005515

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
maker-scaffold44875_len6800_cov83_single-augustus-gene-0.7-mRNA-1	BAYENV (pre.); BAYENV (hat.)	tyrosine decarboxylase	F:GO:0030170; P:GO:0006520; F:GO:0016831
maker-scaffold6770_len59888_cov84-augustus-gene-0.55-mRNA-1	BAYENV (pre.); BAYENV (hyd.)	terminal uridylyltransferase 4 isoform X3	F:GO:0003676; F:GO:0008270; F:GO:0016779
snap_masked-scaffold10378_len9837_cov66-processed-gene-0.7-mRNA-1	BAYENV (pre.); BAYENV (pro.)	hypothetical protein T03_17706	
snap_masked-scaffold19225_len16205_cov80-processed-gene-0.15-mRNA-1	BAYENV (pre.)	Myosin- partial	C:GO:0016459; F:GO:0003774
snap_masked-scaffold27749_len52370_cov82-processed-gene-0.40-mRNA-1	BAYENV (pre.)	---NA---	
snap_masked-scaffold2803_len16312_cov62-processed-gene-0.12-mRNA-1	BAYENV (pre.); BAYENV (hyd.); BAYENV (pro.)	---NA---	

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
snap_masked-scaffold32254_len2672_cov45-processed-gene-0.4-mRNA-1	BAYENV (pre.)	---NA---	
snap_masked-scaffold74258_len18563_cov83_single-processed-gene-0.9-mRNA-1	BAYENV (pre.)	---NA---	
augustus_masked-scaffold12414_len1097_cov37_single-processed-gene-0.0-mRNA-1	BAYENV (hyd.)	estrogen receptor beta	C:GO:0005634; F:GO:0003677; F:GO:0003707; P:GO:0006355; P:GO:0043401
augustus_masked-scaffold23515_len12823_cov74-processed-gene-0.0-mRNA-1	BAYENV (hyd.)	SWI SNF complex subunit SMARCC2 isoform X2	F:GO:0003677; C:GO:0090544; P:GO:0006338; F:GO:0005515; P:GO:0006357

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
augustus_masked-scaffold5760_len49789_cov79-processed-gene-0.5-mRNA-1	BAYENV (hyd.)	---NA---	
genemark-scaffold21237_len13443_cov45-processed-gene-0.1-mRNA-1	BAYENV (hyd.); BAYENV (hat.)	U5 small nuclear ribonucleo 200 kDa helicase	F:GO:0003676; F:GO:0005524; F:GO:0005515
genemark-scaffold9416_len27382_cov82-processed-gene-0.7-mRNA-1	BAYENV (hyd.)	---NA---	C:GO:0016592; F:GO:0001104; P:GO:0006357
maker-scaffold10118_len41310_cov87-augustus-gene-0.43-mRNA-1	BAYENV (hyd.)	neuroblastoma-amplified sequence	P:GO:0006890
maker-scaffold1331_len17087_cov60-augustus-gene-0.15-mRNA-1	BAYENV (hyd.)	RING-H2 finger ATL33	F:GO:0008270; F:GO:0005515
maker-scaffold17291_len68892_cov79-augustus-gene-0.52-mRNA-1	BAYENV (hyd.)	BMA-LIM- isoform e	F:GO:0008270

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
maker-scaffold21377_len14810_cov61-augustus-gene-0.11-mRNA-1	BAYENV (hyd.)	AP-2 complex subunit alpha-2 isoform X1	P:GO:0016192; C:GO:0030131; F:GO:0008565; F:GO:0005488; P:GO:0006886
maker-scaffold31415_len67326_cov84-augustus-gene-0.65-mRNA-1	BAYENV (hyd.)	probable phospholipid-transporting ATPase IF	F:GO:0000287; F:GO:0005524; P:GO:0006812; P:GO:0015917; C:GO:0016021; F:GO:0004012
maker-scaffold3955_len20842_cov63-augustus-gene-0.14-mRNA-1	BAYENV (hyd.)	E3 ubiquitin- ligase TRIM9	F:GO:0008270; F:GO:0005515; C:GO:0005622

