Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Scienze della Terra, della Vita e dell'Ambiente

Ciclo 29

Settore Concorsuale di afferenza: 05/B1 Zoologia e Antropologia

Settore Scientifico disciplinare: BIO/05 Zoologia

TITOLO TESI

Differential gene expression of dorsal pictorial ornaments and pigmentation in skates (Rajidae, Chondrichthyes)

Presentata da: Alice Ferrari

Coordinatore Dottorato
Prof. Barbara Mantovani

Relatore
Prof. Fausto Tinti

Co-Relatori
Prof. Walter Salzburger

Alessia Cariani, PhD

Esame finale anno 2017

ABSTRACT

Approximately twenty years have passed since the beginning of concentrated investigations on the evolution and ecology of skates. The evidence generated thus far suggested that this monophyletic group have experienced multiple, parallel adaptive radiations at a regional scale, which contributed to the delineation of strong phylogeographical signal since the Cretaceous. This background represented the guiding light of the work described in this thesis, where two main themes were developed. The first one focused on the investigation of *Raja miraletus* L. species complex through the analysis of genetic variation derived from both mtDNA and nuDNA. The results presented herein assessed the presence of a restricted gene flow and different degree of divergence between the South African and Mediterranean samples, ascribing these patterns to oceanographic discontinuities. Despite the high species diversity characterising the Family, most Rajidae show a stable gross morphology and peculiar dorsal pigmentation patterns, which may have been implicated in cryptic speciation. Nonetheless, the adaptive value and the genetic basis of these traits remain poorly investigated. To fill this gap, this thesis also describes the application of RNA-sequencing technology on recently diverged skate species with sibling and sister phylogenetic relationships. Therefore, the second goal of this research consisted in investigating the molecular basis of pigmentation in five non-model species. To this end, the transcriptome profiling of different skin tissues was performed using the Illumina platform, whereas longer sequencing data were obtained from *R. miraletus* multiple organs using the Ion Torrent technology. After the assembly of a reference transcriptome and the mapping of Illumina reads, the Differential Gene Expression between skin tissues across species was performed, revealing the expression of transcripts mainly related to metabolic process and catalytic activity in which pigmentary genes appeared involved. This work could be considered the basis for future studies aiming to disentangle how pigmentary traits evolved in skates and other chondrichtyans, to evaluate whether the same or alternative traits have been used in parallel adaptations to similar environments and to understand if these traits follow species divergence or hybridisation.

TABLE OF CONTENTS

Chapter 1. General introduction

1.1. Thesis objectives and overview

This dissertation focuses on Rajidae (Chondrichthyes), a family of cartilaginous fish inhabiting the continental shelf and slope of oceans worldwide and whose evolutionary history is far from being fully elucidated.

The goals of this PhD research can be divided in two main, although interconnected, themes. First, the detection of genetic diversity, phylogeographic patterns and population structure of *Raja miraletus* L. complex aimed to highlight the cryptic evolution hidden under the differentiation of at least three lineages inhabiting the Mediterranean Sea and off the Atlantic coast of Africa.

Second, this research aimed to identify the molecular basis underpinning skates' pigmentation along with other genes connected to skin functions (i.e. collagen structures, metabolism, mucus production and immune response). Most of all, the work aimed to identify transcripts related to pigmentation, to perform the Differential Gene Expression (DGE) analysis of differently pigmented and non-pigmented skin and eventually compile a transcript catalogue for this tissue. The surplus value of this research was adding a tile to the genomic resources currently available for five non-model species and lastly, opening the door to the investigation on the evolution of pigmentary genes and their patterns of adaptation to specific ecological conditions (e.g. as an effect of convergence or the repeated evolution of similar phenotypes serving the same ecological function in two or more taxa). Specific research tasks and questions were separately addressed in each chapter.

Chapter 1 begins with a brief introduction to the species concept and the undertones assumed within this thesis, putting in the spotlight those taxa which are hidden under the same nominal species and highly contribute to enrich the biodiversity inventory. Chapter 1 thus focuses on cryptic speciation in bony fish and chondrichthyans, reporting the most interesting and significant cases.

In Chapter 2 the case of the brown skate *Raja miraletus* is described as an additional evidence of cryptic evolution within a species complex. The genetic variability and differentiation between different geographical populations was investigated at the mitochondrial and nuclear DNA level. The existence of at least three different lineages was discussed, supporting the identification of the South African brown skate as a resurrected species, *Raja ocellifera*, included in the nominal *Raja miraletus*

since 1967.

Chapter 3 illustrated the path leading to the Illumina sequencing of different skin tissues belonging to five different species of genus *Raja*, which are phylogenetically related by sibling and sister relationships and display different dorsal patterning, ascribable to their adaptation to the benthic life.

Chapter 4 focused on the production of long sequencing reads for the reconstruction of a draft transcriptome for *R. miraletus*, from the sampling phase to the library preparation and Ion Torrent sequencing. The transcriptome has been then used during the mapping of the Illumina reads obtained from skin tissues, to lastly perform the Differential Expression analysis (DE), as described in chapter 5.

Chapter 5 introduced to the main critical points and issues especially related to RNA-seq data analysis. Details about the methods chosen are discussed. This final section includes the main results of DE analysis between different skin tissues across five target species and provides a first overview on the functional role of the identified genes. Finally, suggestions are given for future experiments that can validate and expand on the knowledge presented here. In fact, results of this work could be considered the starting point for a deeper and wider exploration of chondrichthyan pigmentary genes.

1.2. Species and cryptic diversity

While reading '*How many species are there on Earth?*' today, the 3 to 5 million species estimated in 1988 (May 1988) appear far from the most recent biodiversity inventory, where about 8.7 million eukaryotic species are predicted to be inhabiting the planet. It has also been appraised that 86% of terrestrial and 91% of marine species have not been described yet (Mora et al. 2011). How long should it take to fill this gap in our knowledge and how should we proceed in those under-sampled regions characterized by high species richness or simply not physically reachable? While answering these questions, Mora et al. (2011) raised an interesting paradox: considering 6,200 species described per year, 24.8 new species described per taxonomist and assuming that these values, together with the current rate of extinction, will remain constant in time and among taxa, characterizing new species may take as long as 1,200 years and the industry of 303,000 taxonomists, meaning that many species will become extinct before we would know they walked the earth or 'swam the oceans' (Mora et al.2011). Nonetheless, it is plausible that this filling-the-gap-issue would be simplified if considering the integration of DNA-based technologies to traditional taxonomy (Cariani et al. 2017; Landi et al. 2014; Costa et al. 2012).

In general, underestimating diversity and overestimating species ranges are the main obstacles to the compilation of a comprehensive biodiversity inventory, especially if build on 'taxonomic lumping', i.e. the practice of combining morphologically similar forms from different geographic regions under a single species name. As a matter of fact, this procedure could concern subspecies, geographic forms, morphotypes, and other nominally intraspecific taxa that have been recognized on the basis of divergence in particular traits (Funk & Omland, 2003). According to Funk & Omland (2003), the monophyletic origin of one or more intraspecific taxa within a nominal species could be a symptom of genetic isolation of these taxa from other 'conspecifics'. Therefore, when reproductive isolation is placed between distinct but sympatric clades, we are likely in front of different species, but when distinct clades are also geographically separated, then the evidence of reproductive compatibility is not fully ascertainable and the decision of whether to recognize them as separate species becomes more difficult.

The two centuries-debate about the definition of species is still open and contributes to a forced, but necessary, categorization of what we simply consider as the fundamental taxonomic unit of classification (Leliaert et al. 2014; De Queiroz 2007; Brookfield 2002). Reaching the species essence on its whole has always involved in some degree both biologist and philosophers (Hull 1965; Sobert,

2008; Claridge et al. 1997), who questioned on species ontological status as natural kinds, individuals or sets of organisms or focused on membership and explanatory requirements described by the puzzling Relational Essentialism (Malpas, J. 'Donald Davidson', The Stanford Encyclopaedia of Philosophy, 2012). The choice of species concept when describing biodiversity is fundamental for our knowledge of diversity and distribution and affects our understanding of species evolutionary history. The Morphological Species Concept (MSC) is certainly the oldest theory defining species according to their phenotypic differences (Cracraft 1983), although not the most accurate as described later. In this thesis, the species denotation should be then interpreted mainly according to the Phylogenetic (PSC; Cracraft 1989) and Biological (BSC; Mayr 1963) species concepts, as a group of organisms bound by a unique ancestry that can successfully interbreed and produce fertile offspring (Freeland & Boag 2016; Cracraft 1983; Bush, 2016; Mayr 1942). This choice is concordant with Avise and Ball work (1990) who, in the effort to integrate speciation concepts, advanced the use of multiple independent loci in species definition. The authors suggested that the time necessary to major phylogenetic distinctions to accumulate concordantly at independent loci is presumably adequate for intrinsic reproductive barriers to form. Consequently, populations showing concordant and reciprocally monophyletic patterns over those loci are hence estimable as single taxonomic units. In addition, the Phylogeographic facet will be fundamental as well throughout this thesis, since it adds a timescale to the understanding of population structure and genealogies, reproductive isolation of population units and speciation events.

In general, morphologically well-differentiated species are often characterized by similar genetic traits due to rapid divergence mechanisms at times accompanied by incomplete lineage sorting or hybridization (i.e. rapid adaptive radiations in African cichlids, Böhne et al. 2016 ; Ivory et al. 2016; Malinsky & Salzburger 2016 or in lizards, Feiner, 2016; Medina et al. 2016; Barley et al. 2013). The widespread detection of an alternative paradigm, where evolutionary lineages are discriminated genetically, but not morphologically, can be an issue for species identification and a contribution to the underestimation of biodiversity in systems involving 'nonadaptive radiations' (Rundell & Price 2009; Jockusch & Wake 2002; Gittemnerger 1991) and 'cryptic evolution' in species complexes (Clare 2011; Funk et al. 2012; Pfenninger & Schwenk 2007; Stuart et al. 2006).

Anyone googling 'cryptic species on the web, would hunt approximately 4,290,000 entries. A total of 983 academic papers are available on public portals, corresponding to about one tenth of the records obtained using any scientific literature database (e.g. ISI Web of Knowledge roughly counts 9,681 entries, among articles, reviews, proceeding papers, meeting abstracts and editorial materials).

Beside the meaning of these fast-changing numbers, what is the significance of cryptic diversity?

Cryptic species can be defined as two or more distinct but morphologically similar species that were ranked and hidden as a single nominal one (Bickford et al. 2007; Pfenninger & Schwenk 2007). The discovery of this phenomenon boosted with the advent of PCR-based approaches, in particular of DNA barcoding (Hebert et al. 2004), which has shown to be helpful in species diagnosis, even between closely related species (Moore 2016; Avise & Walker 1999; Hebert et al. 2003). This approach highly contributed in revealing the astonishing diversity around us and arose both questions and criticisms about cryptic taxa acknowledgment, frequency, age, distribution and boundaries.

Moreover, these species are the most studied and criticized since two decades. Particular caution should be paid when observing deep genetic divergences within species, which do not always imply cryptic speciation events (Horne et al. 2008). Conversely, when deep evolutionary splits are in agreement across both mitochondrial and nuclear DNA markers (henceforth mtDNA and nuDNA), are geographically isolated or show morphological and behavioural dissimilarities (Victor 2015; Knudsen & Clements 2012) then the signal of a cryptic speciation event might be loud and clear. As a matter of fact, an increasing number of DNA-based studies suggest that morphologically similar species can be the product of ancient speciation events (Beheregaray & Caccone 2007).

Despite criticisms regarding the 'taxonomic inflation' issue and other controversies (Trontelj & Fiser 2009; Meiri & Mace 2007; Rubinoff & Holland 2005) cryptic diversity has been richly documented among marine invertebrates (Knowlton 2009) as foraminifera (de Vargas et al. 1999), crustaceans (Baldanzi et al. 2016; Pilgrim et al. 2013; Navarro-Barranco et al. 2013; Schiffer & Herbig 2016), annelids (Grosemans et al. 2016; Johnson et al. 2016; Kawauchi & Giribet 2014), mollusc (Huelsken et al. 2013; Barco et al. 2013; Jörgeret et al. 2012) and echinoderms (Boissin et al. 2011), which show paradoxical wide distribution, but theoretical low dispersal capacity.

More and more case studies are focusing also on vertebrate hidden species among ascidians (Kwan et al. 2014; Pérez-Portela et al. 2013), teleost (Nirchio et al. 2016; Fernandez-Silva et al. 2015; Bradbury et al. 2014; Corander et al. 2013), amphibians (Funk et al. 2012; Elmer et al. 2007), birds (Friesen et al. 1996), reptiles (Ukuwela et al. 2014) and, unexpectedly, on mammalian megafauna (Clare 2011; Brown et al. 2007; Dalebout et al. 2002; Garcia-Rodriguez et al. 1998). Their discovery and cataloguing are fundamental and critical for their conservation and habitat management, legal protection and distribution of limited resources (Crozie et al. 2005; Agapow et al. 2004; Daugherty et al. 1990).

The striking underestimate of global species diversity is particularly important because it can occur in biodiversity conservation hotspots (Brown & Diesmos 2009), but it can't be excluded in other realms, as the marine one, where speciation events and different evolutionary radiations are more and more extensively documented (Henriques et al. 2016; Hodge & Bellwood 2016; Pereira et al. 2016; Siqueira et al. 2016; Frable et al. 2015; Henriques et al. 2015; Craig et al. 2009; Neilson & Stepien, 2009; Ward et al. 2008; Knowlton 2000; Palumbi 1992).

According to Bickford et al. (2007), the basis of the un-correlation between morphological change and species delimitation should be sought in two main recurrent conditions: the lack of visual communication among conspecifics and the conservation of the external gross morphology.

Regarding the first condition, a nonvisual language is likely to be intensified in the marine realm, where organism interact via chemical or electric signals (Klimley 2013; Feulner et al. 2009; Sisneros & Tricas 2002; Sisneros & Tricas 2002; Knowlton 2000). These features could be considered the key of survival for those species arisen so recently that most morphological traits have not yet diverged but, on the other hand, the conservation of external morphology could be the result of stabilizing selection on important adaptive traits in species occupying similar habitat types and ecological niches (Losos & Glor 2003; Schonrogge et al. 2002). Some authors also refer to morphological and ecological stasis as the result of a 'living-fossil' condition, being the latter typically characterized by bradytely (Eldredge & Stanley 1984) and a relatively unmodified gross morphology associated with molecular divergences (Erdmann et al. 1999; Bowen et al. 1993). Considered the rarity of living fossils, in most cases a strong selection may be driving the conservation of external morphological traits, even if the evidence of this phenomenon is not always easily ascertainable. Furthermore, the maintenance of similar external traits can vary a lot on the evolutionary scale, meaning that in some cases a consistent species diagnosis based on morphological characters is not possible, whereas in other cases morphological differentiation appears to have occurred in minor measure, involving only slight differences in body size, for instance.

Among tropical fish, for example, the molecular divergence between evolutionary lineages can be congruent with the observed morphological distinction, e.g. the commercially valuable *Cephalopholis hemistiktos* (Rüppell, 1830), the yellow-fin hind (Priest et al. 2016), while other species show a morphological stability on a strikingly wide range. An example of species differentiation driven by stabilizing selection is the tropical and subtropical genus *Albula* (Scopoli, 1777) which has diverged 3–20 Mya because of the balancing selective force acting in sand habitats (Andrews et al. 2016; Galdino Brandão et al. 2016; Henriques et al. 2016; Henriques et al. 2015; Colborn, et al. 2001),

considered ecologically homogenous all over the world (Bowen & Karl 2006).

Theoretically, a high number of cryptic events should be detected in the richest hotspots for biodiversity. This is the case of the Coral Triangle, where the genus *Eviota* (O.P. Jenkins, 1903) has started developing 23 Mya, after the geological formation of the hotspot (Renema et al. 2008) and evolved during the Pleistocene in more recent species complexes characterized by different regional colour morphs and at least six genetic lineages per complex (Tornabene e t al. 2015).

From one end to another, extreme habitats can also be the nest of cryptic evolution. Around Antarctica, the peculiar geological, climatic and oceanographic conditions fostered the endemic marine fauna to exhibit a higher than expected species richness, including cryptic species (Strugnell & Allcock 2013; Allcock & Strugnell 2012; Thatje 2012; Thornhill et al. 2008; Wilson et al. 2007; Allcock et al. 2011). The yellow-fin rockcod *Lepidonotothen nudifrons* (Lönnberg, 1905) is one of them, having demonstrated to comprise two genetically distinct and geographically isolated partitions that are morphologically identical because of a speciation event fitting into adaptive radiation theory (Dornburg et al. 2016). The authors found that the hidden diversity of this species could be a result of the frequent glacial cycles occurred over the past 5 million years, which could have forced niches' shifts and/or restricted gene flow over long distances (Dornburg et al. 2016) instead of providing ecological opportunities for phenotypic change likewise other Antarctic fauna (Near et al. 2012; O'Loughlin et al. 2011). According to the authors, the lack of morphological diversification involves the 'flexible-stem model' proposed by West-Eberhard (2003), which states that phenotypic diversification requires developmental plasticity: the more flexible is the ancestral population in terms of ecology, natural selection and adaptive capacity, the more phenotypic diversity would be characterizing the radiating lineage in function of time and population size. More specifically, large populations persisting over long timescales would be more prone to develop a higher measure of flexibility, necessary for further phenotypic change (West-Eberhard 2003).

Cryptic evolutionary mechanism has been shown to occur into another extreme habitat which has not been extensively sampled: the deep sea. Independent lineages were recorded in genus *Etelis* (Couvier, 1828), the Deepwater snappers (Andrews et al. 2016), where the nominal *E. carbunculus* was indeed composed of two partitioned species, co-occurring in North-Western Australia, Cocos Island, New Caledonia, Tonga, Wallis and Futuna, Fiji and Samoa. Concordantly with extensive morphological similarity, the lineages have diverged 4-5 Mya in an historic period comprising a reduced water flow between the separated Indian and Pacific Basins (16–8 Mya) and the flow fluctuations occurred during the Pleistocene (2.6–0.12 Mya; Andrews et al. 2016).

According to the reported examples, the role and importance of paleoclimatic history and hydro-geographic discontinuities in driving genetic diversity, population divergences and species demography became evident. But what could we expect from ancient and, in some cases, slow-mutating species (Martin & Palumbi 1993)?

1.3. Cryptic lineages in chondrichthyans

The evolutionary history of cartilaginous fish dates back to the early Devonian, about 400 Mya and 200 Ma before the origin of teleost, as revealed by the fossil records collected between Europe, Lebanon and North Africa (Maisey 2012). Despite their ancient origin, this class is far from being strictly primitive. On the contrary, chondrichthyans evolved sophisticated reproductive strategies and means of communication. Elasmobranchs like mammals, perform courtship, internal fertilization and females experience gestation, during which an extended development may be necessary to enhance the mediation of complex social and sexual behaviour (Demski 1990a). Tricas et al. (1995) have also described the role of the electro-sensory system and the neuroendocrine system while localizing the sexual partner (e.g. *Urolophus halleri* produces a weak electric signal that is rhythmically modulated by the movements of the spiracles and gills during ventilation. The signal is detected by other buried skates via the ampullae of Lorenzini). Nevertheless, their slow growth and longevity (Cailliet et al. 1995; Camhi et al. 1998), low recovery capacity, low fecundity and late maturity (i.e. between 10 and 20 years for males and 11.5 to 30 years for females of *Etmopterus baxteri*; Irvine et al. 2006) are well known among marine biologists and conservationists: most species spend relatively long gestation periods (approximately one year) and produce few egg-capsules per season (Dudgeon et al. 2012; Pratt & Carrier 2001). This K-selected life history trait makes this Class particularly vulnerable to anthropogenic stressors and environmental change (Cariani et al. 2017; Ball et al. 2016; Kousteni et al. 2015; Griffiths et al. 2010; Richards et al. 2009; Corrigan et al. 2008; Abercrombie et al. 2005), and because of their intrinsic sensitivity and susceptibility to overfishing and by-catch, erasing any taxonomic uncertainty is fundamental to their survival and conservation (Bonello et al. 2016).

It has been estimated that more than 1100 cartilaginous fish species (sharks, skates, rays and chimaeras) inhabit the oceans worldwide, while the Mediterranean Sea alone hosts 89 species (about 7% of the global diversity; Cariani et al. 2017). Speaking of cryptic diversity, could we predict a similar scenario as the one observed among marine teleost? This is certainly the case.

Despite fewer cases than among other taxa, molecular approaches have been used to distinguish

cryptic species for elasmobranch groups as well (Pavan-Kumar et al. 2014; Ovenden et al. 2011). Angel sharks *Squatina* spp. (Solé-Cava & Levy 1987), thresher sharks *Alopias* spp. (Eitner, 1995), hound sharks *Mustelus* spp. (Gardner et al. 2002; (Heemstra 1997), shovelnose guitarfish *Rhinobatos* spp. (Sandoval-Castillo et al. 2004), catsharks *Galeus* spp. (Castilho et al. 2007), wobbegong sharks (Corrigan et al. 2008), the spotted eagle ray *Aetobatus* (Richards et al. 2009) and Atlantic skates genus *Dipturus* (Griffiths et al. 2010; Iglésias et al. 2010) have all demonstrated to hide parallel or intricate genealogies driven by hydrographic barriers and paleo-geological events. Below, two interesting case studies are reported and one more has been developed and investigated within this PhD research (see Chapter 2).

Among Squaliformes, the deep-sea lanternsharks genus *Etmopterus* are not commercially interesting, but they represent a significant by-catch component of deep-water trawl and longline fishery (Clarke et al. 2005; Compagno et al. 2005; Jakobsdóttir 2001; Wetherbee & Nichols 2000; Wetherbee et al. 1996). The taxonomic confusion characterising the genus still endangers species conservation and reflects an even more tangled pattern of interspecific and intraspecific relationships within at least six nominal species. As other relatives (e.g. Schaaf da Silva & Ebert 2006; Ward et al. 2008; 2005), lanternsharks display an unclear cryptic diversity within the Southern Hemisphere (Straube et al. 2011). Paleoclimatic events mainly related to the deep-sea cooling (e.g. temperatures decreased from 12°C down to 4.5°C about 34 Mya; Zachos et al. 2001) and ice sheet development on Antarctica were the cause of its separation from the surrounding continents. The opening of the Tasman and Drake passages allowed circumpolar circulation and thermal isolation (Dingle & Lavelle 2000).

Many authors have demonstrated the existence of a cryptic Atlantic lineage related to the cosmopolitan and endangered *Sphyrna lewini* (Griffith & Smith 1834), the scalloped hammerhead shark (Abercrombie et al. 2005; Duncan et al. 2006). This species shows a deep divergence between the Atlantic and Indo-Pacific monophyletic populations according to mtDNA and morphological differentiation as well (Quattro et al. 2006). The corresponding two morphs previously described by Springer (1941) as *S. diplana* (from Atlantic Ocean) and *S. lewini* (from Indo-Pacific region) were later synonymized via 'taxonomic lumping' by Fraser-Brunner (1950) and Gilbert (1967). A third, deeper and isolated lineage was also detected along the coasts from North Carolina to Florida, which earned the name of Carolina hammerhead, aka *S. gilberti* sp. Nov (Quattro et al. 2013) and, more recently, along the Western Atlantic Coasts (Pinhal et al. 2011), widening the distributional area of the species of more than 7,000 km. The uplift of the Isthmus of Panama (3.1-3.5 Mya; Coates et al. 2004; Avise

et al. 2000; Coates & Obando 1996) seems to have influenced the phylogeographic history of the scalloped hammerhead shark (Duncan et al. 2006; Quattro et al. 2006), as well as many other terrestrial, freshwater and marine species (Andrews et al. 2016; Hodge & Bellwood 2016; Schiffer & Herbig 2016; Bowen & Karl 2006; Colborn et al. 2001; Bermingham et al. 1997), including batoids (Richards et al. 2009).

In the next chapter the case of the *Raja miraletus* species complex will be reported as a further exemplification of hidden diversity among the family Rajidae.

Chapter 2. The name game: when evolution hides behind unchanging morphology

2.1. Introduction

Skates (Rajoidei, Rajiformes) are marine and brackish elasmobranchs distributed worldwide that paradoxically exhibit an extraordinary species diversity and high degree of endemism paired with high levels of morphological and ecological stasis (Ebert & Compagno 2007). They are bottom dweller organisms that generally inhabit the sandy habitats of continental plates and shelves, which are frequently up to 1,500 m in depth (Ebert & Compagno 2007). Instead of visual signals, they use a highly efficient electro-sensory system for mating recognition (Tricas et al. 1995) and detection of prey (Camperi et al. 2007). Skates are oviparous and lay very large and benthic egg capsules, usually with a single embryo (Chiquillo et al. 2014). Species diversity and zoogeography of skates are well described in the regional shelf areas that have been surveyed intensely by scientific trawling programs (e.g. the north-eastern Atlantic-Mediterranean [Stehmann & Bürkel 1984] and the south-eastern Atlantic and western Indian Oceans [Compagno & Ebert 2007]). Skate faunas of the north-eastern Atlantic and Mediterranean shared several nominal species with the central and south African faunas (Froese & Pauly 2017). Among them, the brown skate (or twineye skate) *Raja miraletus* L. 1758 is distributed more widely than expected given the limited potential of dispersal conditioned by a relatively dominant residential behaviour of adults and juveniles (Neat et al. 2015; Frisk, et al. 2014; Wearmouth & Sims 2009; Hunter et al. 2005a ; 2005b) and the lack of egg dispersal (Musick & Ellis 2005). The brown skate is a small-sized, fairly fecund species that is reported to be distributed from northern Portugal, the whole Mediterranean and along the western and south-eastern coasts of Africa (Compagno & Ebert 2007; Stehmann & Bürkel 1984). This skate exhibits a pronounced benthic ecology, with depth range from shallow waters to ~450 m, but with most records from 10m to 150m on sandy and hard bottoms (Serena & Mancusi 2010; Compagno & Ebert 2007) and a generalist feeding behaviour (Kadri et al. 2014; Šantić et al. 2012). Because of its high and stable abundance over its distribution, the small body size and the early maturation (age at maturity estimated at 2.7 years; Tsikliras & Stergiou 2014), the brown skate is considered resilient to exploitation and assessed as Least Concern in the Red List (IUCN 2009). *Raja miraletus* also exhibits high levels of stasis of the external rough morphology along with its range, with two very distinctive tricolored (blue, black and yellow) bright eyespots on the upper ochre-brownish surface at the base of pectoral fins (Compagno et al. 1989; Stehmann & Bürkel 1984).

Unexpectedly, the flattened variation of external rough distinctive features contrasted with the pronounced differences in several fine internal and external morphometric and meristic characters. This contrast had started and conducted a "name game" for *R. miraletus*, namely a taxonomic and evolutionary game whose moves aim to lay upon taxonomic boundaries onto natural evolution in this enigmatic skate. The first "move" of the game was played by McEachran et al. (1989) that observed a significant variation of morphometric and meristic characters over its range and recognized at least three distinct groups of brown skates in the Mediterranean, West and South Africa. The Mediterranean and South African samples resulted the most differentiated, while those from West Africa (i.e. including samples from Mauritania-Senegal, Gulf of Guinea-equatorial Africa and Angola) were similar to each other. Among the West African samples, the Angolan brown skates were the most distinct, even if for some characters displayed intermediate variation between the Mediterranean and South African specimens (McEachran et al. 1989). The subtle variation of characters amongst groups and the cline trend of some of them, lead McEachran et al. (1989) to consider *R. miraletus* a polymorphic species with at least three parapatric or allopatric populations (i.e. located in the Mediterranean, West Africa and South Africa). Two oceanographic fronts with steep thermal gradients in the Eastern Atlantic, namely the upwelling areas at Cape Blanc (21°N) and Cape Frio (18°S) may act as barriers to maintain the partial reproductive separation among these populations (McEachran et al. 1989). The second "move" of the game has been played 20 years later by Naylor et al. (2012) and Caira et al. (2013) who provided a preliminary evidence of cryptic speciation in *R. miraletus* by integrating results from mitochondrial DNA analysis, morphology and host-parasite relationships from specimens collected in Senegal and South Africa. These authors suggested the existence of at least three clusters within *R. miraletus* and the species was updated to a complex of species (Naylor et al. 2012). Two clusters are sympatric in Senegal and the third was located in South Africa. Specimens of each clade host different Diphyllidean flatworm species of the genus *Echinobothrium* (Caira et al. 2013) and, even if additional material should be collected to formally describe these taxa, these independent data strongly corroborate cryptic speciation of *R. miraletus* (Caira et al. 2013; Naylor et al. 2012). A third "move" of this taxonomic and evolutionary game has been recently played by Last and Séret (2016) that, according to deep morphometric and meristic analyses of specimens that were mtDNA-typed (Naylor et al. 2012 and unpublished data), recognized *R. miraletus* as a species complex of at least four valid species: 1) the northernmost *R. miraletus*, occurring in the Mediterranean and adjacent North Eastern Atlantic waters, 2) the southernmost *R. ocellifera* occurring, in the Atlantic, off South Africa and Namibia and, in the Indian

Ocean, from False Bay to Durban, 3) the central African *R. parva* sp. nov., distributed from Senegal to Angola and 4) a still not described species, occurring from Mauritania to Senegal where it is therefore sympatric with *R. parva*. This latter taxon can be identified by divergent mtDNA haplotype at the NADH2 gene (Naylor et al. 2012).

The advent of high-throughput DNA technologies and the launch of global biodiversity assessments (e.g. the DNA barcoding based on mtDNA universal markers as the fragment of the cytochrome oxidase subunit I, COI; Hebert et al. 2003), are providing raw data to go deeply in determining taxonomic, ecological and evolutionary issues of cryptic (*sensu* Bickford et al. 2007) and sibling species, where the term 'sibling' connotes species with a recent common ancestry, implying a sister-species relationship (Knowlton 1986) and even more challenging conservation issues (Bickford et al. 2007). However, the use of molecular methods coupling mitochondrial and nuclear DNA markers better resolved species boundaries as well as gene introgression/hybridization phenomena in marine fish and in closely related elasmobranches (Frodella et al. 2016; Arlyza et al. 2013; Morgan et al. 2012; Pasolini et al. 2011).

This study aims to play again "the name game" by measuring genetic variation at both mtDNA (i.e. the COI barcode sequence marker) and nuDNA markers (i.e. eight EST-linked polymorphic microsatellite loci; El Nagar et al. 2010) in 323 individuals exhibiting the distinctive phenotype of the "*R. miraletus*" species complex (Compagno et al. 1989; Stehmann & Bürkel 1984). Individuals were collected from Mediterranean, North-eastern Atlantic, Senegal, Angola and South Africa, thus likely representing most of the diversity of the "*Raja miraletus* species complex" at multiple taxonomic levels. With these data, we tested the hypothesis that restricted gene flow and genetic divergence within this species complex are associated with climatic/oceanographic discontinuities (e.g. oceanographic fronts, depth barriers and environmentally unsuitable habitats) and parallelise morphological and parasite variation assessed independently (Caira et al. 2013; McEachran et al. 1989). Furthermore, we inferred the time of the most recent common ancestor and reconstructed the relative historical demography of the *R. miraletus* species complex to estimate the timeframe of populations and species' evolutionary dynamics.

2.2. Materials and methods

2.2.1. Sampling

Specimens and tissues were collected from Mediterranean individuals caught during scientific research programs. No specific approval of this vertebrate work is required since the individuals sampled in this study were obtained for scientific and commercial activities. A total number of 323 brown skates were collected from 2000 to 2014 (Table S1 in Appendix I). Most of the individuals were collected by international scientific trawl survey campaigns carried out in South Africa (Africana cruises), Angola (Nansen cruises), the whole Mediterranean Sea (MEDITS scientific surveys) and national scientific trawl surveys (the Italian GruND; Relini 2000; the Portuguese scientific surveys of the *Instituto Português de Investigação Marítima*). Additional samples were provided by contracted commercial fishermen (Senegal, Levantine Sea and Israel) or collected at local fish markets (Algeria). Therefore, sampling covered most of the wide geographical distribution of *Raja miraletus* (Figure 1). All individuals were easily assigned to *R. miraletus* on the basis of their very distinctive morphotype and species-specific diagnostic characters (Stehmann & Burkel 1984; Compagno et al. 1989). Fin clips and muscle tissues were cut from each individual using sterile tweezers and clippers, transferred to a clean tube filled with 96 % ethanol and stored at -20°C for subsequent DNA analyses.

2.2.2. DNA isolation and PCR amplification

Total genomic DNA (gDNA) was extracted from about 20 mg of fin clip or muscle tissue using the Invisorb® Spin Tissue Mini Kit (Stratec®molecular) according to manufacturer's protocol (http://www.stratec.com/en/molecular/Products_Molecular/Genomic_DNA/Invisorb_Spin_Tissue_Mini_Kit/Invisorb_Tissue_Mini.php).

2.2.3. mtDNA COI sequence

From the extracted gDNA, a fragment of the mitochondrial COI gene of about 650bp was amplified using the COI-3 primer cocktail described by Ivanova et al. (2007). The PCR reactions were performed in 50µL total volume containing 4µL of pure gDNA corresponding to ~25ng, 10µL of buffer (1X), 5µL of $MgCl_2$ (2.5mM), 0.5µL of dNTP mix (0.1mM each), 1 µL of each primer (0.2mM) and 0.25U of Taq DNA Polymerase. All PCR reagents were supplied by Promega®, except for primers, supplied by

LifeTechnologies®. The cycle was planned with the following thermal profile: 94°C for 2 min, 35 cycles of 94°C for 30 sec, 52°C for 40 sec, 72°C for 1 min and a final extension step at 72°C for 10 min and it was performed on a Biometra T-Gradient Thermocycler. The PCR products were electrophoresed on agarose gel at 2 % and amplicons were stored at -20ºC until shipping to Macrogen Europe (Amsterdam, the Netherlands). Individual Sanger sequencing was carried out on Applied Biosystems 3730xl DNA Analyser.

2.2.4. Expressed Sequence Tag-linked microsatellite loci

Eight Expressed Sequence Tag-linked microsatellite loci (henceforth EST-SSRs) developed from the little skate *Leucoraja erinacea* and cross-amplifying in several skate species (El Nagar et al. 2010) were optimized for cross-amplification in the *R. miraletus* species complex. The EST-SSR PCR reactions were performed in a 10µL total volume containing 3µL of gDNA corresponding to ~20ng, 2µL of PCR Buffer (1X), 0.5µL of $MgCl_2$ (1.25mM), 0.8µL of dNTP mix (0.05mM each), 0.5µL of each primer (0.5µM; 0.25µM for LERI 26) and 1U of Taq DNA Polymerase. All PCR reagents were supplied by Promega®, except the primers which were from LifeTechnologies®. DNA amplifications were run on a Biometra T-Gradient Thermocycler as follows: after an initial denaturation at 94°C for 3 min, amplification was performed with 30 cycles consisting of denaturation at 94°C for 30 sec, annealing at 53°C for 30 sec, extension at 72°C for 30 sec, followed by a final extension at 72°C for 10 min. Amplicons were electrophoresed on 2.5 % agarose gel and stored at -20ºC until shipping to Macrogen Europe (Amsterdam, the Netherlands). Individual genotyping was performed on ABI3100 Genetic Analyser (Applied Biosystems), using labelled forward primers and LIZ HD500 (Applied Biosystems) as internal size standard.

2.2.5. Data analysis

A total of 275 COI sequence electropherograms was manually edited and aligned by CLUSTAL W software (Thompson et al. 1994) and incorporated into MEGA v.6.0 (Kumar, Stecher & Tamura 2015). The correct amino acidic translation was assessed to exclude the presence of stop codons and sequencing errors (Moulton et al. 2010). For each individual, consensus COI sequences were first compared with published sequences from both the NCBI (http://www.ncbi.nlm.nih.gov/genbank/) and the Barcode of Life Data System (Ratnasingham et al. 2007; BOLD at http://www.boldsystems.org) on-line databases through the BLAST algorithm

(http://blast.ncbi.nlm.nih.gov/Blast.cgi) in order to rule out any error due to mishandling of samples on board or during the laboratory activities. Additional homologous COI sequences of *R. miraletus* were retrieved from both on-line databases selecting, when accessible, records from different geographical origins: South Africa, Sicilian Channel, Aegean Sea and Israel (Table S1). The retrieved sequences were aligned with those newly generate and a final dataset of 306 COI homologous sequences was obtained.

The number of polymorphic sites (S), the number of haplotypes (H), the haplotype diversity (Hd), the nucleotide diversity (π; Nei 1987) and their standard deviations were calculated using DNASP v.5.0 (Librado et al. 2009). Genetic distances were assessed using MEGA. The same software package was also used to estimate the best evolutionary substitution model following the corrected Akaike Information Criterion (AICc; Akaike 1981). The haplotype frequencies were estimated using ARLEQUIN v.3.5.2.2. (Excoffier et al. 2010).

A total of 256 chromatograms for each of the eight EST-SSR loci were obtained and manually inspected using GENEMAPPER v.5.0 (Applied Biosystems). The allele calling and binning have been performed with GENEMAPPER v.5.0. ML-NullFreq (Kalinowski et al. 2006) and FreeNA (Chapuis & Estoup, 2007) were used to test for the presence of stuttering, large allele dropout and null allele's artefacts. The multilocus EST-SSR genotypes were analysed using GENETIX v.4.05 (Belkhir et al. 2004) to estimate the observed (H_O) and expected heterozigosity (H_E) and number of alleles (N_A). Jackknifing over loci was performed to assess the single-locus effects on Weir & Cockerham's F-statistics estimators. The deviation from the Hardy–Weinberg equilibrium (HWE) and Linkage Disequilibrium (LD) was investigated using GENEPOP on the web v.4.2 (Rousset 2008). The allelic richness (A_r) and the inbreeding coefficient (Fis) were estimated using FSTAT v.2.9.3.2 (Goudet 2002).

The phylogenetic relationships among individual haplotypes were inferred by parsimony network analysis implemented in the software HAPLOVIEWER (http://www.cibiv.at/~greg/haploviewer) and the *dnapars* program of the PHYLIP package v.3.6 (Felsenstein 2005). The graphical representation of the resulting network has been modified with Adobe Photoshop.

The average genetic distances observed within and between the two identified Central-Southern African and the NE Atlantic-Mediterranean clades of *R. miraletus* were calculated with MEGA using the Tamura-Nei (1993)'s model implemented therein and compared with the range of COI genetic distances estimated among other congeneric species. Homologous COI sequences of *Raja straeleni, Raja microocellata, Raja asterias, Raja brachyura, Raja clavata, Raja montagui, Raja polystigma, Raja*

radula and Raja undulata were retrieved from on-line databases (NCBI and BOLD) and inter-species distances were calculated applying the Tamura-Nei model.

The population connectivity of *R. miraletus* was investigated by estimates of Fst and Φ_{ST} values using ARLEQUIN with 10,000 permutations, $p<0.05$. The Tamura and Nei (1993) substitution model was applied to the mtDNA dataset to estimate Φ_{ST} values. The virtual spatial differentiation and genetic relationships among geographical population samples were assessed through Principal Coordinate Analysis (PCoA) and were conducted on genotypic and haploid genetic distance matrixes. PCoA plots were generated using the packages 'ade4' (Dray & Dufour 2007) and 'ape' (Paradis et al. 2004) in R environment 3.0.2 (R Core Team 2013). The genetic heterogeneity among the geographical samples was also assessed by the hierarchical analysis of molecular variance (AMOVA, Excoffier et al. 1992). Significance was assessed using a null distribution of the test statistic generated by 10,000 random permutations of the individuals in the samples. The significance threshold of the pairwise comparisons ($p<0.05$) was adjusted with the sequential Bonferroni correction for multiple simultaneous comparisons (Rice 1989) implemented in the software SGoF+ v.3.8 (Carvajal-Rodriguez & de Uña-Alvarez 2011). Different grouping of the geographical samples was tested, following both a priori subdivisions according to geographical areas of collection and PCoA groups.

In order to unravel the individual-based genetic clustering, the EST-SSR dataset was analysed using the Bayesian algorithm implemented in STRUCTURE v.2.3.4 (Falush et al. 2007). This analysis was carried out assuming an admixture ancestry model with the geographical origin of samples as prior information (LOCPRIOR models), associated with a correlated allele frequencies model. For each simulation of K (1-20), five independent replicates were run, setting a burn-in of 200,000 iterations and 500,000 iterations for the Markov Chain Monte Carlo (MCMC) simulation. The most likely value for K based on the STRUCTURE output was determined by both by plotting ln Pr (X|K) vs K (Pritchard et al., 2000) and assessing the ΔK statistic (Evanno et al. 2005) based on a rate of change in the log probabilityof the data as implemented in STRUCTURE HARVESTER (Earl & von Holdt 2012). The results were averaged over multiple runs using CLUMMP (Jakobsson and Rosenberg, 2007) and displayed using DISTRUCT (Rosenberg, 2004) to obtain a bar plot representing the membership coefficients for each individual in each geographical population.