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
maker-scaffold7638_len74487_cov80-augustus-gene-0.56-mRNA-1	BAYENV (hyd.)	apoptosis-inducing factor mitochondrial	F:GO:0016491; F:GO:0050660; F:GO:0046983; P:GO:0045454; P:GO:0055114
snap_masked-scaffold138_len22458_cov56-processed-gene-0.10-mRNA-1	BAYENV (hyd.); BAYENV (hat.)	hepatocyte nuclear factor 3-beta-like	C:GO:0005667; F:GO:0003700; P:GO:0006355; F:GO:0043565; P:GO:0006355
snap_masked-scaffold21377_len14810_cov61-processed-gene-0.6-mRNA-1	BAYENV (hyd.)	---NA---	
snap_masked-scaffold2345_len24629_cov66-processed-gene-0.12-mRNA-1	BAYENV (hyd.)	---NA---	
snap_masked-scaffold23515_len12823_cov74-processed-gene-0.8-mRNA-1	BAYENV (hyd.)	---NA---	

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
snap_masked-scaffold26526_len6308_cov119-processed-gene-0.3-mRNA-1	BAYENV (hyd.)	Kinase suppressor of Ras 2	P:GO:0035556
snap_masked-scaffold3316_len54093_cov76-processed-gene-0.42-mRNA-1	BAYENV (hyd.)	ankyrin-3- partial	F:GO:0005515
snap_masked-scaffold34622_len3477_cov45-processed-gene-0.2-mRNA-1	BAYENV (hyd.)	polycystin-1-like isoform X2	F:GO:0005515; C:GO:0016020
snap_masked-scaffold654_len169985_cov81-processed-gene-0.114-mRNA-1	BAYENV (hyd.)	---NA---	
snap_masked-scaffold654_len169985_cov81-processed-gene-0.150-mRNA-1	BAYENV (hyd.)	---NA---	
snap_masked-scaffold70385_len24210_cov75-processed-gene-0.15-mRNA-1	BAYENV (hyd.)	---NA---	
augustus_masked-scaffold13471_len64125_cov81-processed-gene-0.6-mRNA-1	BAYENV (pro.)	ribonuclease kappa-B-like	P:GO:0051252; P:GO:0009303; F:GO:0004521

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
augustus_masked-scaffold16670_len17668_cov76-processed-gene-0.0-mRNA-1	BAYENV (pro.)	---NA---	
augustus_masked-scaffold23115_len40371_cov78-processed-gene-0.3-mRNA-1	BAYENV (pro.)	---NA---	
augustus_masked-scaffold2556_len23011_cov68-processed-gene-0.4-mRNA-1	BAYENV (pro.)	CREB-binding -like isoform X2	F:GO:0004402; C:GO:0005667; F:GO:0008270; P:GO:0042967; F:GO:0003712; P:GO:0006355; P:GO:0006355; C:GO:0000123
augustus_masked-scaffold35995_len16129_cov75-processed-gene-0.3-mRNA-1	BAYENV (pro.)	Apoptosis-stimulating of p53 1	F:GO:0005515

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
augustus_masked-scaffold4469_len52091_cov83-processed-gene-0.3-mRNA-1	BAYENV (pro.)	acid-sensing ion channel 2	P:GO:0006814; F:GO:0005272; C:GO:0016021
augustus_masked-scaffold5316_len17262_cov77-processed-gene-0.2-mRNA-1	BAYENV (pro.)	---NA---	
augustus_masked-scaffold56872_len4744_cov49-processed-gene-0.1-mRNA-1	BAYENV (pro.); BAYENV (hat.)	kyphoscoliosis peptidase-like	F:GO:0008270; F:GO:0046872
augustus_masked-scaffold9166_len20044_cov73-processed-gene-0.1-mRNA-1	BAYENV (pro.)	DNA-directed RNA polymerase II subunit RPB2	F:GO:0003677; P:GO:0006206; F:GO:0003899; F:GO:0032549; P:GO:0006351; P:GO:0006144; C:GO:0005730
genemark-scaffold20036_len5260_cov55-processed-gene-0.1-mRNA-1	BAYENV (pro.)	zinc finger 423 homolog isoform X2	F:GO:0003676; F:GO:0046872

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
genemark-scaffold2809_len115511_cov82-processed-gene-0.43-mRNA-1	BAYENV (pro.)	ATP-dependent DNA helicase Q4	F:GO:0003676; F:GO:0005524
genemark-scaffold3326_len7895_cov59-processed-gene-0.1-mRNA-1	BAYENV (pro.)	acetyl- carboxylase 1 isoform X1	F:GO:0005524; F:GO:0003989; P:GO:0006633; P:GO:0006090; C:GO:0009317
genemark-scaffold85069_len38045_cov83_single-processed-gene-0.6-mRNA-1	BAYENV (pro.)	myosin-IIIb isoform X2	C:GO:0016459; F:GO:0005524; F:GO:0005515; F:GO:0003774
genemark-scaffold9636_len4626_cov57-processed-gene-0.4-mRNA-1	BAYENV (pro.)	probable G- coupled receptor No18	F:GO:0004930; P:GO:0007186; C:GO:0016021
maker-scaffold29906_len31076_cov80-augustus-gene-0.25-mRNA-1	BAYENV (pro.)	protocadherin-1 isoform X2	F:GO:0005509; P:GO:0007156; C:GO:0005886

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
maker-scaffold3406_len22921_cov75-augustus-gene-0.24-mRNA-1	BAYENV (pro.)	nipped-B B isoform X1	P:GO:0007064; P:GO:0010468; F:GO:0003682; C:GO:0000785
maker-scaffold3918_len81077_cov82-augustus-gene-0.99-mRNA-1	BAYENV (pro.)	speedy A	C:GO:0005634; F:GO:0019901; P:GO:0045737; P:GO:0007140; P:GO:0008284; P:GO:0000082; P:GO:0006974
maker-scaffold6500_len58321_cov77-snap-gene-0.60-mRNA-1	BAYENV (pro.)	interferon-inducible GTPase 5-like	F:GO:0005525; C:GO:0016020
maker-scaffold9598_len134620_cov83-augustus-gene-0.120-mRNA-1	BAYENV (pro.)	very long-chain specific acyl- mitochondrial-like	F:GO:0003995; P:GO:0006118; F:GO:0050660; P:GO:0055114

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
maker-scaffold9850_len50016_cov84-snap-gene-0.45-mRNA-1	BAYENV (pro.)	pre-rRNA processing FTSJ3	C:GO:0005634; F:GO:0008649; P:GO:0031167
snap_masked-scaffold2556_len23011_cov68-processed-gene-0.14-mRNA-1	BAYENV (pro.)	---NA---	
snap_masked-scaffold29906_len31076_cov80-processed-gene-0.19-mRNA-1	BAYENV (pro.)	---NA---	
snap_masked-scaffold3326_len7895_cov59-processed-gene-0.4-mRNA-1	BAYENV (pro.)	---NA---	
snap_masked-scaffold3918_len81077_cov82-processed-gene-0.60-mRNA-1	BAYENV (pro.)	---NA---	
snap_masked-scaffold4469_len52091_cov83-processed-gene-0.48-mRNA-1	BAYENV (pro.)	---NA---	
snap_masked-scaffold56872_len4744_cov49-processed-gene-0.3-mRNA-1	BAYENV (pro.); BAYENV (hat.)	---NA---	

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
snap_masked-scaffold6876_len69886_cov83-processed-gene-0.51-mRNA-1	BAYENV (pro.)	---NA---	
snap_masked-scaffold9166_len20044_cov73-processed-gene-0.13-mRNA-1	BAYENV (pro.)	---NA---	
snap_masked-scaffold9598_len134620_cov83-processed-gene-0.71-mRNA-1	BAYENV (pro.)	---NA---	
augustus_masked-scaffold11329_len46584_cov79-processed-gene-0.1-mRNA-1	BAYENV (hat.)	---NA---	
augustus_masked-scaffold1152_len15314_cov63-processed-gene-0.0-mRNA-1	BAYENV (hat.)	pre-mRNA-processing-splicing factor 8	F:GO:0017070; F:GO:0030623; F:GO:0005515; C:GO:0005681; P:GO:0000398

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
augustus_masked-scaffold122774_len770_cov73_single-processed-gene-0.0-mRNA-1	BAYENV (hat.)	inositol 1,4,5-trisphosphate receptor isoform X1	C:GO:0016020; F:GO:0005220; P:GO:0070588; C:GO:0005783
augustus_masked-scaffold134414_len14929_cov83_single-processed-gene-0.1-mRNA-1	BAYENV (hat.)	hypothetical protein LOTGIDRAFT_129214	F:GO:0003824; P:GO:0008152
augustus_masked-scaffold13725_len21899_cov79-processed-gene-0.4-mRNA-1	BAYENV (hat.)	ATP-binding cassette sub-family E member 1	F:GO:0016887; F:GO:0005524
augustus_masked-scaffold147_len65044_cov80-processed-gene-0.0-mRNA-1	BAYENV (hat.)	probable ubiquitin carboxyl-terminal hydrolase FAF-X isoform X1	P:GO:0016579; F:GO:0005515; F:GO:0036459; P:GO:0006511
augustus_masked-scaffold23664_len560_cov60_single-processed-gene-0.0-mRNA-1	BAYENV (hat.)	hypothetical protein	