The phylogenetic relationships between the Central-Southern African and NE Atlantic-Mediterranean COI lineages was estimated using a Bayesian coalescent approach, implemented in BEAST v.1.8.3 (Drummond & Rambaut 2012). Sequences of *R. undulata*, the closest related species to *R. miraletus*, were used as outgroup. The Bayesian reconstruction was obtained using the Hasegawa, Kishino and

Yano (HKY+G) model of evolution (Hasegawa et al. 1985), as the most appropriated model inferred by MEGA software, a strict molecular clock model, the Yule Process as species tree prior, the Piecewise linear and constant root as population size prior. To ensure convergence of the posterior distributions, an MCMC run of 60,000,000 generations sampled every 1,000 generations with the first 25 % of the sampled points removed as burn-in was performed. We analysed the log file using TRACER V.1.6 (Rambaut & Drummond 2007) to calculate the robustness of the posterior distributions for all parameters and recover average divergence time and 95% confidence intervals. The plausible trees obtained with BEAST were summarized using the program TREEANNOTATOR and the resulting phylogenetic relationships among population samples and the posterior probabilities at nodes were visualized with FigTree v.1.4.2 (Rambaut 2014).

2.3. Results

2.3.1. Genetic diversity

The COI dataset counted 306 sequences over 23 geographic population samples, while the EST-SSR dataset was made up of a total of 256 individuals distributed in 20 geographic population samples (Figure 1 and Table S1)

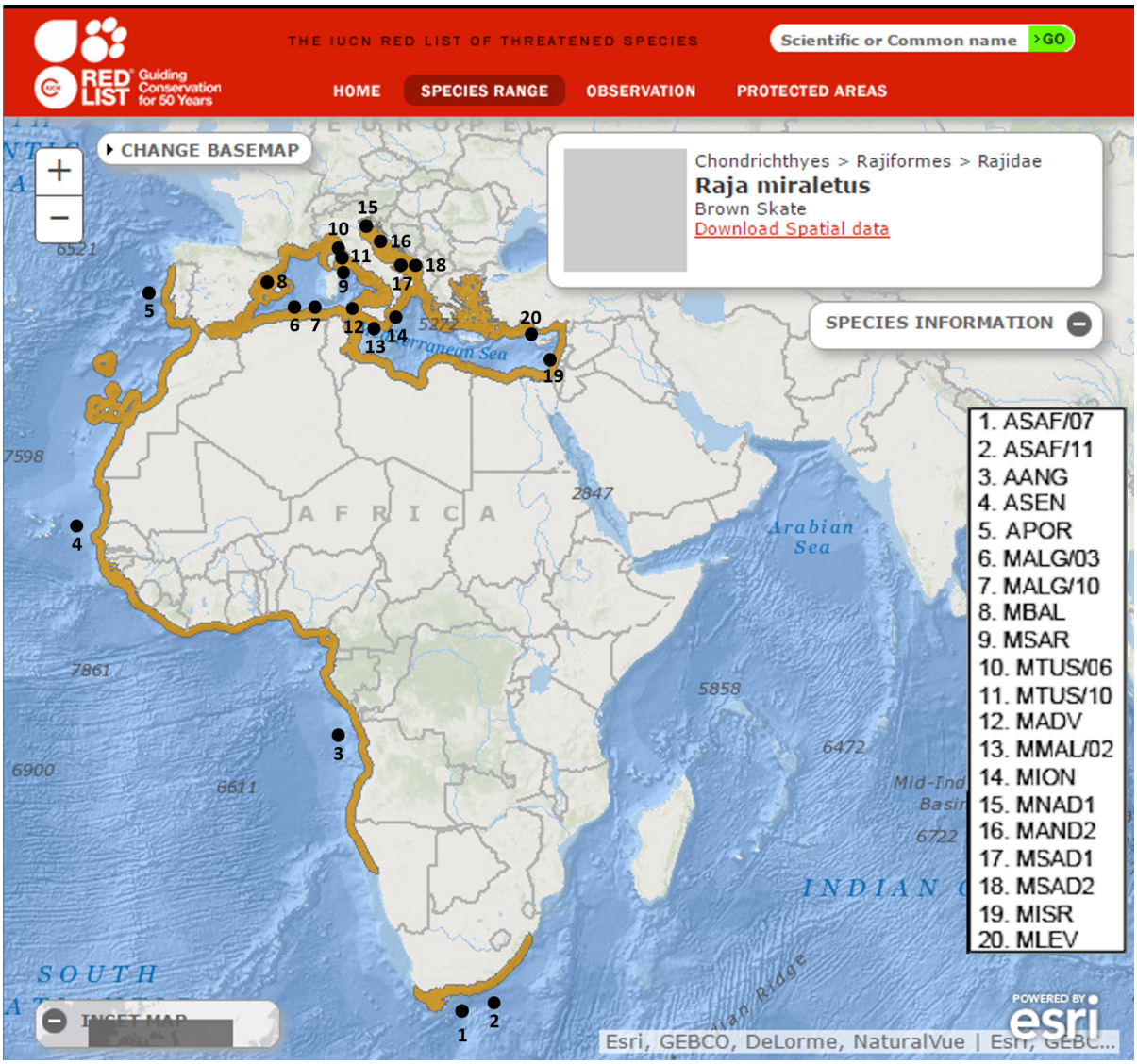


Figure 1 Geographical distribution of the sampling locations for the European (north-eastern Atlantic and Mediterranean) and African *Raja miraletus* (aka *R. ocellifera* from south-eastern Atlantic and western Indian Ocean). Acronyms are used as in Table S1. Sampling locations are overlayered to the IUCN distribution map of *R. miraletus* available at http://maps.iucnredlist.org/map.html?id=161599.

The final COI alignment consisted of 529 nucleotide positions and counted 76 variable sites (14.3%) and 64 parsimony informative sites (12.1%). The mitochondrial gene polymorphism showed low nucleotide diversity (π) and very high haplotype diversity (Hd). AANG sample was the most polymorphic (Hd = 0.858 ± 0.041 SD, π= 0.02543 ± 0.00380 SD, K = 13.453; Table S5 in Appendix I).

The average Tamura-Nei genetic distances (D_{TN}) among geographical samples of the NE Atlantic-Mediterranean were extremely low (D_{TN} =0.0025 ± 0.0011 SE; Table S4 in Appendix I) while those observed among geographical samples of the Central-Southern Africa were an order of magnitude higher (mean TN = 0.0188 ± 0.0031 SE). The D_{TN} between NE Atlantic-Mediterranean and Central-Southern Africa samples were much higher (mean D_{TN} = 0.0734 ± 0.0115 SE) and slightly greater than those estimated from the pairwise comparison with the outgroup *R. undulata* (D_{TN} ranged from 0.0364 to 0.0697; Table S4 in Appendix I).

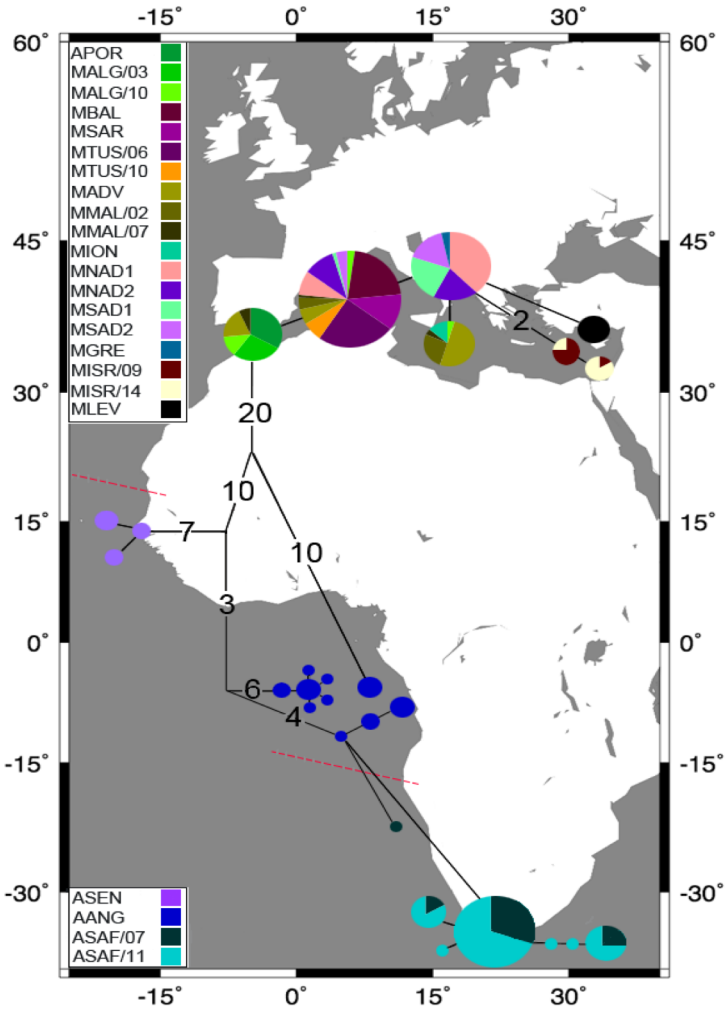


Figure 2 Parsimony network of the COI haplotypes of the European and African *R. miraletus*. The confidence interval was at 95%. The size of the circles is proportional to the number of individuals that shared that haplotype. The haplotypes are indicated by codes as given in Table S1 in Appendix I. Dashed red lines correspond to the principal oceanographic fronts of Cape Blanc (21°N) and Cape Frio (18°S). For graphical reasons only the most frequent COI haplotypes found in the Mediterranean Sea were reported.

Thirty-eight haplotypes were found and none of which was shared between samples from the North-Eastern Atlantic-Mediterranean and Central-Southern African (Fig. 2 and Table S5 Appendix I). The parsimony network of the COI haplotypes (Fig. 2) identified two main haplogroups, which are differentiated by at least 30 mutations and correspond to the Central-Southern African and the NE Atlantic-Mediterranean samples. The former haplogroup included 22 haplotypes that grouped into four largely differentiated geographic clusters occurring in Senegal, Angola and Angola/South Africa. The Senegalese cluster, which coincided with the ASEN sample (N =5), showed three slightly differentiated private haplotypes. On the contrary, the Angolan sample (AANG, N = 28) showed strongly differentiated haplotypes grouped in two endemic Angolan subclusters together and in a third cluster shared with the South African samples (ASAF, total N = 40).

The NE Atlantic-Mediterranean haplogroup included 16 weakly divergent haplotypes (Fig. 2 and Table S5 Appendix I). Four of them were shared by several samples and areas: i) the haplotype Hap_23 was shared by Portuguese, Algerian and Sicilian Channel samples; ii) the most frequent Hap_24 was shared by 12 samples from Algeria, Balearic Islands, Sardinia, Sicilian Channel, Tuscany and Adriatic Sea; iii) the H_26 was shared by samples from Algeria, Sicilian Channel and Ionian Sea; iv) the Hap_31 was shared by Adriatic and Greek samples. In contrast, three endemic haplotypes characterized the Eastern Mediterranean samples of the Israeli coasts and Levantine Sea.

The summary statistics of the eight polymorphic microsatellite loci per geographical sample and over all the loci considered is shown in the Table S6 in the Appendix I. The number of alleles (N_A) ranged from six (LERI 40) to 15 (LERI 27). After Bonferroni correction, no significant LD was detected between any pairs of locus and the average mean observed and expected heterozygosity (H_O/H_E) at eight loci was 0.2595/0.3920. After applying the Bonferroni correction, significant HWE departures were found over all loci in all samples, apart from ASEN, APOR, MALG/03, MMAL/02, MION and MLEV. The Portuguese sample was monomorphic at five loci (LERI 26, LERI 34, LERI 63, LERI 40 and LERI 44). Ml-Nullfreq and FreeNA results detected the presence of null alleles at loci LERI 40, LERI 50 and LERI 44. Nevertheless, we did not exclude any of them, since Jackknife analysis didn't reveal outliers over the confidence interval.

2.3.2. Population connectivity

Because the small sample size affecting some geographical areas and the subsequent decrease in power of the analyses, an appropriate caution should be applied while interpreting the results obtained here. However, similar experimental designs and analytical approaches proved geographical population structure and genetic differentiation at multiple taxonomic levels in other skate species (Pasolini et al. 2011; Plank et al. 2010; Chevolot et al. 2006a).

The PCoA (Principal Coordinate Analysis) based on the haploid COI genetic distance (Figure 3a) was performed to assess the virtual spatial differentiation among geographical populations. The PCo1 strongly separated the Central-Southern African samples from those of the NE Atlantic-Mediterranean Sea. Within the former group, PCo1 also differentiated the AANG from ASAF while the PCo2 markedly separated the Senegalese sample (ASEN). The ASAF temporal replicates were in

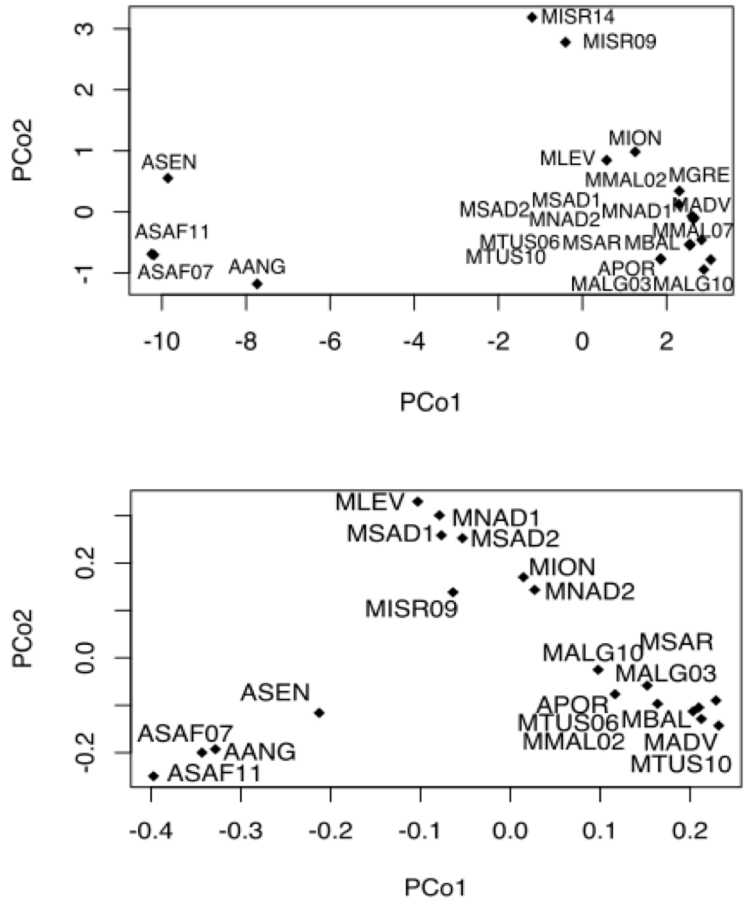


Figure 3a Left side - Plot of the PCoA carried out on COI haplotype genetic distance matrix over all geographical and temporal samples of the *Raja miraletus* species complex. **Figure 3b** Right side - Plot of the the PCoA carried out on genotypic EST-SSR genetic Rey's distance matrix over all geographical samples of the *Raja miraletus* species complex.

contrast not differentiated. Within NE Atlantic-Mediterranean Sea, the PCo2 separated geographical samples according to a longitudinal gradient with the Israeli temporal samples (MISR/09 and MISR/14) as the most distinct samples followed by the Levantine and Ionian Sea samples as intermediately differentiated. Similarly, the PCoA based on the genetic distance matrix computed on the EST-SSR data (Figure 3b) mainly separated Central-Southern African and NE Atlantic-Mediterranean samples. A smaller level of differentiation than that revealed by COI data was observed between ASAF and AANG samples. In contrast, microsatellite data confirmed the genetic distinction of the Senegalese sample. The EST-SSR genetic variation also confirmed the longitudinal pattern of differentiation of the NE Atlantic-Mediterranean samples with a marked separation of Portuguese and Central Western Mediterranean samples (APOR, MALG/03, MALG/10, MBAL, MSAR, MTUS/06, MTUS/10, MADV and MMAL/02) from those of the Central-Eastern and Eastern Mediterranean (MION, MNAD1, MNAD2, MSAD1, MSAD2 and MLEV and MIRS/09).

The genetic differentiation among 23 samples based on COI data showed highly significant pairwise Φst values, even after the Bonferroni correction was applied (Table S7 in Appendix I). High levels of differentiation were observed between the Atlantic African and NE Atlantic-Mediterranean samples, but also between Western and Eastern Mediterranean. Accordingly, the EST-SSR data showed a similar pattern of genetic differentiation even after the Bonferroni correction was applied (Table S8

in Appendix I). No significant differentiation was detected between temporal replicates from the same geographical areas, either with mitochondrial and nuclear data.

The hierarchical AMOVA performed on the COI dataset and testing five sample groupings, confirmed that 89.81% of the genetic variation was significantly attributable to differences between the two main groups (AMOVA 1, Table S3). In contrast, such proportion of differentiation between two groups was much lower based on EST-SSRs loci (26.07%, P = 0.0006). However, the proportion of genetic variation among populations within groups was high and indicated a significant genetic heterogeneity within groups (23.93%, P = 0). The grouping tested in the AMOVA 5 (seven groups) better explained the total mitochondrial genetic variation among samples with a very low proportion of the genetic variation among populations within groups over that among groups (1%). On the contrary, the lowest proportion of the genetic variation among populations within groups over that among groups with the EST-SSR data was obtained in the AMOVA 3 (16.9%) with five groups separating the three Central Southern African samples, the NE Atlantic and Western Mediterranean samples (including those of the Sicilian Channel) and the Eastern Mediterranean samples (including the Ionian sample).

The analysis of STRUCTURE outputs did not provide a clear indication of the most likely number of clusters, therefore results from K=2 to K=7 were assessed (Figure 4). The barplot of the clustering K=2 revealed the separation of the samples from NE Atlantic-Mediterranean Sea and Central-Southern Africa (Figure 4) with an admixed genetic composition of the Senegalese individuals. The clustering K=3 and K=4 further discriminated between samples from Western and Eastern Mediterranean as well as the Angolan sample from those of the South Africa. The clustering K = 5, corresponding to the best grouping revealed by AMOVAs (Table S3) contributed to differentiate the samples from Sicilian Channel (MADV and MMAL/02) from the other NE Atlantic and Western Mediterranean samples while any relevant separation within the Eastern Mediterranean was appreciated. This AMOVA revealed that AANG displayed an intermediate genetic composition between the South African and Senegalese clusters and that the Senegalese genetic component was also exhibited by two South African individuals. The clustering K=6-7 did not contribute a better resolution of population structure

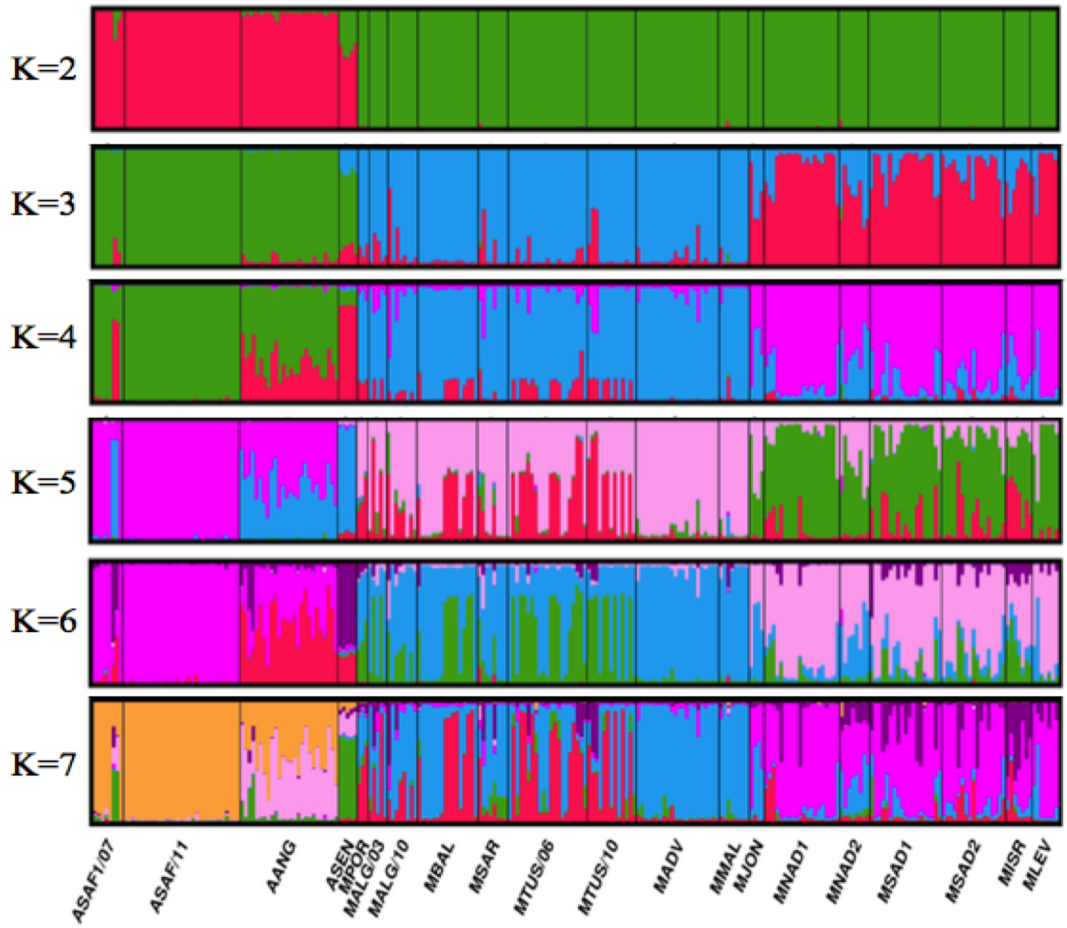
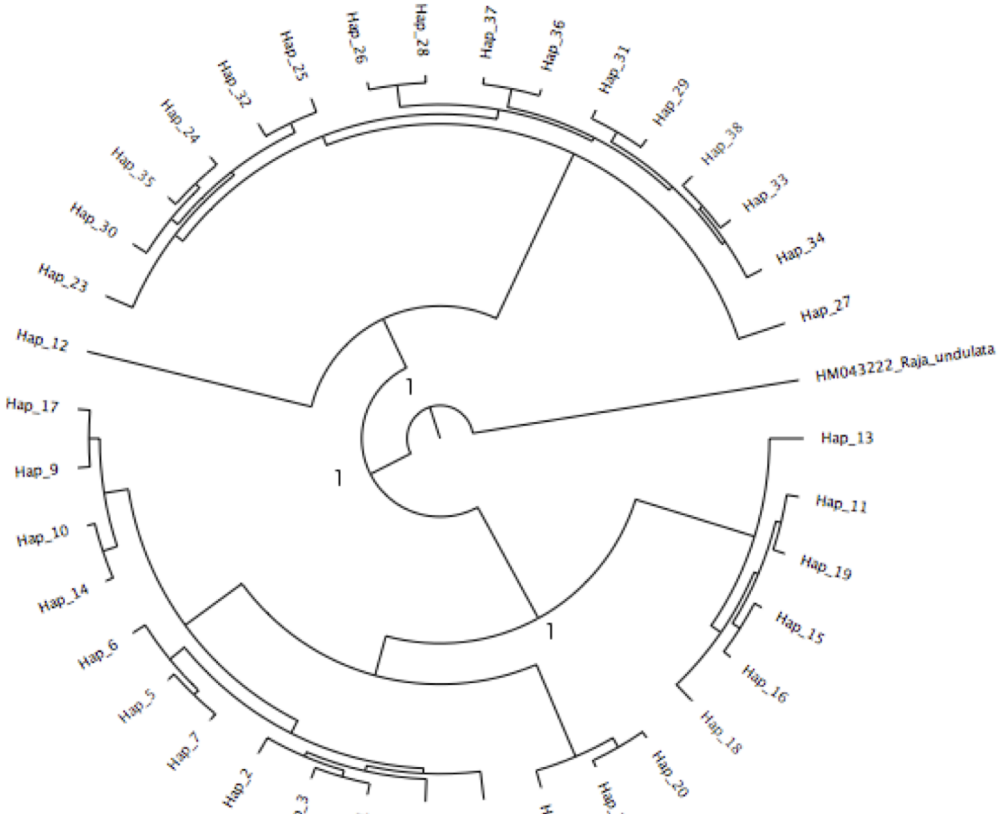


Figure 4 Barplots of the STRUCTURE analysis of *Raja miraletus* species complex based on the complete EST-SSRs dataset and with estimated K=2-7. On the horizontal axis are reported the geographic samples, while on the vertical axis is reported the percentage of individual membership to a given genetic cluster. Single vertical bar represents one individual. Fractions of colour of a bar represent the estimated membership to a certain genetic group of that individual. Vertical black lines separate different sampling locations as set *a priori*.

Bayesian approach using MCMC simulation was used to test any speciation signal (Yule process) between the Central-Southern African and NE-Atlantic-Mediterranean lineages (Figure 5, see Table S5 in Appendix I for haplotype distribution among samples). All effective sample size (ESS) values exceeded 200, indicating a solid evaluation of all parameters. The model based on the substitution rate estimated for mtDNA showed a clear separation between the Central-Southern African (from Hap_1 to Hap_23) and the NE-Atlantic-Mediterranean haplotypes (Hap_24 to Hap_30; Fig 4), with exception of the Hap_12 (an Angolan haplotype) that clustered together with the haplotypes of the NE-Atlantic-Mediterranean lineage. The phylogenetic relationships among lineages and haplotypes were congruent with the relationships obtained with the parsimony network results (Figure 2).

Furthermore, within the main Central-Southern African lineage, three clusters of haplotypes were reconstructed with high posterior probability (p=1): the most basal cluster formed by six Angolan haplotypes (Hap_11, Hap_13, Hap_15, Hap_16, Hap_18 and Hap_19), a Senegalese cluster (Hap_20, Hap_21 and Hap_22), and a third Angolan/South African cluster formed by all the South African haplotypes (Hap_1-9) and the Angolan Hap_14 and Hap_17.

Figure 5 Phylogenetic relationships of *Raja miraletus* species complex based on the COI haplotypes. Refer to Tab S3 for the distribution of the COI haplotypes among *R. miraletus* geographical samples. *Raja undulata* was used as an outgroup for the analysis. Numbers near nodes represent the posterior probability value.



2.4. Discussion and conclusions

The evolutionary and phylogenetic history of skates has been recently described with the increased popularity of DNA-based molecular techniques. Many authors have demonstrated the correlations between Earth's paleo-climatic events or the occurrence of oceanographic barriers and the great levels of diversification of skates at multiple taxonomic levels (Naylor et al. 2012; Pasolini et al. 2011; Iglesias et al. 2010; Compagno & Ebert, 2007; Chevolot et al. 2006; Valsecchi et al. 2005). In parallel, the ability of integrated methodological approaches combining morphological and molecular data in detecting taxonomic and geographic species boundaries of skates have amplified the number of studies and the amount of knowledge supporting species' risk assessment and conservation action plans (Dulvy & Reynolds 2009; Iglesias et al. 2010; Griffiths et al. 2010; Cannas et al. 2010; Frodella et al. 2016; Cariani et al. 2017). The combination between massive sampling and DNA barcoding of skates carried out prevalently in the Atlantic and in the Mediterranean over broad geographical ranges have revealed the occurrence of several cryptic species and species complexes behind the high level of morpho-anatomical and ecological stasis (Iglesias et al. 2010; Naylor et al. 2012; Frodella et al. 2016; Cariani et al. 2017). The recent achievement that *Raja miraletus*, behind a distinctive and quite constant pattern of the dorsal body side, has a status of species complex (i.e. three valid species *R. miraletus*, *R. ocellifera*, *R. parva* sp. nov. together with a not yet described taxon only identified by the mtDNA haplotype at the NADH2 gene; Naylor et al. 2012; Last & Séret 2016), skyrocketed the interest for an evolutionary appraisal based on a deeper and more extensive analysis using massive sampling and nuclear/mitochondrial combined genetic data.

This study, based on an unprecedented number of individual specimens of the *Raja miraletus* species complex collected from the areas where the four taxa occurred and on the coupling of the sequence variation of the universal COI barcode and the allele frequency of eight polymorphic EST- linked microsatellite loci, provided advances in resolving the "name game" in this intriguing and enigmatic skate.

Mitochondrial and microsatellite data I obtained consistently agreed in genetically defining the taxonomic and geographical boundary of *Raja miraletus* L. 1758 which is distributed in the whole Mediterranean Sea and in the adjacent NE Atlantic Ocean, at least in the Portuguese coastal waters. The great divergence of this taxon shown by a Tamura-Nei genetic distance estimated between the NE Atlantic-Mediterranean clade and the Central-Southern African clade (0.073) greater than the corresponding pairwise interspecific estimates obtained among several congeneric species spoke in

favour of a specific level of differentiation. In the NE Atlantic and Mediterranean area, such species resulted structured in at least two main populations genetically well differentiated: the "western population" inhabiting the NE Atlantic coastal waters and those of the Western Mediterranean from Gibraltar Strait to the Sicilian Channel and the "eastern population", inhabiting the Eastern Mediterranean Sea from the Ionian and Adriatic Sea to the easternmost Israeli coasts. The transition area in the Sicilian Channel that separates Western and Eastern Mediterranean biogeographic sectors (Bianchi 2007; Coll et al. 2010) showed a slight genetic differentiation of the brown skates living in this area within the Western Mediterranean. The genetic similarity of temporal replicates tested only some geographical areas (Algeria, Tuscany, Sicilian Channel and Israel) over a time range corresponding to two generations of *R. miraletus* speaks in favour of an interannual stability of population genetic structure. Even if the evident unbalance in the sampling design of *R. miraletus* between Mediterranean and adjacent NE Atlantic, there was no differentiation between populations of *R. miraletus* from the two areas. This suggested that the Gibraltar strait does not represent an effective geographical barrier to gene flow between Atlantic and Mediterranean populations. On this debate, controversial patterns were obtained in the thornback skate *R. clavata* by Chevolot et al. (2006) and Pasolini et al. (2011). Whilst Chevolot et al. (2006) detected spatial genetic differentiation and restricted gene flow between NE Atlantic and Mediterranean thornback skates by analysing the genetic variation at the cytochrome b and species-specific microsatellite loci, Pasolini et al. (2011) did not detect substantial divergence between population samples of the same species collected from the two areas using more polymorphic loci such as the mitochondrial non-coding region D-loop and genomic markers as the Amplified Fragment Length Polymorphisms (AFLPs). Using the D-loop sequence marker and only two samples of small size (N = 14), Griffiths et al. (2010) also detected a weak but significant divergence and restricted or null gene flow between longnosed skates *Dipturus oxyrhinchus* collected from the NE Atlantic (Norway, Rockall) and Western Mediterranean (Balearic Islands). This comparative data led to suppose that the Strait of Gibraltar cannot always represent a barrier to gene flow for skates and other marine organisms, rather than an accession gate to ancient *refugia* (Patarnello et al. 2007).

The marked population genetic structure of *R. miraletus* detected within the Mediterranean was never detected in other *Raja* species by previous population genetic studies carried out with similar sampling design and genetic markers (i.e. uniparental polymorphic mitochondrial DNA sequences, biparental genomic and nuclear loci). Within the Mediterranean *R. clavata*, Chevolot et al. (2006) have analysed the variation of the mtDNA cytochrome b and species-specific microsatellite loci in

three geographical samples (Corsica, Adriatic Sea and Black Sea) detecting any significant structure. Pasolini et al. (2011) found a weak but detectable genetic divergence only of the thornback skates collected from the Eastern Mediterranean, without any restriction of gene flow between western and eastern Mediterranean samples. Recently, Frodella et al. (2016) using the same markers I used in *R. miraletus*, have detected genetic homogeneity and high level of genetic connectivity among Western and Central Mediterranean population samples of the spotted skate *R. polystigma*, a small-sized species endemic of the basin. Only a weak but detectable divergence was shown by the Adriatic deme because a private fixed COI haplotype characterized by a single nucleotide mutation. Such divergence is quite similar to that observed in the *R. miraletus* collected from the Eastern Mediterranean (Israeli and Levantine coastal waters) which showed fixed private haplotypes slightly differentiated by 1-3 nucleotide mutations from the closest haplotype distributed prevalently in the population samples of the Adriatic and Aegean Seas. Subtle genetic divergence in marine fish populations with shallow evolutionary histories can be better assessed by mtDNA markers than nuclear ones (Hoarau et al. 2004). Haploid maternal inheritance of mtDNA can lead to smaller effective population size (Birky et al. 1989) and thus faster genetic drift. Population structure within the Mediterranean could be related to bathymetry and hydrogeological fronts or discontinuities. The shallow bathymetry characterizing the Southwest part of the basin would likely enhance the transition of brown skate, except for the Ionian area from which the divergence pattern displayed a shift. The area ranging from the easternmost part of Sicily and the adjacent geo-morphological depression of the Calabrian Arc (down to 3,000m of depth) is dominated from cyclonic/anti-cyclonic inversions of water masses. The combination of these environmental features could have driven the differentiation of the Eastern Mediterranean samples.

The “name game” appeared more complicated to be resolved in the Central-Southern African taxa of the R*. miraletus* species complex than in the NE Atlantic and Mediterranean. Mitochondrial and nuclear markers contributed differently to define taxonomic and geographical species boundaries. It has to be considered that molecular taxonomic methods using both nuclear and mitochondrial data have proven useful in assessing relationships between pairs of morphologically similar taxa (Morgan et al. 2012, Arlyza et al. 2013) and that the maternal inheritance of mtDNA in Vertebrates combined with ecological and behavioural processes such as natal homing and phylopatry can lead to profound discordance between genetic patterns of structuring obtained with mitochondrial DNA markers with respect to those obtained with biparentally-inherited nuclear DNA markers on the same individuals (see for an example Pardini et al. 2001). According to the concept of biological species, speciation is

the acquisition of reproductive isolation by natural populations (Mayr 1942, 1963). Thus, nuclear DNA markers are the best candidate tools to assess reproductive isolation and species boundaries in bisexual organisms while mitochondrial DNA markers are more devoted to resolve phylogenetic relationships and evolutionary histories among taxa.

The EST-SSR genetic variation of Central-Southern African taxa of the R*. miraletus* species complex disentangled by both individual-based and allele frequency-based statistical methods of analysis indicated the occurrence in this group of at least three taxa whose genetic differentiation seemed to correspond to the species level. Despite the paucity of individuals analysed (N = 5), the Senegalese genetic cluster (ASEN) displayed a deep nuclear genetic divergence as revealed by STRUCTURE and distance-based analyses (e.g. PCoA, Fst and phylogenetic reconstruction). A great mtDNA and phylogenetic divergence of this cluster fully overlapped with the reproductive isolation of this cluster. Presently, I cannot indicate that this Senegalese genetic cluster of the *R. miraletus* species complex corresponded to *R. parva* sp. nov. reported and described by Last & Séret (2016) because I did not perform deep morphological analyses of the five specimens included in this analysis and no mtDNA sequences have been deposited in the public repositories by Last & Séret (2016). However, it is likely that the Senegalese genetic cluster I targeted and characterized could correspond to the new species reported by Last & Séret (2016) *R. parva* (the African brown skate), which spread in Senegal, Liberia and Angola and likely wider within the West Africa. However, since Last & Séret (2016) reported that a further taxon, identified as *Raja* cf. *miraletus* and characterized by a broader disc, more broadly pointed snout, larger spiracles, and a slightly longer and broader tail, co-occurred in Senegal and Mauritania, additional comparative molecular and morphological investigations are needed to solve the intriguing evolutionary histories of such taxa.

A second taxon differentiated at the species level identified by the microsatellite-based analyses is the southernmost genetic cluster formed by the South African samples collected in the South Coast (Western Indian Ocean). The great number of specimens of the ASAF population samples here analysed (N = 39) and the marked interannual genetic stability tested over a time range corresponding to two generations (2007 and 2011) make robust the evidence that such genetic cluster is assigned to *Raja ocellifera* Regan 1906, a species that has been recently resurrected by Last & Séret (2016). The EST-SSR results of the STRUCTURE analysis and PCoA as well as those of the AMOVAs 3 and 4 carried out splitting the South African population samples from the Angolan individuals clearly indicated the genetic uniqueness of this cluster at the nuclear loci. The nuclear uniqueness and divergence of this cluster at the nuclear loci is fully coherent with the mitochondrial

uniqueness given by the occurrence of a private geographical lineage formed by nine closely-related haplotypes (Hap_1 – Hap_9).

The third taxon identified by the genetic analyses based on EST-SSR data corresponded to the genetic cluster formed by the Angolan individuals that was well defined by both STRUCTURE analysis and PCoA. The Angolan genetic cluster has however an admixed genetic composition intermediate between the Senegalese and South African clusters. Phylogenetic trees and haplotype network clearly revealed that the Angolan sample possessed a highly heterogeneous mitochondrial gene pool (k = 13.453; Table S2) including at least three extremely differentiated COI lineages: the former lineage is represented by the Hap_12 exhibited by four individuals that is more related to the Mediterranean COI lineage than to the other two Angolan lineages; the second lineage (formed by the most frequent Hap_16 and other five satellite weakly-differentiated haplotypes) resulted the most ancient being at the basal position in the phylogenetic tree; the third lineage (formed by the most frequent Hap_10 and other two haplotypes) was the more recent and showed great sequence similarity with the South African lineage. Therefore, the Angolan taxon exhibited contradictorily genetic homogeneity at the nuclear loci together with extraordinary mitochondrial heterogeneity. Such pattern could have been arisen by repeated secondary contacts in this area and gene introgression events that occurred in the past.

Whether these three taxa can be elevated to the rank of sibling species or maintained at the status conspecific polymorphic populations as suggested at the beginning of the name game by McEachran et al. (1989) is a debate that needs to be solved by integrating at the individual level the mitochondrial DNA data I obtained with those obtained by Naylor et al. (2012) and Last & Séret (2016). At any rate, the amount of genetic divergence among them expressed as COI genetic distance within the Central-Southern African cluster (D_{TN} = 0.0188; Table S4 in the Appendix I) fell down into the lower end of the range of the pairwise interspecific genetic distances among congeneric species (Table S4 in the Appendix I; *R. clavata* vs *R. straeleni*: 0.015 – *R. asterias* vs *R. microocellata*: 0.088). It has to be noticed that the lower end of this range was occupied by interspecific genetic distance of allopatric and parapatric sibling species (i.e. cryptic sister species; sensu Bickford et al. 2007) as *R. clavata* and *R. straeleni* (D_{TN} = 0.015) and *R. polystigma* and *R. montagui* (D_{TN} = 0.023). It is likely that the low level of mtDNA sequence divergence of sibling skate species is linked to the slowed down mtDNA substitution rate in the cartilaginous fishes (Martin and Palumbi 1993).

The playing of “the name game” led to a systematic scenario in which the Central and Southern African species *R. parva*, *R. ocellifera* and the *Raja* cf *miraletus* taxa from Senegal and Mauritania (Last

& Séret 2016) and from the Angola (this study) together with the more differentiated NE Atlantic-Mediterranean *R. miraletus* can be considered a specious complex of parapatric/sympatric sibling species (i.e. the concept of super-species namely a monophyletic group of allopatric or nearly allopatric taxa that are known or believed to have evolved to the species level; Rensch 1929; Mayr 1931). According to Amadon (1966) this super-species can be named in the Linnean nomenclature as *Raja* [*miraletus*] and it includes the following species *R.* [*miraletus*] *miraletus*, *R.* [*miraletus*] *parva*, *R.* [*miraletus*] *ocellifera* and the *Raja* [*miraletus*] cf *miraletus* taxa from Senegal and Mauritania and Angola.

The evolutionary history and cladogenetic events of this super-species need to be still completely disentangled and unravelled. Within an evolutionary trajectory, they likely represent hierarchical stages of increasing complexity of the geographical speciation process. In this species complex, stages of this hierarchical series ranged from panmictic populations to recently diverged sibling species, with morphologically identical populations spread over a continuous range (Mayr 1954; Palumbi and Lessios 2005). Within the evolutionary trajectories experienced by the skates of the genus *Raja*, geographical sibling species could be intermediate frames in the micro-evolutionary animation that was proposed by Ernst Mayr to model geographical speciation (Mayr 1954). Such a continuum in the process of geographical speciation begins with polytypic species that inhabit a continuous range and ends with super-species, which correspond to a group of geographical species that show complete reproductive isolation with respect to gross morphological traits. The abundance of cryptic species in the elasmobranches, marine fish that intrinsically exhibit biological and reproductive traits that enhance species vulnerability and risk of extinction (Dulvy et al. 2003), poses relevant questions for the conservation of genetic and species diversity (Bickford et al. 2007; Dulvy et al. 2000; Dulvy and Reynolds 2009; Iglésias et al. 2010).