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
augustus_masked-scaffold32615_len5806_cov64-processed-gene-0.0-mRNA-1	BAYENV (hat.)	plexin-D1 isoform X3	F:GO:0017154; P:GO:0071526
augustus_masked-scaffold40852_len16878_cov80-processed-gene-0.1-mRNA-1	BAYENV (hat.)	ATP synthase mitochondrial F1 complex assembly factor 2	P:GO:0043461
augustus_masked-scaffold4390_len5468_cov60-processed-gene-0.2-mRNA-1	BAYENV (hat.)	Mediator of RNA polymerase II transcription subunit 31	C:GO:0016592; P:GO:0006355; F:GO:0001104
augustus_masked-scaffold53578_len2112_cov34-processed-gene-0.0-mRNA-1	BAYENV (hat.)	---NA---	
augustus_masked-scaffold6238_len20409_cov82-processed-gene-0.1-mRNA-1	BAYENV (hat.)	signal recognition particle subunit SRP68	F:GO:0030942; P:GO:0006614; F:GO:0008312; F:GO:0005047; C:GO:0005786

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
genemark-scaffold131895_len5770_cov90_single-processed-gene-0.3-mRNA-1	BAYENV (hat.)	vacuolar sorting-associated partial	C:GO:0030904; P:GO:0015031; F:GO:0008565; P:GO:0042147
genemark-scaffold2888_len28769_cov78-processed-gene-0.10-mRNA-1	BAYENV (hat.)	---NA---	
genemark-scaffold3985_len38543_cov65-processed-gene-0.13-mRNA-1	BAYENV (hat.)	arginine vasopressin receptor partial	F:GO:0004930; P:GO:0007186; C:GO:0016021
maker-scaffold10086_len11099_cov49-augustus-gene-0.10-mRNA-1	BAYENV (hat.)	histone-arginine methyltransferase CARM1	F:GO:0008168; P:GO:0006479
maker-scaffold10472_len17195_cov84-augustus-gene-0.16-mRNA-1	BAYENV (hat.)	Cilia- and flagella-associated 58	C:GO:0005615
maker-scaffold3147_len92768_cov81-augustus-gene-0.92-mRNA-1	BAYENV (hat.)	mucin-17 isoform X1	

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
maker-scaffold5248_len36051_cov70-snap-gene-0.46-mRNA-1	BAYENV (hat.)	radial spoke head 14 homolog	F:GO:0005488
maker-scaffold661_len35413_cov82-snap-gene-0.27-mRNA-1	BAYENV (hat.)	---NA---	
maker-scaffold6779_len77775_cov82-augustus-gene-0.85-mRNA-1	BAYENV (hat.)	centrosomal of 97 kDa	F:GO:0005515
snap_masked-scaffold103554_len21087_cov72-processed-gene-0.19-mRNA-1	BAYENV (hat.)	coronin-7 isoform X1	F:GO:0005515
snap_masked-scaffold11329_len46584_cov79-processed-gene-0.30-mRNA-1	BAYENV (hat.)	---NA---	
snap_masked-scaffold1152_len15314_cov63-processed-gene-0.6-mRNA-1	BAYENV (hat.)	---NA---	
snap_masked-scaffold13725_len21899_cov79-processed-gene-0.14-mRNA-1	BAYENV (hat.)	hypothetical protein BN1708_000423	

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
snap_masked-scaffold147_len65044_cov80-processed-gene-0.38-mRNA-1	BAYENV (hat.)	---NA---	
snap_masked-scaffold23664_len560_cov60_single-processed-gene-0.2-mRNA-1	BAYENV (hat.)	---NA---	
snap_masked-scaffold3147_len92768_cov81-processed-gene-0.51-mRNA-1	BAYENV (hat.)	---NA---	
snap_masked-scaffold32615_len5806_cov64-processed-gene-0.6-mRNA-1	BAYENV (hat.)	---NA---	
snap_masked-scaffold5248_len36051_cov70-processed-gene-0.36-mRNA-1	BAYENV (hat.)	---NA---	
snap_masked-scaffold53578_len2112_cov34-processed-gene-0.3-mRNA-1	BAYENV (hat.)	PRDM9 zinc finger domain partial	F:GO:0003676; F:GO:0046872
snap_masked-scaffold7727_len7764_cov58-processed-gene-0.2-mRNA-1	BAYENV (hat.)	---NA---	

Table A.5. (continued)

Gene name	Software (Parameter)	Description	Gene ontology number
snap_masked-scaffold82913_len3970_cov38-processed-gene-0.4-mRNA-1	BAYENV (hat.)	RNA-directed DNA polymerase from mobile element jockey-like	