The Quaternary and present oceanographic discontinuities that occur along the western African continental shelf (e.g., Cape Blanc and the Angola–Benguela Front; Gasse et al. 2008) might contribute to the maintenance of low or null levels of gene flow between these closely related siblings. The African Atlantic marine faunas can be subdivided into three zoogeographic provinces by two main oceanographic discontinuities, which correspond to steep thermal gradients in the upwelling areas of Cape Blanc at 21°N and of Cape Frio at 18°S (Briggs 1974). The oceanographic discontinuity at Cape Blanc (Mauritania) acts as a physical barrier against the southwards dispersal of the endemic *R.* [*miraletus*] *miraletus* from the Mediterranean and northeastern Atlantic shelves. Here the inter-tropical Canary/Angola currents inflowing from northeast could have thus influenced

also the diversification of the Senegalese *R.* [*miraletus*] *parva* and likely of the *Raja* [*miraletus*] cf *miraletus* (sensu Last & Séret 2016) whose migration northward would be opposed by the intermittent Cape Blanc upwelling area. The upwelling at Cape Frio and the Angola–Benguela front are the southern limit of the tropical skate faunas (Hulley 1972). In particular, the Angola–Benguela Current region ranges from Cape Agulhas to Cape Frio, where the upstreaming flow meets the down-streaming Angola current. The deriving front would have likely reduced with the gene flow of *R.* [*miraletus*] *ocellifera* and *R.* [*miraletus*] *cf miraletus* (sensu this study).

The accumulated genetic differences among the *Raja* [*miraletus*] species, and their parallel morphological stasis might be the evolutionary outputs of stabilizing selection that has tended to conserve a well-adapted phenotype across the wide-ranging distribution of the clade (Williamson 1987), in relation to the stasis of marine communities on evolutionary time scales (Colborn et al. 2001; Jackson and Sheldon 1994). From Chapter 3 the methods aiming to identify the candidate genes regulating peculiar phenotypic traits will be described and discussed, setting the basis for addressing the investigations of stabilizing selection and phenotypic stasis in *R.* [*miraletus*] *miraletus* and *R.* [*miraletus*] *ocellifera* and in other related skate species.

Chapter 3. Retrieving skates' skin transcripts by Illumina sequencing

3.1. Introduction

Skates' species diversity can be considered as striking as their tendency to maintain analogous ecological and morphological traits at the evolutionary scale. As well as other batoids, skates abandoned the body conformation typical of ancestral Neoselachians for a depressed, rounded pectoral disk supported to the snout tip by fin radials, a short tail, reduced caudal and dorsal fins (Aschliman et al. 2012) and a pigmentation associable to aposematism, signalling, mimicry and, most likely, camouflage strategies. Probably, the minor use of visual signals in favour of an efficient electrosensory system for conspecific recognition (Tricas et al. 1995) and prey detection (Camperi et al. 2007) could explain the lack of exaggerated phenotypes, even if some exceptions are observable in nature (e.g. the blue-spotted ribbontail ray, *Taeniura lymma* Forsskål, 1175).

The extremely pronounced conservatisms of benthic habits and the preference for soft bottoms characterizing this Order drove to develop similar synapomorfies (e.g. oviparous development, alar/malar thorns in mature males, cartilage modification and specialization; Mceachran et al. 1998) and to follow multiple parallel adaptive radiations at a regional scale (Valsecchi et al. 2005), responsible for strong phylogeographical and population structuring (Frodella et al. 2016; Pasolini et al. 2011; Griffiths et al. 2010; Chevolot et al. 2007; Chevolot et al. 2006).

Aside from the mere records of albinism, leucism (i.e. abnormal integumentary pigmentation with normal iris pigmentation) and piebaldism (i.e. localized patches of reduced pigmentation interspersed with the normal patterning; Acevedo et al. 2009; Bechtel 1995; Ebert 1985) observed among different elasmobranch orders (Ball et al. 2016; Bigman et al. 2015; Wakida-Kusunoki 2015; Diatta et al. 2013; Hoare et al. 2009; Bottaro et al. 2008; Sandoval-Castillo et al. 2006; Clark 2002; Ishihara et al. 2001), no studies have been conducted on the genetic base of elasmobranch pigmentation, which can be quite variable according to latitude, age, habitat and feeding behaviour (Ari 2014; Visconti et al. 1999). Among skates the adaptive value of some phenotypic traits as, for instance, peculiar ornaments of the dorsal part of the body, remains poorly understood. Therefore, the relationship existing between skin pictorial motifs (e.g. horizontal stripes, dots, pseudo-eyespots, eyespots, undulate bands) and their evolutionary or adaptive purpose remains enigmatic, although their involvement in camouflage and aposematism systems seemed to be the most probable.

Understanding how these traits evolved represented the biggest questions here. Coupling genomics

and evolutionary biology would likely be a significant success in understanding how morphological variation, molecular patterning and genetic factors could have influenced the evolution of different colour-morphs. Here, we investigated for the first time some of the pigmentary features specifically found among non-model skate species, representing an interesting ecological model system.

Taking advantage of improvements in genome research such as next generation RNA sequencing technology (RNA-seq) we attempted the identification of novel transcripts of pigmented and non-pigmented skin tissues of recently diverged species with sibling and sister species phylogenetic relationships (*Raja clavata* L.*, R. straeleni* Poll, 1951 and *R. asterias* Delaroche, 1809 respectively) and in species that are easily recognized by specific patterning traits and show strong monophyly (the Mediterranean *R. miraletus* L. and the South African *R. ocellifera* Regan, 1906).

While model species are considered precious proxies for a wide fan of known pathways and traits, non-model ones are more flexible in letting explore the molecular basis of specific traits in greater detail (Parsons & Albertson 2013). For their intrinsic nature, this sort of organisms brings bigger limitations in terms of genomic resources, making the challenge even harder. Fortunately, the advancing technology of Next Generation sequencing (henceforth NGS), the progressive lowering of costs of some of the currently available platforms and the more and more robust computational tools highly simplify and encourage their employment (Grabherr et al. 2011; Wheat 2010; Bräutigam et al. 2008).

NGS technologies generally differ in the average length and number of reads generated (Holt & Jones 2008). Short-read technologies (30 to 125bp reads) usually produce more sequence data per run (Million reads) than long-read technologies (200 to 500bp reads; Mardis 2008b). At the same time, the formers are less successful in assembling large, complex genomes or transcriptomes (Whiteford et al. 2005). Thus, long-reads obviate this limitation, whereas short-read technologies remain suitable for experimental projects where a reference genome or transcriptome is already available for their mapping (Schuster 2008; Whiteford et al. 2005).

As in this case, many evolutionary questions need experimental designs based on highly variable genetic markers. Among these, coding sequences can be easily detected in mRNA transcripts (retro-transcribed in cDNA), usually not as complex to assemble as genomic sequences. Furthermore, transcripts are non-uniformly represented in an organism: transcripts of highly expressed genes are expected to be highly abundant, whereas those of lower expressed genes are less represented. For this reason, RNA-seq cannot provide uniform coverage.

In general, comparative transcriptomics based on quantitative data favour the exploration of interesting changes in gene expression and transcript complexity in a way that even families of genes underpinning phenotypes can be identified. For this reason, the Illumina sequencing technology was used to produce short reads containing transcripts belonging to three different skin tissues, corresponding to different body areas carrying different patterning. Their mapping to a newly-reconstructed reference transcriptome of *Raja miraletus* will allow to compare skin tissues transcripts and measure their Differential Gene Expression (DGE) across differently pigmented and non-pigmented tissues, among five different species.

3.1.1. Skin structure in elasmobranchs

Linking teleost and cartilaginous fish skin organization is certainly a natural and effective basis for the comparison of structures and mechanisms. Nonetheless, many differences can be highlighted and many similarities to other less obvious taxa can be found. Besides the fundamental role of skin in stabilizing and protecting the body from external stimuli, environmental threats, mechanical abrasion and pathogen attacks (Tsutsui et al. 2009), many aspects of this natural barrier are still poorly known in rays and skates. For instance, the meaning of different number of secretory cells per unit area in relation to the cell layer, the lack of sex-related epidermal thickness variations (Hay et al. 1976) or, the meaning of the absence of mesenchymal tissues and dermal endothelium (i.e. a single cell sheet between the dermis and hypodermis; Whitear 1986b) are pending questions.

In certain measure, chondrichthyan skin appears simpler, although effective in many aspects, in comparison to other Class where keratin gene products are much more numerous (Froschauer et al. 2006).

Despite slight differences in some components clearly detectable between Selachii and Batoidea subdivisions, three are the main layers characterizing elasmobranchs' skin. Hypodermis is the deepest skin layer, located between the body musculature and the dermal *stratum compactum*. It is recognizable by varying sizes of blood and lymphatic vessels and fat cells that, depending on the food abundance, are variably interspersed with loose connective tissue.

Dermis' strong mechanical power relies on high amounts of collagenous fibres and fibre bundles (fibroblasts and fibrocytes) characterizing its *stratum compactum* and contributing, for example, to a faster swimming capacity (Meyer & Seegers 2012). Other free cells as melanocytes, macrophages and mast-cells contribute to pigmentation, anti-inflammation and immune responses. Among

batoids an inconspicuous melanocyte layer constitutes the dermis, while cells of haemolymphatic capillaries show short protrusions of the internal cell membrane towards the lumen, filled with mitochondria. These structures seem to be related to the maintenance of low Na+ and Cl- concentration in body fluids in relation to the surrounding tissues, to regulate the intensity of osmotic exchange and save precious energy when necessary (Meyer & Seegers 2012).

The epidermis is composed in turn of three more layers. The monolayer *stratum basale* is in perpetual mitotically active status. Cells normally have a large nucleus with one to two nucleoli and a homogeneous cytoplasm containing mitochondria, a very dense network of cytofilaments, a rough endoplasmatic reticulum with numerous closely contacting *cisternae*, a low number of free ribosomes and, especially in batoids, several electron-dense vesicles. The *stratum spinosum* is characterized by many highly-connected cells and long protrusions of melanocytes filled with melanin granules and reaching the *stratum superficiale*. The latter is constituted of secretory cells, which are typically fewer in skates. Among sharks, many columnar-like secretory cells predominate the epidermal layer, while these cells are fewer and with an irregular distribution among skates and rays. Epidermis is also crossed by free nerve fibres and is characterised by many mucus cells, appearing as large balls (Meyer & Seegers 2012), lymphocytes, macrophages and several types of granular leucocytes.

The most important feature distinguishing elasmobranchs (in sharks and in most skates and rays) is the presence of ectomesodermal placoid scales enriching the skin. These structures are firmly anchored in the dermal *stratum compactum* by strong collagenous fibres bundles and they are not imbricated. Scales body coverage highly depends on the swimming strategy: shark-swimming elasmobranchs (sharks, guitarfish and sawfish) are characterised by scales covering the entire body, while among other batoid taxa, the undulatory/oscillatory movement contributed to maintain a sparse scales coverage, mainly concentrated on the tail. In some species, these features are completely absent (Myliobatiformes), but a thick mucus layer seems to compensate this lack (Marshall 1978).

Some secondary structures likely derived from enlarged placoid scale are generally called ‘thorns’. They are very common in skate species and they typically show a secondarily thickened, acellular bony base and a narrow, hook-shaped and curved crown (Reif 1979). Nuchal and scapular thorns are present in both males and females as well as some median row spines distributed along the disc. Males’ alar thorns were deeply investigated by McEachran (1984; 1977). Because of their shape, position and orientation they are involved into courtship and mating, while anchoring to the female

pelvic fins. For this reason, alar spines are considered secondary sexual characters and maturity indicators (Bigelow & Schroeder 1953).

Similarly, malar thorns are seated lateral to the orbits and spiracles in maturing male skates and their orientation on the dorsal disk is oblique. Some authors attempted to use skates' alar and malar thorns orientation and squamation to unravel the phylogenetic and population structure between two main clades, although no evidence of differentiation was found (McEachran & Konstantinou 1996). On the contrary, these and other 21 meristic characters contributed to delineate at least three different clades of brown skate *Raja miraletus* inhabiting the Southern and Atlantic African Coasts and the Mediterranean Sea (McEachran et al. 1989).

Within this tissue hides an extraordinary sophisticated sensory system constituted by pit organs, the lateral line organ and the ampullae of Lorenzini (Meyer & Seegers 2012).

Pit organs are sensory units set in correspondence of bigger or modified placoid scales, near the lateral line (Peach & Rouse 2004; Maruska 2001; Tester & Nelson 1967). They are connected to cranial nerves and cover both dorsal and ventral body portions in shark, even if they are highly reduced in skates and rays ventral part because of adaptation of living close to the sea-floor (Peach & Marshall 2009, 2000; Peach & Rouse 2004). These organs are involved in capturing the water motion and velocity, generated by water flow, predators, conspecifics or distortions in the surrounding water while the animal is swimming.

The same function is carried out by the bilaterally symmetrical network of lateral line canals (Maruska & Tricas 2004), constituted by sub-epidermal dorsal pored and ventral non-pored canals. Neuromasts are the sensory units at the basis of this system and are constituted by hair and supportive cells covered by a *cupula* (Kasumyan 2003; Maruska 2001). The vesicles of Savi are specialized units of the lateral line organ peculiar of skates, rays and benthic sharks. They are set in the ventral portion of the body and have a mechanic-tactile functions, which enable the consciousness of the skin surface displacements.

Differently, the ampullae of Lorenzini allow the caption of electric fields. Discovered by Stefano Lorenzini in 1678, these canals are used during navigation and the detection of mates, predators or preys (Wilkens 2005; Tricas et al. 1995). Each ampulla system lies in the hypodermis in form of connected alveoli and consists of each somatic pore leading to a jelly-filled canal. Ampullae cluster together in capsules of connective tissue allowing the suppression of interference from the animal's own electric field (Meyer & Seegers 2012). The interior portion of the alveoli is also covered by

sensory epithelium cells developing from the neural crest. Neurons are connected to this system as afferents and are in contact with multiple receptor cells.

As it has been briefly described here, the complexity of skin structure is fundamental to its functions. At the same time, this condition has been an obstacle to the laboratory praxis on one hand and it has been recalled within the data analyses aiming to rule out the genetic characterization of skin features on the other.

3.1.2. Skin ornaments and pigmentation structures

Differently from endotherms, teleost and cartilaginous fish display several classes of pigment cells, or chromatophores, that retain their pigments intracellularly (Bagnara & Matsumoto, 2006; Clark, 2002). In these groups, body patterning relies on the distribution, density and aggregation of these different integumentary cell-units. Among fish in particular, the colours we perceive are the result of both light absorption and reflection capability of pigments contained in different kind of these cells: melanophores contain black eumelanin pigment, erythrophores and xanthophores are filled with yellow-red carotenoid or pteridine pigments and cyanophores contain blue pigment. The typical metallic iridescence depends on light reflection of the iridophores or leucophores, containing purine crystals (Leclercq et al. 2010; Braasch et al. 2010; Goda & Fujii 1995; Taylor & Bagnara 1972). Fish can synthesise eumelanin from tyrosine, while carotenoids need to be provided by the diet (Sefc et al. 2014).

This complex variety of colour structure seems to be less known among cartilaginous fish, where the physiological colour control is more similar to amphibians rather than teleost, since no nervous control is involved (Hogben 1936; Wykes 1936). Hogben (1936) identified in *Mustelus canis* (Mitchhill, 1835) three kinds of chromatophores, one seated in the epidermis, one, larger, in the dermis and xantophores. It has been demonstrated that the pituitary gland secretions (e.g. alpha-MSH, Melanophores Stimulating Hormone or 'darkening hormone'; Gelsleichter 2012; Kemp 1999; Visconti & Castrucci 1993) are responsible of the regulation of body darkening and paling in *Leucoraja erinacea* and other cartilaginous fish after an exposure of nine, twelve up to twenty-four hours to different light conditions (Claes & Mallefet 2010; Parker 1993; Bagnara & Hadley 1973). Furthermore, only melanocytes where previously described among class Chondrichthyes (Meyer & Seegers 2012; Kemp 1999).

In many skate and ray species, the dorsal body colour is uniformly black (e.g. genus *Mobula*) or brown and often enriched with spots, dots, eyespots and pseudo-eyespots, undulated bars and blotches. Thus, skin ornaments or pictorial motifs are variable in shape, colour and position, depending not only on the geography (geographic colour pattern differentiation is usually accompanied by genetic differentiation, even if not all species with distinct population structure show phenotypic variation e.g. Koblmüller et al. 2011; Duftner et al. 2006), but also age (Ari, 2014) and behaviour. As well as for teleost and other vertebrates, patterning in skates depends on pigments and structural colours (Sefc et al. 2014) and, differently from other species, no sexual dimorphism is reflected by pigmentary traits.

Among the skate species considered in this dissertation, a considerable variety of dorsal ornaments and colours are observable. *Raja clavata* dorsal side hues from grey to dark brown with either a dark, marbled appearance or with numerous dark spots and brownish blotches (i.e. larger and irregular pigmented patches with ill-defined boundaries; Santos et al. 2016). Specimens may also display a larger black eyespot at each pectoral fin base. Dark and light bands alternate on the tail, while the ventral skin is white (Serena & Mancusi 2010).

The sibling *R. straeleni* is brown to grey dorsally, with small to large black spots, whorls and blotches, and a larger eyespot at each pectoral fin base. The underside is plain white. Juveniles lack the dark dorsal spots (Froese et al. 2014).

The sister *R. asterias* is brown-yellowish with lighter, yellowish small blotches asymmetrically distributed and numerous dark dots. Ventral side is also white (Serena & Mancusi 2010).

A very different patterning is showed by *R. miraletus* and the South African *R. ocellifera*, that are characterized by a sibling relationship and a well-recognizable and strikingly similar pigmentation: on the upper ochre-brownish surface scattered with dark spots (Stehmann & Bürkel 1984; Compagno et al. 1989), two well-defined bright blue-purple eyespots stand clearly out at the base of both pectoral fins. However, whereas in *R. miraletus* specimens from Senegal these spots are surrounded by a black and then a yellow ring, in *R. ocellifera* from South Africa blue spots are encircled by a ring of tight brown spots and then a ring of tight yellow/white spots (Caira et al. 2013). Ocelli in the Mediterranean brown skate are slightly smaller than Angolan or Senegalese specimens (Last & Séret 2016). Furthermore, the presence of these features is not associated with body development or reproductive stages. Ventral side is white (Serena & Mancusi 2010).

The pictorial motifs described here are the main elements of interest in this part of the dissertation.

3.1.3. Pigmentary genes

Entering the field of skin pigmentation is a challenge. Vertebrate's pigmentation is regulated by multiple genes. Some of them are involved in pigment development, others are components of melanosomes and their precursors, or are responsible for melanosome construction and transport. Among vertebrates more than 370 genes are involved in colorations and 128 of them regulate the pigment synthesis in zebrafish (Braasch et al. 2010), while some of them underpin interesting patterning in teleosts (Santos et al. 2016, Sefc et al. 2014). More in general, pigments synthetized by an organism are mainly regulated by the interaction of two units: the melanocortin-1 receptor (MC1R) and the Agouti. The mechanism is virtually simple: the signalling activity of the transmembrane MC1R at the surface of melanocytes favour the production of brown-black eumelanin. Alternatively, the binding of the antagonist Agouti, causes a switch into the synthesis of yellow-red phaeomelanin. Actually, the genetic basis of this system is more than complex, thus it represents the utmost question for many Evolutionary Biologists.

As it has been demonstrated by Manceau et al. (2010), different pigmentary genes can produce similar patterning among closely related species (Manceau et al. 2010). At the same time, an exact amino acid mutation in the same gene (e.g. Arginine to Cysteine amino acid) can lead to similar phenotypes in highly distant species, like the Alabama beach mice (*Peromyscus polionotus ammobates* Bowen, 1969) and mammoth (Römpler et al. 2006). Again, a mutation in different genes can regulate very different functions, but lead to similar patterning (Arendt & Reznick 2008).

The MC1R generally controls the synthesis of eumelanin through the regulation of the expression of different genes belonging to the melanin pathway (e.g. Proopiomelanocortin or POMC, Prohormone convertase PCSK1/3, PCSK2 and ASIP; Vachtenheim & Borovansky 2010) and generally controls the eumelanin-pheomelanin switch. Mutations at these genes can affect MC1R location, its affinity to bind at other ligands (i.e. alpha-Melanin Stimulating Hormone and ASIP) and the G protein coupling activity (Dessinioti et al. 2011). Alterations of the G protein can subsequently influence the expression and activity of downstream genes involved in the melanin pathway as well. For instance, in the European population of barn owls (*Tyto alba* Scopoli, 1769), MC1R is polymorphic for the amino acid 126, where a valine-to-isoleucine substitution explains about 30% of variation in ventral coloration (Burri et al. 2016; San-Jose et al. 2013).

The Agouti gene is antagonist to MC1R which is known for producing a transcriptional isoform expressed in the ventral skin and associated with dorsal–ventral differences in pigmentation and cryptic adaptation to the environment (Barsh 1996; Bultman et al. 1994; Vrieling et al. 1994).

The totality of pigmentary genes has been categorized according to their function (as. melanophore development, components of melanosomes, melanosome biogenesis and transport, regulation of melanogenesis, systemic effect, xantophore development, pteridine synthesis, iridophore development and unidentified function). The complexity of these functions and structures in bony fish is strongly related to the two rounds of whole genome duplication that involved vertebrate groups, amplified by the migration of more pigment cells from the neural crest. This developmental trait led teleost to display approximately 30% more pigmentary genes than other vertebrates (Braasch et al. 2009, 2007).

According to the recent Santos et al. (2016), 29 genes seem to be involved into the evolution of egg-spots in *Astatotilapia burtoni* (Günter, 1894). Among these genes, the endothelin B receptor (EDNRB), the transcription factors SOX9 and MITF (microphthalmia) and the cell-adhesion molecule binding periostin-like isoform X2 (POSTN) can be involved in egg-spot morphogenesis on male's anal fin. This peculiar trait is likely responsible for the breeding behaviour of the species and are involved in sexual selection.

Starting from a different point of view, totally unbound from sexual behaviour in favour of cryptism or aposematism mechanisms, the recovery of similar proteins and genes is expected among skates' skin samples, after comparing different body areas carrying (or not) eyespots and pseudo-eyespots.

3.2. Materials and methods

3.2.1. The development of a sampling protocol and experimental scheme

The sampling protocol was established starting from standard guidelines for gene expression profiling (i.e. FishPopTrace Project; https://fishpoptrace.jrc.ec.europa.eu). To ensure a sterilized environment during the whole procedure, a careful preparation of materials needed to be carried out before this phase. In particular, 2 ml Nuclease-free microtubes with cap and O-ring were previously labelled with Sample ID (code and replicate 'a' or 'b') and filled with 1.2 ml of RNA*later*® Stabilization Solution (ThermoFisher Scientific). A minimum of 10 volumes of RNAlater® reagent (10µl/1mg of tissue) is recommended by the user guide (https://tools.thermofisher.com/content/sfs/manuals/7020M.pdf).

The working surface and the surgical instruments (scissors, tweezers and scalpel blades) were washed and cleaned with denaturised Ethanol and then with RNase*Zap*® RNase Decontamination Solution (ThermoFisher Scientific). The operator had to wear clean surgical gloves and rub gloved-hands up with RNAse Zap® if necessary. Instruments were cleaned up and sterilized after the sampling of each tissue and clean gloves were replaced after sampling each specimen.

Tissues used in this study were collected from Mediterranean and South African individuals caught during scientific research programs. No specific approval of this vertebrate work is required since the individuals sampled in this study were obtained for scientific and commercial activities.

Withering specimens were delicately washed and quickly scrubbed with clean paper towel to get rid of eventual organic and inorganic materials and sediment. Then, they were accommodated on an aluminium layer sprayed with RNAse Zap® and placed on ice. Biometric measures were taken (total length, disk length, disc width and weight) and sex was determined. Skin slices, less than 0.5 cm thick, were collected in correspondence of the dorsal colour motif, in the adjacent area where skin was uniformly pigmented and the non-pigmented ventral skin. Samples were taken symmetrically from each body side, trimming fat and muscle tissue and transferring them to the corresponding tube. Tubes were first stored at 4°C overnight, then at −20°C until further processing to maximize RNA fixation. Samples were transported in dry ice during their transfer to the laboratory.

Three skin samples differing in pigmentation patterns (the dorsal ornament, the adjacent uniformly pigmented skin and the un-pigmented ventral skin) were collected from skate species caught during scientific and commercial surveys carried out in the Adriatic Sea, Sicilian Channel and Tyrrhenian Sea between August and December 2014 during the SoleMon (FAO-Adriamedn Project) and MEDITS scientific surveys (http://www.sibm.it/SITO%20MEDITS/principaleprogramme.html) or commercial vessels. While samples from around Italy were personally collected, samples of *R. ocellifera* and *R. straeleni* from off South Africa were gathered in May 2015 during the Afrikaan cruise, see (Ebert & Compagno 2007), carried out by the Department of Agriculture and Fisheries of the South African Government of Cape Town (see Fig S1, S2 and Table S9 in Appendix II for details).

The initial intent of boarding in North-Eastern Atlantic and include interesting species in the experimental design (i.e. *Raja montagui* Delaroche, 1809 or *R. undulata* Lacepède, 1802) was not supported because of the lack of a Basic Sea Survival Certificate (http://www.rya.org.uk/wheresmynearest/Pages/CourseDetail.aspx?code=BSSCC).

Based on the available specimens collected during the sampling phase, an experimental design was ideated to perform the Illumina sequencing of differently pigmented tissues and compare group of

sibling species with similar patterning (e.g. *Raja clavata, R. straeleni* vs *R. miraletus, R. ocellifera*) and sister species showing different pictorial ornaments (*R. asterias* vs *R. clavata* and *R. straeleni*). A total of five individuals per species, three skin tissues per individual, were included in the experimental phase (Fig 6).

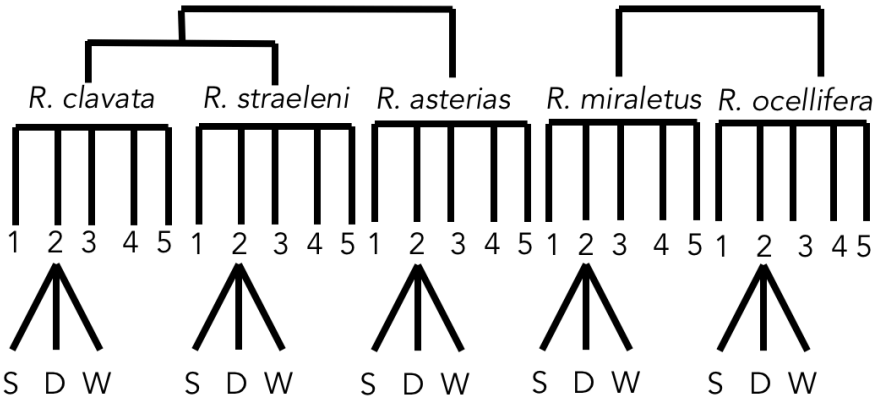


Figure 5 The experimental scheme for Illumina sequencing. S dorsal spot. D dorsal uniform matrix. W ventral white skin.

3.2.2. RNA isolation

3.2.2.1. RNA extraction using TRIzol® Reagent

The first RNA extraction attempts were performed at the GenoDREAM lab of Ravenna (Dept. Biological, Geological & Environmental Sciences, University of Bologna).

Standard TRIzol® extraction protocol (ThermoFisher Scientific) was initially tested on *Raja asterias.* Skin tissues were carefully trimmed from muscle and collagenous tissues and homogenized with a T10 basic ULTRA-TURRAX® (Ika, Germany) following the manufacturer's instructions (https://tools.thermofisher.com/content/sfs/manuals/trizol_reagent.pdf). RNA was suspended in 25µl and its purity measured using a MBA200 spectrophotometer (Perkin Elmer).

3.2.2.2. RNA extraction using the Direct-zol™ RNA MiniPrep Kit

Using *R. asterias* individual replicates, the Direct-zol™ RNA MiniPrep Kit (ZymoResearch, Germany) was also tested (http://www.zymoresearch.com/downloads/dl/file/id/467/r2050i.pdf) after tissue

homogenation with a T10 basic ULTRA-TURRAX® as well. RNA was suspended in 25µl and its purity measured using a MBA200 spectrophotometer.

3.2.2.3. RNA extraction using the Maxwell® 16MDx Instrument

Skin samples of *Leucoraja melitensis* were used for testing the Maxwell® 16MD Instrument (Promega) for nucleic acids extraction. After homogenizing tissues with scissors and mortar and pestle, the extraction phase was prepared as indicated by the manufacturer (https://ita.promega.com/-/media/files/resources/protocols/technical-bulletins/101/maxwell-16-total-rna-purification-kit-protocol.pdf).
The whole procedure was completely automatized. RNA purity was measured using a MBA200 spectrophotometer.

3.2.2.4. RNA extraction using the BeadBeater

The following steps were carried out at the Zoological Institute (Dept. of Evolutionary Biology, University of Basel) under the supervision and support Salzburger group.

Isolation of RNA from *Leucoraja melitensis* skin tissues was carried out testing Santos et al. (2014) protocol, developed and optimised on cichlid fins and currently in use at the Salzburger lab (http://www.salzburgerlab.org). Basically, RNA extraction was performed according to the TRIzol® protocol after incubating the dissected tissues in 750µl of TRIzol® at 4°C overnight or, alternatively, for 8–16h to increase the RNA yield after long-term storage. Tissues were then homogenized at 6s/m for 1 minute with a BeadBeater (FastPrep-24, MP Biomedicals, France) and Zirconia-silica beads (1.0mm diameter; BioSpec Products). This step was repeated when necessary. Subsequent DNase treatment was performed with DNase I (Sigma-Aldrich). RNA quantity and quality was determined with a NanoDrop ND-1000 spectrophotometer (ThermoFisher Scientific).

3.2.2.5. Combining RNA isolation protocols

Since the RNA yield conspicuously increased with the use of the zirconia-silica spheres and BeadBeater as homogenizers, the Direct-zol™ RNA MiniPrep kit was tested again. An optimal method

for isolating good quality RNA was finally established combining Santos et al. (2014) and the Direct-zol™ protocols, and applied to five main species, on a total of 132 individuals.

Key steps of the combined procedure involved a first purification of homogenised samples (with 10mm of zirconia-silica beads at 4s/m for 1 minute) with 200µl of Chloroform without Isoamyl alcohol (officially: 2:10 Chloroform:TRIzol) in order to get rid of proteins, pigments and other interfering substances. After centrifuging, the aqueous phase was collected and transferred to the corresponding 1.5ml Lo-bind Eppendorf tube filled with 400µl of fresh 99% EtOH. The cocktail obtained was then transferred to the spin columns and hence, the Direct-zol™ protocol was followed, adding two more washing steps with Washing Buffer (WB) to increase RNA purity. Genomic DNA was also carefully eliminated from RNA preparations by on-column digestion using DNase I set as the manufacturers recommended in their protocol. RNA concentration and purity were measured using a NanoDrop spectrophotometer, whereas RNA integrity was assessed according to the intensity and shape of 28S and 18S rRNA peaks on the Bioanalyzer 2100 using the Agilent RNA 6000 Nano Kit (Agilent Technologies, Germany). The five best RNA extractions per species, including the dorsal ornament, dorsal uniform matrix and the un-pigmented ventral skin, were chosen for library preparation and following Illumina sequencing (refer to Table S10 in Appendix for codes-conversion). The choice was strictly based on sample concentration (ng/µl), the 260/280 and 260/230 absorbance ratios which are considered as proxies of RNA purity, and RNA integrity, indicated by the BioAnalyzer as a RNA integrity number (hereafter RIN).

3.2.3. Library preparation and sequencing

Library preparation and sequencing were performed at the Department of Biosystems Science and Engineering of Basel (ETH Zürich; https://www.bsse.ethz.ch).

A total amount of 200ng of high-quality RNA was used for preparing libraries from 75 samples. The TruSeq Stranded RNA Library Prep HS Kit (Illumina, San Diego, USA) was applied following the manufacturer's instructions (http://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/samplepreps_truseq/truseqstrandedtotalrna/truseq-stranded-total-rna-sample-prep-guide-15031048-e.pdf). In general, this kit is particularly effective in eliminating rRNA by a poly-A enrichment, even with low quality or low abundant RNA (10-100ng). Target enrichment is fundamental when working on eukaryotic organisms, since it produces sufficient mRNA population and separates mRNA from rRNA

contaminants. Despite the verified efficiency of the kit, some biases toward the 3′-end of the expressed transcripts can be encountered because of random hexamer priming (Hansen et al. 2010). Library quality was measured on a 2200 TapeStation Instrument (Agilent Genomics) and multiplexed applying appropriate double indices which were incorporated and sequenced as a separate read for each flow cell cluster. The 25-sample libraries were pooled, aliquoted and randomly distributed across each of three lanes of the flow cell. Sequencing was performed on an Illumina HiSeq2500 sequencer using the 125bp single-end mode, which enables about 250 Million reads of output. Sequence reads were then demultiplexed by index using CASAVA v.1.8.2 software (Wilson-Leedy & Ingermann 2007). The quality checking of raw reads was then performed using the FastQC algorithm (http:// www. bioinformatics. babraham. ac. uk/ projects/ fastqc/). FastQ files from individual lanes were merged for each sample and were employed as input for the following analyses.

3.2.4. Raw reads quality check

In bioinformatics, the raw read quality is the cornerstone for obtaining reasonable data, avoiding false positives and bias which are likely related to the previous steps of the experimental phase. Despite the de novo assembly generally requires much cleaner reads than alignment to a reference genome or transcriptome, the best practice is to perform a read filtering step and discard or trim those reads containing adapters and low quality bases. Illumina reads are known to have a higher error rate as moving towards the 3′-end of the transcript, so if a drop-in quality is detected within the read, trimming the rest of the read off is necessary.

The sequencing reads, along with the corresponding base call qualities (i.e. the probability of a correct call) are displayed as a FASTQ file (which has the extension '.fastq' or '.fq'). These FastQ files contain a four-line record for each read, including its nucleotide sequence, a '+' sign separator and a corresponding ASCII string of quality characters, each one corresponding to an integer ***i*** ranging from -5 to 41 (depending on the version of Illumina software used for base-calling).

One of the tools providing quality control checks on raw sequence data is FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), which quickly identifies any issue or problem that must be considered before proceeding with further analysis.

A typical FastQC report consist of twelve sections (see http://www.bioinformatics.babraham.ac.uk/projects/fastqc/Help/3%20Analysis%20Modules/ for details).

The ‘Basic Statistics’ module reports the original filename of the file analysed, the actual base calls, the ASCII encoding of quality values used, the total number of processed sequences, the flagged ones to be removed from all analyses, the sequence length and the overall %GC of all bases in all sequences.

The ‘Per Base Sequence Quality’ displays the range of quality values across all bases at each position in the file. In the plot, the central red line is the median value, the yellow box represents the inter-quartile range (25-75%), the upper and lower bars represent the 10% and 90% points, whereas the blue line represents the mean quality. The y-axis on the graph shows the quality scores and the graph background colours divide the y axis into very good quality calls (in green), calls of reasonable quality (in orange), and calls of poor quality (in red).

The ‘Per Tile Sequence Quality’ encodes the flowcell tile from which each read came. Using cold and hot colour scales the module shows the deviation from the average quality for each tile and to check any losses in quality associated with only one part of the flowcell.

The ‘Per Sequence Quality Scores’ allows to identify subsets of sequences with low quality values.

The ‘Per Base Sequence Content’ plots the proportion of each base position in a file for which each of the four DNA bases (G, A, T, C) has been called. Because of the stochastic character of libraries, we would expect to be little or no difference between the bases of a sequence run.

The ‘Per Sequence GC Content’ quantifies the content of GC bases across the whole length of each sequence and compares it to a modelled normal distribution.

The ‘Per Base N Content’ indicates N substitutions to conventional base call, plotting the percentage of N base calls at each position.

The ‘Sequence Length Distribution’ identifies reads of varying lengths, rather than uniform, and shows the distribution of fragment sizes in the file considered.

The ‘Duplicate Sequence Levels’ section counts the degree of duplication for every sequence in a library and creates a plot showing the proportion of the library which is composed of duplicates. In the plot, a blue line represents the full sequence set, whereas the red line represents the duplicated sequences. In general, the level of coverage of the target sequence is high when a low level of duplication occurred. On the contrary, high level of duplication indicates enrichment bias (e.g. PCR over amplification). In a properly diverse library most sequences should fall into the far left of the plot in both the red and blue lines. More specific enrichments of subsets, or the presence of low complexity contaminants, will tend to produce spikes towards the right of the plot. When the amount of overrepresented sequence is high the red line is flattened, as a symptom of over-sequencing.

In the ‘Overrepresented Sequences’ module highly repeated base sequences (more than 0.1% of the total) are listed. Those sequences can either be highly biologically significant, or a symptom of library contamination.

The ‘Adapter Content’ is tightly linked to the overrepresented sequences as well and usually includes the presence of read-through adapter sequences at the end of the sequence.

This module is also correlated to the last one, the ‘Kmer Content’, showing an enrichment in k-mers (e.g. substring of length k) in any exactly duplicated sequences. This can happen when working with very long sequences with poor quality, which brings to random sequencing errors causing a dramatic reduction in the counts of exactly duplicated sequences, or when a partial sequence appears at different places within your sequence and won't be seen either by the ‘Per Base Content’ plot or the ‘Duplicate Sequence Levels’ modules.

The FastQC reports analysis were the meter of evaluation prior to the next step. The trimming of adapters and poor quality reads was performed using Trimmomatic v.0.36 (Bolger 2014).

3.3. Results

Three skin samples differing in pigmentation patterns (presence and absence of spot, dorsal uniform matrix and un-pigmented ventral skin) were collected from 92 specimens belonging to eight species: *Raja asterias, R. clavata, R. miraletus, R. ocellifera, R. straeleni, R. polystigma, R. brachyura* and *Leucoraja melitensis*. Of these, only *R. asterias, R. clavata, R. miraletus, R. ocellifera,* and *R. straeleni* were included in the experimental design (Figure 6) after evaluating both biological/evolutionary questions and experimental efforts.

Although the encouraging increase in the quantity of RNA extracted owed to the key-role of bead-beating, Santos’s protocol (2014) provided suboptimal absorbance ratios (max 260/280=1.78; max 260/230=1.60) and unsatisfying purity curves, independently from the kind of tissue (Table S10).

Only after the optimisation of a combined protocol coupling Santos’ (2014) with the Direct-zol™ kit, a total of 132 RNA samples, equally distributed among dorsal pigmented ornament (eyespot or pseudo-eyespot), dorsal pigmented uniform matrix and un-pigmented ventral skin, were successfully extracted from 44 individuals (see Table S10 in Appendix). A total of 75 differently pigmented samples considered optimal across five individuals per species were chosen for further experimental steps.

After the successful preparation of Illumina libraries and their sequencing, 75 FastQ files and their FastQC reports were obtained.

From a quantitative point of view, the Illumina sequencing of skin tissues was satisfying, since the expected output (250M read per lane) was abundantly reached. A total of 251,334,280 reads were produced in lane 1, 324,868,379 in lane 2 and 327,659,344 reads were generated in lane 3.

Sequencing was also highly performing from a qualitative point of view as well. According to the FastQC reports, no flagged sequences were identified as poor in quality across samples and main errors or warnings concerned the 'per base sequence content', the 'adapter content' and the 'k-mer content' analyses. The amount of overrepresented sequence (and adapters among them) seems to be the common denominator at the basis of warnings across these analysis modules. Indeed, if any evidence of overrepresented sequences, including rRNA or adapter dimers in a sample, as in this case, then a bias could weigh on the overall base composition and be the cause of an anomalous k-mer enrichment. After trimming and cleaning sequences from adapters, low quality reads and overrepresented sequences, between 0.28% (sample R14W) and 54.41% (sample R55S) of the reads per sample were discarded (see Table S12 in Appendix II and compare Figures S3 and S4 where the 'Per base sequence quality' plots for R55S before and after trimming are reported). The surviving reads recovered per each tissue were considered of high quality, therefore ideal for their mapping against the reference trancriptome and perform the DGE analysis (Chapter 5).

3.4. Discussion

Deep RNA sequencing has opened a new horizon for understanding gene expression in all kinds of tissues belonging to model and non-model species. The study of elasmobranch skin is still underestimated despite its multifunctional relevance for the individual and despite the unrevealed evolutionary meaning that some traits could carry (i.e. pigmentation, dermal denticle development and so on). Thus, applying the Illumina technology, we obtained good quality sequences from of differently pigmented and un-pigmented skate skin.

Considering the results obtained so far within this work, the path leading to high-quality sequences should be discussed. Alike the totality of NGS-based research, the key of success measured in high quality reads and the absence of contaminants, is mainly represented by the quality of input sample. Here, the difficulty in obtaining optimal RNA principally lied in two main factors: the homogenation methodology and the pigment content. The first condition highly depended on the roughness of skin, enhanced by the presence of dermal denticles and the high abundance of collagen fibres. In second instance the quality of dorsal pigmented tissues was slightly lower than the ventral, non-pigmented

ones. The hypothesis is that melanin or other co-extracted contaminants may have prevented the correct estimation of RNA purity, and therefore quality, interfering with both ND spectrophotometer and BioAnalyzer Instrument. Previous studies focussing on the efficiency of PCR and Real-Time PCR in presence of different concentration of co-extracted inhibitors identified collagen, and most of all melanin, as affecting the primer elongation (Alaeddini 2012) and reverse transcription (Eckhart et al. 2000). In the case presented here, no effects were observed on library preparation, rather than on the extraction step. Furthermore, melanin has been deeply investigated using spectrophotometry (Kalleberg et al. 2015; Treesirichod et al. 2014), suggesting that some other contaminant could have affected quality and quantity measurements. Overall, the combined protocol designed for RNA isolation have revealed to be effective in recovering enough good-quality RNA, as demonstrated by the purity and quality values (i.e. absorbance ratios and RINs) recorded (Table S10 in Appendix).

The success of Illumina sequencing applied on skate skin was demonstrated by both FastQC reports and the quantity of reads obtained per lane (more than 250M). For what concerned the high amount of overrepresented sequences, Trimming was determinant and in few cases, severe (e.g. R55S, tissue carrying an eyespot in *Raja clavata* has been trimmed for more than 54%), although the quality of input RNA was high (i.e. for R55S absorbance rations were 260/280=1.94, 260/230=2.11, RIN 7.10 and conc. 169ng/ul). Their clipping contributed to avoid the inclusion of misleading and biased information in downstream analysis

Overall, the Illumina sequencing has revealed to be highly performing on skate's skin tissues and sequences were favourably used for DE analysis.

Chapter 4. Retrieving a reference transcriptome for *Raja miraletus* by Ion Torrent sequencing

4.1. Introduction

In recent times, the progressive expansion and availability of genomic resources seem to have reached epic proportions, although this did not always coincide with a significant decrease in sequencing costs or, most of all, in time consumed for computational data analysis (e.g. genome annotation). The early and less onerous microarray technology, for instance, is based on the heterologous hybridisation of specific probes and transcript sequences between different species (Beneš & Muckenthaler 2003) and have been alternatively used in measuring gene expression. Despite its adaptability to a wide range of molecules (e.g. DNA, RNA, protein, lipids and carbohydrates), limitations of microarrays are represented by the divergence between the sequences used (Machado et al. 2009), the small number of detected genes and data interpretation (Bier & Kleinjung 2001).

With the increasing popularity of RNA-sequencing as an effective tool for exploring alternative splicing, alternative Transcription Start Site Selection (TSS), comparative transcriptomics and alternative polyadenylation, both model and non-model species started to be deeply analysed with equivalent data processing effort (Wang et al. 2009a). The turning point of RNA-seq is its power of providing genome-wide deep sequencing of RNA transcripts within a sample to single base resolution, even with low amounts of input RNA, and identifying gene sequences and polymorphisms without needing a priori information. Using RNA-seq, dynamic gene expression profiles can be described, small portion of genome activity analysed and transcriptome composition and complexity can be untangled (Wang et al. 2009).

At the same time, the analysis of RNA-seq data for non-sequenced species has been gradually improved using alternative, although effective, approaches that enable the reconstruction of a reference transcriptome by assembling raw transcripts which are annotated by homology searching among the available genomic options. Reference transcriptomes can be assembled using a genome or a transcriptome from a closely-related species (no more than 100Mya according to Hornett & Wheat 2012) as a guide ('genome-guided'), or performing a reference sequence-independent assembly ('de novo'; Garber et al. 2011). In both cases, the amount of homology found in those transcripts has been demonstrated to be inversely proportional to the divergence among sequences (Colgan et al. 2011; Kawahara-Miki et al. 2011).

Previous studies have demonstrated that the quality of an assembly can be highly variable, independently from the platform chosen for RNA-sequencing (Lu et al. 2013; Vijay et al. 2013; Hornett & Wheat 2012) and that direct mapping-based methods can be highly effective in identifying genes involved in a trait of interest in non-model species.

The assembly of a transcriptome can be definitively useful to discover novel transcripts and previously unknown genes or explore gene expression level, but all these opportunities naturally come along with some issues. Thus, given the importance of comparative analyses in exploring the molecular basis and evolution of biological traits, maximising gene detection rates and minimising false positives before analysing and interpreting expression profiles is recommended, or better, mandatory.

Among the main issues, the non-uniformity of sequence coverage should be considered. Lowly expressed genes may be partially covered by a few reads and be hardly represented in full length, leading to a high number of gaps. In second instance, the handling of alternative spliced isoforms is complex, since one gene may have several isoforms and, in general, short reads do not allow the exact assignment of isoforms to their origin. Furthermore, alternative splice variants often share exons and distinct transcripts can be difficult to resolve (Pyrkosz et al. 2013). Thirdly, homologous and repeated sequences can be a source of ambiguity in the assembly and lastly, the discrimination of exons and introns originated from incompletely spliced RNA precursors is hard to accomplish. The comparison between the currently available assemblers carried out on both simulated and real sequencing data (Lu et al. 2013), revealed that the most flexible and efficient tool in overcoming most of these issues, discovering more loci and identifying full-length transcripts is Trinity (Haas et al., 2013; Grabherr et al. 2011).

Trinity was specifically created for transcriptome reconstruction starting from short reads and it is composed of three main modules: Inchworm, Chrysalis and Butterfly. The Inchworm module assembles reads in full-length contigs (i.e. contiguous sequences) for one major isoform and reports only unique portions of minor spliced variants. Chrysalis create clusters of Inchworm contigs according to a defined and sufficient k-mer overlap and builds de Bruijn graphs for each cluster, where reads are modelled as overlapping sub-sequences (k-mers) and graph nodes can be traversed to generate transcripts. Butterfly simplifies the graphs, treating each of them separately, then groups them with original reads, counting full-length alternatively splice isoforms and separates transcripts from paralogous genes. The final assembly that will be used for further analysis is in .'fasta' format.

Here, Trinity tool v.2.3.2. was employed to assemble a reference transcriptome for the brown skate, *R. miraletus* L., representing one the target species of the initial experimental phase and one of the most striking cases of cryptic evolution among Rajidae. Taking advantage of the available genome of the little skate (*Leucoraja erinacea*) a genome-guided assembly was performed to likely strengthen the power of mapping and the resolution of DE analysis.

4.2. Materials and methods

4.2.1. Sampling

Tissues used in this phase were collected from the Adriatic Sea during the SoleMon scientific research program carried out during November 2015. No specific approval of this vertebrate work is required since the individuals sampled in this study were obtained for scientific purposes.
After the preparation and cleaning of the materials needed (2 ml Nuclease-free tubes with cap and O-ring labelled with Sample ID and filled with 1.2 ml of RNAlater®, scissors, tweezers and scalpel blades) and the sterilization of the working bench, specimens were delicately washed and quickly scrubbed with clean paper towel to get rid of eventual organic and inorganic residuals. Then, they were accommodated on an aluminium layer sprayed with RNAse Zap® and placed on ice. Principle biometric measures were taken (total length and weight) and sex was determined. The main organs (skin from different body portions, heart, liver, brain, gonad, muscle, eye, dorsal fin, jaw bone and gill) were isolated, collected in replicate and transferred to the corresponding tube. Tubes were first stored at 4°C overnight, then at −20°C until further processing to maximize RNA fixation. Samples were transported in dry ice during their transfer to the laboratory.
Based on the number of specimens collected during the sampling phase, an experimental design was ideated to perform the Ion Torrent sequencing of all tissues and organs, with the aim of building a reference transcriptome for one of the species of interest.

4.2.2. RNA isolation

RNA isolation was performed using the combined protocol developed for the Illumina sequencing. RNA concentration was verified on NanoDrop ND-1000 spectrophotometer, whereas its integrity was checked on a Bioanalyzer 2100 with the Agilent RNA 6000 Nano Kit.

4.2.3. Ion Torrent library preparation and sequencing

After the dilution to an equimolar concentration of 2000ng/50µl, samples were poly-A enriched with the Dynabeads mRNA DIRECT Micro Purification kit (ThermoFisher Scientific). The amount of selected mRNA was checked on a Bioanalyzer 2100 with the Agilent RNA 6000 Pico Kit.

Genomic libraries for strand specific reads were prepared with the Ion Total RNA-Seq Kit v2 (ThermoFisher Scientific). The choice of the kit was driven by the read length promisingly produced after sequencing (400bp) that should improve the quality of the assembled reference and reduce errors related to the mapping step (Pyrkosz et al. 2013). This issue will be discussed in chapter 5.

Library preparation started with an optimized enzymatic fragmentation step performed at 37°C for 5 sec. Incubations were carried out in a Veriti®384-Well Thermal Cycler (Applied Biosystems). Samples were vacuum-concentrated at 30°C for about 9 min in a Concentrator Plus instrument (Eppendorf). Hybridisation between fragmented mRNA and adapters was performed incubating the cocktail at 65°C for 10 min, then at 30°C for 5 min. The following ligation step was performed at 30°C for 2 hours with the lid temperature turned off. Retro-transcription was carried out at 70°C for 10 min and after adding a 10X Superscript III enzyme an incubation at 42°C for 30 min followed. After a further cDNA purification step, two Ion Xpress RNA 3' barcodes were added to male's samples, while a different barcode was assigned to the female's ovary.

The successive cDNA amplification was run as indicated by the manufacturer instructions (https://tools.thermofisher.com/content/sfs/manuals/MAN0010655_IonTotalRNASeqKit_v2_Whole Transcriptom_Libr_QR.pdf). Barcoded libraries were first purified and then quantified on a Bioanalyzer 2100 with the Agilent High Sensitivity DNA Kit. After dilution at an equimolar concentration, libraries were pooled and used for the Template preparation.

This intermediate step was carried out using the Ion PGM Template OT2 400 kit (https://tools.thermofisher.com/content/sfs/manuals/MAN0007219_Ion_PGM_Template_OT2_400_Kit_QR.pdf).

Template was then sequenced on the Ion Torrent Ion PGM™ Sequencer using the Ion PGM HI-Q View Sequencing kit. The chip used for sample loading was an Ion 318 Chip Kit v2 (ThermoFisher Scientific). A total of three runs were performed at different library concentrations ([125pM], [200pM] and [100pM]). Reads were automatically partitioned by the Ion Torrent software (Torrent suite v.4.4) into BAM and FASTQ files based on tags (reads without tags were discarded). The output BAM and FastQ

files were downloaded from the Ion Torrent server and the raw reads contained in the FastQ files were used as input for FastQC tool for quality check.
Clipping of adapters and low quality reads were performed using Trimmomatic v.0.36 (Bolger 2014).

4.2.4. The assembly

The genome belonging to the little skate (*Leucoraja erinacea*) and available at http://skatebase.org, resulted a useful resource while assembling a reference transcriptome for the brown skate. The assembly was performed using the 'genome guided' function included in Trinity tool (Grabherr et al. 2011) as follows.
The little skate' genome was first indexed using the bowtie2-build command included in Bowtie2 v.2.2.9. (Langmead & Salzberg 2012). Male's clipped reads were first aligned to *L. erinacea* genome using STAR v.2.5. (Dobin et al. 2012) without GFF3 or GTF annotations or annotated junctions and allowing one multiple alignment per read. The final alignment resulting in a '.bam' file was then used in Trinity, as a frame, for grouping the overlapping reads into clusters. Differently from other approaches (e.g. Cufflinks), transcripts were build according to the actual read sequences rather than the exact matching with genome sequences. This resulted particularly helpful, since *R. miraletus* divergence from genus *Leucoraja* has been estimated at approximately 35Mya (according to mDNA; Tinti et al. in prep).
The assembly was then performed with default k-mer=25 and maximum intron length of 10kb. The quality of the assembly was evaluated blasting it to the UniProt knowledgebase (Universal Protein Resource http://www.uniprot.org). To this end the BLASTx software was used (Altschul et al. 1990) and the `--evalue` option was set as stringent (E: 1e-20).
To estimate the transcript abundance, the alignment-based RSEM tool v.1.2.31. was applied to each sample individually (RNA-Seq by Expectation Maximization; Li & Dewey 2011). This method performs transcript-level estimates of the count of RNA-Seq fragments that were derived from each transcript, considering the transcript length, the total number of reads mapping to any transcripts and reporting the normalized expression metrics as 'fragments per kb transcript length per million fragments mapped' (FPKM) or 'transcripts per million transcripts' (TPM). Low expressed genes and transcripts were filtered out using the `filter_low_expr_transcripts.pl` setting the minimum expression level required across any sample at 1 and the minimum percent of dominant isoform expression at 1.

Before moving towards results, it should be specified that from this point of the dissertation, results will be described and discussed using the 'gene' and 'transcript' terminology. With 'gene' Trinity identifies the most dominant resolved pathway of a given de-Bruijn graph produced by Chrysalis module. Therefore, its acceptation is quite far from the molecular definition of hereditary unit. Differently, with the 'transcript' term all the isoforms for that 'gene' were defined.

4.2.5. Functional annotation of transcripts

A first attempt of functional annotation of transcripts was carried out using the Trinotate protocol available at http://trinotate.github.io. After the initialization of the 'Trinotate.sqlite' database, consisting with the loading of the brown skate reference transcriptome in '.fasta' format, the loading of gene/transcript relationship and the loading Trasndecoder-predicted candidate coding regions and ORFs (http://transdecoder.github.io). The 'population' of the Trinotate.sqlite inventory was then performed loading protein information derived from a less stringent BLASTx search on Swissprot database (`--evalue`=1e-5) and protein domains identified by HMMER tool v.3.1. (http://hmmer.org) across the PFAM protein families (http://pfam.xfam.org).
Then, the annotation report of genes and functions likely matching to the brand-new gene and isoform sequences was produced and used as an input for the implementation of Go-seq v3.4. Bioconductor package (http://www.bioconductor.org/packages/release/bioc/html/goseq.html) for the Gene Ontology (GO) assignment and mapping as implemented by Trinity tool.

4.3. Results

Two mature individuals, one male and one female, of brown skate were sampled during the scientific survey and used for the experimental phase (see Table S13 in Appendix III for sampling details).
A total of 24 RNA extractions were performed (Table S14). Those belonging to the male were considered the best ones in term of quantity and quality. Of these, eleven were chosen for Ion Torrent library preparation. The RNA extracted from the female's ovary was also included in the library preparation and individually indexed (sample 29T).
Library preparation and sequencing involved a total of twelve organs and tissues. After the template preparation and sequencing, Ion Torrent run reports were produced. The run performed with template with concentration 125pM showed 68% of enriched Ion Sphere Particles (hence ISP)

Loading and produced 4,011,859 raw reads. The enrichment (the positive cohesion between ISP and template) was 100% and polyclonal value (the percentage of beads estimated to have more than one template for ISP) was median low (36%), as well as the low-quality reads and the adapter dimer percentage (6.4% and 4.7%, respectively). Both fragment tests were high (TF_C and TF_1>90%), indicating a technically well performing run.

The run performed with template [200pM] showed 67% of ISP Density and produced 3,765,129 raw reads. The enrichment was 100% and polyclonal percentage was slightly higher (41%). Low-quality reads and the adapter dimer percentage were 5% and 4.2%, respectively). Both fragment tests were high (97% TF_C and TF_1 93% respectively), indicating a satisfying run.

The last run with template [100pM] showed 64% of ISP Density and produced 3,071,821 raw reads. The enrichment was 100% as well and polyclonal value was high (43%). The amount of low quality reads was higher (9.8%) and the adapter dimer percentage low 4.2%. Both fragment tests were high (97% TF_C and TF_1 91% respectively), indicating well performing runs. Overall, Ion Torrent sequencing can be considered sub-optimal in terms of ISP density and number of usable reads, ranging from 53% to 43%.

Overall, the desired read length of 400bp was reached, but the median read length obtained was much shorter. The maximum read count was observed between 100bp and 200bp. Also, the mean read length produced for the ovary tissue was shorter than the pooled male's tissues and the number of total bases synthetized was much lower (refer to Ion Torrent run reports in Appendix III).

After raw reads quality check, FastQC tool did not reveal any adapter contamination. The amount of overrepresented sequences was low, but many low-quality bases affected the terminal part of the read. As a result, sequence quality slightly decreased as moving from the 5' towards the 3' position. After the removal of about 15% of low quality reads, the ovary sample resulted too poor for further analyses. A total of 165,919 sequences were recovered from the first run [125pM], 145,324 sequences were recovered from the second [200pM] and 128,821 sequences were retained from the third one [100pm]. On the contrary, the pooled male's sequences produced at each run were considered satisfying in relation to the number of reads retained after trimming (approximately 2%), the low number of overrepresented sequences and the mitigation of the slight decreasing trend in sequence quality after trimming. For these reasons, suitable in both quantity and quality for the reference transcriptome reconstruction, thus employed for the assembly.

The assembly produced 118,429 putative transcripts and 116,136 putative 'genes' with median contig length of longest isoform per gene counting 379 nucleotides. The first line of the reference

transcriptome is reported here as an example of a Trinity output. The accession ID 'TRINITY_GG_1_c0_g1_i1' indicates the method chosen to perform the assembly (GG, genome-guided), the read cluster '_c0', the gene '_g1' and the relative isoform '_i1'. Furthermore, the length of the transcript in bases and the path employed for resolving the de Brujin graph are reported. The nucleotide sequence follows. Trinity generates thousands of these clusters differing in '_g' or '_i'.

>TRINITY_GG_1_c0_g1_i1 len=252 path=[1:0-251] [-1, 1, -2]
GAGGATAACTACACCTGTAACGCTGGCAACACGGCAGCTTCACCTCCATGACATTCACCCACACACCGGTGT
CCAAACACAGCGCGAGAGTGAGTGAGTGAGTGAGGCTCACACGTTAACGTGGACTTGCCCAGGGCAGCAG
CTGCTGGAGACTCACAACTTCCACACGACTCCAGCCGCCAACCAGCTTCCTCCACAGCATCCTGTCTGTGTCT
GATCCCACCTCTGTATCTC

A total of 300 proteins matched a transcript by >80% and <= 90% of their protein lengths, 728 proteins were represented by nearly full-length transcripts (>80% alignment coverage) and 428 proteins were covered by more than 90% of their protein lengths, as inferred from BLASTx quality assessment. The length of the N50 contiguity index, usually employed as an indicator of transcript fragmentation, was 460 meaning that at least half of the assembled bases are in contigs of at least 460 nucleotides in length.

After the creation of a gene-list based on the known features involved in pigmentation at any level (from the pigment synthesis, the melanosome construction to systemic effects) the presence of candidate genes was assessed performing a raw count of BLASTx and PFAM matchings. This approach highlighted the presence in the assembly of 150 transcripts ascribable to 51 different pigmentary proteins and genes (Figure 7; Table S15 in Appendix III). Among the most represented gene-complex the XDH (Xanthine dehydrogenase/oxidase) and the transmembrane receptor EDNRB were found. The former is an enzyme involved in yellow pteridine pigment synthesis, while the latter is involved in the melanocyte differentiation and many other molecular and biological functions (i.e. nervous system development, positive regulation of cellular proliferation, neural crest cell migration).

The Mediator of RNA polymerase Transcription (MED12) gene was also observed, and it is known for its role in the regulated transcription of all RNA polymerase II-dependent genes. In *Danio rerio* for instance, MED12 is required for the development of the body axis, brain, ear, kidney, forelimb, neural crest and for pigmentation. Among its multiple functions, it appears the responsible of iridophores

differentiation. It also acts as a coactivator for SOX9A and/or SOX9B promoting the expression of several neuronal determination genes, also reported by BLASTx results.

Two more genes, tightly related to each other, were identified as contributors to pigmentation regulation: TYRP1 and TYRP2. The tyrosinase-related protein 1 catalyses the oxidation of 5,6-dihydroxyindole-2-carboxylic acid (DHICA) into indole-5,6-quinone-2-carboxylic acid. It also regulates or, at least seems to influence, the type of melanin synthesized. Also to a lower extent, it is capable of hydroxylating tyrosine and producing melanin. According to the GO database, its biological meaning is related to both melanocyte differentiation and melanosome organization. Similarly, TYRP2 tautomerase, enhance the conversion of L-dopachrome into DHICA and it is involved in the regulation of both eumelanin and phaeomelanin levels. Also, it plays a positive regulation of neuroblast proliferation and division.

The Melanocyte protein (PMEL) plays a central role in the biogenesis of melanosomes in *H. sapiens*. As a matter of fact, it is involved in the maturation of melanosomes from stage I to II. It can also be involved in immunity response and its homologues have been found from other mammals, to birds (i.e. chicken) and reptiles.

The Microphthalmia-associated transcription factor (MITF), besides being a promoter of cell differentiation and survival, it is a regulator of TYRP1, thus it is fundamental for pigmentation, differentiation of neural crest-derived melanocytes, mast cells, osteoclasts and optic cup-derived retinal pigment epithelium.

Lastly, the Mast/stem cell growth factor receptor Kit (or simply KIT) is certainly one of the most powerful feature influencing biological and molecular functions in human as in amphibians. Although this power it was scarcely represented within the assembly.

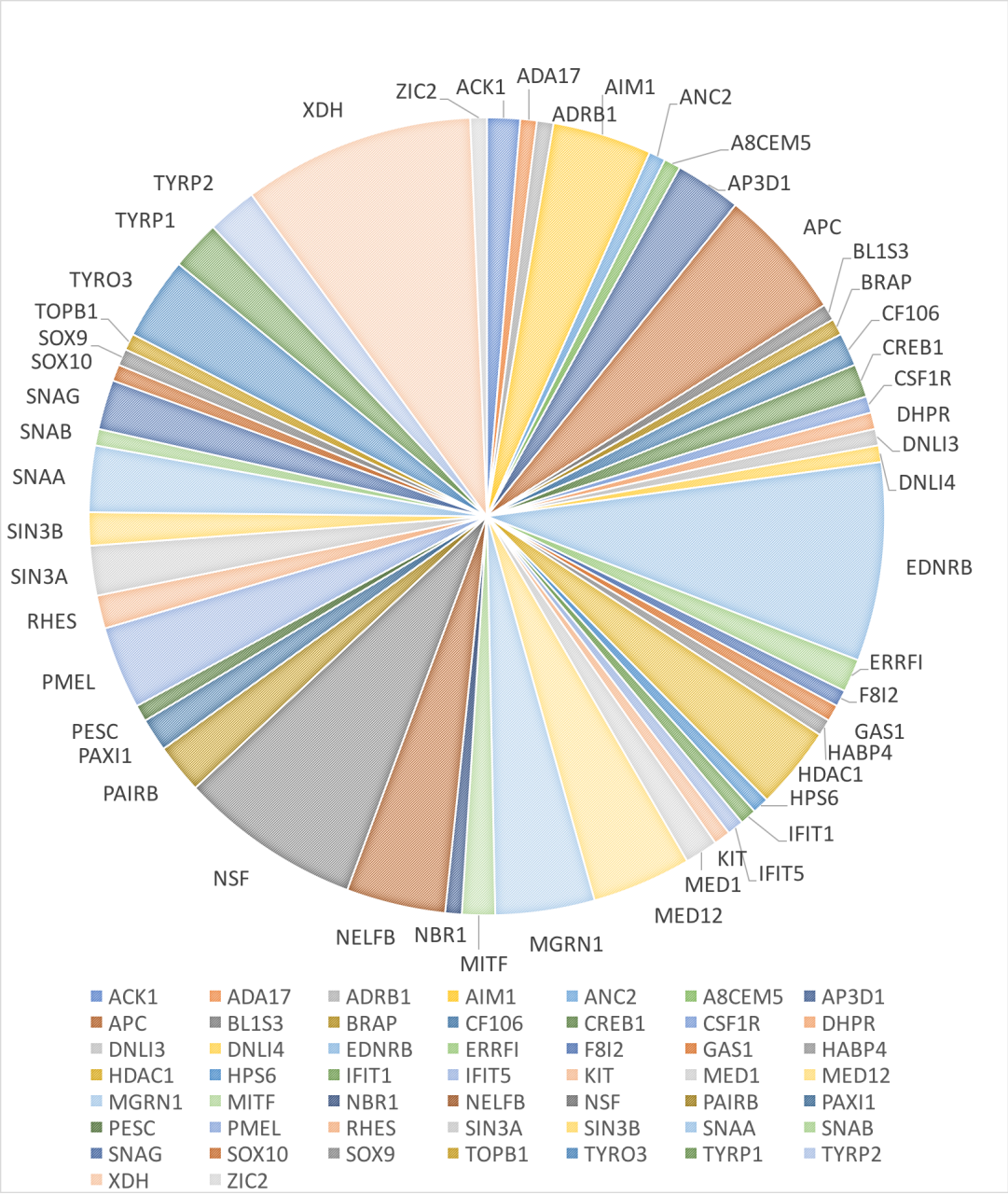


Figure 7 Putative pigmentary genes populating the assembly according to BLASTx best matchings against UniProt and PFAM knowledgebase. Refer to Tab S15 in Appendix III for gene acronyms and functions.

For what concerns the RSEM evaluation of transcripts quality, the number of expressed genes by at least one TPM in any of the considered samples amounted to about 18k (Figure 8). This number describes the number of genes that are best supported by the expression data. After filtering, 90.79% of total transcripts (117,349/129,254) was used for the mapping of Illumina reads obtained from differently pigmented skin tissues.

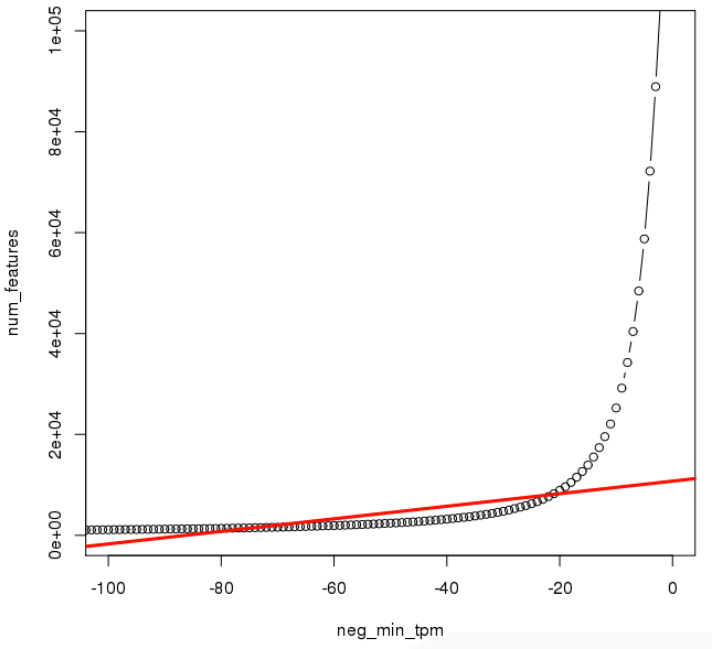


Figure 8 Graphic representation of the number of expressed 'genes' recovered with the transcriptome assembly as a function of minimum TPM threshold. A linear regression in red indicates the approximate number of expressed features

4.4. Discussion

The Ion Torrent platform is known for its effectiveness in generating longer reads than Illumina (up to 500bp with the IA 500 Kit, Thermofisher Scientific) in relatively short time, even if with lower output per run in terms of raw reads produced (Liu et al. 2012). Therefore, the expected read length was determining in the choice of the platform. In general, the longer reads are generated within sequencing, the less fragmented and incomplete should be the resulting assembly (Vijai et al. 2013). The unbalanced read output produced by the mentioned techniques was also observed within this work, where the Illumina experimental data exceed the 250 Million reads per lane, while the Ion Torrent sequencing produced about 4 Million reads per run.

The experiment described here resulted extremely laborious in terms of protocol optimisation for Library and Template preparations. Considering the low IPS loading success, likely depending on the Template preparation (i.e. emulsion PCR) and the read length reached, Total RNA-seq was slightly underperforming.

Nevertheless, the quality check of raw reads assessed with FastQC revealed a sufficient quality of sequences in the pooled male's samples which were thus used for the assembly of a reference transcriptome for *Raja miraletus*.

In general, the assembly showed a high number of putative genes and transcripts, but a short median contig length (379 nucleotides). This condition could depend on two main factors. One is related to the sequencing quality, in terms of both raw reads and sequencing depth. A high coverage is always desirable to better capture reads representing the moderately to lowly expressed transcripts. When the latter are abundant, further sequencing would likely help the building of complex transcriptomes. For this reason, the assembly was performed with the awareness that the coverage reached within the Ion Torrent sequencing was suboptimal.

The second factor could be related to the construction of contigs itself. This condition highly depends on the assembler employed. Trinity tool was chosen for its effectiveness in assembling short reads compared to other assemblers (Celaj et al. 2014; Clarke et al. 2013; Grabherr et al. 2011). Furthermore, this tool has demonstrated its power in previous studies where Ion Torrent sequence data were employed (Amin et al. 2014). The amount of the N50 contiguity index obtained here was consistent with the median contig length. Different results were obtained by Amin et al. (2014), who assembled the transcriptome of the gastropod *Nerita melanotragus* (E. A. Smith, 1884) recovering N50=258 and 293 nucleotides of contig length from 1,000,000 Ion Torrent reads (133bp of average

length). However, N50 measures are widely considered misleading for transcriptome assemblies since transcript length can be highly heterogeneous and the transcriptome highly complex (Vijai et al. 2013).

Results obtained blasting the brand-new reference transcriptome of the brown skate were encouraging, since more than 50% of the transcripts was assigned to a putative function. In addition, 150 transcripts were matching with 51 genes related to pigmentary regulation, melanocyte development and differentiation in different model organisms, although with a certain amount of redundancy. This could be related to the high number of genes generated within the assembly, paired with the genome size of the analysed species and its complexity. For *R. miraletus* the genome size has been estimated with C-value= 3.80 pg (Stingo et al. 1989), corresponding to about 3716 Mb.

The UniProt knowledgebase was chosen for the preliminary investigation of the putative pigmentary genes populating the assembly since it represents one of the most comprehensive, verified and up-to date sequence database containing information about genes and their functions across the whole living model and non-model organisms. Considered the scarce or fragmented genomic resources available for chondrichthyans, this approach was specifically chosen to obtain an overview on the whole characterization of the assembly. Performing a BLASTx analysis on more closely-related and almost entirely annotate model species (e.g. *Danio rerio*) or species showing high rate of gene homology and similar skin structures (e.g. *Xenopus tropicalis* or *Salmo salar)* could likely improve the percentage of identity, alignment and E-values between the newly obtained transcripts and putative pigmentary genes already known for their involvement in phenotypic traits as eyes-spots, egg-dummies and blotches (e.g. *Astatotilapia burtoni* or *Haplochromis elegans* Treewavas, 1933). Lastly, the use of the Non-Redundant NCBI database (i.e. REFseq https://www.ncbi.nlm.nih.gov/refseq/) could also improve the functional annotation of unassigned transcripts.

Chapter 5. Differentially expressed genes in skate skin

5.1. Introduction

With the term of gene expression is defined the process by which the nucleotide sequence of a gene encodes for a product. Identifying the different patterns of gene expression according to the experimental condition or phenotype is one of the increasing interests among evolutionary biologists. These aims could be also considered delicate; many factors should be considered while analysing differentially expressed genes (DEGs). The most crucial among them are the normalization of expression values and the accuracy employed while detecting differential expression, especially in dynamic systems like RNA-sequencing.

Within these systems, DEGs are identified and extracted based on a combination of expression level threshold and expression score cut-off, generated by statistical models and the expression level of each replicate or sample is measured by the number of sequenced fragments hitting the transcript, which are expected to correlate directly with its abundance level (Rapaport et al. 2013).

In addition, two more factors should be considered as risky. On one hand, the overall library preparation methodology and quality could be an important factor of bias (Roberts et al. 2011; Hansen et al. 2010; Li et al. 2010). On the other hand, including too few replicates into the RNA-seq experimental design to reduce sequencing costs could also affect the resolution of the whole analysis.

The advent of RNA-seq led to a renewal of data analysis methods aiming to statistically quantify gene expression, compare it between samples and wiggle out of the limitations of other platforms (e.g. microarrays; Wilhelm & Landry 2009; Wang et al. 2009). Traditionally, statistical methods applied to NGS rely on the use of biological or technical replicates to estimate variability in the data. Popular methods for analysing RNA-Seq data with replicates include edgeR (Robinson et al. 2010), DESeq2 (Love et al. 2014), Cuffdiff (Trapnell et al. 2012) and the recent NPEBSeq (Bi & Davuluri 2013).

One of the most delicate steps following the reference assembly and prior to DE measurement, is mapping short reads (35-125bp). Therefore, when transcripts are similar in length, it is not always possible to uniquely assign them to a specific gene and the handling of those reads mapping equally well to multiple locations in the transcriptome (from 10% to 50%; Li et al. 2011; Turro et al. 2011) becomes the critical point (Oshlak et al. 2010; Hashimoto et al. 2009). Indeed, consistent errors generally come along with the mapping phase, mostly when the reference is incomplete (Pyrkosz et al. 2013). This condition leads to unprecise transcript expression levels. Furthermore, it has been

described that alternative splice variants tend to increase the false positive rate of mapping (Mortazavi et al. 2008).

Many aligners have been developed based on fast indexing algorithms, while keeping an eye on these issues (e.g. STAR, TopHat, Bowtie), although no consensus about the best performing has been reached yet and still, some studies have indicated that the accurate allocation of all multimapping reads cannot be performed because some splice variants are linear combinations of other splice variants (Xia et al. 2011; Lacroix 2008).

The most used approach in NGS for hits counting usually considers the total number of reads overlapping the exons of a gene. Nevertheless, a partition of reads will always map outside the boundaries of the known exons (Pickrell et al. 2010). A reliable alternative strategy considers the whole length of a gene, treating also those reads derived from introns. Besides, if correctly treated in the mapping step, spliced reads can be useful to estimate the abundance of different splicing isoforms of a gene (Gatto et al. 2014), especially in RNA experiments where repetitive regions are part of the biological variability.

Once having acknowledged all the limitations, a satisfactory subset of aligned reads can be identified and grouped considering all sub-sequences of a certain length k (k-mers) and used them to compute counts and to obtain an estimate of expression levels.

Recalling the very final aim of this research, the DGE of differently pigmented tissues (i.e. eyespot or pseudo-eyespot dorsal ornament, the dorsal pigmented skin not carrying the ornament and white ventral portion) across species could help to evaluate whether the same or alternative traits have been used in parallel adaptations to similar environments. The hypothesis here is that the well-defined and similar eyespots characterizing the siblings *Raja miraletus* and *R. ocellifera* are controlled by the same genetic components and might display analogous expression profiles. Secondly, these genetic traits are expected to be different in *R. asterias,* and the sibling *R. clavata* and *R. straeleni*, displaying different dorsal ornaments, in form of pseudo-eyespots.

Within this last chapter, the methods chosen for the performance of the Differential Gene Expression analysis of differently pigmented tissues according to these rationales and the functional meaning of the DEGs identified have been described and discussed.

5.2. Materials and methods

5.2.1. Differential gene expression analysis

Bowtie2 v.2.3.0 aligner (Love et al. 2014) was chosen for the mapping phase, since it has been optimized for aligning long reads (more than 100bp) with relatively fast computational processing (Love et al. 2014).

The single library of reads corresponding to differently pigmented and non-pigmented tissues were mapped to *R. miraletus* reference transcriptome with default `–k`=1. The option to report up to one valid alignment per read was also set. The `––best` option was chosen to report best singleton alignments per stratum (i.e. number of mismatches), and `–m` option was set at 200, aiming to suppress those alignments for a read exceeding the value.

The abundance of the aligned transcripts within each one of the 75 libraries of trimmed reads was recovered using RSEM (Relative Estimated RNA-seq fragment) v.1.2.31, launching the `align_and_estimate.pl script` included in Trinity. The gene and isoform outputs created for each individual at the three-level skin tissue (D dorsal matrix, S spot and W ventral white skin) were then used to calculate the 'gene.count.matrix' and 'trans.count.matrix' respectively, using the `abundance_estimates_to_matrix.pl` plugin also included in Trinity. Matrices' function consisted in counting the raw and normalized read expected for each sample. The 'gene matrix' counts were used for the DE analysis, since the 'trans matrix' was expected to contain too many isoforms that might contribute to noisy estimations. No further filtering of low expressed transcripts was applied prior to DE analysis, to avoid the loss of biologically relevant transcripts.

For DE analysis DESeq2 v.3.4. Bioconductor package was chosen, as the more accurate and conservative between other tools (e.g. edgeR; Rapaport et al. 2013) which can identify differentially expressed transcripts and cluster them together based on expression profiles. This tool has been designed with a higher filtering power of low count genes, flagging those with large outlier counts and preventing the overestimation of variance among groups through the dispersion measurement in relation to normalized data (Love et al. 2014). Furthermore, DESeq2 has been programmed to recognize the noise derived from counting, sampling noise or Poisson dispersion, which are intrinsic of variable data counting and are introduced by both technical and biological variance.

The pairwise DE analysis among samples was carried out using DESeq2 as implemented in Trinity tool.

The input 'count.matrix' of read counts within S (spot), D (dorsal matrix) and W (white ventral portion) across five species were separately used as input, thus the DE analysis run in parallel for the three conditions considered.

The TMM normalized expression matrix obtained as a further RSEM output for each condition (S, D and W), was then employed along with the DE analysis. These matrices represented a crucial tool in the identification of differentially expressed genes, because they were based on the Trimmed Mean of M-Values normalization method (TMM), which assumes that most genes, are common among the compared samples, but not differentially expressed. Thence, on the basis of 'TMM.EXPR.matrix' transcripts mostly differentially expressed across the samples were extracted and clustered according to the most significant p-value cut-off for FDR (i.e. the ratio between the number of false positives and the total number of regulated gene) and fold-changes (i.e. the ratio between gene expression values). Genes having P-values at most 1e-3 and being at least four-fold differentially expressed were extracted and divided in subsets of up-regulated and down-regulated traits for each couple of comparisons. While the MA (with M representing the log ratio and A the mean average) and Volcano scatter plots showing the differences in DE between species were recovered by the first step of the analysis, the amount of up and down-regulated genes and their correlation across species were generated within the extraction of most differently expressed genes.

The software GFOLD v.1.1.4. (Feng et al. 2012) was used for a further DE analysis based on fold change instead of P-value. The cut-off limit chosen for the analysis was 0.01. This enabled the pairwise comparison of each tissue (spot, dorsal matrix and white ventral portion) within the same individual. Then, transcripts occurring at the species level for each tissue were used for GO annotation in order to identify enzymes, domains and protein regions matching with the DEGs observed.

5.3. Results

The percentage of alignment obtained when mapping the filtered Illumina reads against the brown skate transcriptome ranged from 68% to 76%. Of those, about 38% mapped more than one time.

The DE analysis carried out on separate skin tissues (spot, dorsal uniform matrix and ventral portion) across species highlighted a total of 7,548 genes differentially expressed in spot samples (S) across all species, 7,415 in pigmented but uniform dorsal portion (D) and 8,354 genes were differentially expressed in the ventral portion (W). The distribution of DE genes per tissues and couple of species

is reported in Table S16 in Appendix IV.

The MA plot generated within the DE analysis (as the one reported as an example in Figure 9) provided a useful overview for the pairwise comparisons among individuals. Within these plots each gene was represented by a dot. The x-axis reported the average expression over the mean of normalized counts (A-values), while the y-axis was the fold change between treatments (M-values). Similarly, the 'Volcano' plot summarized both fold-change and FDR significance (an example of Volcano plot is also reported in Figure 9). In particular, these scatter plots were built on the negative log10-trasformed e-values for FDR (y-axis) against the log2 fold-change (x-axis). The log2 of the fold-change was applied in two directions (i.e. up and down). Data points located at the top and far from either the left or the right side represented values with large magnitude fold changes and high a statistical significance (low FDR).

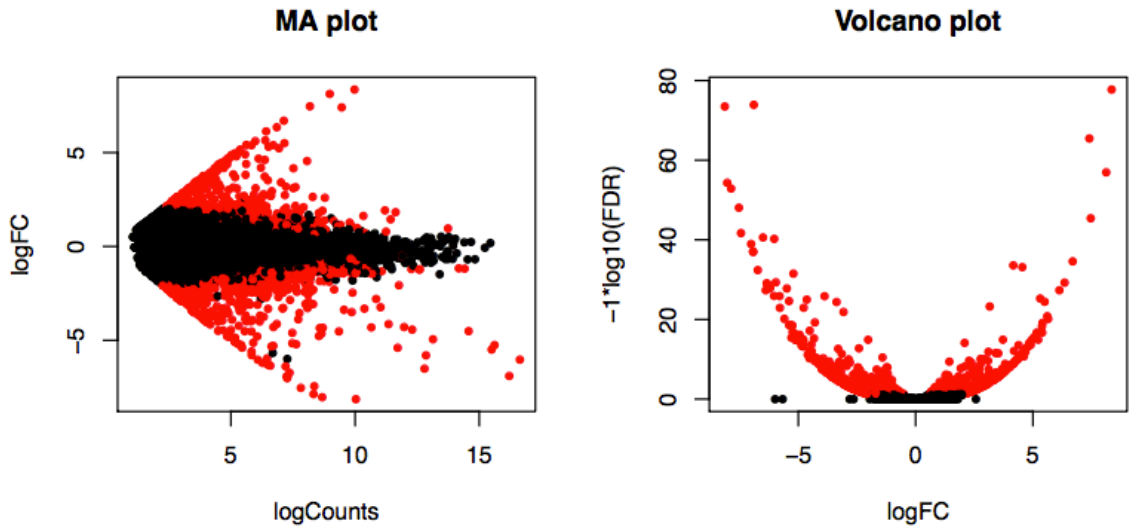


Figure 9 MA and Volcano plots obtained from the DE analysis. The example refers to *R. miraletus and R. ocellifera* comparison at the eyespot level. In MA plot, red data point indicated that the adjusted p-value was less than 0.05. In the volcano plot, the red points indicate genes of interest that displayed both large-magnitude fold-changes as well as high statistical significance (FDR<0.05).

To visually explore the degree of similarity across groups of closely related genes, heatmap representations were employed. As a two-dimensional plot, heatmaps quantitatively and qualitatively reflected the original experimental observations. The rows and columns of the gene-matrix for each species were hierarchically and independently rearranged into clusters based on Euclidean distances. Genes with similar expression patterns resulted adjacent. Furthermore, the computed dendrogram resulting from the clustering was automatically added as a summary of the relationships among genes (see Figures 10-12). The up (in yellow) and down-regulated genes (in purple) observed per replicate are reported in the species-heatmaps below. The expression values are plotted in log2 space and the mean expression value for each gene was subtracted from each of its expression values in that row (i.e. mean-centred).

Results presented in Figure 10 highlighted a clear DE profile between spot samples (S) clustering within species. In particular, *R. asterias*, *R. clavata* and *R. straeleni*, which carry a pseudo-eyespot on the dorsal surface of the body, grouped together following the expected trend reported by the dendrogram at the top of the heatmap. The siblings *R. miraletus* and *R. ocellifera,* characterized by a bright blue eyespot at the base of each pectoral fin, displayed similar levels of expression, although some slight differences can be detected in the top-right portion of the heatmap.

The heatmap reported in Figure 11, was much more informative than expected. As a matter of fact, lower levels of DE were supposed to characterize the dorsal uniform matrix across species. *R. asterias*, which appeared to have up-regulated genes in middle-left part of the heatmap, to a greyish-marbled skin in *R. clavata* and *R. straeleni*. Surprisingly, the comparison between *R. miraletus* and *R. ocellifera* highlighted a stronger DE signal, despite their almost overlapping patterning.

Results obtained analysing the DE within ventral skin samples across species were interesting as well. Samples appeared to cluster into species as in previous comparisons, highlighting a stronger difference in expression levels between the siblings *R. miraletus* and *R. ocellifera* (Figure 12).

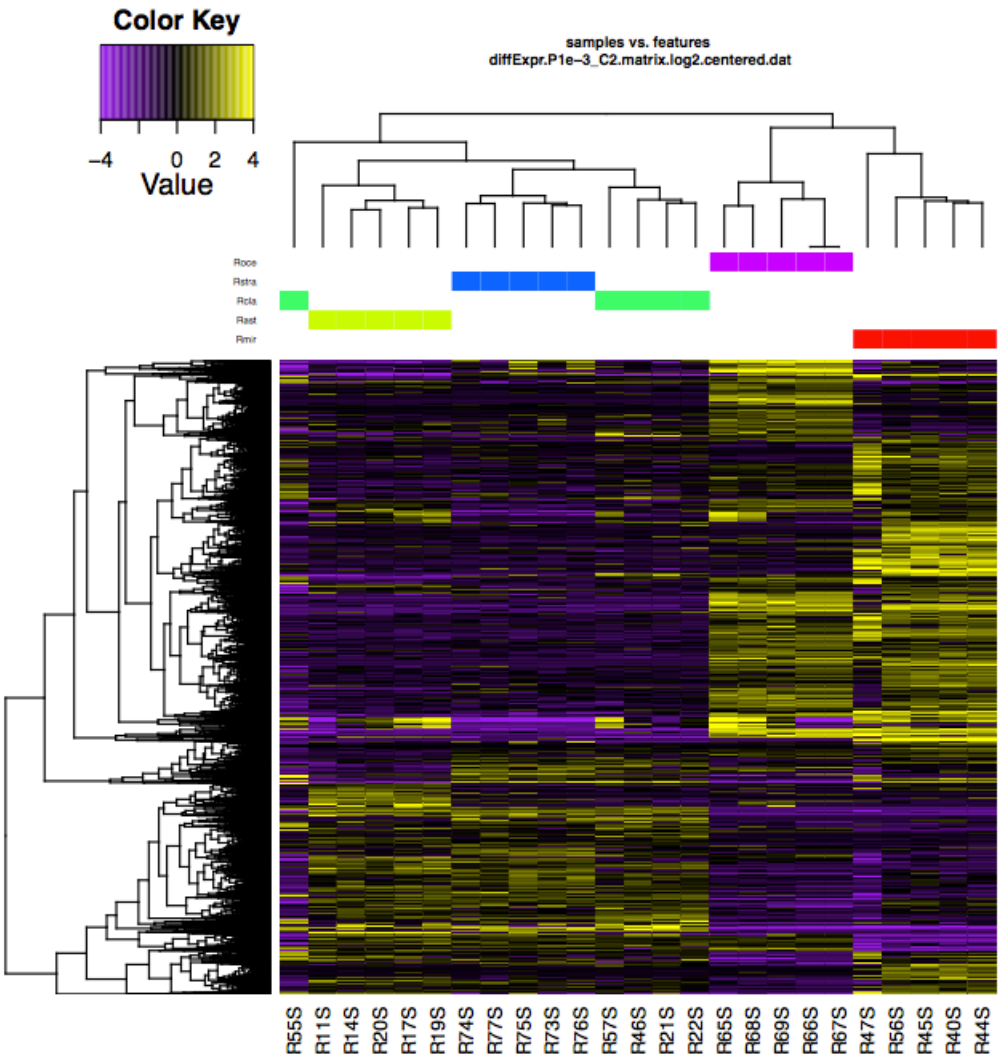


Figure 10 Heatmap representing the up-regulated (in yellow) and down-regulated genes (in purple) for dorsal skin tissues carrying an eyespot or a pseudo-eyespot across species. Their correlation is described at both individual level (clusters on the top of the map) and at the gene level (clusters on the left of the graph). RAST *Raja aterias*, RCLA *R. clavata*, RSTRA *R. straeleni*, RMIR *R. miraletus*, ROCE *R. ocellifera*.

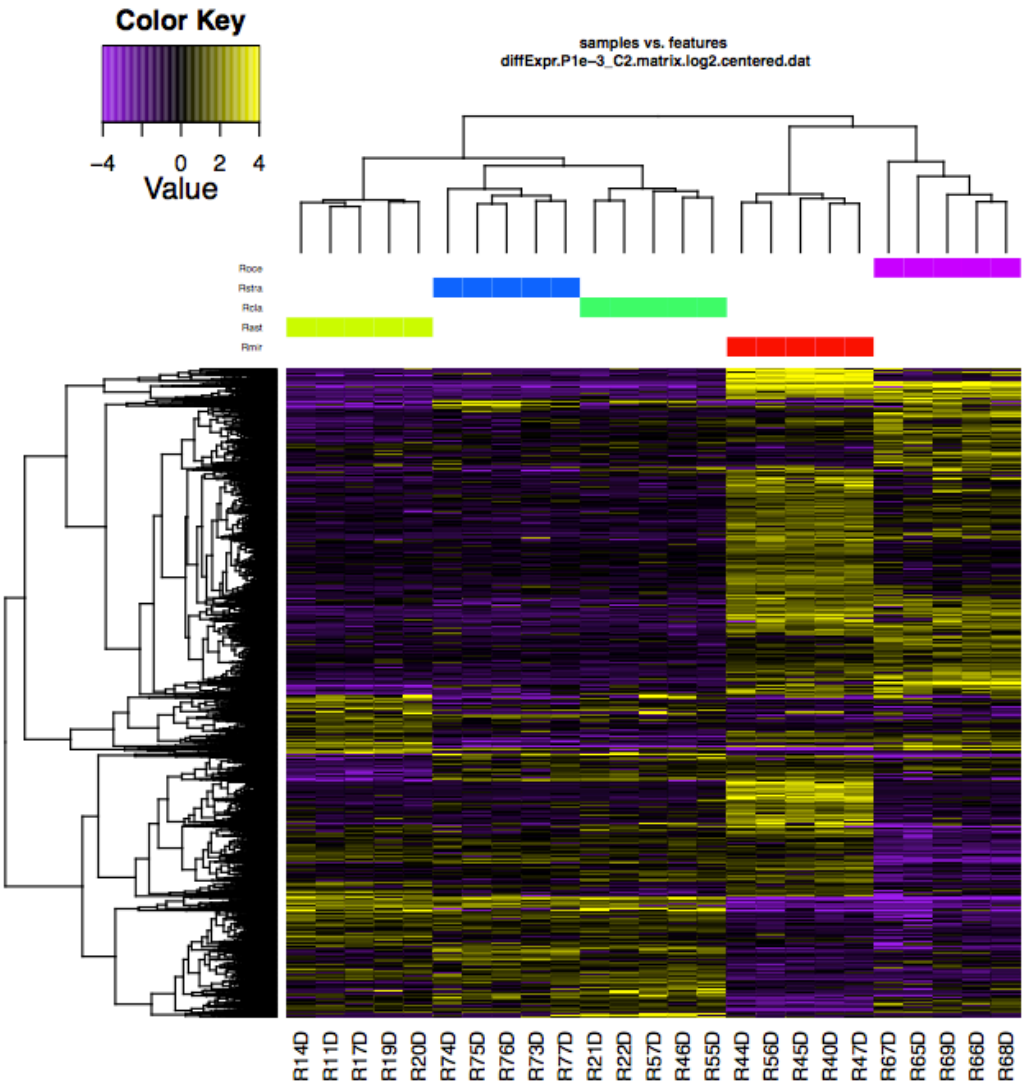


Figure 11 Heatmap representing the up-regulated (in yellow) and down-regulated genes (in purple) for dorsal skin tissues without ornaments across species. Their correlation is described at both individual level (clusters on the top of the map) and at the gene level (clusters on the left of the graph). RAST *Raja aterias*, RCLA *R. clavata*, RSTRA *R. straeleni*, RMIR *R. miraletus*, ROCE *R. ocellifera*.

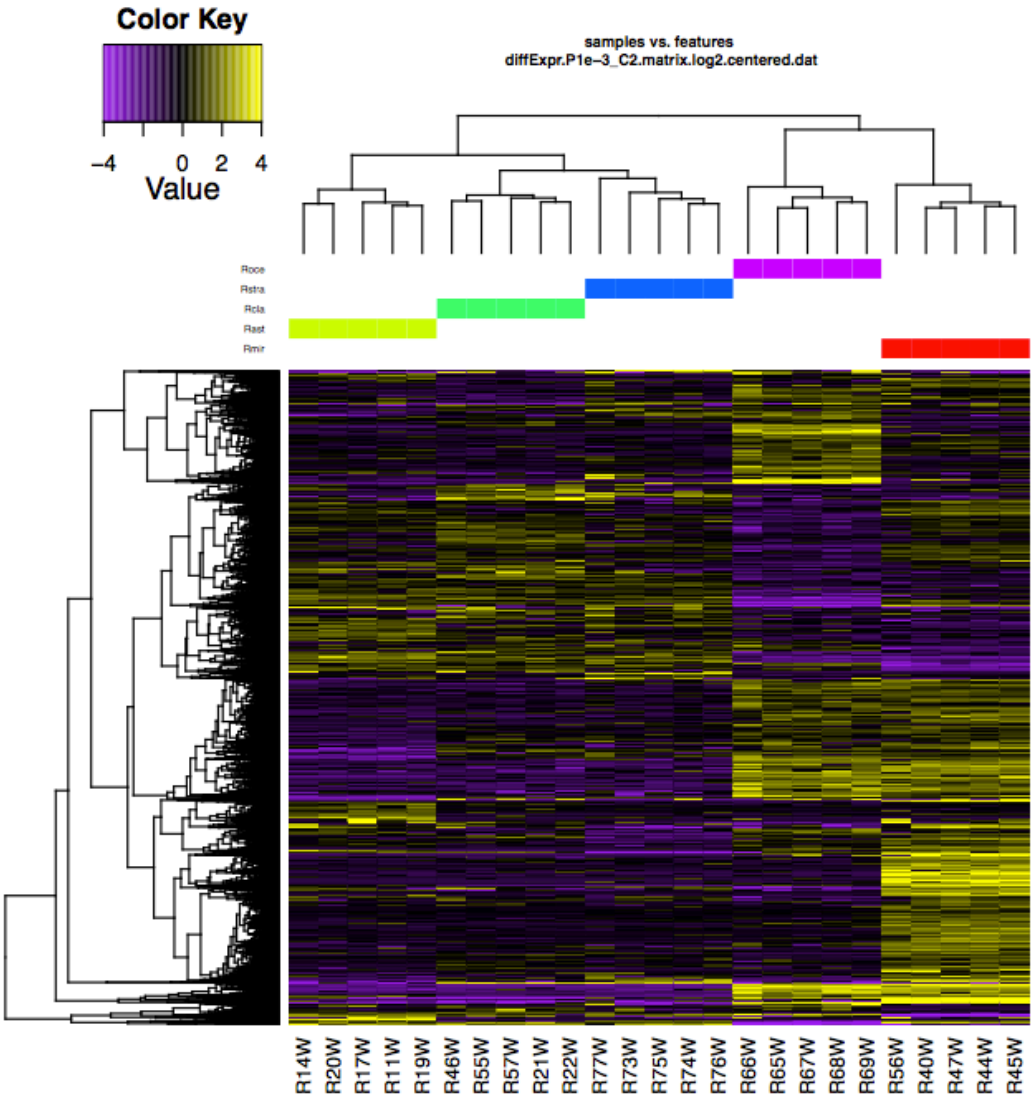


Figure 12 Heatmap representing the up-regulated (in yellow) and down-regulated genes (in purple) for vwntral white skin tissues across species. Their correlation is described at both individual level (clusters on the top of the map) and at the gene level (clusters on the left of the graph). RAST *Raja aterias*, RCLA *R. clavata*, RSTRA *R. straeleni*, RMIR *R. miraletus*, ROCE *R. ocellifera*.

In order to compare DEGs between differently pigmented tissues within the same individual a DE analysis without biological replicates was performed as implemented in GFOLD. The number of DE transcripts recovered is summarised in Tab S17 in Appendix. Among them, many of the candidate genes were related to pigmentation as reported in Tables 1 and 2.

Table 1 List of domains characterizing the DEGs between the dorsal uniform matrix (without ornament) and the white ventral body portion identified with GO. The groupings reported here were chosen according to the species clustering (dendrograms) displayed in Figure 9, 10 an 11. Functions in bold correspond to putative proteins and domains known for their involvement in pigmentary mechanism.

	Under-expressed PFAM domain	*Over-expressed* PFAM domain
Common to all species	Fasciclin domain Zinc finger, C2H2 type	T-box Homeobox domain
R. asterias *R. clavata* *R. straeleni*	Intermediate filament protein Cystine-knot domain	
R. asterias *R. clavata*	von Willebrand factor type D domain Mucin-2 protein WxxW repeating region C8 domain	
R. asterias *R. straeleni*	7 transmembrane receptor (rhodopsin family)	
R. clavata *R. straeleni*	EF-hand domain	
R. miraletus *R. ocellifera*	Kazal-type serine protease inhibitor domain Tissue inhibitor of metalloproteinase Calcium-binding EGF domain Leucine rich repeat	Thioredoxin-like Peptidase family M3
R. asterias	Sodium neurotransmitter symporter family Lectin C-type domain	C2 domain Pre-pro-megakaryocyte potentiating factor precursor (Mesothelin)
R. clavata	Inosine-uridine preferring nucleoside hydrolase Trypsin Inhibitor like cysteine rich domain	Cadherin prodomain like Galactosyltransferase Cadherin domain
Raja miraletus	Fibronectin type II domain Extracellular link domain AMP-binding enzyme Hemopexin Scramblase WIF domain Protein tyrosine kinase	Utp14 protein Pou domain - N-terminal to homeobox domain Phosphorylase superfamily DZF domain SNF2 family N-terminal domain RNA polymerase II elongation factor ELL Pentaxin family Olfactomedin-like domain THRAP3/BCLAF1 family Tetratricopeptide repeat Matrixin RIO1 family Helicase conserved C-terminal domain Molybdopterin oxidoreductase Glutathione S-transferase, C-terminal domain GDP-mannose 4,6 dehydratase
Raja ocellifera	von Willebrand factor type C domain Fz domain Complement Clr-like EGF-like von Willebrand factor type A domain Nucleoside diphosphate kinase TB domain Follistatin/Osteonectin-like EGF domain	Plectin repeat

	Laminin G domain
	Human growth factor-like EGF
	Leucine rich repeat N-terminal domain
	Caldesmon
	Sushi repeat (SCR repeat)
	Laminin EGF domain
	Fibrillar collagen C-terminal domain
	Subtilase family
Raja straeleni	'Paired box' domain
	Hyaluronidase
	Plectin repeat

Table 2 List of domains characterizing the DE genes between the dorsal ornament (eyespot and pseudo-eyespot) and the white ventral body portion. The groupings reported here were chosen upon the species clustering (dendrograms) displayed in Figure 9, 10 an 11. Functions in bold correspond to putative proteins known for their involvement in pigmentary mechanism.

	Under-expressed PFAM domain	*Over-expressed* PFAM domain
Common to all species	**Zinc finger, C2H2 type**	von Willebrand factor type D domain Mucin-2 protein WxxW repeating region C8 domain
R. asterias *R. clavata*		**Armadillo/beta-catenin-like repeat**
R. asterias *R. straeleni*		Cystine-knot domain
R. miraletus *R. ocellifera*	Dapper Fz domain Collagen triple helix repeat (20 copies) Olfactomedin-like domain Fibrinogen beta and gamma chains, C-al b Hemopexin **Endothelin family** WIF domain Leucine rich repeat Leucine rich repeats (6 copies) ADP-ribosylation factor family von Willebrand factor type C domain Zinc carboxypeptidase PAP2 superfamily CUB domain Kazal-type serine protease inhibitor Animal haem peroxidase von Willebrand factor type A domain SOUL heme-binding protein Tissue inhibitor of metalloproteinase Leucine rich repeat N-terminal domain	**Agouti protein** Peptidase family M3
R. ocellifera *R. straeleni*	**7 transmembrane receptor (rhodopsin**	
R. asterias	Sodium:neurotransmitter symporter	Pre-pro-megakaryocyte potentiating factor precursor (Mesothelin)
R. clavata	Inosine-uridine Tropomyosin Immunoglobulin I-set domain Calsequestrin Intermediate filament protein Immunoglobulin domain IQ calmodulin-binding motif Troponin EF-hand domain Nucleoside diphosphate kinase ATP guanido-phosphotransferase, C-al catalytic domain **Protein kinase domain** Ependymin ATP guanido-phosphotransferase, N-al domain	Tetratricopeptide repeat Lipoxygenase Cadherin cytoplasmic region Cadherin domain

	Myosin head (motor domain)
	EF-hand domain pair

R. miraletus	Secreted phosphoprotein 24 (Spp-24) cystatin-like domain	Utp14 protein
	EGF-like domain	Putative peptidoglycan binding domain
	2OG-Fe(II) oxygenase superfamily	Phosphorylase superfamily
	Helix-loop-helix DNA-binding domain	DZF domain
	lactate/malate dehydrogenase, NAD binding domain	SNF2 family N-terminal domain
	lactate/malate dehydrogenase, alpha/beta C-terminal domain	RNA polymerase II elongation factor ELL
	Scramblase	Pentaxin family
	Myosin N-terminal SH3-like domain	THRAP3/BCLAF1 family
	Protein phosphatase 2C	Matrixin
	zinc-finger of a C2HC-type	RIO1 family
	Homeobox domain	Molybdopterin oxidoreductase
	Adenosine/AMP deaminase	Helicase conserved C-terminal domain
	Calcium-activated chloride channel	
	VWA domain containing CoxE-like protein	
	Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase	
R. ocellifera	Fibronectin type II domain	Microtubule binding
	Putative peptidoglycan binding domain	DAN domain
	Extracellular link domain	A-macroglobulin complement component
	UNC-6/NTR/C345C module	
	Connexin	
	Complement Clr-like EGF-like	
	F5/8 type C domain	
	Hairy Orange	
	DAN domain	
	PDGF/VEGF domain	
	short chain dehydrogenase	
	Low-density lipoprotein receptor repeat class B	
	Thrombospondin type 3 repeat	
	Glutathione S-transferase, N-terminal domain	
	Laminin G domain	
	Microfibril-associated glycoprotein (MAGP)	
	Human growth factor-like EGF	
	Fibrillar collagen C-terminal domain	
	Reeler domain	
	Glutathione S-transferase, C-terminal domain	
	BTB/POZ domain	
	Cadherin domain	
	Kunitz/Bovine pancreatic trypsin inhibitor domain	
	TB domain	
	Follistatin/Osteonectin-like EGF domain	
	Myelin proteolipid protein (PLP or lipophilin)	
	Nucleotidyltransferase	
	C1q domain	
	Serpin (serine protease inhibitor)	
	Matrixin	
	Amidohydrolase family	
	Sushi repeat (SCR repeat)	
	EF hand	
	Scavenger receptor cysteine-rich domain	
	Ras family	
R. straeleni	'Paired box' domain	DAN domain
	Plectin repeat	A-macroglobulin complement component

The comparison of functions and domains between the spotted and non-spotted skin tissues revealed that many of them were more abundant among downstream genes expressed between spot and ventral portion (Figure 13). These entries were mostly related to Immunoglobulin I-set domain, Connexin, Collagen triple helix repeat, Olfactomedin-like domain, Fibrinogen beta and gamma chains, Endothelin family, Myosin N-terminal SH3-like domain, Microfibril-associated glycoprotein (MAGP), zinc-finger of a C2HC-type, Homeobox domain, Protein kinase domain, Myelin proteolipid protein (PLP or lipophilin) and Ras family.

Peculiar of the comparison between dorsal matrix and white surface were instead the AMP-binding enzyme, Hyaluronidase, Laminin EGF domain, Mucin-2 protein, Calcium-binding EGF domain and the Protein tyrosine kinase.

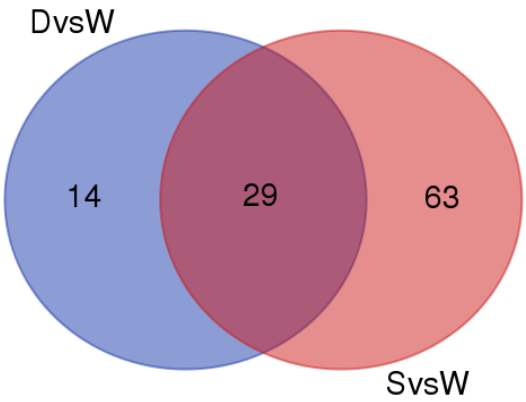


Figure 13 Venn diagram comparing domains and functions controlled by down-regulated genes in both dorsal matrix (D) and eyespots (S) across species.

Some of the functions regulated by upstream genes were the Armadillo/beta-catenin-like repeat and the Agouti protein for what concerned the spot-ventral portion, while T-Box and Homeobox were expressed only in the comparison dorsal matrix-ventral portion.

Figure 14 Venn diagram comparing functions controlled by up-regulated genes in both dorsal matrix (D) and eyespots (S) across species.

5.4. Discussion and conclusion

Understanding the molecular basis of phenotypic variation in relation to evolutionary questions is one of the most interesting challenge in evolutionary biology.

Taking advantage of the RNA-seq technology we could investigate for the first time the molecular basis underpinning peculiar pigmentation traits found in the dorsal surface of five skate species: eyespots and pseudo-eyespots. With the final aim of identifying the candidate genes involved in skates' skin pigmentation, the differences in gene expression between differently pigmented areas of both dorsal and ventral portions were estimated. The rationale at the basis of this research, was that similar pattering should share the same genetic basis and similar levels of expression.

Eyespots have been rarely studied, except for taxonomical purpose, nevertheless, they could play an interesting role in aposematism and camouflage strategies, which have been deeply studied in other organisms in relation to survival and increasing of fitness (Kavanagh et al. 2016; Nokelainen et al. 2016).

The differentially expressed (DE) genes and transcripts recovered when comparing the same skin tissue across all the species considered here demonstrated that the experiment was successful. Also, a conspicuous number of genes were DE between spot and non-spot skin tissues.

Overall, the GO analysis highlighted many molecular pathways that differed significantly between the replicates, some of which were already known for their role in melanogenesis and melanosome differentiation (i.d. Endothelin family and Agouti signalling protein). The families and domains identified are known to include processes of development of pigmentation patterns (i.e. Homeobox domain, Laminine G domain and RAS family). Eyespot or pseudo- eyespot development at the embryotic or adult stage, are supposed to rely on pigment production, activated via membrane receptor activity (Santos., 2016). These features were also retrieved from the PFAM and GO-analysis of the reference transcriptome.

The findings described here appeared close to results obtained by similar experiments conducted on teleost (Santos et al., 2016; Baldo et al. 2011), where the surplus value consisted in the identification of uncharacterized genes under positive selection. Here, the high redundancy of the transcripts identified did not allow the punctual assessment of genes potentially lineage-specific. Nevertheless, the first inspection of those DE genes between spot and non-spot tissues in comparison to the unpigmented ventral surface suggested that the strategy underlying the experiment could be considered a valid approach to identify those candidate genes likely involved in evolutionary

scenarios. For instance, further investigations could focus on the calculation of evolutionary rates of those DE genes between eyespot and pseudo-eyespots and determine if they are evolving with different rates. Considered the phenotypic stasis of the species considered, the stabilizing selection could likely be the driving force that contributed to conserve a well-adapted phenotype across the wide-ranging distribution of the clade (Williamson 1987).

References

Abercrombie, D. L., Clarke, S. C., & Shivji, M. S. (2005). Global-scale genetic identification of hammerhead sharks: Application to assessment of the international fin trade and law enforcement. *Conservation Genetics*, *6*(5), 775–788. https://doi.org/10.1007/s10592-005-9036-2

Alaeddini, R. (2012). Forensic implications of PCR inhibition—a review. *Forensic Science International: Genetics*, *6*(3), 297-305.

Allcock, A. L., Barratt, I., Eléaume, M., Linse, K., Norman, M. D., Smith, P. J., … Strugnell, J. M. (2011). Cryptic speciation and the circumpolarity debate: A case study on endemic Southern Ocean octopuses using the COI barcode of life. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *58*(1–2), 242–249. https://doi.org/10.1016/j.dsr2.2010.05.016

Allcock, A. L., & Strugnell, J. M. (2012). Southern Ocean diversity: New paradigms from molecular ecology. *Trends in Ecology and Evolution*, *27*(9), 520–528. https://doi.org/10.1016/j.tree.2012.05.009

Amadon D: The Superspecies Concept. Systematic Biology 1966, 15(3):245-249.

Amin, S., Prentis, P. J., Gilding, E. K., & Pavasovic, A. (2014). Assembly and annotation of a non-model gastropod (Nerita melanotragus) transcriptome: a comparison of de novo assemblers. *BMC research notes*, *7*(1), 488.

Anders, S. and Huber, W. Differential expression analysis for sequence count data," *Genome Biol*, vol. 11, no. 10, p. R106, 2010.

Andrews, K. R., Williams, A. J., Fernandez-Silva, I., Newman, S. J., Copus, J. M., Wakefield, C. B., … Bowen, B. W. (2016). Phylogeny of deepwater snappers (Genus Etelis) reveals a cryptic species pair in the Indo-Pacific and Pleistocene invasion of the Atlantic. *Molecular Phylogenetics and Evolution*, *100*, 361–371. https://doi.org/10.1016/j.ympev.2016.04.004

Arendt, J. & Reznick, D. 2008 Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? Trends Ecol. Evol. 23, 26–32. (doi:10.1016/ j.tree.2007.09.011)

Arlyza IS, Shen KN, Solihin DD, Soedharma D, Berrebi P, Borsa P: Species boundaries in the Himantura uarnak species complex (Myliobatiformes: Dasyatidae). Molecular Phylogenetics and Evolution 2013, 66(1):429-435.

Aschliman, N. C., Nishida, M., Miya, M., Inoue, J. G., Rosana, K. M., & Naylor, G. J. P. (2012). Body plan convergence in the evolution of skates and rays (Chondrichthyes: Batoidea). *Molecular*

Phylogenetics and Evolution, *63*(1), 28–42. https://doi.org/10.1016/j.ympev.2011.12.012

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, *215*(3), 403-410.

Avise, J. C., Nelson, W. S., Bowen, B. W., & Walker, D. (2000). Phylogeography of colonially nesting seabirds, with special reference to global matrilineal patterns in the sooty tern (Sterna fuscata). *Molecular Ecology*, *9*(11), 1783–1792. https://doi.org/10.1046/j.1365-294X.2000.01068.x

Avise, J. C., & Walker, D. (1999). Species realities and numbers in sexual vertebrates : Perspectives from an asexually transmitted genome. *Proceedings of the National Academy of Sciences*, *96*(February), 992–995.

Avise, J. C., and R. M. Ball. Principles of genealogical concordance in species concepts and biological taxonomy. *Oxford surveys in evolutionary biology* 7.7 (1990)

Bagnara, J. T., & Hadley, M. E. (1973). Chromatophores and colour change. *The comparative physiology of animal pigmentation*.

Baldanzi, S., Gouws, G., Barker, N. P., & Fratini, S. (2016). Molecular evidence of distinct evolutionary units in the sandhopper Talorchestia capensis (Amphipoda, Talitridae) along South African coasts. *Hydrobiologia*, *779*(1), 35–46. https://doi.org/10.1007/s10750-016-2797-4

Ball, R. E., Serra-Pereira, B., Ellis, J., Genner, M. J., Iglésias, S., Johnson, A. F., … Griffiths, A. M. (2016). Resolving taxonomic uncertainty in vulnerable elasmobranchs: are the Madeira skate (Raja maderensis) and the thornback ray (Raja clavata) distinct species? *Conservation Genetics*, *17*(3), 565–576. https://doi.org/10.1007/s10592-015-0806-1

Barco, A., Houart, R., Bonomolo, G., Crocetta, F., & Oliverio, M. (2013). Molecular data reveal cryptic lineages within the northeastern Atlantic and mediterranean small mussel drills of the Ocinebrina edwardsii complex (Mollusca: Gastropoda: Muricidae). *Zoological Journal of the Linnean Society*, *169*(2), 389–407. https://doi.org/10.1111/zoj.12069

Barley, A. J., White, J., Diesmos, A. C., & Brown, R. M. (2013). The challenge of species delimitation at the extremes: Diversification without morphological change in philippine sun skinks. *Evolution*, *67*(12), 3556–3572. https://doi.org/10.1111/evo.12219

Beheregaray, L. B., & Caccone, A. (2007). Cryptic biodiversity in a changing world. *Journal of Biology*, *6*(4), 9. https://doi.org/10.1186/jbiol60

Belkhir, K., Borsa, P., Chikhi, L., Goudet, J., & Bonhomme, F. (2004). GENETIX 4.05, Windows TM software for population genetics. Laboratoire génome, populations, interactions, CNRS UMR, 5000.

Beneš, V., & Muckenthaler, M. (2003). Standardization of protocols in cDNA microarray analysis. *Trends in Biochemical Sciences*, *28*(5), 244–249. https://doi.org/10.1016/S0968-0004(03)00068-9

Bermingham, E. (1997). Fish biogeography and molecular clocks: perspectives from the Panamanian Isthmus. Molecular Systematicsof Fishes, 113-128.

Bi, Y. and Davuluri, R.V. Npebseq: nonparametric empirical bayesian-based procedure for differential expression analysis of rna-seq data," *BMC bioinformatics*, vol. 14, no. 1, p. 262, 2013.

Bianchi, C.N. Biodiversity issues for the forthcoming tropical Mediterranean Sea. Hydrobiologia (2007) 580: 7. doi:10.1007/s10750-006-0469-5

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., … Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends in Ecology {&} Evolution*, *22*(3), 148–155. https://doi.org/10.1016/j.tree.2006.11.004

Bier, F. F., & Kleinjung, F. (2001). Feature-size limitations of microarray technology - A critical review. *Analytical and Bioanalytical Chemistry*, *371*(2), 151–156. https://doi.org/10.1007/S002160101003

Bigelow, H. B., & Schroeder, W. C. (1953). Fishes of the Gulf of Maine (Vol. 53, p. 588). Washington, DC: US Government Printing Office.

Birky CW, Fuerst P, Maruyama T (1989) Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. Genetics 121: 613−627

Böhne, A., Wilson, C. A., Postlethwait, J. H., & Salzburger, W. (2016). Variations on a theme: Genomics of sex determination in the cichlid fish Astatotilapia burtoni. *BMC Genomics*, *17*(1), 883. https://doi.org/10.1186/s12864-016-3178-0

Boissin, E., Stöhr, S., & Chenuil, A. (2011). Did vicariance and adaptation drive cryptic speciation and evolution of brooding in Ophioderma longicauda (Echinodermata: Ophiuroidea), a common Atlanto-Mediterranean ophiuroid? *Molecular Ecology*, 4737–4755. https://doi.org/10.1111/j.1365-294X.2011.05309.x

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, btu170.

Bonello, J., Bonnici, L., Ferrari, A., Cariani, A., & Schembri, P. J. (2016). Not all that clear cut: intraspecific morphological variability in Squalus blainville (Risso, 1827) and implications for identification of the species. *Journal of the Marine Biological Association of the United Kingdom*,

1–12. https://doi.org/10.1017/S0025315415001915

Bottaro, M., Ferrando, S., Gallus, L., Girosi, L., & Vacchi, M. (2008). First record of albinism in the deep-water shark Dalatias licha. *Marine Biodiversity Records*, *1*(2002), e10. https://doi.org/10.1017/S1755267205001156

Bowen, B. W., & Karl, S. a. (2006). Resolving Lineages and Taxonomy of Bonefi shes (Albula spp .). *Most*, (1758), 147–155.

Bowen, B. W., Nelson, W. S., & Avise, J. C. (1993). A molecular phylogeny for marine turtles : Trait mapping , rate assessment , and conservation relevance, *90*(June), 5574–5577.

Braasch, I., Brunet, F., Volff, J. N., & Schartl, M. (2010). Pigmentation Pathway Evolution after Whole-Genome Duplication in Fish. *Genome Biology and Evolution*, *1*(0), 479–493. https://doi.org/10.1093/gbe/evp050

Bradbury, I. R., Bowman, S., Borza, T., Snelgrove, P. V. R., Hutchings, J. A., Berg, P. R., … Bentzen, P. (2014). Long distance linkage disequilibrium and limited hybridization suggest cryptic speciation in Atlantic cod. *PLoS ONE*, *9*(9). https://doi.org/10.1371/journal.pone.0106380

Briggs JC: Marine Zoogeography. New York: McGraw-Hill Book Co.; 1974.

Brookfield, H., Padoch, C., Parsons, H., & Stocking, M. (2002). Cultivating biodiversity: understanding, analysing and using agricultural diversity. ITDG Publishing.

Brown, D. M., Brenneman, R. A., Koepfli, K.-P., Pollinger, J. P., Milá, B., Georgiadis, N. J., … Wayne, R. K. (2007). Extensive population genetic structure in the giraffe. *BMC Biology*, *5*, 57. https://doi.org/10.1186/1741-7007-5-57

Brown, R. M., Siler, C. D., Diesmos, A. C., & Alcala, A. C. (2009). Philippine frogs of the genus Leptobrachium (Anura; Megophryidae): phylogeny-based species delimitation, taxonomic review, and descriptions of three new species. Herpetological Monographs, 23(1), 1-44.

Bultman, S. J., Klebig, M. L., Michaud, E. J., Sweet, H. O., Davisson, M. T. & Woychik, R. P. 1994 Molecular analysis of reverse mutations from nonagouti (a)to black-and-tan (a(t)) and white-bellied agouti (Aw) reveals alternative forms of agouti transcripts. Genes Dev. 8, 481–490. (doi:10.1101/gad.8.4.481)

Burri R, Antoniazza S, Gaigher A et al. (2016) The genetic basis of color-related local adaptation in a ring-like colonization around the Mediterranean. Evolution, 70, 140–153.

Bush, G. L. (2016). Modes of Animal Speciation Author (s): Guy L . Bush Published by : Annual Reviews Stable URL : http://www.jstor.org/stable/2096835 REFERENCES Linked references are available on JSTOR for this article : MODES OF ANIMAL, *6*(1975), 339–364.

Cailliet, G. M., Musick, J. A., Simpfendorfer, C. A., & Stevens, J. D. (1995). Chapter 3 Ecology and Life History Characteristics of Chondrichthyan Fish, 12–18.

Caira, J. N., Rodriguez, N., & Pickering, M. (2013). New African Species of *Echinobothrium* (Cestoda: Diphyllidea) and Implications for the Identities of their Skate Hosts. *Journal of Parasitology*, *99*(5), 781–788. https://doi.org/10.1645/13-198.1

Camhi, M., Fowler, S., Musick, J., Bräutigam, A., & Fordham, S. (1998). *Sharks and their Relatives: Ecology and Conservation*. *IUCN/SSC Shark Specialist Group. IUCN, Gland, Switzerland and Cambridge, UK* (Vol. 3). https://doi.org/10.1016/j.wear.2011.03.022

Camperi, M., Tricas, T. C., & Brown, B. R. (2007). From morphology to neural information: The electric sense of the skate. *PLoS Computational Biology*, *3*(6), 1083–1096. https://doi.org/10.1371/journal.pcbi.0030113

Cannas R, Follesa MC, Cabiddu S, Porcu C, Salvadori S, Iglésias SP, Deiana AM, Cau A: Molecular and morphological evidence of the occurrence of the Norwegian skateDipturus nidarosiensis(Storm, 1881) in the Mediterranean Sea. Marine Biology Research 2010, 6(4):341-350.

Carvajal-Rodriguez, A., & de Uña-Alvarez, J. (2011). Assessing significance in high-throughput experiments by sequential goodness of fit and q-value estimation. PLoS One, 6(9), e24700.

Cariani, A., Messinetti, S., Ferrari, A., Arculeo, M., Bonello, J. J., Bonnici, L., … Tinti, F. (2017). Improving the Conservation of Mediterranean Chondrichthyans: The ELASMOMED DNA Barcode Reference Library. *Plos One*, *12*(1), e0170244. https://doi.org/10.1371/journal.pone.0170244

Castilho, R., Freitas, M., Silva, G., Fernandez-Carvalho, J., & Coelho, R. (2007). Morphological and mitochondrial DNA divergence validates blackmouth, Galeus melastomus, and Atlantic sawtail catsharks, Galeus atlanticus, as separate species. *Journal of Fish Biology*, *70*(SUPPL. C), 346–358. https://doi.org/10.1111/j.1095-8649.2007.01455.x

Celaj, A., Markle, J., Danska, J., & Parkinson, J. (2014). Comparison of assembly algorithms for improving rate of metatranscriptomic functional annotation. *Microbiome*, *2*(1), 39.

Chapuis, M. P., & Estoup, A. (2007). Microsatellite null alleles and estimation of population differentiation. Molecular biology and evolution, 24(3), 621-631.

Chevolot, M., Ellis, J. R., Hoarau, G., Rijnsdorp, A. D., Stam, W. T., & Olsen, J. L. (2006). Population structure of the thornback ray (Raja clavata L.) in British waters. *Journal of Sea Research*, *56*(4), 305–316. https://doi.org/10.1016/j.seares.2006.05.005

Chevolot, M., Wolfs, P. H. J., Pálsson, J., Rijnsdorp, A. D., Stam, W. T., & Olsen, J. L. (2007). Population structure and historical demography of the thorny skate (Amblyraja radiata, Rajidae) in the

North Atlantic. *Marine Biology*, *151*(4), 1275–1286. https://doi.org/10.1007/s00227-006-0556-1

Chiquillo, K. L., Ebert, D. A., Slager, C. J., & Crow, K. D. (2014). The secret of the mermaid's purse: Phylogenetic affinities within the Rajidae and the evolution of a novel reproductive strategy in skates. *Molecular Phylogenetics and Evolution*, *75*(1), 245–251. https://doi.org/10.1016/j.ympev.2014.01.012

Claes, J. M., & Mallefet, J. (2010). The lantern shark's light switch: turning shallow water crypsis into midwater camouflage. *Biology Letters*, *6*(April), 685–687. https://doi.org/10.1098/rsbl.2010.0167

Clare, E. L. (2011). Cryptic species? patterns of maternal and paternal gene flow in eight Neotropical bats. *PLoS ONE*, *6*(7). https://doi.org/10.1371/journal.pone.0021460

Claridge, M. F., Dawah, H. A., & Wilson, M. R. (1997). Species: the units of biodiversity. Chapman and Hall Ltd.

Clark, S. (2002). First report of albinism in the white-spotted bamboo shark, Chiloscyllium plagiosum (Orectolobiformes: Hemiscyllidae), with a review of reported color aberrations in elasmobranchs. *Zoo Biology*, *21*(6), 519–524. https://doi.org/10.1002/zoo.10068

Clarke, M. W., Borges, L., & Officer, R. A. (2005). Comparisons of trawl and longline catches of deepwater elasmobranchs west and north of Ireland. *Journal of Northwest Atlantic Fishery Science*, *35*(April), 429–442. https://doi.org/10.2960/J.v35.m516

Clarke, K., Yang, Y., Marsh, R., Xie, L., & Zhang, K. K. (2013). Comparative analysis of de novo transcriptome assembly. *Science China. Life sciences*, *56*(2), 156

Coates, A. G., & Obando, J. A. (1996). The geologic evolution of the Central American Isthmus. Evolution and environment in tropical America, 21-56.

Colborn, J., Crabtree, R. E., Shaklee, J. B., Pfeiler, E., & Bowen, B. W. (2001). The evolutionary enigma of bonefishes (Albula spp.): cryptic species and ancient separations in a globally distributed shorefish. *Evolution*, *55*(4), 807–820. https://doi.org/10.1111/j.0014-3820.2001.tb00816.x

Coll M, Piroddi C, Steenbeek J, Kaschner K, Ben Rais Lasram F, Aguzzi J, Ballesteros E, Bianchi CN, Corbera J, Dailianis T et al: The Biodiversity of the Mediterranean Sea: Estimates, Patterns, and Threats. PLoS ONE 2010, 5(8):e11842.

Colgan, T. J., Carolan, J. C., Bridgett, S. J., Sumner, S., Blaxter, M. L., & Brown, M. J. (2011). Polyphenism in social insects: insights from a transcriptome-wide analysis of gene expression in the life stages of the key pollinator, Bombus terrestris. *BMC genomics*, *12*(1), 623

Compagno, L.J.V., Dando, M. & Fowler, S. (2005). Collins Field Guide: Sharks of the World. Harper Collins, London.

Compagno, L. J., Ebert, D. A., & Smale, M. J. (1989). Guide to the sharks and rays of southern Africa.

Corander, J., Majander, K. K., Cheng, L., & Merilä, J. (2013). High degree of cryptic population differentiation in the Baltic Sea herring Clupea harengus. *Molecular Ecology*, *22*(11), 2931–2940. https://doi.org/10.1111/mec.12174

Corrigan, S., Huveneers, C., Schwartz, T. S., Harcourt, R. G., & Beheregaray, L. B. (2008). Genetic and reproductive evidence for two species of ornate wobbegong shark Orectolobus spp. on the Australian east coast. *Journal of Fish Biology*, *73*(7), 1662–1675. https://doi.org/10.1111/j.1095-8649.2008.02039.x

Costa, F. O., Landi, M., Martins, R., Costa, M. H., Costa, M. E., Carneiro, M., … Carvalho, G. R. (2012). A Ranking System for Reference Libraries of DNA Barcodes: Application to Marine Fish Species from Portugal. *PLoS ONE*, *7*(4), e35858. https://doi.org/10.1371/journal.pone.0035858

Cracraft, J. (1983). Species concepts and speciation analysis. Current ornithology, 1(4), 159-187.

Cracraft, J. (1989). Speciation and its ontology: the empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. Speciation and its Consequences, 28-59.

Craig, M. T., Graham, R. T., Torres, R. A., Hyde, J. R., Freitas, M. O., Ferreira, B. P., … Robertson, D. R. (2009). How many species of goliath grouper are there? Cryptic genetic divergence in a threatened marine fish and the resurrection of a geopolitical species. *Endangered Species Research*, *7*(3), 167–174. https://doi.org/10.3354/esr00117

Crozier, R. H., Dunnett, L. J., & Agapow, P. M. (2005). Phylogenetic biodiversity assessment based on systematic nomenclature. *Evol. Bioinform. Online*, *1*, 11–36.

Dalebout, M. L., Mead, J. G., Baker, C. S., Baker, A. N., & Van Helden, A. L. (2002). A new species of beaked whale Mesoplodon perrini sp. n. (Cetacea: Ziphiidae) discovered through phylogenetic analyses of mitochondrial DNA sequences. *Marine Mammal Science*, *18*(3), 577–608. https://doi.org/10.1111/j.1748-7692.2002.tb01061.x

Daugherty, C. H., Cree, A., Hay, J. M. and Thompson, M. B. (1990). Neglected taxonomy and continuing extinction of the tuatara (Sphenodon). *Nature*, *347*, 177–179.

Demski, L.S. 1990a. Neuroendocrine mechanism controlling the sexual development and behavior of sharks and rays. J. Aquariculture and Aquatic Sci. 5: 53–67.

De Queiroz, K. De. (2007). Species concepts and species delimitation. *Systematic Botany*, *56*(6), 879–

886. https://doi.org/10.1080/10635150701701083

Dessinioti C, Antoniou C, Katsambas A, Stratigos AJ (2011) Melanocortin 1 receptor variants: functional role and pigmentary associations. Photochemistry and Photobiology, 87, 978–987.

de Vargas, C., Norris, R., Zaninetti, L., Gibb, S. W., & Pawlowski, J. (1999). Molecular evidence of cryptic speciation in planktonic foraminifers and their relation to oceanic provinces. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(6), 2864–2868. https://doi.org/10.1073/pnas.96.6.2864

Dingle, R. V., & Lavelle, M. (2000). Antarctic Peninsula Late Cretaceous-Early Cenozoic pal┐oenvironments and Gondwana pal┐ogeographies. *Journal of African Earth Sciences*, *31*, 91–105. https://doi.org/10.1016/S0899-5362(00)00075-0

Dornburg, A., Federman, S., Eytan, R. I., & Near, T. J. (2016). Cryptic species diversity in sub-Antarctic islands: A case study of Lepidonotothen. *Molecular Phylogenetics and Evolution*, *104*, 32–43. https://doi.org/10.1016/j.ympev.2016.07.013

Dray, S., and Dufour A.B. The ade4 package: implementing the duality diagram for ecologists.Journal of statistical software 22.4 (2007): 1-20.

Dudgeon, C. L., Blower, D. C., Broderick, D., Giles, J. L., Holmes, B. J., Kashiwagi, T., … Ovenden, J. R. (2012). A review of the application of molecular genetics for fisheries management and conservation of sharks and rays. *Journal of Fish Biology*, *80*(5), 1789–1843. https://doi.org/10.1111/j.1095-8649.2012.03265.x

Duftner, N., Sefc, K. M., Koblmüller, S., Nevado, B., Verheyen, E., Phiri, H., & Sturmbauer, C. (2006). Distinct population structure in a phenotypically homogeneous rock-dwelling cichlid fish from Lake Tanganyika. *Molecular Ecology*, *15*(9), 2381–2395. https://doi.org/10.1111/j.1365-294X.2006.02949.x

Dulvy NK, Metcalfe JD, Glanville J, Pawson MG, Reynolds JD: Fishery stability, local extinctions, and shifts in community structure in skates. Conservation Biology 2000, 14(1):283-293.

Dulvy NK, Reynolds JD: BIODIVERSITY Skates on thin ice. Nature 2009, 462(7272):417-417.

Dulvy NK, Sadovy Y, Reynolds JD: Extinction vulnerability in marine populations. Fish and Fisheries 2003, 4(1):25-64.

Duncan, K. M., Martin, A. P., Bowen, B. W., & De Couet, H. G. (2006). Global phylogeography of the scalloped hammerhead shark (Sphyrna lewini). *Molecular Ecology*, *15*(8), 2239–2251. https://doi.org/10.1111/j.1365-294X.2006.02933.x

Earl, D. A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation genetics resources, 4(2), 359-361.

Eckhart, L., Bach, J., Ban, J., & Tschachler, E. (2000). Melanin binds reversibly to thermostable DNA polymerase and inhibits its activity. *Biochemical and biophysical research communications*, *271*(3), 726-730.

Eitner, B. J. (1995). Systematics of the genus Alopias (Lamniformes: Alopiidae) with evidence for the existence of an unrecognized species. Copeia, 562-571.

Eldredge, N., & Stanley, S. M. (Eds.). (2012). Living fossils. Springer Science & Business Media.

El Nagar, A., McHugh, M., Rapp, T., Sims, D. W., & Genner, M. J. (2010). Characterisation of polymorphic microsatellite markers for skates (Elasmobranchii: Rajidae) from expressed sequence tags. *Conservation Genetics*, *11*(3), 1203–1206. https://doi.org/10.1007/s10592-009-9919-8

Elmer, K. R., Dávila, J. a, & Lougheed, S. C. (2007). Cryptic diversity and deep divergence in an upper Amazonian leaflitter frog, Eleutherodactylus ockendeni. *BMC Evolutionary Biology*, *7*, 247. https://doi.org/10.1186/1471-2148-7-247

Erdmann, M. V, Caldwell, R. L., Jewett, S. L., & Tjakrawidjaja, A. (1999). The second recorded living coelocanth from north Sulawesi. *Environmental Biology of Fishes*, *54*(175), 445–451. Retrieved from d:%5CJournal PDFs%5CEnvironmental Biology of Fishes PDFs%5CEnvBiolFish54-445-451.pdf

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14, 2611–2620

Falush, D., Stephens, M., & Pritchard, J. K. (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. Molecular ecology notes, 7(4), 574-578.

Feiner, N. (2016). Accumulation of transposable elements in Hox gene clusters during adaptive radiation of Anolis lizards. *Proceedings of the Royal Society of London B: Biological Sciences*, *283*(1840).

Felsenstein, J. (1989). PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistic*, *5*, 164–166.

Felsenstein, J. (2005). PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Feng, J., Meyer, C. A., Wang, Q., Liu, J. S., Liu, X.S., Zhang, Y. (2012) GFOLD: a generalized fold change for ranking differentially expressed genes from RNA-seq data. *Bioinformatics* 2012

Fernandez-Silva, I., Randall, J. E., Coleman, R. R., Dibattista, J. D., Rocha, L. A., Reimer, J. D., … Bowen,

B. W. (2015). Yellow tails in the Red Sea: Phylogeography of the Indo-Pacific goatfish Mulloidichthys flavolineatus reveals isolation in peripheral provinces and cryptic evolutionary lineages. *Journal of Biogeography*, *42*(12), 2402–2413. https://doi.org/10.1111/jbi.12598

Feulner, P. G. D., Plath, M., Engelmann, J., Kirschbaum, F., & Tiedemann, R. (2009). Electrifying love: electric fish use species-specific discharge for mate recognition. *Biology Letters*, *5*(2), 225–8. https://doi.org/10.1098/rsbl.2008.0566

Froese, R., & Pauly, D. (2010). Fishbase World Wide Web electronic publications. Available in: www. fishbase. org. Accessed, 27.

Froschauer, A., Braasch, I., & Volff, J. N. (2006). Fish genomes, comparative genomics and vertebrate evolution. Current Genomics, 7(1), 43-57.

Frable, B. W., Wolfe Wagman, D., Frierson, T. N., Aguilar, A., & Sidlauskas, B. L. (2015). A new species of Sebastes (Scorpaeniformes: Sebastidae) from the northeastern Pacific, with a redescription of the blue rockfish, S. mystinus (Jordan and Gilbert, 1881). *Fish. Bull*, *113*, 355–377. https://doi.org/10.7755/FB.113.4.1

Fraser-Brunner, A. (1950). A synopsis of the hammerhead sharks (Sphyrna), with description of a new species. Records of the Australian Museum, 22(3), 213-219.

Freeland, J. R., & Boag, P. T. (2016). The Mitochondrial and Nuclear Genetic Homogeneity of the Phenotypically Diverse Darwin ' s Ground Finches Author (s): Joanna R . Freeland and Peter T . Boag Published by : Society for the Study of Evolution Stable URL : http://www.jstor.org/stable/26409, *53*(5), 1553–1563.

Friesen, V. L., Piatt, J. F., & Baker, A. J. (1996). Evidence from cytochrome b sequences and allozymes for a'new'species of alcid: the long-billed murrelet (Brachyramphus perdix). Condor, 681-690.

Frisk, M. G., Jordaan, A., & Miller, T. J. (2014). Moving beyond the current paradigm in marine population connectivity: Are adults the missing link? *Fish and Fisheries*, *15*(2), 242–254. https://doi.org/10.1111/faf.12014

Frodella, N., Cannas, R., Velonà, A., Carbonara, P., Farrell, E. D., Fiorentino, F., … Cariani, A. (2016). Population connectivity and phylogeography of the Mediterranean endemic skate Raja polystigma and evidence of its hybridization with the parapatric sibling R. Montagui. *Marine Ecology Progress Series*, *554*, 99–113. https://doi.org/10.3354/meps11799

Froschauer A, Braasch I, Volff JN. 2006. Fish genomes, comparative genomics and vertebrate evolution. Curr Genomics. 7:43–57.

Funk, D. J., & Omland, K. E. (2003). Frequency, Causes, and Consequences, with Insights from Animal

Mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics*, *34*(1), 397–423. https://doi.org/10.1146/annurev.ecolsys.34.011802.132421

Funk, W. C., Caminer, M., & Ron, S. R. (2012). High levels of cryptic species diversity uncovered in Amazonian frogs. *Proceedings. Biological Sciences / The Royal Society*, *279*(1734), 1806–14. https://doi.org/10.1098/rspb.2011.1653

Galdino Brandão, J. H. S., de Araújo Bitencourt, J., Santos, F. B., Watanabe, L. A., Schneider, H., Sampaio, I., & Affonso, P. R. A. de M. (2016). DNA barcoding of coastal ichthyofauna from Bahia, northeastern Brazil, South Atlantic: High efficiency for systematics and identification of cryptic diversity. *Biochemical Systematics and Ecology*, *65*, 214–224. https://doi.org/10.1016/j.bse.2016.02.012

Garcia-Rodriguez, A. I., Bowen, B. W., D, D., A, M. A., M, M., A, M.-O. R., … M, M. P. (1998). Phylogeography of the West Indian manate.pdf. *Molecular Ecology*, *7*, 1137–1149.

Gardner, M. G., & Ward, R. D. (2002). Taxonomic affinities within Australian and New Zealand Mustelus sharks (Chondrichthyes: Triakidae) inferred from allozymes, mitochondrial DNA and precaudal vertebrae counts. Copeia, 2002(2), 356-363.

Gatto A, Torroja-Fungairiñ˜o C, Mazzarotto F, et al. Fine- Splice, enhanced splice junction detection and quantification: a novel pipeline based on the assessment of diverse RNA-Seq alignment solutions. Nucleic Acids Res 2014;42:e71.

Gelsleichter, J. 2004 Hormonal regulation of elasmobranch physiology. In Biology of sharks and their relatives (eds J. Carrier, J. Musick & M. Heithaus), pp. 287–323. Boca Raton, FL: CRC Press.

Gilbert, C. (1967). A taxonomic synopsis of the hammerhead sharks (Family Sphyrnidae). Sharks, skates and rays, 69-78.

Gittemnerger, E. (1991). What about non adaptive radiation? *Biological Journal of the Linnean Society*, *43*(4), 263–272. https://doi.org/10.1111/j.1095-8312.1991.tb00598.x

Goda, M., & Fujii, R. (1995). Blue Chromatophores in Two Species of Callionymid Fish. *Zoological Science*. https://doi.org/10.2108/zsj.12.811

Goudet, J. (1995). FSTAT (version 1.2): a computer program to calculate F-statistics. Journal of heredity, 86(6), 485-486.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–52. https://doi.org/10.1038/nbt.1883

Griffiths, A. M., Sims, D. W., Cotterell, S. P., El Nagar, A., Ellis, J. R., Lynghammar, A., … Genner, M. J.

(2010). Molecular markers reveal spatially segregated cryptic species in a critically endangered fish, the common skate (Dipturus batis). *Proceedings. Biological Sciences / The Royal Society*, *277*(1687), 1497–1503. https://doi.org/10.1098/rspb.2009.2111

Grosemans, T., Morris, K., Thomas, W. K., Rigaux, A., Moens, T., & Derycke, S. (2016). Mitogenomics reveals high synteny and long evolutionary histories of sympatric cryptic nematode species. *Ecology and Evolution*, *6*(6), 1854–1870. https://doi.org/10.1002/ece3.1975

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., … Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, *8*(8), 1494–1512. https://doi.org/10.1038/nprot.2013.084

Hansen, K. D., Brenner, S. E., & Dudoit, S. (2010). Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic Acids Research*, *38*(12), 1–7. https://doi.org/10.1093/nar/gkq224

Hashimoto, T., De Hoon, M. J., Grimmond, S. M., Daub, C. O., Hayashizaki, Y., & Faulkner, G. J. (2009). Probabilistic resolution of multi-mapping reads in massively parallel sequencing data using MuMRescueLite. *Bioinformatics*, *25*(19), 2613-2614.

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H., & Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(41), 14812–14817. https://doi.org/10.1073/pnas.0406166101

Hebert, P. D. N., Ratnasingham, S., & Jeremy, R. (2003). Barcoding animal life : cytochrome c oxidase subunit 1 divergences among closely related species, (figure 1), 96–99. https://doi.org/10.1098/rsbl.2003.0025

Heemstra, P. C. (1997). A review of the smooth-hound sharks (genus Mustelus, family Triakidae) of the western Atlantic Ocean, with descriptions of two new species and a new subspecies. *Bulletin of Marine Science*, *60*(3), 894–928.

Henriques, R., Potts, W. M., Sauer, W. H. H., & Shaw, P. W. (2015). Incipient genetic isolation of a temperate migratory coastal sciaenid fish (Argyrosomus inodorus) within the Benguela Cold Current system. *Marine Biology Research*, *11*(January 2015), 423–429. https://doi.org/10.1080/17451000.2014.952309

Henriques, R., Potts, W. M., Sauer, W. H., Santos, C. V., Kruger, J., Thomas, J. A., & Shaw, P. W. (2016). Molecular genetic, life-history and morphological variation in a coastal warm-temperate

sciaenid fish: evidence for an upwelling-driven speciation event. *Journal of Biogeography*, *43*(9), 1820–1831. https://doi.org/10.1111/jbi.12829

Hoare, D. (2009). First record of albinism in the Lesser-spotted Dogfish (Scyliorhinus canicula (L., 1758))(Chondrichthyes: Scyliorhinidae) and a review of albinism in other elasmobranchs. The Irish Naturalists' Journal, 115-118.

Hodge, J. R., & Bellwood, D. R. (2016). The geography of speciation in coral reef fishes: the relative importance of biogeographical barriers in separating sister-species. *Journal of Biogeography*, *43*(7), 1324–1335. https://doi.org/10.1111/jbi.12729

Horne, J. B., van Herwerden, L., Choat, J. H., & Robertson, D. R. (2008). High population connectivity across the Indo-Pacific: Congruent lack of phylogeographic structure in three reef fish congeners. *Molecular Phylogenetics and Evolution*, *49*(2), 629–638. https://doi.org/10.1016/j.ympev.2008.08.023

Hornett, E. a, & Wheat, C. W. (2012). Quantitative RNA-Seq analysis in non-model species: assessing transcriptome assemblies as a scaffold and the utility of evolutionary divergent genomic reference species. *BMC Genomics*, *13*, 361. https://doi.org/10.1186/1471-2164-13-361

Huelsken, T., Keyse, J., Liggins, L., Penny, S., Treml, E. A., & Riginos, C. (2013). A Novel Widespread Cryptic Species and Phylogeographic Patterns within Several Giant Clam Species (Cardiidae : Tridacna) from the Indo-Pacific Ocean. *PLoS ONE*, *8*(11). https://doi.org/10.1371/journal.pone.0080858

Hull, D. (1965), "The Effect of Essentialism on Taxonomy: 2000 Years of Stasis," British Journalfor the Philosophy of Science 15: 314-16; 16: 1-18.

Hunter, E., Buckley, a. a., Stewart, C., & Metcalfe, J. D. (2005a). Migratory behaviour of the thornback ray, Raja clavata , in the southern North Sea. *Journal of the Marine Biological Association of the UK*, *85*, 1095. https://doi.org/10.1017/S0025315405012142

Hunter, E., Buckley, a. a., Stewart, C., & Metcalfe, J. D. (2005b). Repeated seasonal migration by a thornback ray in the southern North Sea. *Journal of the Marine Biological Association of the UK*, *85*(January 2000), 1199. https://doi.org/10.1017/S0025315405012300

Iglésias, S. P., Toulhoat, L., & Sellos, D. Y. (2010). Taxonomic confusion and market mislabelling of threatened skates: important consequences for their conservation status. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *20*(3), 319–333. https://doi.org/10.1002/aqc.1083

Irvine, S. B., Stevens, J. D., & Laurenson, L. J. B. (2006). Comparing external and internal dorsal-spine bands to interpret the age and growth of the giant lantern shark, Etmopterus baxteri

(Squaliformes: Etmopteridae). *Environmental Biology of Fishes*, *77*(3–4), 253–264. https://doi.org/10.1007/s10641-006-9130-4

Ishihara, H., Homma, K., & Nakamura, R. (2001). The occurrence of albinism in individuals of the manta ray and Japanese common skate found in the western Pacific. IOP Diving News, 12(7), 6-20.

Ivory, S. J., Blome, M. W., King, J. W., McGlue, M. M., Cole, J. E., & Cohen, A. S. (2016). Environmental change explains cichlid adaptive radiation at Lake Malawi over the past 1.2 million years. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(42), 11895–11900. https://doi.org/10.1073/pnas.1611028113

Jakobsdóttir, K. B. (2001). Biological aspects of two deep-water squalid sharks: Centroscyllium fabricii (Reinhardt, 1825) and Etmopterus princeps (Collett, 1904) in Icelandic waters. *Fisheries Research*, *51*(2–3), 247–265. https://doi.org/10.1016/S0165-7836(01)00250-8

Jakobsson, M., and Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23, 1801–1806.

Jockusch, E. L., & Wake, D. B. (2002). Falling apart and merging: Diversification of slender salamanders (Plethodontidae: Batrachoseps) in the American West. *Biological Journal of the Linnean Society*, *76*(3), 361–391. https://doi.org/10.1046/j.1095-8312.2002.00071.x

Johnson, N. D., Sanders, C., Maiorova, A., & Schulze, A. (2016). Cryptic species in Pacific sipunculans (Sipuncula: Phascolosomatidae): east-west divergence between non-sister taxa. *Zoologica Scripta*, *45*(4), 455–463. https://doi.org/10.1111/zsc.12158

Jörger, K. M., Norenburg, J. L., Wilson, N. G., & Schrödl, M. (2012). Barcoding against a paradox? Combined molecular species delineations reveal multiple cryptic lineages in elusive meiofaunal sea slugs. *BMC Evolutionary Biology*, *12*, 245. https://doi.org/10.1186/1471-2148-12-245

Kadri, H., Marouani, S., Bradai, M. N., & Bouaïn, A. (2014). Food habits of the brown ray Raja miraletus (Chondrichthyes: Rajidae) from the Gulf of Gabès (Tunisia). *Marine Biology Research*, *10*(4), 426–434. https://doi.org/10.1080/17451000.2013.797583

Kalleberg, K., Nip, J., & Gossage, K. (2015). Multispectral imaging as a tool for melanin detection. *Journal of Histotechnology*, *38*(1), 14-21.

Kasumyan, A.O. and Marusov, E.A.,Behavioral Responses to Alimentary Signals in the Carp Cyprinus carpio in the Norm and after Chronic Anosmia, J. Ichthyology, 2005b, vol. 45, supl. 2, pp. S315–S323.

Kawahara-Miki, R., Sano, S., Nunome, M., Shimmura, T., Kuwayama, T., Takahashi, S., ... & Kono, T. (2013). Next-generation sequencing reveals genomic features in the Japanese quail. *Genomics*, *101*(6), 345-353.

Kavanagh, P. H., Shaw, R. C., & Burns, K. C. (2016). Potential aposematism in an insular tree species: are signals dishonest early in ontogeny?. *Biological Journal of the Linnean Society*, *118*(4), 951-958.

Kawauchi, G. Y., & Giribet, G. (2014). Sipunculus nudus Linnaeus, 1766 (Sipuncula): Cosmopolitan or a group of pseudo-cryptic species? An integrated molecular and morphological approach. *Marine Ecology*, *35*(4), 478–491. https://doi.org/10.1111/maec.12104

Kemp, N. E. (1999). Integumentary system and teeth. *Sharks, Skates and Rays. The Biology of Elasmobranch Fishes*, 43-68

Klimley, A. P. (2013). The biology of sharks and rays. University of Chicago Press.

Knowlton, N. (2000). Molecular genetic analyses of species boundaries in the sea. *Hydrobiologia*, *420*, 73–90.

Knowlton, N. (2009). Sibling Species in the Sea. *Source : Annual Review of Ecology and Systematics*, *24*(1993), 189–216.

Knowlton, N. (1986). Cryptic and sibling species among the decapod Crustacea. Journal of crustacean biology, 356-363.

Knudsen, S. W., & Clements, K. D. (2013). Kyphosus gladius, a new species of sea chub from Western Australia (Teleostei: Kyphosidae), with comments on Segutilum klunzingeri Whitley. Zootaxa, 3599(1), 1-18.

Koblmüller, S., Salzburger, W., Oberm??ller, B., Eigner, E., Sturmbauer, C., & Sefc, K. M. (2011). Separated by sand, fused by dropping water: Habitat barriers and fluctuating water levels steer the evolution of rock-dwelling cichlid populations in Lake Tanganyika. *Molecular Ecology*, *20*(11), 2272–2290. https://doi.org/10.1111/j.1365-294X.2011.05088.x

Kousteni, V., Kasapidis, P., Kotoulas, G., & Megalofonou, P. (2015). Strong population genetic structure and contrasting demographic histories for the small-spotted catshark (Scyliorhinus canicula) in the Mediterranean Sea. *Heredity*, *114*(3), 333–43. https://doi.org/10.1038/hdy.2014.107

Kwan, J. C., Tianero, M. D. B., Donia, M. S., Wyche, T. P., Bugni, T. S., & Schmidt, E. W. (2014). Host control of symbiont natural product chemistry in cryptic populations of the tunicate Lissoclinum patella. *PLoS ONE*, *9*(5). https://doi.org/10.1371/journal.pone.0095850

Lacroix V (2008) Esact transcriptome reconstruction from short sequence reads, Springer. WAB08: Proceedings of the 8th international workshop on Algorithms in Binformatics edition, pp. 500–563.

Landi, M., Dimech, M., Arculeo, M., Biondo, G., Martins, R., Carneiro, M., … Costa, F. O. (2014). DNA barcoding for species assignment: The case of Mediterranean marine fishes. *PLoS ONE*, *9*(9). https://doi.org/10.1371/journal.pone.0106135

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods*, *9*(4), 357-359.

Last, P. R., & Séret, B. (2016). A new Eastern Central Atlantic skate Raja parva sp. nov. (Rajoidei: Rajidae) belonging to the Raja miraletus species complex. *Zootaxa*, *4147*(4), 477–489. https://doi.org/10.11646/zootaxa.4147.4.8

Leliaert, F., Verbruggen, H., Vanormelingen, P., Steen, F., López-Bautista, J. M., Zuccarello, G. C., & De Clerck, O. (2014). DNA-based species delimitation in algae. *European Journal of Phycology*, *49*(2), 179–196. https://doi.org/10.1080/09670262.2014.904524

Leclerc, E. A., Huchenq, A., Mattiuzzo, N. R., Metzger, D., Chambon, P., Ghyselinck, N. B., ... & Guerrin, M. (2009). Corneodesmosin gene ablation induces lethal skin-barrier disruption and hair-follicle degeneration related to desmosome dysfunction. *Journal of cell science*, *122*(15), 2699-2709

Li, J., Jiang, H. and Wong, W. “Method modeling non-uniformity in short-read rates in rna-seq data,” *Genome Biol*, vol. 11, no. 5, p. R25, 2010.

Li B, Dewey C. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 2011;12:323.

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., ... & Law, M. (2012). Comparison of next-generation sequencing systems. *BioMed Research International*, *2012*.

Losos, J. B., & Glor, R. E. (2003). Phylogenetic comparative methods and the geography of speciation. *Trends in Ecology and Evolution*, *18*(5), 220–227. https://doi.org/10.1016/S0169-5347(03)00037-5

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, *15*(12), 550.

Lu, B. X., Zeng, Z. B., & Shi, T. L. (2013). Comparative study of de novo assembly and genome-guided assembly strategies for transcriptome reconstruction based on RNA-Seq. *Science China Life Sciences*, *56*(2), 143–155. https://doi.org/10.1007/s11427-013-4442-z

Lundstrom, H. M., & Bard, P. (1932). Hypophysial control of cutaneous pigmentation in an elasmobranch fish. *The Biological Bulletin*, *62*(1), 1-9.

Machado, H. E., Pollen, A. A., Hofmann, H. A., & Renn, S. C. (2009). Interspecific profiling of gene expression informed by comparative genomic hybridization: a review and a novel approach in African cichlid fishes. *Integrative and comparative biology*, icp080.

Maisey, J. G. (2012). What is an "elasmobranch"? The impact of palaeontology in understanding elasmobranch phylogeny and evolution. *Journal of Fish Biology*, *80*(5), 918–951. https://doi.org/10.1111/j.1095-8649.2012.03245.x

Malinsky, M., & Salzburger, W. (2016). Environmental context for understanding the iconic adaptive radiation of cichlid fishes in Lake Malawi. *Proceedings of the National Academy of Sciences*, *113*(42), 201614272. https://doi.org/10.1073/pnas.1614272113

Manceau, M., Domingues, V. S., Linnen, C. R., Rosenblum, E. B., & Hoekstra, H. E. (2010). Convergence in pigmentation at multiple levels: mutations, genes and function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1552), 2439–2450. https://doi.org/10.1098/rstb.2010.0104

Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics. *Trends in genetics*, *24*(3), 133-141.

Mark O'Loughlin, P., Paulay, G., Davey, N., & Michonneau, F. (2011). The Antarctic region as a marine biodiversity hotspot for echinoderms: Diversity and diversification of sea cucumbers. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *58*(1–2), 264–275. https://doi.org/10.1016/j.dsr2.2010.10.011

Marshall, W. S. (1978). On the involvement of mucous secretion in teleost osmoregulation. Canadian Journal of Zoology, 56(5), 1088-1091.

May, R. M. (1988). How Many Species Are There on Earth? *Science*, *241*(4872), 1441–1449. https://doi.org/10.1126/science.241.4872.1441

Mayr, E. (1963). Animal species and evolution. *The Eugenics Review*. https://doi.org/10.1016/0169-5347(94)90187-2

Mayr, E. (1942). Systematics and the origin of species, from the viewpoint of a zoologist. Harvard University Press.

Mceachran, J. D., Dunn, K. A., & Mceachran, J. D. (1998). Phylogenetic Analysis of Skates, a Morphologically Conservative Clade of Elasmobranchs (Chondrichthyes: Rajidae). *COPEIA*, *2*, 271–290.

Mceachran, J. D., Seret, B., & Miyake, T. (1989). Morphological Variation within Raja miraletus and Status of R . ocellifera (Chondrichthyes , Published by : American Society of Ichthyologists and Herpetologists (ASIH) Stable URL : http://www.jstor.org/stable/1445490 Accessed : 04-03-2016 10 : 51 UTC, *3*, 629–641.

McEachran, J. D. (1977). Variation in Raja garmani and the status of Raja lentigenosa (Pisces: Rajidae). Bulletin of Marine Science 27, 423–439.

McEachran, J. D. (1984). Anatomical investigations of the New Zealand skates Bathyraja asperula and B. spinifera, with an evaluation of their classification within the Rajoidei (Chondrichthyes). Copeia, 45-58.

Medina, I., Losos, J. B., & Mahler, D. L. (2016). Evolution of dorsal pattern variation in Greater Antillean Anolis lizards. *Biological Journal of the Linnean Society*. https://doi.org/10.1111/bij.12881

Meiri, S., & Mace, G. M. (2007). New taxonomy and the origin of species. *PLoS Biology*, *5*(7). https://doi.org/10.1371/journal.pbio.0050194

Meyer, W., & Seegers, U. (2012). Basics of skin structure and function in elasmobranchs: A review. *Journal of Fish Biology*, *80*(5), 1940–1967. https://doi.org/10.1111/j.1095-8649.2011.03207.x

Moore, W. S. (2016). Inferring Phylogenies from mtDNA Variation : Mitochondrial-Gene Trees Versus Nuclear- Gene Trees Author (s): William S . Moore Published by : Society for the Study of Evolution Stable URL : http://www.jstor.org/stable/2410325 JSTOR is a not-for-profit s, *49*(4), 718–726.

Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G. B., & Worm, B. (2011). How Many Species Are There on Earth and in the Ocean ?, *9*(8), 1–8. https://doi.org/10.1371/journal.pbio.1001127

Mortazavi A, Williams BA, McCue K, Schaeffer L,Wold B (2008) Mapping and quantifying mam- malian transcriptomes by RNA-Seq. Nat Methods 5: 621–628.

Musick, J. A., & Ellis, J. K. (2005). Reproductive evolution of chondrichthyans. Reproductive biology and phylogeny of Chondrichthyes, 3, 45-79.

Navarro-Barranco, M. P. C. C., Guerra-garcı, M., & Jose, M. R. (2013). Long-distance dispersal , low connectivity and molecular evidence of a new cryptic species in the obligate rafter Caprella andreae Mayer , 1890 (Crustacea : Amphipoda : Caprellidae). *Helgol Mar Res*, *67*, 483–497. https://doi.org/10.1007/s10152-012-0337-9

Naylor, G. J. P., Caira, J. N., Jensen, K., Rosana, K. A. M., White, W. T., & Last, P. R. (2012). A DNA Sequence–Based Approach To the Identification of Shark and Ray Species and Its Implications

for Global Elasmobranch Diversity and Parasitology. *Bulletin of the American Museum of Natural History*, *367*, 1–262. https://doi.org/10.1206/754.1

Near, T. J., Eytan, R. I., Dornburg, a., Kuhn, K. L., Moore, J. a., Davis, M. P., … Smith, W. L. (2012). Resolution of ray-finned fish phylogeny and timing of diversification. *Proceedings of the National Academy of Sciences*, *109*(34), 13698–13703. https://doi.org/10.1073/pnas.1206625109

Neat, F., Pinto, C., Burrett, I., Cowie, L., Travis, J., Thorburn, J., … Wright, P. J. (2015). Site fidelity, survival and conservation options for the threatened flapper skate (Dipturus cf. intermedia). *Aquatic Conservation: Marine and Freshwater Ecosystems*, *25*(1), 6–20. https://doi.org/10.1002/aqc.2472

Nei, M. (1972). Genetic Distance between Populations. American Naturalist 106, 283–292.

Neilson, M. E., & Stepien, C. A. (2009). Evolution and phylogeography of the tubenose goby genus Proterorhinus (Gobiidae: Teleostei): Evidence for new cryptic species. *Biological Journal of the Linnean Society*, *96*(3), 664–684. https://doi.org/10.1111/j.1095-8312.2008.01135.x

Nirchio, M., Oliveira, C., & Siccha-Ramirez, Z. (2016). Cryptic Caribbean species of Scorpaena (Actinopterygii: Scorpaeniformes) suggested by cytogenetic and molecular data. *Journal of Fish*, 1947–1957. https://doi.org/10.1111/jfb.13089

Nokelainen, O., & Stevens, M. (2016). Camouflage. *Current Biology*, *26*(14), R654-R656.

Norman, J. R. (1937). Coast fishes. University Press.

Oshlack, A., Robinson, M. D., & Young, M. D. (2010). From RNA-seq reads to differential expression results. *Genome biology*, *11*(12), 220.

Ovenden, J. R., Morgan, J. A. T., Street, R., Tobin, A., Simpfendorfer, C., Macbeth, W., & Welch, D. (2011). Negligible evidence for regional genetic population structure for two shark species Rhizoprionodon acutus (Rüppell, 1837) and Sphyrna lewini (Griffith & Smith, 1834) with contrasting biology. *Marine Biology*, *158*(7), 1497–1509. https://doi.org/10.1007/s00227-011-1666-y

Palumbi, S. (1992). Marine speciation on a small planet. *Trends in Ecology & Evolution (Personal Edition)*, *7*(4), 114–118. https://doi.org/10.1016/0169-5347(92)90144-Z

Paradis, E., Claude, J., & Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. Bioinformatics, 20(2), 289-290.

Pasolini, P., Ragazzini, C., Zaccaro, Z., Cariani, A., Ferrara, G., Gonzalez, E. G., … Tinti, F. (2011). Quaternary geographical sibling speciation and population structuring in the Eastern Atlantic skates (suborder Rajoidea) Raja clavata and R. straeleni. *Marine Biology*, *158*(10), 2173–2186.

https://doi.org/10.1007/s00227-011-1722-7

Pavan-Kumar, A., Gireesh-Babu, P., Babu, P. P. S., Jaiswar, A. K., Hari Krishna, V., Prasasd, K. P., … Lakra, W. S. (2014). Molecular phylogeny of elasmobranchs inferred from mitochondrial and nuclear markers. *Molecular Biology Reports*, *41*(1), 447–457. https://doi.org/10.1007/s11033-013-2879-6

Pereira, R. J., Barreto, F. S., Pierce, N. T., Carneiro, M., & Burton, R. S. (2016). Transcriptome-wide patterns of divergence during allopatric evolution. *Molecular Ecology*, *25*(7), 1478–1493. https://doi.org/10.1111/mec.13579

Pérez-Portela, R., Arranz, V., Rius, M., & Turon, M. (2013). Cryptic speciation or global spread? The case of a cosmopolitan marine invertebrate with limited dispersal capabilities. *Scientific Reports*, *3*, 3197. https://doi.org/10.1038/srep03197

Pfenninger, M., & Schwenk, K. (2007). Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. *BMC Evolutionary Biology*, *7*, 121. https://doi.org/10.1186/1471-2148-7-121

Pickrell JK, Marioni JC, Pai AA, et al. Understanding mechanisms underlying human gene expression variation with RNA sequencing. Nature 2010;464:768–72

Pilgrim, E. M., Blum, M. J., Reusser, D. A., Lee, H., & Darling, J. A. (2013). Geographic range and structure of cryptic genetic diversity among Pacific North American populations of the non-native amphipod Grandidierella japonica. *Biological Invasions*, *15*(11), 2415–2428. https://doi.org/10.1007/s10530-013-0462-7

Pinhal, D., Shivji, M. S., Vallinoto, M., Chapman, D. D., Gadig, O. B. F., & Martins, C. (2011). Cryptic hammerhead shark lineage occurrence in the western South Atlantic revealed by DNA analysis. *Marine Biology*, *159*(4), 829–836. https://doi.org/10.1007/s00227-011-1858-5

Plank, S. M., Lowe, C. G., Feldheim, K. A., Wilson, R. R. & Brusslan, J. A. (2010). Population genetic structure of the round stingray Urobatis halleri(Elasmobranchii: Rajiformes) in southern California and the Gulf of California. Journal of Fish Biology 77, 329–340.

Pratt, H. L., & Carrier, J. C. (2001). A review of elasmobranch reproductive behavior with a case study on the nurse shark, Ginglymostoma cirratum. *Environmental Biology of Fishes*, *60*(1–3), 157–188. https://doi.org/10.1023/A:1007656126281

Priest, M. A., Dibattista, J. D., Mcilwain, J. L., Taylor, B. M., Hussey, N. E., & Berumen, M. L. (2016). A bridge too far: Dispersal barriers and cryptic speciation in an Arabian Peninsula grouper (Cephalopholis hemistiktos). *Journal of Biogeography*, *43*(4), 820–832.

https://doi.org/10.1111/jbi.12681

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics 155, 945–959.

Pyrkosz, A. B., Cheng, H., & Brown, C. T. (2013). RNA-Seq Mapping Errors When Using Incomplete Reference Transcriptomes of Vertebrates. *arXiv:1303.2411*, 1–17. Retrieved from http://arxiv.org/abs/1303.2411

Quattro, J. M., Driggers, W. B. I., Grady, J. M., Ulrich, G. F., & Roberts, M. a. (2013). Sphyrna gilberti sp. nov., a new hammerhead shark (Carcharhiniformes, Sphyrnidae) from the western Atlantic Ocean . *Zootaxa*, *3702*(2), 159. https://doi.org/10.11646/zootaxa.3702.2.5

Quattro, J. M., Stoner, D. S., Driggers, W. B., Anderson, C. A., Priede, K. A., Hoppmann, E. C., … Grady, J. M. (2006). Genetic evidence of cryptic speciation within hammerhead sharks (Genus Sphyrna). *Marine Biology*, *148*(5), 1143–1155. https://doi.org/10.1007/s00227-005-0151-x

Rambaut, A., & Drummond, A. (2007). Tracer version 1.6 [computer program]; 2007.

Rambaut, A. (2014). FigTree 1.4. 2 software. Institute of Evolutionary Biology, Univ. Edinburgh.

Rapaport, R. Khanin, Y. Liang, M. Pirun, A. Krek, P. Zumbo, C. E. Mason, N. D. Socci, and D. Betel, "Comprehensive evaluation of differential gene expression analysis methods for rna-seq data," *Genome biology*, vol. 14, no. 9, p. R95, 2013.

Renn SCP, Aubin-Horth N, Hofmann HA (2004) Biologically meaningful expression profiling across species using heterologous hybridization to a cDNA microarray. BMC genomics, 5, 42-55.

Reif, W. E. (1979). Morphogenesis and histology of large scales in batoids (Elasmobranchii). Palaeontologische Zeitschrift 53, 26–37.

Renema, W., Bellwood, D. R., Braga, J. C., Bromfield, K., Hall, R., Johnson, K. G., … Pandolfi, J. M. (2008). Hopping hotspots: Global shifts in marine biodiversity. *Science*, *321*(5889), 654–657. https://doi.org/10.1126/science.1155674

Richards, V. P., Henning, M., Witzell, W., & Shivji, M. S. (2009). Species delineation and evolutionary history of the globally distributed spotted eagle ray (Aetobatus narinari). *Journal of Heredity*, *100*(3), 273–283. https://doi.org/10.1093/jhered/esp005

Roberts, A. C., Trapnell, C. J., Donaghey, J. J., Rinn, J.L. and Pachter, L. "Improving rna-seq expression estimates by correcting for fragment bias," *Genome biology*, vol. 12, no. 3, p. R22, 2011.

Robinson, M. D., McCarthy, D. J. and Smyth, G. K. edgeR: a bioconductor package for differential expression analysis of digital gene expression data," *Bioinformatics*, vol. 26, no. 1, pp. 139–140, 2010.

Römpler, H., Rohland, N., Lalueza-Fox, C., Willerslev, E., Kuznetsova, T., Rabeder, G., ... & Hofreiter, M. (2006). Nuclear gene indicates coat-color polymorphism in mammoths. *Science*, *313*(5783), 62-62.

Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Molecular ecology resources, 8(1), 103-106.

Rosenberg, N. A. (2004). DISTRUCT: a program for the graphical display of population structure. Molecular Ecology Notes 4, 137–138.

Rubinoff, D., & Holland, B. S. (2005). Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. *Systematic Biology*, *54*(6), 952–961. https://doi.org/10.1080/10635150500234674

Rundell, R. J., & Price, T. D. (2009). Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends in Ecology and Evolution*, *24*(7), 394–399. https://doi.org/10.1016/j.tree.2009.02.007

Sandoval-Castillo, J., Rocha-Olivares, A., Villavicencio-Garayzar, C., & Balart, E. (2004). Cryptic isolation of Gulf of California shovelnose guitarfish evidenced by mitochondrial DNA. *Marine Biology*, *145*(5), 983–988. https://doi.org/10.1007/s00227-004-1378-7

San-Jose, L.M., Granado-Lorencio, F., Sinervo, B., Fitze, P.S., 2013. Iridophores and not carotenoids account for chromatic variation of carotenoid-based coloration in common lizards (Lacerta vivipara). Am. Nat. 181, 396–409

Šantić, M., Rađa, B., & Pallaoro, A. (2012). Feeding habits of small-spotted catshark (Scyliorhinus canicula Linnaeus, 1758) from the eastern central Adriatic Sea. *Marine Biology Research*, *8*(10), 1003–1011. https://doi.org/10.1080/17451000.2012.702912

Schiffer, P. H., & Herbig, H. G. (2016). Endorsing Darwin: global biogeography of the epipelagic goose barnacles Lepas spp. (Cirripedia, Lepadomorpha) proves cryptic speciation. *Zoological Journal of the Linnean Society*, *177*(3), 507–525. https://doi.org/10.1111/zoj.12373

Schonrogge, K., Barr, B., Wardlaw, J., Napper, E., Gardner, M., Breen, J., … Thomas, J. A. (2002). When rare species become endangered: Cryptic speciation in myrmecolphilous hoverflies. *Journal of the Linnean Society*, *75*, 291–300. https://doi.org/10.1046/j.1095-8312.2002.00019.x

Schuster, S. C. (2008). Next-generation sequencing transforms today's biology. *Nature methods*, *5*(1), 16.

Sefc, K. M., Brown, A. C., & Clotfelter, E. D. (2014). Carotenoid-based coloration in cichlid fishes. *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology*, *173*, 42–51. https://doi.org/10.1016/j.cbpa.2014.03.006

Serena, F., & Mancusi, C. (2010). Field identification guide to the skates (Rajidae) of the Mediterranean Sea ..., (January 2010). https://doi.org/10.13140/2.1.2414.9764

Siqueira, A. C., Oliveira-Santos, L. G. R., Cowman, P. F., Floeter, S. R., & Algar, A. (2016). Evolutionary processes underlying latitudinal differences in reef fish biodiversity. *Global Ecology and Biogeography*, *25*(12), 1466–1476. https://doi.org/10.1111/geb.12506

Sisneros, J. a., & Tricas, T. C. (2002). Ontogenetic Changes in the Response Properties of the Peripheral Electrosensory System in the Atlantic Stingray *(Dasyatis sabina)*. *Brain, Behavior and Evolution*, *59*(3), 130–140. https://doi.org/10.1159/000064160

Sisneros, J. A., & Tricas, T. C. (2002). Neuroethology and life history adaptations of the elasmobranch electric sense. *Journal of Physiology Paris*, *96*(5–6), 379–389. https://doi.org/10.1016/S0928-4257(03)00016-0

Sobert, E. (2008). Evolution , Population Thinking , and Essentialism Author (s): Elliott Sober Source : Philosophy of Science , Vol . 47 , No . 3 (Sep ., 1980), pp . 350-383 Published by : The University of Chicago Press on behalf of the Philosophy of Science Associati, *47*(3), 350–383.

Solé-Cava, A. M., & Levy, J. A. (1987). Biochemical Evidence for a Third Species of Angel Shark off the East Coast of South America. *Biochemical Systematics and Ecology*, *15*(1), 139–144. https://doi.org/10.1016/0305-1978(87)90093-7

Springer, S. (1941). A new species of hammerhead shark of the genus Sphyrna. In Proceedings of the Florida Academy of Sciences (Vol. 5, pp. 46-53).

Stingo, V., Capriglione, T., Rocco, L., Improta, R., and Morescalchi, A. (1989). Genome size and A-T rich DNA in selachians. *Genetica* 79: 197-205

Straube, N., Kriwet, J., & Schliewen, U. K. (2011). Cryptic diversity and species assignment of large lantern sharks of the Etmopterus spinax clade from the Southern Hemisphere (Squaliformes, Etmopteridae). *Zoologica Scripta*, *40*(1), 61–75. https://doi.org/10.1111/j.1463-6409.2010.00455.x

Strugnell, J. M., & Allcock, A. L. (2013). Southern Ocean evolution in a global context: a molecular viewpoint. In Adaptation and Evolution in Marine Environments, Volume 2 (pp. 35-53). Springer Berlin Heidelberg.

Stuart, B. L., Inger, R. F., & Voris, H. K. (2006). High level of cryptic species diversity revealed by sympatric lineages of Southeast Asian forest frogs. *Biology Letters*, *2*(3), 470–474. https://doi.org/10.1098/rsbl.2006.0505

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2015). MEGA6: Molecular Evolutionary Geneteics Analysis Version 6.0. 2013. Molecular Biology and Evolution.

Taylor, J. D., & Bagnara, J. T. (1972). Dermal Chromatophoies. American Zoologist, 12(1), 43-62.

Team, R. C. (2000). R language definition. Vienna, Austria: R foundation for statistical computing. Tester, A. L., & Nelson, G. J. (1967). Free neuromasts (pit organs) in sharks. Sharks, skates and rays, 503-531.

Tester, A. L., & Nelson, G. J. (1967). Free neuromasts (pit organs) in sharks. Sharks, skates and rays, 503-531.

Thatje, S. (2012). Effects of capability for dispersal on the evolution of diversity in antarctic benthos. *Integrative and Comparative Biology*, *52*(4), 470–482. https://doi.org/10.1093/icb/ics105

Thompson, J. D., Gibson, T., & Higgins, D. G. (2002). Multiple sequence alignment using ClustalW and ClustalX. Current protocols in bioinformatics, 2-3.

Thornhill, D. J., Mahon, A. R., Norenburg, J. L., & Halanych, K. M. (2008). Open-ocean barriers to dispersal: A test case with the Antarctic Polar Front and the ribbon worm Parborlasia corrugatus (Nemertea: Lineidae). *Molecular Ecology*, *17*(23), 5104–5117. https://doi.org/10.1111/j.1365-294X.2008.03970.x

Tornabene, L., Valdez, S., Erdmann, M., & Pezold, F. (2015). Support for a “Center of Origin” in the Coral Triangle: Cryptic diversity, recent speciation, and local endemism in a diverse lineage of reef fishes (Gobiidae: Eviota). *Molecular Phylogenetics and Evolution*, *82*(PA), 200–210. https://doi.org/10.1016/j.ympev.2014.09.012

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L. and Pachter, L. Differential analysis of gene regulation at transcript resolution with rna-seq *Nature biotechnology*, vol. 31, no. 1, pp. 46–53, 2012.

Treesirichod, A., Chansakulporn, S., & Wattanapan, P. (2014). Correlation between skin color evaluation by skin color scale chart and narrowband reflectance spectrophotometer. *Indian journal of dermatology*, *59*(4), 339.

Trontelj, P., & Fiser, C. (2009). Cryptic species diversity should not be trivialised. *Systematics and Biodiversity*, *7*(1), 1–3. https://doi.org/Doi 10.1017/S1477200008002909

Tsikliras, A. C., & Stergiou, K. I. (2014). Age at maturity of Mediterranean marine fishes. *Mediterranean Marine Science*, *16*(1), 5–20. https://doi.org/10.12681/mms.659

Tsutsui, S., Iwamoto, K., Nakamura, O., & Watanabe, T. (2007). Yeast-binding C-type lectin with opsonic activity from conger eel (Conger myriaster) skin mucus. *Molecular immunology*, *44*(5), 691-702.

Turro E, Su S, Goncalves Aˆ , et al. Haplotype and isoform specific expression estimation using multi-mapping RNA- seq reads. Genome Biol 2011;12:R13.

Ukuwela, K. D. B., de Silva, A., Mumpuni, Fry, B. G., & Sanders, K. L. (2014). Multilocus phylogeography of the sea snake Hydrophis curtus reveals historical vicariance and cryptic lineage diversity. *Zoologica Scripta*, *43*(5), 472–484. https://doi.org/10.1111/zsc.12070

Vachtenheim J, Borovansky J (2010) "Transcription physiology" of pigment formation in melanocytes: central role of MITF. Experimental Dermatology, 19, 617–627.

Valsecchi, E., Pasolini, P., Bertozzi, M., Garoia, F., Ungaro, N., Vacchi, M., … Tinti, F. (2005). Rapid Miocene-Pliocene dispersal and evolution of Mediterranean rajid fauna as inferred by mitochondrial gene variation. *Journal of Evolutionary Biology*, *18*(2), 436–446. https://doi.org/10.1111/j.1420-9101.2004.00829.x

Victor, B.C., 2015. How many coral reef fish species are there? Cryptic diversity and the new molecular taxonomy. In: Mora, C. (Ed.), Ecology of Fishes on Coral Reefs. Cambridge University Press, Cambridge, pp. 76–87.

Vijay, N., Poelstra, J. W., Künstner, A., & Wolf, J. B. W. (2013). Challenges and strategies in transcriptome assembly and differential gene expression quantification. A comprehensive in silico assessment of RNA-seq experiments. *Molecular Ecology*, *22*(3), 620–634. https://doi.org/10.1111/mec.12014

Visconti, M. A., Ramazini, G. C., Camargo, C. R., & Castrucci, A. M. L. (1999). Elasmobranch color change : A short review and novel data on hormone regulation Elasmobranch Color Change : A Short Review and Novel Data on Hormone Regulation. *Journal of Experimental Zoology*, *284*(February), 485–491. https://doi.org/10.1002/(SICI)1097-010X(19991001)284

Vrieling, H., Duhl, D. M., Millar, S. E., Miller, K. A. & Barsh, G. S. 1994 Differences in dorsal and ventral pig- mentation result from regional expression of the mouse agouti gene. Proc. Natl Acad. Sci. USA 91, 5667–5671. (doi:10.1073/pnas.91.12.5667)

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews. Genetics*, *10*(1), 57–63. https://doi.org/10.1038/nrg2484

Ward, R. D., Costa, F. O., Holmes, B. H., & Steinke, D. (2008). Dna barcoding of shared fish species from the North Atlantic and Australasia: Minimal divergence for most taxa, but zeus faber and lepidopus caudatus each probably constitute two species. *Aquatic Biology*, *3*(1), 71–78. https://doi.org/10.3354/ab00068

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., & Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(September), 1847–1857. https://doi.org/10.1098/rstb.2005.1716

Wearmouth, V. J., & Sims, D. W. (2009). Movement and behaviour patterns of the critically endangered common skate Dipturus batis revealed by electronic tagging. *Journal of Experimental Marine Biology and Ecology*, *380*(1–2), 77–87. https://doi.org/10.1016/j.jembe.2009.07.035

West-Eberhard, M. J. (2003). Developmental plasticity and evolution. Oxford University Press.

Wetherbee, B. M., Crow, G. L., & Lowe, C. G. (1996). Biology of the Galapagos shark,Carcharhinus galapagensis, in Hawai'i. *Environmental Biology of Fishes*, *45*(3), 299–310. https://doi.org/10.1007/BF00003099

Wetherbee, B. M., & Nichols, P. D. (2000). Lipid composition of the liver oil of deep-sea sharks from the Chatham Rise, New Zealand. *Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology*, *125*(4), 511–521. https://doi.org/10.1016/S0305-0491(00)00154-1

Whitear, M. (1986a). The skin of fishes including cyclostomes: epidermis. In Biology of the Integument, Vol. 2 (Bereiter-Hahn, J., Matoltsy, A. G. & Richards, K. S., eds), pp. 8–38. Berlin: Springer.

Whitear, M. (1986b). The skin of fishes including cyclostomes: dermis, In Biology of the Integument, Vol. 2 (Bereiter-Hahn, J., Matoltsy, A. G. & Richards, K. S., eds), pp. 39–64. Berlin: Springer.

Whitear, M. & Moate, R. (1998). Cellular diversity in the epidermis of Raja clavata (Chondrichthyes). Journal of Zoology 246, 275–285.

Whiteford, N., Haslam, N., Weber, G., Prügel-Bennett, A., Essex, J. W., Roach, P. L., ... & Neylon, C. (2005). An analysis of the feasibility of short read sequencing. *Nucleic acids research*, *33*(19), e171-e171.

Wilhelm, B. T., & Landry, J. R. (2009). RNA-Seq—quantitative measurement of expression through massively parallel RNA-sequencing. *Methods*, *48*(3), 249-257.

Wilkens, L. A., & Hofmann, M. H. (2005). Behavior of animals with passive, low-frequency electrosensory systems. In Electroreception (pp. 229-263). Springer New York.

Wilkens, H., Convergent Adaptations to Cave Life in the Rhamdia laticauda Catfish Group (Pimelodidae, Teleostei), Environm. Biol. Fishes, 2001, vol. 62, pp. 251–261.

Wilson, A. J., Pemberton, J. M., Pilkington, J. G., Clutton-Brock, T. H., Coltman, D. W., & Kruuk, L. E. B. (2007). Quantitative genetics of growth and cryptic evolution of body size in an island population. *Evolutionary Ecology*, *21*(3), 337–356. https://doi.org/10.1007/s10682-006-9106-z

Wilson-Leedy, J. G., & Ingermann, R. L. (2007). Development of a novel CASA system based on open source software for characterization of zebrafish sperm motility parameters. *Theriogenology*, *67*(3), 661-672

Xia Z, Wen J, Chang CC, Zhou X (2011) NSMAP: a method for spliced isoforms identification and quantification from RNA-Seq. BMC Bioinformatics 12: 162.

Zachos, J., Pagani, M., Sloan, L., Thomas, E., & Billups, K. (2001). Trends, Rhythms, and Aberrations in Global Climate 65 Ma to Present. *Source: Science, New Series*, *292*(5517), 686–693. https://doi.org/10.1126/science.1059412

Appendix I Chapter 2

Table S1 Sampling data and locations. The last row refers to geographical samples previously compared from McEachran et al. 1989. 1-Mediterranean group. 2 - Mauritania and Senegal group. 3 - Gulf of Guinea-equatorial African group. 4 - Angolan sample. 5 – South African sample. n.a.: not available. ST: Scientific Trawl survey. CF: Contracted Fishermen. FM: Fishery Market. GB: GenBank Database. BOLD: Barcoding of Life Database.

Sampling area	Area	Year	Sample Code	N	COI	SSRs	Source (Trawl survey program)	McEachran et al. 1989
Atlantic - Indian (A)								
South Africa - South Coast	Western Indian Ocean	2006	ASAF/07	8	5	8	ST (Africana)	5
South Africa - South Coast	Western Indian Ocean	2007		0	5	0	GB	5
South Africa - South Coast	Western Indian Ocean	2011	ASAF/11	32	30	31	ST (Africana)	5
Angola	South Eastern Atlantic	2006	AANG	28	27	26	ST (Nansen)	4
Senegal	Central Eastern Atlantic	2007	ASEN	5	5	5	CF	2
Portugal	North Eastern Atlantic	2007	APOR	3	0	3	ST (IPIMAR)	n.a.
Portugal	North Eastern Atlantic	2005, 2007		0	10	0	GB	n.a.
Mediterranean (M)								
Algeria	Western Mediterranean	2002, 2003	MALG/03	8	8	5	FM (Algiers)	1
Algeria	Western Mediterranean	2009, 2010	MALG/10	9	8	8	FM (Algiers)	1
Balearic Islands	Western Mediterranean	2006	MBAL	19	19	16	ST (MedITS)	1
Sardinia	Western Mediterranean	2002, 2005	MSAR	11	11	8	ST (MedITS; GruND)	1
Tuscany	Western Mediterranean	2005, 2006	MTUS/06	26	22	21	ST (MedITS; GruND)	1

Tuscany	Western Mediterranean	2008, 2010	MTUS/10	16	6	13	ST (MedITS; GruND)	1
Sicilian Channel - Adventura Bank	Western Mediterranean	2014	MADV	22	22	22	ST (MedITS)	1
Sicilian Channel - Maltese Bank	Western Mediterranean	2000, 2002	MMAL/02	16	12	8	ST (MedITS; GruND)	1
Sicilian Channel - Maltese Bank	Western Mediterranean	2007	MMAL/07	0	6	0	ST (MedITS; GruND)	1
Ionian Sea	Eastern Mediterranean	2004	MION	4	3	4	ST (MedITS; GruND)	1
Northern Adriatic Sea - Italian coast	Eastern Mediterranean	2006, 2007	MNAD1	39	31	20	ST (MedITS; GruND)	1
Northern Adriatic Sea - Croatian coast	Eastern Mediterranean	2002, 2004	MNAD2	24	24	8	ST (MedITS; GruND)	1
Southern Adriatic Sea - Italian coast	Eastern Mediterranean	2004	MSAD1	19	16	19	ST (MedITS; GruND)	1
Southern Adriatic Sea - Albanian coast	Eastern Mediterranean	2004	MSAD2	19	13	17	ST (MedITS; GruND)	1
Greece - Aegean coast	Eastern Mediterranean	2014	MGRE	0	2	0	GB	1
Israel	Eastern Mediterranean	2009	MISR/09	8	7	7	CF	1
Israel	Eastern Mediterranean	2012	MISR/14	0	3	0	BOLD	1
Israel	Eastern Mediterranean	2014		0	4	0	GB	1
Levantine Sea	Eastern Mediterranean	2009	MLEV	7	7	7	CF	1

Table S2 Mitochondrial gene polymorphism. *Nh* number of haplotypes, *S* number of polymorphic sites, *Hd* haplotype diversity, π nucleotide diversity, *k* average number of nucleotide differences and *SD* standard deviation. Sample code as given in Table S1.

Sample Code	n	H	S	Hd ± SD	π ± SD	k
ASAF/07	10	4	4	0.644 ± 0.152	0.00181 ± 0.00063	0.956
ASAF/11	30	8	6	0.733 ± 0.066	0.00206 ± 0.00034	1.087
AANG	27	10	40	0.858 ± 0.041	0.02543 ± 0.00380	13.453
ASEN	5	3	1	0.700 ± 0.218	0.00132 ± 0.00041	0.700
APOR	10	1	0	0.000	0.000	0.000
MALG/03	8	1	0	0.000	0.000	0.000
MALG/10	8	4	4	0.750 ± 0.139	0.00250 ± 0.00075	1.321
MBAL	19	1	0	0.000	0.000	0.000
MSAR	11	1	0	0.000	0.000	0.000
MTUS/06	22	2	1	0.091 ± 0.081	0.00017 ± 0.00015	0.091
MTUS/10	6	1	0	0.000	0.000	0.000
MADV	22	3	3	0.654 ± 0.061	0.00277 ± 0.00022	1.463
MMAL/02	12	3	3	0.667 ± 0.091	0.00241 ± 0.00043	1.273
MMAL/14	6	5	5	0.933 ± 0.122	0.00391 ± 0.00076	2.067
MION	3	1	0	0.000	0.000	0.000
MNAD1	31	2	1	0.396 ± 0.078	0.00075 ± 0.00015	0.420
MNAD2	24	5	4	0.630 ± 0.065	0.00143 ± 0.00025	0.757
MSAD1	16	3	2	0.242 ± 0.135	0.00068 ± 0.00040	0.358
MSAD2	13	2	1	0.385 ± 0.132	0.00073 ± 0.00025	0.385
MGRE	2	1	0	0.000	0.000	0.000
MISR/09	7	2	1	0.286 ± 0.196	0.00054 ± 0.00037	0.286
MISR/14	7	2	1	0.476 ± 0.171	0.00090 ± 0.00032	0.527
MLEV	7	1	0	0.000	0.000	0.000

Table S3 Hierarchical AMOVAs performed on the COI dataset and on the EST-SSRs datasets.

AMOVA–Groupings	COI			EST-SSRs			COI	EST-SSRs
	Total variation (%)	ϕ statistics	P	Total variation (%)	F statistics	P	%SC over %CT	%SC over %CT
AMOVA1- two groups: Central-Southern African vs NE Atlantic-Mediterranean Sea								
Among groups	89.81	ϕ_{CT} 0.8981	0.0002±0.0001	26.07	F_{CT} 0.2607	0.0006±0.0002	0.056	0.918
Among populations within groups	4.99	ϕ_{SC} 0.4901	0.0000±0.0000	23.93	F_{SC} 0.3237	0.0000±0.0000		
Within population	5.20	ϕ_{ST} 0.9480	0.0000±0.0000	17.72	F_{IS} 0.3544	0.0000±0.0000		
Within individuals				32.28	F_{IT} 0.6772	0.0000±0.0000		
AMOVA2- four groups: South Africa + Angola vs Senegal vs Portugal + Western Mediterranean vs Eastern Mediterranean								
Among groups	85.69	ϕ_{CT} 0.8569	0.0000±0.0000	40.20	F_{CT} 0.4020	0.0000±0.0000	0.066	0.192
Among populations within groups	5.64	ϕ_{SC} 0.3946	0.0000±0.0000	7.71	F_{SC} 0.1289	0.0000±0.0000		
Within population	8.67	ϕ_{ST} 0.9133	0.0000±0.0000	18.46	F_{IS} 0.3544	0.0000±0.0000		
Within individuals				33.63	F_{IT} 0.6637	0.0000±0.0000		
AMOVA3- five groups: South Africa vs Angola vs Senegal vs Portugal +Western Mediterranean vs Eastern Mediterranean								
Among groups	88.62	ϕ_{CT} 0.8862	0.0001±0.0001	40.32	F_{CT} 0.4021	0.0000±0.0000	0.029	0.169
Among populations within groups	2.56	ϕ_{SC} 0.2249	0.0000±0.0000	6.82	F_{SC} 0.1285	0.0000±0.0000		
Within population	8.82	ϕ_{ST} 0.9118	0.0000±0.0000	18.73	F_{IS} 0.3544	0.0000±0.0000		
Within individuals				34.13	F_{IT} 0.6587	0.0000±0.0000		
AMOVA4- six groups: South Africa vs Angola vs Senegal vs Portugal + Western Mediterranean vs Eastern Mediterranean vs Israel								
Among groups	89.46	ϕ_{CT} 0.8946	0.0001±0.0001	39.82	F_{CT} 0.3982	0.0000±0.0000	0.016	0.172
Among populations within groups	1.43	ϕ_{SC} 0.1357	0.0000±0.0000	6.84	F_{SC} 0.1137	0.0000±0.0000		
Within population	9.11	ϕ_{ST} 0.9089	0.0000±0.0000	18.90	F_{IS} 0.3544	0.0000±0.0000		
Within individuals				34.43	F_{IT} 0.6557	0.0000±0.0000		

AMOVA5- seven groups: South Africa vs Angola vs Senegal vs Portugal + West Mediterranean vs Sicily Channel vs East Mediterranean vs Israel									
Among groups	89.16	ϕ_{CT}	0.8916	0.0000±0.0000	36.12	F_{CT}	0.3613	0.0000±0.0000	0.010	0.227
Among populations within groups	0.93	ϕ_{SC}	0.0856	0.0000±0.0000	8.19	F_{SC}	0.1282	0.0000±0.0000		
Within population	9.91	ϕ_{ST}	0.9001	0.0000±0.0000	19.73	F_{IS}	0.3544	0.0000±0.0000		
Within individuals					35.95	F_{IT}	0.6405	0.0000±0.0000		

Table S4 Tamura-Nei genetic distances observed within and between *Raja* species and *Raja miraletus* geographical samples.

	Raja asterias	*Raja brachyura*	*Raja clavata*	*Raja microocellata*	*Raja montagui*	*Raja polystigma*	*Raja radula*	*Raja straeleni*	*Raja undulata*	*Raja miraletus* C-S Africa	*Raja miraletus* NE Atl-Med
Raja asterias	0.0025 ± 0.0012	0,013	0,011	0,014	0,013	0,012	0,010	0,011	0,013	0,015	0,015
Raja brachyura	0,086	0.0030 ± 0.0015	0,010	0,009	0,011	0,010	0,012	0,011	0,012	0,013	0,015
Raja clavata	0,059	0,051	0.0000 ± 0.0000	0,011	0,010	0,010	0,007	0,005	0,012	0,012	0,015
Raja microocellata	0,088	0,046	0,060	0.0000 ± 0.0000	0,012	0,011	0,011	0,011	0,013	0,013	0,015
Raja montagui	0,083	0,061	0,051	0,064	0.0000 ± 0.0000	0,007	0,012	0,011	0,012	0,012	0,014
Raja polystigma	0,075	0,053	0,051	0,060	0,023	0.0000 ± 0.0000	0,010	0,011	0,012	0,011	0,013
Raja radula	0,049	0,068	0,028	0,065	0,064	0,056	0.0010 ± 0.0010	0,007	0,012	0,013	0,015
Raja straeleni	0,059	0,059	0,015	0,060	0,059	0,056	0,028	0.0019 ± 0.0010	0,012	0,013	0,015
Raja undulata	0,077	0,074	0,072	0,079	0,078	0,071	0,074	0,073	0.0018± 0.0011	0,012	0,014
Raja miraletus C-S Africa	0,105	0,086	0,086	0,095	0,083	0,078	0,090	0,091	0,076	0.0188 ± 0.0032	0,012
Raja miraletus NE Atl-Med	0,107	0,101	0,098	0,100	0,091	0,088	0,096	0,102	0,096	0,073	0.0025 ± 0.0011

Table S5 Distribution of the COI haplotypes in the *Raja miraletus* samples. Samples code given as in Table S1.

	ASAF/07	ASAF/11	AANG	ASEN	APOR	MALG/03	MALG/10	MBAL	MSAR	MTUS/06	MTUS/10	MADV	MMAL/02	MMAL/07	MION	MNAD1	MNAD2	MSAD1	MSAD2	MGRE	MISR/09	MISR/14	MLEV
	N=10	N=30	N=27	N=5	N=10	N=8	N=8	N=19	N=11	N=22	N=6	N=22	N=12	N=6	N=3	N=31	N=24	N=16	N=13	N=2	N=7	N=7	N=7
Hap_1	0.2000	0.2000																					
Hap_2	0.6000	0.4670																					
Hap_3	0.1000	0.1670																					
Hap_4	0.1000																						
Hap_5		0.0333																					
Hap_6		0.0333																					
Hap_7		0.0333																					
Hap_8		0.0333																					
Hap_9		0.0333																					
Hap_10			0.2590																				
Hap_11			0.0370																				
Hap_12			0.1480																				
Hap_13			0.0741																				
Hap_14			0.0741																				
Hap_15			0.0370																				
Hap_16			0.2590																				
Hap_17			0.0370																				
Hap_18			0.0370																				
Hap_19			0.0370																				
Hap_20				0.6000																			
Hap_21				0.2000																			
Hap_22				0.2000																			
Hap_23					1.0000	1.0000	0.5000					0.2730		0.3330									
Hap_24							0.2500	1.0000	1.0000	0.9550	1.0000	0.2270	0.3330	0.1670		0.2580	0.3750	0.0625	0.2310				
Hap_25							0.1250																

Hap_26	0.1250			0.5000	0.5000	0.1670	1.0000										
Hap_27		0.0455															
Hap_28					0.1670												
Hap_29						0.1670											
Hap_30						0.1670											
Hap_31								0.7420	0.5000	0.8750	0.7690	1.0000					
Hap_32									0.0417								
Hap_33									0.0417								
Hap_34									0.0417								
Hap_35										0.0625							
Hap_36													0.8570	0.2860			
Hap_37													0.1430	0.7140			
Hap_38																	1.0000

Table S6 Summary statistics of the SSR polymorphism per geographical sample and over all the loci considered. N number of individuals. A number of alleles. Ar allelic richness. Ho observed heterozygosity. HE expected heterozygosity. FIS value. HWE deviation from Hardy-Weinberg equilibrium. Significant P-values are highlighted in bold (P<0.05), *P significant after sequential Bonferroni. Samples code given as in Table S1.

Locus	ASAF/07	ASAF/11	AANG	ASEN	APOR	MALG/03	MALG/10	MBAL	MSAR	MTUS/06	MTUS/10	MADV	MMAL/02	MION	MNAD1	MNAD2	MSAD1	MSAD2	MISR/14	MLEV
	n =8	n =31	n =26	n =5	n =3	n =5	n =8	n =16	n =8	n =21	n =13	n =22	n =8	n =4	n =8	n =20	n =20	n =20	n =7	n =7
LERI 27	N_A=15																			
N	8	27	25	4	3	5	6	16	8	20	11	22	7	4	8	18	19	18	7	7
A	5	5	13	7	2	2	3	3	4	5	4	6	3	2	4	2	4	5	2	2
Ar	1.7330	1.4860	1.9240	1.9640	1.5330	1.2000	1.6820	1.2800	1.6920	1.5440	1.5710	1.5690	1.4730	1.4290	1.6500	1.3860	1.3660	1.5160	1.2640	1.2640
H_O	0.2500	0.4074	0.6800	1.0000	0.0000	0.2000	0.5000	0.3125	0.2500	0.3500	0.7273	0.6364	0.4286	0.5000	0.7500	0.3889	0.2632	0.4444	0.0000	0.0000
H_E	0.6875	0.4767	0.9056	0.8438	0.4444	0.1800	0.6250	0.2715	0.6484	0.5300	0.5455	0.5558	0.4388	0.3750	0.6094	0.3750	0.3560	0.5015	0.2449	0.2449
FIS	0.6740	0.1640	0.2680	-0.0430	1.0000	0.0000	0.2860	-0.1190	0.6540	0.3620	-0.2900	-0.1220	0.1000	-0.2000	-0.1670	-0.0080	0.2860	0.1420	1.0000	1.0000
NA	0.2765	0.0591	0.1229	0.0000	0.3150	0.0000	0.0759	0.0000	0.2454	0.1392	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0929	0.0005	0.2364	0.2364
HWE	**0.0064***	0.3678	**0.0000***	1.0000	0.1984	-	0.3226	1.0000	**0.0162***	**0.0435***	0.4259	**0.0088***	0.4408	1.0000	1.0000	1.0000	**0.0163***	0.1301	0.0781	0.0750
LERI 26	N_A=12																			
N	7	29	24	5	3	5	8	16	7	20	13	21	8	4	8	20	20	20	7	7
A	5	6	9	4	1	1	2	2	2	3	3	1	1	2	5	3	2	5	3	2
Ar	1.8020	1.7230	1.7070	1.7780	1.0000	1.0000	1.2330	1.2260	1.2640	1.1880	1.1510	1.0000	1.0000	1.4290	1.8080	1.1880	1.2240	1.5580	1.4730	1.2640
H_O	0.1429	0.1724	0.4583	0.6000	0.0000	0.0000	0.0000	0.2500	0.0000	0.2000	0.1538	0.0000	0.0000	0.5000	0.1250	0.2000	0.1500	0.4500	0.4286	0.0000
H_E	0.7449	0.7105	0.6918	0.7000	0.0000	0.0000	0.2188	0.2188	0.2449	0.1838	0.1450	0.0000	0.0000	0.3750	0.7578	0.1838	0.2188	0.5438	0.4388	0.2449
FIS	0.8330	0.7650	0.3560	0.2500	NA	NA	1.0000	-0.1110	1.0000	-0.0630	-0.0210	NA	NA	-0.2000	0.8540	-0.0630	0.3370	0.1970	0.1000	1.0000
NA	0.3599	0.3146	0.1462	0.0379	0.0010	0.0010	0.2238	0.0000	0.2364	0.0000	0.0000	0.0010	0.0010	0.0000	0.3644	0.0000	0.0846	0.0423	0.0000	0.2364
HWE	**0.0001***	**0.0000***	**0.0000***	0.6966	-	-	0.0654	1.0000	0.0782	1.0000	1.0000	-	-	1.0000	**0.0002***	1.0000	0.2447	0.0709	0.4335	0.0769
LERI 24	N_A=12																			
N	7	31	25	5	3	4	6	15	8	19	12	22	8	4	8	19	18	19	7	7
A	3	3	8	4	3	3	2	4	3	3	4	2	2	3	2	4	4	5	3	2
Ar	1.6150	1.4660	1.7800	1.7780	1.8000	1.4640	1.4850	1.6250	1.5670	1.5680	1.5870	1.4850	1.4000	1.7500	1.5250	1.5850	1.5840	1.5900	1.5380	1.5270
H_O	0.2857	0.5484	0.5200	0.8000	0.0000	0.5000	0.3333	0.5333	0.0000	0.0526	0.1667	0.4091	0.5000	0.2500	0.8750	0.3158	0.3889	0.1053	0.1429	0.5714
H_E	0.5714	0.4584	0.7648	0.7000	0.666	0.4063	0.4444	0.6044	0.5313	0.5526	0.5625	0.4742	0.3750	0.6563	0.4922	0.5693	0.5679	0.5748	0.5000	0.4898

FIS	0.5560	-0.1810	0.3380	-0.0320	1.0000	-0.0910	0.3330	0.1520	1.0000	0.9100	0.7250	0.1600	-0.2730	0.7000	-0.7500	0.4670	0.3410	0.8260	0.7500	-0.0910
NA	0.2032	0.0000	0.1581	0.0000	0.4000	0.0000	0.0815	0.0734	0.3553	0.3246	0.2446	0.0455	0.0000	0.2538	0.0000	0.1725	0.1353	0.3060	0.2458	0.0000
HWE	0.1051	0.6248	**0.0051***	0.6936	0.0658	1.0000	1.0000	0.2372	**0.0007***	**0.0000***	**0.0014***	0.6504	1.0000	0.1498	0.1388	**0.0115***	0.0632	**0.0000***	**0.0227***	1.0000
LERI 34	N_A=8																			
N	83	304	265	53	11	52	62	152	83	203	121	222	84	42	81	192	183	183	72	72
A																				
Ar	1.5080	1.5660	1.6700	1.6220	1.0000	1.3560	1.5450	1.5150	1.4330	1.2680	1.0000	1.4060	1.3500	1.5360	1.0000	1.1930	1.4890	1.4460	1.5270	1.4400
H_O	0.3750	0.4333	0.6154	0.4000	0.0000	0.4000	0.3333	0.4000	0.0000	0.1000	0.0000	0.2727	0.2500	0.7500	0.0000	0.1053	0.2222	0.3333	0.2857	0.0000
H_E	0.4766	0.5567	0.6568	0.5600	0.0000	0.3200	0.5000	0.4978	0.4063	0.2612	0.0000	0.3967	0.3281	0.4688	0.0000	0.1884	0.4753	0.4336	0.4898	0.4082
FIS	0.2760	0.2380	0.0830	0.3850	NA	-0.1430	0.4120	0.2290	1.0000	0.6330	NA	0.3330	0.3000	-0.5000	NA	0.4630	0.5530	0.2580	0.4780	1.0000
NA	0.1866	0.0637	0.0674	0.1333	0.0010	0.0000	0.1111	0.0654	0.3116	0.1646	0.0010	0.0988	0.0000	0.0000	0.0010	0.1072	0.1923	0.0833	0.1381	0.3024
HWE	0.1246	0.2842	**0.0457***	0.2408	-	1.0000	0.4748	0.6048	**0.0051***	**0.0102***	-	0.2708	0.2078	1.0000	-	0.1606	**0.0042***	0.2001	0.4404	**0.0203***

LERI 63	N_A=14																			
N	74	312	268	56	31	41	62	162	72	143	133	223	72	32	82	174	195	198	43	74
A																				
Ar	1.7140	1.4550	1.8220	1.8440	1.0000	1.0000	1.4090	1.4980	1.2640	1.3730	1.2180	1.5610	1.4400	1.6000	1.5250	1.7110	1.7580	1.7980	1.4640	1.7800
H_O	0.4286	0.5484	0.7692	0.6000	0.0000	0.0000	0.1667	0.1875	0.0000	0.1429	0.0769	0.4091	0.5714	0.3333	0.8750	0.5294	0.5263	0.3158	0.5000	0.5714
H_E	0.6633	0.4480	0.8062	0.7600	0.0000	0.0000	0.3750	0.4824	0.2449	0.3597	0.2101	0.5486	0.4082	0.5000	0.4922	0.6903	0.7382	0.7770	0.4063	0.7245
FIS	0.4190	-0.2090	0.0650	0.3140	NA	NA	0.6150	0.6310	1.0000	0.6260	0.6570	0.2760	-0.3330	0.5000	-0.7500	0.2620	0.3120	0.6110	-0.0910	0.2840
NA	0.1803	0.0404	0.0304	0.0006	0.0010	0.0010	0.1677	0.2012	0.2364	0.1912	0.1582	0.1099	0.0000	0.1111	0.0000	0.1012	0.1109	0.2661	0.0000	0.0666
HWE	**0.0348***	0.4180	0.2287	0.0516	-	-	0.2720	**0.0313***	0.0791	**0.0408***	**0.0426***	0.0543	1.0000	1.0000	0.1360	0.1586	**0.0101***	**0.0000**	1.0000	0.1370
LERI 50	N_A=8																			
N	83	302	263	43	33	53	73	143	62	184	133	221	81	41	81	163	143	155	54	72
A																				
Ar	1.4330	1.0330	1.1120	1.6070	1.6000	1.6220	1.6150	1.5820	1.1670	1.6160	1.6770	1.0000	1.0000	1.0000	1.0000	1.4920	1.5610	1.5400	1.7330	1.4950
H_O	0.2500	0.0333	0.1154	0.7500	0.3333	0.0000	0.0000	0.0000	0.1667	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1333	0.2000	0.1429
H_E	0.4063	0.0328	0.1102	0.5313	0.5000	0.5600	0.5714	0.5612	0.1528	0.5988	0.6509	0.0000	0.0000	0.0000	0.0000	0.4766	0.5408	0.5222	0.6600	0.4592

FIS	0.4400	0.0000	-0.0270	-0.2860	0.5000	1.0000	1.0000	1.0000	0.0000	1.0000	1.0000	NA	NA	NA	NA	1.0000	1.0000	0.7600	0.7500	0.7270
NA	0.1465	0.0001	0.0000	0.0000	0.0008	0.3659	0.3687	0.3615	0.0000	0.3766	0.3948	0.0010	0.0010	0.0010	0.0010	0.3371	0.3552	0.2580	0.2863	0.2222
HWE	0.1449	-	1.0000	1.0000	0.1930	**0.0173***	**0.0035***	**0.0000***	-	**0.0000***	**0.0000***	-	-	-	-	**0.0000***	**0.0000***	**0.0000***	**0.016***	0.1051
LERI 40	N_A=6																			
N A	7 2	31 2	26 2	5 2	3 1	4 1	8 2	15 2	6 2	19 2	13 2	22 1	7 2	4 1	8 1	18 2	19 3	17 4	4 3	7 2
Ar	1.4950	1.0320	1.2080	1.5330	1.0000	1.0000	1.3250	1.1290	1.1670	1.2730	1.2710	1.0000	1.2640	1.0000	1.0000	1.2030	1.4210	1.4800	1.6070	1.1430
H_O	0.1429	0.0323	0.0769	0.8000	0.0000	0.0000	0.1250	0.0000	0.1667	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1176	0.5000	0.1429
H_E	0.4592	0.0317	0.2041	0.4800	0.0000	0.0000	0.3047	0.1244	0.1528	0.2659	0.2604	0.0000	0.2449	0.0000	0.0000	0.1975	0.4100	0.4654	0.5313	0.1327
FIS	0.7270	0.0000	0.6350	-0.6000	NA	NA	0.6320	1.0000	0.0000	1.0000	1.0000	NA	1.0000	NA	NA	1.0000	1.0000	0.7600	0.2000	0.0000
NA	0.1677	0.0001	0.1469	0.0000	0.0010	0.0010	0.1641	0.1703	0.0000	0.2460	0.2435	0.0010	0.2364	0.0010	0.0010	0.2131	0.3100	0.2392	0.0000	0.0000
HWE	0.1068	-	**0.017***	0.4300	-	-	0.1988	**0.0341***	-	**0.0004***	**0.005***	-	0.0750	-	-	**0.0026***	**0.0000***	**0.0000***	0.4246	-
LERI 44	N_A=14																			
N A	4 3	16 2	12 1	2 2	1 1	3 3	7 4	16 1	6 1	21 1	12 1	21 7	8 6	4 2	8 3	16 3	17 4	16 8	5 4	5 1
Ar	1.4640	1.0630	1.0000	1.6670	1.0000	1.7330	1.7580	1.0000	1.0000	1.0000	1.0000	1.5150	1.7170	1.4290	1.4920	1.4860	1.6080	1.6980	1.6440	1.0000
H_O	0.2500	0.0625	0.0000	0.0000	0.0000	0.6667	0.4286	0.0000	0.0000	0.0000	0.0000	0.4286	0.7500	0.0000	0.6250	0.4375	0.4118	0.3750	0.8000	0.0000
H_E	0.4063	0.0605	0.0000	0.5000	0.0000	0.6111	0.7041	0.0000	0.0000	0.0000	0.0000	0.5023	0.6719	0.3750	0.4609	0.4707	0.5900	0.6758	0.5800	0.0000
FIS	0.5000	0.0000	NA	1.0000	NA	0.1110	0.4550	NA	NA	NA	NA	0.1710	-0.0500	1.0000	-0.2960	0.1030	0.3290	0.4710	-0.2800	NA
NA	0.0009	0.0000	0.0010	0.3333	0.0010	0.0000	0.1616	0.0010	0.0010	0.0010	0.0010	0.0000	0.0000	0.2903	0.0000	0.0207	0.1165	0.1652	0.0000	0.0010
HWE	0.1445	-	-	0.3320	-	1.0000	0.1619	-	-	-	-	**0.0344***	**0.0229***	0.1429	1.0000	1.0000	0.1040	**0.0000***	1.0000	-
All loci																				
A_{mean}	3.5000	3.2500	6.1250	3.8750	1.6250	2.0000	2.5000	2.3750	2.3750	3.0000	2.6250	2.8750	2.6250	1.8750	2.3750	2.8750	3.5000	5.3750	3.0000	2.1250
Ar_{mean}	1.5955	1.3530	1.5279	1.7241	1.2416	1.2969	1.5065	1.3569	1.3193	1.3538	1.3094	1.3170	1.3305	1.3966	1.3750	1.4055	1.5014	1.5783	1.5313	1.3641
$H_{O\,mean}$	0.2656	0.2798	0.4044	0.6188	0.0417	0.2208	0.2359	0.2104	0.0729	0.1057	0.1406	0.2695	0.3125	0.2917	0.4063	0.2471	0.2453	0.2844	0.3571	0.1786
$H_{E\,mean}$	0.5519	0.3469	0.5175	0.6344	0.2014	0.2597	0.4679	0.3451	0.2977	0.3440	0.2968	0.3097	0.3084	0.3438	0.3516	0.3939	0.4871	0.5618	0.4814	0.3380
FIS_{mean}	0.5740	0.2100	0.2380	0.1940	0.8570	0.2810	0.5550	0.4190	0.7840	0.7070	0.5560	0.1530	0.0580	0.3000	-0.0900	0.3980	0.5180	0.5160	0.3490	0.5290
HWE	**0.0000***	**0.0000***	**0.0000***	1.5895	0.0627	0.4225	**0.0180***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0045***	0.1312	0.8090	**0.0046***	**0.0000***	**0.0000***	**0.0000***	0.0545	**0.0089***

Table S7 Pair-wise Φst values (below the diagonal) and associated significance (above the diagonal). Significant P-values are highlighted in bold (P<0.05), *P significant after sequential Bonferroni. Samples code given as in Table S1.

		W Indian		SE Atlantic	CE Atlantic	NE Atlantic	Western Mediterranean									Eastern Mediterranean								
		ASAF/07	ASAF/11	AANG	ASEN	APOR	MALG/03	MALG/10	MBAL	MSAR	MTUS/06	MTUS/10	MADV	MMAL/02	MMAL/14	MION	MNAD1	MNAD2	MSAD1	MSAD2	MGRE	MISR/09	MISR/14	MLEV
W Ind	ASAF/07	*	0,7061	**0.0000***	**0.0001***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0,0049**	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0,0098**	**0.0000***	**0.0000***	**0.0000***
	ASAF/11	-0,0252	*	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***
SE Atl	AANG	**0.3142***	**0.4187***	*	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0001***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0029***	**0.0000***	**0.0000***	**0.0000***
CE Atl	ASEN	**0.9428***	**0.93301***	**0.5001***	*	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0020***	**0.0000***	**0.0000***	**0,0039**	**0,0166**	**0.0000***	**0.0000***	**0.0000***	**0.0001***	**0,0352**	**0.0000***	**0.0000***	**0.0020***
NE Atl	APOR	**0.9865***	**0.9767***	**0.7262***	**0.9944***	*	0,9990	**0,0244**	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0,0068**	**0.0049***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0,0293**	**0.0000***	**0.0000***	**0.0000***
W Med	MALG/03	**0.9848***	**0.9754***	**0.7123***	**0.9934***	0,0000	*	0,0615	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0020***	**0.0000***	**0,0195**	**0.0029***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0,0264**	**0.0000***	**0.0000***	**0.0000***
	MALG/10	**0.9690***	**0.9686***	**0.7088***	**0.9723***	**0,2862**	0,2449	*	**0.0000***	**0.0029***	**0.0000***	0,0938	0,0596	**0.0029***	0,6523	**0,0283**	**0.0000***	**0,0029**	**0.0000***	**0.0001***	0,1006	**0.0000***	**0.0000***	**0.0001***
	MBAL	**0.9912***	**0.9816***	**0.7778***	**0.9968***	**1.0000***	**1.0000***	**0.4164***	*	0,9990	0,9990	0,9990	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0020***	**0.0000***	**0.0000***	**0.0000***
	MSAR	**0.9875***	**0.9779***	**0.7394***	**0.9950***	**1.0000***	**1.0000***	**0.3044***	0,0000	*	0,9990	0,9990	**0.0039***	**0.0000***	**0.0012***	**0.0049***	**0.0000***	**0.0020***	**0.0000***	**0.0000***	**0,0039**	**0.0000***	**0.0000***	**0.0000***
	MTUS/06	**0.9903***	**0.9816***	**0.7879***	**0.9952***	**0.9400***	**0.9360***	**0.3994***	-0,0069	-0,0359	*	0,9990	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0078***	**0.0000***	**0.0000***	**0.0000***
	MTUS/10	**0.9831***	**0.9747***	**0.7056***	**0.9922***	**1.0000***	**1.0000***	0,1931	0,0000	0,0000	-0,0845	*	**0,0234**	**0,0059**	0,2148	**0.0068***	**0,0010**	**0,0176**	**0.0000***	**0.0029***	**0,0322**	**0.0000***	**0.0000***	**0.0001***
	MADV	**0.9642***	**0.9659***	**0.7718***	**0.9667***	**0.4824***	**0.4602***	0,1371	**0.4048***	**0.3383***	**0.4105***	**0,2753**	*	0,1660	0,2744	0,2373	**0.0000***	**0.0016***	**0.0000***	**0.0049***	0,3740	**0.0000***	**0.0000***	**0.0000***
	MMAL/02	**0.9696***	**0.9693***	**0.7387***	**0.9723***	**0.7258***	**0.7024***	**0.3660***	**0.6468***	**0.5629***	**0.6408***	**0,4799**	0,0462	*	0,0508	0,5498	**0.0000***	**0.0001***	**0.0001***	**0.0009***	0,0596	**0.0000***	**0.0000***	**0.0000***
	MMAL/14	**0.9620***	**0.9662***	**0.6951***	**0.9637***	**0,4227**	**0,3735**	-0,0611	**0.3632***	**0.2394***	**0.3537***	0,1143	0,0190	0,1937	*	0,0840	**0.0088***	0,0537	**0.0020***	**0,0098**	0,4932	**0.0000***	**0.0000***	**0.0001***
E Med	MION	**0,9791**	**0.9723***	**0.6827***	**0,9886**	**1.0000***	**1.0000***	**0,6030**	**1.0000***	**1.0000***	**0.9599***	**1,0000**	0,1939	0,0233	0,3882	*	**0.0000***	**0.0000***	**0.0000***	**0.0059***	0,1162	**0.0068***	**0,0098**	**0.0098***
	MNAD1	**0.9858***	**0.9802***	**0.8154***	**0.9893***	**0.8341***	**0.8262***	**0.5194***	**0.6859***	**0.6420***	**0.6673***	**0.6037***	**0.2956***	**0.4116***	**0.3198***	**0.7292***	*	0,1719	0,3066	0,9990	0,9990	**0.0000***	**0.0000***	**0.0000***
	MNAD2	**0,9780**	**0.9746***	**0.7890***	**0.9815***	**0.70600***	**0.6908***	**0.3222***	**0.4338***	**0.3704***	**0.4278***	**0,3109**	**0.2089***	**0.3159***	0,1382	**0.6007***	0,0177	*	0,0576	0,3731	0,6231	**0.0000***	**0.0000***	**0.0000***
	MSAD1	**0.9843***	**0.9775***	**0.7664***	**0.9894***	**0.8866***	**0,8773**	**0.5416***	**0.8235***	**0.7779***	**0.7899***	**0.7333***	**0.2926***	**0.3843***	**0.3384***	**0.7517***	0,0089	0,0950	*	0,6201	0,9990	**0.0000***	**0.0000***	**0.0000***
	MSAD2	**0.9830***	**0.9763***	**0.7518***	**0.9886***	**0.8774***	**0.8658***	**0.4604***	**0.7909***	**0.7324***	**0.7463***	**0.6730***	**0.2363***	**0.3405***	**0,2398**	**0.7494***	-0,0556	0,0039	-0,0298	*	0,9990	**0.0000***	**0.0000***	**0.0000***
	MGRE	**0,9771**	**0.9720***	**0.6650***	**0,9863**	**1,0000**	**1,0000**	0,3722	**1.0000***	**1,0000**	**0.9185***	**1,0000**	0,1217	0,1724	0,0363	1,0000	-0,1016	-0,0250	-0,2872	-0,1607	*	**0,0205**	**0,0205**	**0,0205**
	MISR/09	**0.9826***	**0.9761***	**0.7328***	**0.9895***	**0.9723***	**0.9681***	**0.7779***	**0.9769***	**0.9656***	**0.9573***	**0.9506***	**0.6738***	**0.7300***	**0.6945***	**0.9327***	**0.8448***	**0.7642***	**0.8559***	**0.8527***	**0,8880**	*	0,0977	**0.0000***
	MISR/14	**0.9809***	**0.9756***	**0.7329***	**0.9868***	**0.9593***	**0.9532***	**0.7866***	**0.9671***	**0.9513***	**0.9519***	**0.9304***	**0.7046***	**0.7504***	**0.7119***	**0,9056**	**0.8613***	**0.7897***	**0.8640***	**0.8585***	**0,8536**	0,4167	*	**0.0001***
	MLEV	**0.9851***	**0.9767***	**0.7277***	**0.9933***	**1.0000***	**1.0000***	**0.7328***	**1.0000***	**1.0000***	**0.9658***	**1.0000***	**0.5632***	**0.6441***	**0.6139***	**1.0000***	**0.7589***	**0.6516***	**0.7951***	**0.7999***	**1,0000**	**0.9546***	**0.9359***	*

Table S8 Pair-wise Fst values (below the diagonal) and associated significance (above the diagonal). Significant P-values are highlighted in bold (P<0.05), *P significant after sequential Bonferroni. Samples code given as in Table S1.

		W Indian		SE Atlantic	CE Atlantic	NE Atlantic	Western Mediterranean								Eastern Mediterranean						
		ASAF/07	ASAF/11	AANG	ASEN	APOR	MALG/03	MALG/10	MBAL	MSAR	MTUS/06	MTUS/10	MADV	MMAL/02	MION	MNAD1	MNAD2	MSAD1	MSAD2	MISR/14	MLEV
W Ind	ASAF/07	*	0.1524	0.2038	**0.0178***	**0.0048***	**0.0006***	**0.0005***	**0.0000***	**0.0035***	**0.0000***	**0.0000***	**0.0000***	**0.0002***	**0.0082***	**0.0160***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0005***
	ASAF/11	0.0802	*	**0.0000***	**0.0117***	**0.0003***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.00099***	**0.0067***	**0.0000***	**0.0000***	**0.0000***	**0.0001***	**0.0000***
SE Atl	AANG	0.0457	**0.1965***	*	**0.0019***	**0.0001***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0001***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***
CE Atl	ASEN	**0.1926***	**0.2246***	**0.2559***	*	**0.0198***	**0.0062***	**0.0146***	**0.0004***	**0.0137***	**0.0001***	**0.0003***	**0.0000**	**0.0004***	**0.0071***	0.1738	**0.0015***	**0.0032***	0.0660	0.0687	**0.0445***
NE Atl	APOR	**0.5409***	**0.4944***	**0.5134***	**0.4339***	*	0.9999	0.9999	0.5704	0.7785	0.6497	0.9999	0.9999	0.9999	**0.0290***	0.1260	**0.0005***	**0.0021***	**0.0012***	**0.0077***	**0.0090***
W Med	MALG/03	**0.5952***	**0.5183***	**0.5392***	**0.5139***	0.0000	*	0.9999	0.5322	0.4875	0.4220	0.9999	0.9999	0.9999	**0.0085***	**0.0207***	**0.0000***	**0.0001***	**0.0001***	**0.0042***	**0.0011***
	MALG/10	**0.5190***	**0.4684***	**0.4903***	**0.3903***	-0.0260	0.0252	*	0.9999	0.9999	0.8716	0.7511	0.2631	0.9999	**0.0052***	**0.0299***	**0.0000***	**0.0000***	**0.0000***	**0.0012***	**0.0003***
	MBAL	**0.5697***	**0.5012***	**0.5249***	**0.4554***	-0.0190	0.0231	-0.0492	*	0.3129	0.7832	0.4702	0.0822	0.2547	**0.0001***	**0.0004***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***
	MSAR	**0.4496***	**0.4115***	**0.3988***	**0.3725***	0.0857	0.1444	0.0095	0.0244	*	0.2488	0.2827	0.1175	0.4694	**0.0076***	0.0859	**0.0000***	**0.0000***	**0.0002***	**0.00198***	**0.0007***
	MTUS/06	**0.6103***	**0.5199***	**0.5455***	**0.5186***	-0.0173	0.0226	-0.0408	-0.0243	0.0309	*	0.7082	0.0767	0.2894	**0.0002***	**0.0002***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***
	MTUS/10	**0.6074***	**0.5212***	**0.5444***	**0.5289***	-0.0721	-0.0301	-0.0198	-0.0056	0.0447	-0.0212	*	0.5359	0.4976	**0.0014***	**0.0012***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***
	MADV	**0.7755***	**0.6080***	**0.6399***	**0.7507***	0.0000	0.0000	0.1089	0.0725	0.3112	0.0634	-0.0138	*	0.9999	**0.0002***	**0.0001***	**0.0000***	**0.0000***	**0.0000***	**0.0000***	**0.0000***
	MMAL/02	**0.6524***	**0.5438***	**0.5675***	**0.5937***	0.0000	0.0000	0.0667	0.0529	0.2000	0.0495	-0.0034	0.0000	*	**0.0020***	**0.0069***	**0.0000***	**0.0000***	**0.0000***	**0.0004***	**0.0002***
E Med	MION	**0.3553***	**0.3770***	**0.3895***	**0.3538***	**0.6778***	**0.7428***	**0.6072***	**0.6380***	**0.5826***	**0.6756***	**0.6864***	**0.8973***	**0.8012***	*	0.3910	**0.0001***	**0.0001***	**0.0000***	**0.0118***	**0.0149***
	MNAD1	**0.1644***	**0.1945***	**0.2326***	0.1177	0.2829	**0.3462***	**0.2831***	**0.3448***	0.2130	**0.3695***	**0.3646***	**0.5622***	**0.4121***	0.0761	*	**0.0000***	**0.0002***	**0.0014***	**0.0294***	**0.0084***
	MNAD2	**0.6137***	**0.5327***	**0.5295***	**0.4598***	**0.8303***	**0.8432***	**0.7828***	**0.7801***	**0.7652***	**0.7979***	**0.8141***	**0.8979***	**0.8587***	**0.7527***	**0.5093***	*	0.7516	**0.0201***	0.1930	0.4823
	MSAD1	**0.5847***	**0.5166***	**0.5117***	**0.4148***	**0.7896***	**0.8053***	**0.7448***	**0.7473***	**0.7232***	**0.7669***	**0.7807***	**0.8714***	**0.8239***	**0.7187***	**0.4726***	-0.0162	*	**0.0411***	0.3014	0.4781
	MSAD2	**0.3241***	**0.3525***	**0.3168***	0.1122	**0.5402***	**0.5696***	**0.5143***	**0.5510***	**0.4691***	**0.5788***	**0.5779***	**0.6865***	**0.6025***	**0.4690***	**0.2445***	**0.0817***	**0.0662***	*	0.4846	0.1973
	MISR/14	**0.3766***	**0.3840***	**0.3709***	0.1699	**0.6063***	**0.6601***	**0.5722***	**0.6106***	**0.5374***	**0.6503***	**0.6541***	**0.8273***	**0.7154***	**0.4910***	**0.2157***	0.0625	0.0193	-0.0053	*	0.5846
	MLEV	**0.4886***	**0.4545***	**0.4449***	**0.3265***	**0.8178***	**0.8472***	**0.7523***	**0.7613***	**0.7370***	**0.7917***	**0.8064***	**0.9312***	**0.8765***	**0.6317***	**0.3579***	0.0201	0.0286	0.0495	0.0250	*

Appendix II. Chapter 3

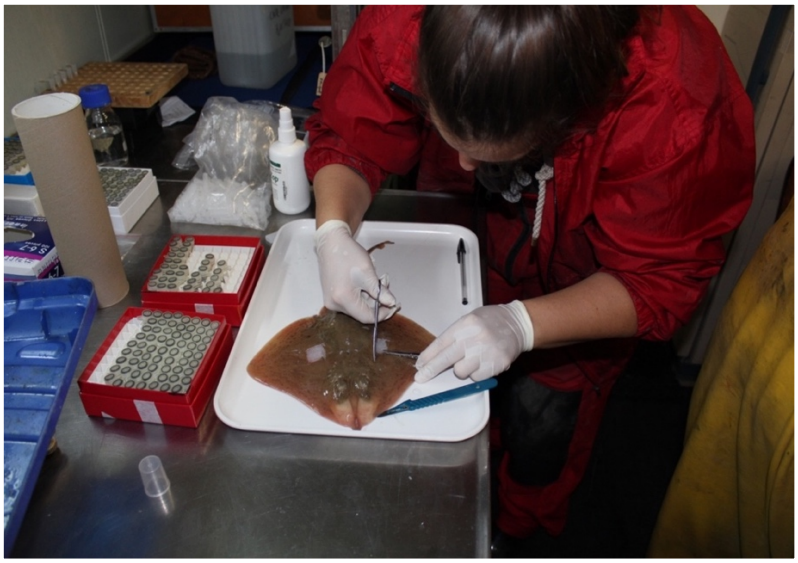


Figure S1 On board sampling procedure (sampling *Raja miraletus* during SoleMon 2014).

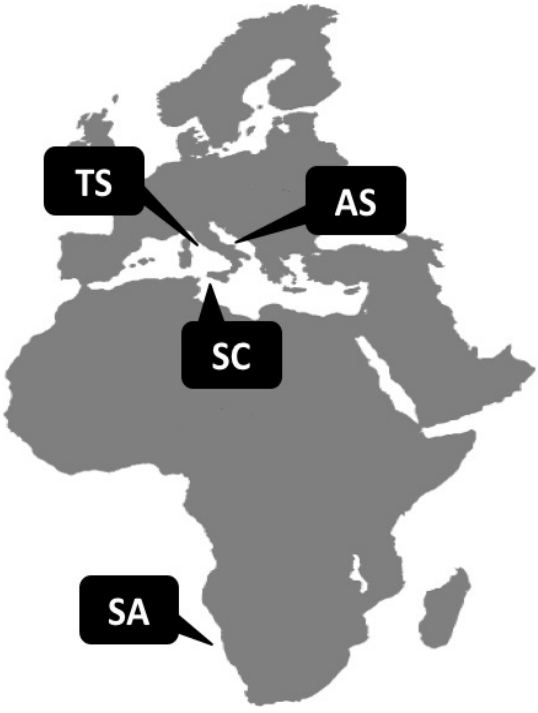


Figure S2 Sampling locations. AS Adriatic Sea, SC Sicilian Channel, TS Tyrrhenian Sea, SA South Africa.

Table S9 Sampling details for each individual. Na not available information.

Sample_ID	Site	Date	Haul	Lat_D_S	Lon_D_S	Depth (m)	Species	Weight (g)	Length (cm)	Disc lenght (cm)	Disc width (cm)	Sex
RAJ01	Chioggia	03/03/2014	land	land	land	Na	*Raja asterias*	Na	60	Na	Na	M
RAJ02	Viareggio	25/07/2014	3	Na	Na	Na	*Raja asterias*	Na	50.5	Na	Na	M
RAJ03	Viareggio	25/07/2015	4	43°48'42''N	10°02'819''E	32	*Raja asterias*	Na	53	Na	Na	M
RAJ04	Viareggio	25/07/2016	5	43°48'85''N	10°01'63''E	35	*Raja asterias*	Na	52	Na	Na	M
RAJ05	Viareggio	25/07/2017	6	43°48'80''N	10°01'00''E	37	*Raja asterias*	Na	50.5	Na	Na	M
RAJ06	Viareggio	25/07/2018	6	43°48'80''N	10°01'00''E	37	*Raja asterias*	Na	53	Na	Na	F
RAJ07	Viareggio	25/07/2019	7	43°48'756''N	10°02'44''E	32	*Raja asterias*	Na	54.5	Na	Na	F
RAJ08	GSA17	16/10/2014	120b	Na	Na	Na	*Raja asterias*	Na	32	Na	Na	F
RAJ09	GSA17	16/10/2014	120b	Na	Na	Na	*Raja asterias*	Na	31	Na	Na	M
RAJ10	GSA17	16/10/2014	122	Na	Na	Na	*Raja asterias*	Na	27	Na	Na	M
RAJ11	GSA17	16/10/2014	123	Na	Na	Na	*Raja asterias*	Na	29	Na	Na	M
RAJ12	GSA17	16/10/2014	123	Na	Na	Na	*Raja asterias*	Na	31.5	Na	Na	F
RAJ13	GSA17	18/10/2014	131	Na	Na	Na	*Raja clavata*	Na	21	Na	Na	M
RAJ14	GSA17	11/10/2014	34	43°24'627"N	13°52'078"E	18	*Raja asterias*	Na	49	Na	Na	F
RAJ15	Viareggio	25/07/2014	1	43°48'56''N	10°04'415''E	26	*Raja asterias*	Na	58	Na	Na	F
RAJ16	GSA17	11/10/2014	33	43°20'317"N	13°44'222"E	8	*Raja asterias*	Na	27.7	Na	Na	M
RAJ17	GSA17	11/10/2014	33	43°20'317"N	13°44'222"E	8	*Raja asterias*	Na	29	Na	Na	F
RAJ18	GSA17	11/10/2014	31	43°34'263"N	13°37'565"E	10	*Raja asterias*	Na	31.5	Na	Na	M
RAJ19	GSA17	11/10/2014	31	43°34'263"N	13°37'565"E	10	*Raja asterias*	Na	34	Na	Na	F
RAJ20	GSA17	13/11/2014	50	43°04'619"N	14°01'448"E	24	*Raja asterias*	Na	33.5	Na	Na	M
RAJ21	GSA17	20/11/2014	53	44°34'697"N	13°32'562"E	40	*Raja clavata*	2702	73.1	37.5	54	F2
RAJ22	GSA17	20/11/2014	53	44°34'697"N	13°32'562"E	40	*Raja clavata*	1124	58.3	27.8	39.2	F1
RAJ23	GSA17	20/11/2014	22	44°25'110"N	13°38'506"E	52	*Raja clavata*	2322	72.5	35.3	46.4	M3
RAJ24	GSA17	21/11/2014	36	43°40'475"N	14°50'656"E	85	*Raja miraletus*	143	33.2	15.3	19.7	F1
RAJ25	GSA17	21/11/2014	36	43°40'475"N	14°50'656"E	85	*Raja miraletus*	40	23.7	6.5	14.5	F1

RAJ26	GSA17	21/11/2014	36	43°40'475"N	14°50'656"E	85	*Raja miraletus*	285	39.2	20.5	24	F2
RAJ27	GSA17	21/11/2014	32	44°00'685"N	14°30'762"E	72.5	*Raja miraletus*	214	35.4	17.5	22.2	M2
RAJ28	GSA17	21/11/2014	32	44°00'685"N	14°30'762"E	72.5	*Raja miraletus*	265	48.2	20.2	24	F
RAJ29	GSA17	15/11/2014	38	43°22'266"N	14°08'540"E	68	*Raja asterias*	Na	52.5	Na	Na	F
RAJ30	GSA17	15/11/2014	35	43°25'494"N	14°13'055"E	75	*Raja asterias*	Na	48	Na	Na	M
RAJ31	GSA17	21/11/2014	32	44°00'685"N	14°30'762"E	72.5	*Raja miraletus*	37	21.9	10.2	13.3	M1
RAJ32	GSA17	21/11/2014	32	44°00'685"N	14°30'762"E	72.5	*Raja miraletus*	57	24.7	11.2	15.6	M1
RAJ34	GSA17	27/11/2015	14	44°45'768"N	13°05'320"E	40.6	*Raja clavata*	504	46	22	29.9	M1
RAJ35	GSA17	27/11/2015	14	44°45'768"N	13°05'320"E	40.6	*Raja clavata*	68	26.1	11.7	15.6	F
RAJ36	GSA17	27/11/2015	14	44°45'768"N	13°05'320"E	40.6	*Raja clavata*	61	21.1	11.2	15	F
RAJ37	GSA17	27/11/2015	13	44°36'022"N	12°41'893"E	33.2	*Raja clavata*	60	22.5	10.7	14.1	M1
RAJ39	GSA16	02/12/2014	34	37°42'28"N	12°21'38"E	146	*Leucoraja melitensis*	Na	39.5	Na	Na	M
RAJ40	GSA16	02/12/2014	34	37°42'28"N	12°21'38"E	146	*Raja miraletus*	Na	44	Na	Na	M
RAJ41	GSA16	02/12/2014	34	37°42'28"N	12°21'38"E	146	*Leucoraja melitensis*	Na	41	Na	Na	F
RAJ42	GSA16	02/12/2014	35	37°42'07"N	12°10'54"E	200	*Leucoraja melitensis*	Na	38	Na	Na	F
RAJ43	GSA16	02/12/2014	35	37°42'07"N	12°10'54"E	200	*Leucoraja melitensis*	Na	34.5	Na	Na	F
RAJ44	GSA16	02/12/2014	36	37°39'97"N	11°54'76"E	98	*Raja miraletus*	Na	33	Na	Na	F
RAJ45	GSA16	02/12/2014	36	37°39'97"N	11°54'76"E	98	*Raja miraletus*	Na	34	Na	Na	M
RAJ46	GSA16	03/12/2014	40	37°19'04"N	11°46'15"E	124	*Raja clavata*	Na	64	Na	Na	F
RAJ47	GSA16	03/12/2014	42	37°13'79"N	12°03'26"E	82	*Raja miraletus*	Na	36	Na	Na	M
RAJ48	GSA16	03/12/2014	42	37°13'79"N	12°03'26"E	82	*Raja miraletus*	Na	39	Na	Na	F
RAJ49	GSA16	03/12/2014	42	37°13'79"N	12°03'26"E	82	*Raja miraletus*	Na	37	Na	Na	F
RAJ50	GSA16	03/12/2014	42	37°13'79"N	12°03'26"E	82	*Raja miraletus*	Na	36	Na	Na	F
RAJ51	GSA16	03/12/2014	42	37°13'79"N	12°03'26"E	82	*Raja clavata*	Na	38	Na	Na	F
RAJ52	GSA16	03/12/2014	43	37°13'35"N	12°06'77"E	81	*Raja polystigma*	Na	32	Na	Na	M

RAJ53	GSA16	03/12/2014	43	37°13'35"N	12°06'77"E	81	*Leucoraja melitensis*	Na	34.5	Na	Na	F
RAJ54	GSA16	04/12/2014	45	36°55'96"N	12°18'54"E	120	*Leucoraja melitensis*	Na	39	Na	Na	F
RAJ55	GSA16	04/12/2014	46	36°57'66"N	12°28'82"E	118	*Raja clavata*	Na	49	Na	Na	F
RAJ56	GSA16	04/12/2014	46	36°57'66"N	12°28'82"E	118	*Raja miraletus*	Na	26	Na	Na	F
RAJ57	GSA16	05/12/2014	53	37°00'34"N	13°00'84"E	269	*Raja clavata*	Na	32	Na	Na	F
RAJ58	GSA16	07/12/2014	58	37°21'60"N	12°18'94"E	74	*Raja brachyura*	Na	35	Na	Na	M
RAJ59	GSA16	07/12/2014	60	37°34'33"N	12°17'92"E	66	*Raja clavata*	Na	37	Na	Na	F
RAJ60	GSA16	07/12/2014	60	37°34'33"N	12°17'92"E	67	*Raja clavata*	Na	36	Na	Na	F
RAJ61	SC	16/04/2015	D00504	34° 41.06'S	021° 18.97'E	68	*R. ocellifera*	370	39	27	27	F
RAJ62	SC	16/04/2015	D00505	34° 30.65'S	021° 23.94'E	64	*R. ocellifera*	260	34	23	23	F
RAJ63	SC	16/04/2015	D00505	34° 30.65'S	021° 23.94'E	64	*R. ocellifera*	720	46	32	32	F
RAJ64	SC	16/04/2015	D00505	34° 30.65'S	021° 23.94'E	64	*R. ocellifera*	1025	52	35	35	F
RAJ65	SC	16/04/2015	D00505	34° 30.65'S	021° 23.94'E	64	*R. ocellifera*	270	35	25	25	M
RAJ66	SC	16/04/2015	D00505	34° 30.65'S	021° 23.94'E	64	*R. ocellifera*	265	36	23	23	M
RAJ67	SC	16/04/2015	D00505	34° 30.65'S	021° 23.94'E	64	*R. ocellifera*	355	40	37	37	M
RAJ68	SC	16/04/2015	D00505	34° 30.65'S	021° 23.94'E	64	*R. ocellifera*	230	33	23	23	M
RAJ69	SC	30/04/2015	D00559	33° 44.94'S	026° 04.88'E	33	*R. ocellifera*	357	36	26	26	F
RAJ70	SC	30/04/2015	D00559	33° 44.94'S	026° 04.88'E	33	*R. ocellifera*	175	30	21	21	M
RAJ71	SC	30/04/2015	D00559	33° 44.94'S	026° 04.88'E	33	*R. ocellifera*	101	26	18	18	M
RAJ72	SC	30/04/2015	D00559	33° 44.94'S	026° 04.88'E	33	*R. ocellifera*	75	24	16	16	M
RAJ73	SC	04/12/2015	D00488	35° 19.62'S	020° 12.87'E	136	*Raja straeleni*	955	51.5	37	37	M
RAJ74	SC	04/12/2015	D00488	35° 19.62'S	020° 12.87'E	136	*Raja straeleni*	1575	60	43	43	M
RAJ75	SC	04/12/2015	D00488	35° 19.62'S	020° 12.87'E	136	*Raja straeleni*	1450	52	39	39	F
RAJ76	SC	04/12/2015	D00488	35° 19.62'S	020° 12.87'E	136	*Raja straeleni*	2970	69	53	53	F
RAJ77	SC	15/4/2015	D00499	35° 11.55'S	020° 24.61'E	115	*Raja straeleni*	2401	69	48	48	M
RAJ78	SC	15/4/2015	D00499	35° 11.55'S	020° 24.61'E	115	*Raja straeleni*	2305	67	47	47	M
RAJ79	SC	15/4/2015	D00499	35° 11.55'S	020° 24.61'E	115	*Raja straeleni*	1935	60	46	46	F

RAJ80	SC	15/4/2015	D00499	35° 11.55'S	020° 24.61'E	115	*Raja straeleni*	2035	66	45	45	M
RAJ81	SC	15/4/2015	D00499	35° 11.55'S	020° 24.61'E	115	*Raja straeleni*	2580	69	47	47	M
RAJ82	SC	22/04/2015	D00520	35° 56.90'S	021° 48.66'E	178	*Raja straeleni*	2690	67	50	50	F
RAJ83	SC	22/04/2015	D00520	35° 56.90'S	021° 48.66'E	178	*Raja straeleni*	1426	58	41	41	M
RAJ84	SC	22/04/2015	D00520	35° 56.90'S	021° 48.66'E	178	*Raja straeleni*	1992	62	47	47	F
RAJ85	GSA9	13/07/2015	104	43°11'661''N	09°51'28"E	188	*Raja polystigma*	80	27	Na	Na	M1
RAJ86	GSA9	13/07/2015	104	43°11'661''N	09°51'28"E	188	*Raja polystigma*	160	33.5	Na	Na	M1
RAJ87	GSA9	13/07/2015	104	43°11'661''N	09°51'28"E	188	*Raja polystigma*	140	31	Na	Na	F1
RAJ88	GSA9	13/07/2015	104	43°11'661''N	09°51'28"E	188	*Raja polystigma*	180	37	Na	Na	M1
RAJ89	GSA9	13/07/2015	104	43°11'661''N	09°51'28"E	188	*Raja polystigma*	160	31	Na	Na	F1
RAJ90	GSA9	13/07/2015	104	43°11'661''N	09°51'28"E	188	*Raja polystigma*	180	34.5	Na	Na	F1
RAJ91	GSA9	13/07/2015	104	43°11'661''N	09°51'28"E	188	*Raja polystigma*	120	30	Na	Na	F1
RAJ92	GSA9	13/07/2015	104	43°11'661''N	09°51'28"E	188	*Raja polystigma*	60	24.5	Na	Na	M1
RAJ93	GSA9	13/07/2015	104	43°11'661''N	09°51'28"E	188	*Raja polystigma*	60	25.5	Na	Na	F1
RAJ94	GSA9	13/07/2015	104	43°11'661''N	09°51'28"E	188	*Raja polystigma*	80	26.5	Na	Na	M1

Tab S10 List of RNA extractions and details about homogenation method, protocol used and measures of purity and RIN (when performed, i.e. na not available because not measured, N/A not detected by the Instrument).

Sample ID	Tissue	Extraction tube ID	Homogenization	beating m/s x time	Prurification	Initial centrifuge	DNAse	Washes	Other	Elution	ng/ul	260/280	260/230	RIN
RAJ01 Raja asterias	pseudo eyespot	R01_Sa	UTurrax+750ul trizol		Standard Trizol	yes	yes	Standard Trizol		25ul	49.00	na	na	na
	matrix	R01_Da	UTurrax+750ul trizol		Standard Trizol	yes	yes	Standard Trizol		25ul	62.00	na	na	na
	ventral	R01_Wa	UTurrax+750ul trizol		Standard Trizol	yes	yes	Standard Trizol		25ul	40.00	na	na	na
RAJ01 Raja asterias	pseudo eyespot	R01_Sb	UTurrax+750ul trizol		Zymo	yes	yes	kit		25ul	350.00	na	na	na
	matrix	R01_Db	UTurrax+750ul trizol		Zymo	yes	yes	kit		25ul	323.00	na	na	na
	ventral	R01_Wb	UTurrax+750ul trizol		Zymo	yes	yes	kit		25ul	400.00	na	na	na
RAJ39 Leucoraja melitensis	eyespot	R39_Sa	UTurrax+750ul trizol		Standard Trizol	yes	yes	Standard Trizol		25ul	25.00	na	na	na
	matrix	R39_Da	UTurrax+750ul trizol		Standard Trizol	yes	yes	Standard Trizol		25ul	48.00	na	na	na
	ventral	R39_Wa	UTurrax+750ul trizol		Standard Trizol	yes	yes	Standard Trizol		25ul	53.00	na	na	na
RAJ39 Leucoraja melitensis	eyespot	R39_Sb	UTurrax+750ul trizol		Zymo	yes	yes	kit		25ul	291.00	na	na	na
	matrix	R39_Db	UTurrax+750ul trizol		Zymo	yes	yes	kit		25ul	77.00	na	na	na
	ventral	R39_Wb	UTurrax+750ul trizol		Zymo	yes	yes	kit		25ul	140.00	na	na	na
RAJ39 Leucoraja melitensis	eyespot	R39_Sc	Scissors, pestle		Maxwell16					60ul	40.00	na	na	na
	matrix	R39_Dc	Scissors, pestle		Maxwell16					60ul	56.00	na	na	na
	ventral	R39_Wc	Scissors, pestle		Maxwell16					60ul	4.00	na	na	na
	eyespot	1	beads+750ul trizol O/N	6 m/s 60''	beads normal	no	no	1ml etoh 75% once		50ul	344.10	1.75	1.60	na

RAJ41 Leucoraja melitensis	matrix	2	beads+750ul trizol O/N	6 m/s 60''	beads normal	no	no	1ml etoh 75% once		50ul	249.90	1.78	1.28	na
	ventral	3	beads+750ul trizol O/N	6 m/s 60''	beads normal	no	no	1ml etoh 75% once		50ul	299.80	1.73	1.58	na
RAJ42 Leucoraja melitensis	eyespot	4a	beads+750ul trizol O/N	6 m/s 60''	Zymo	no	yes	kit		30ul	8.20	1.81	0.43	na
	eyespot	4b	beads+750ul trizol O/N	6 m/s 60''	Zymo	no	yes	kit		30ul	3.60	1.81	1.20	na
	matrix	5a	beads+750ul trizol O/N	6 m/s 60''	Zymo	no	yes	kit		30ul	3.60	1.51	1.04	na
	matrix	5a	beads+750ul trizol O/N	6 m/s 60''	Zymo	no	yes	kit		30ul	4.30	1.29	1.18	na
	ventral	6a	beads+750ul trizol O/N	6 m/s 60''	Zymo	no	yes	kit		30ul	21.80	2.03	2.04	na
	ventral	6b	beads+750ul trizol O/N	6 m/s 60''	Zymo	no	yes	kit		30ul	25.60	1.77	1.86	na
RAJ43 Leucoraja melitensis	eyespot	7	beads+1ml trizol O/N	6 m/s 60''	beads	yes	no	double, 1ml+500ul etoh 75%		50ul	444.65	1.71	1.40	na
	matrix	8	beads+1ml trizol O/N	6 m/s 60''	beads	yes	no	double, 1ml+500ul etoh 75%		50ul	182.41	1.71	1.49	na
	ventral	9	beads+1ml trizol O/N	6 m/s 60''	beads	yes	no	double, 1ml+500ul etoh 75%		50ul	60.50	1.86	1.86	na
RAJ53 Leucoraja melitensis	eyespot	10a	beads+1ml trizol O/N	6 m/s 60''	Zymo	yes	yes	kit	350 trizol+350 abs EtOH	30ul	89.73	1.90	1.86	na
	eyespot	10b	beads+1ml trizol O/N	6 m/s 60''	Zymo	yes	yes	kit	650ul trizol+200ul chlorf+400ul abs EtOH	30ul	345.40	1.86	1.59	na
	matrix	11a	beads+1ml trizol O/N	6 m/s 60''	Zymo	yes	yes	kit	350ul trizol+350ul abs EtOH	30ul	36.09	1.96	2.05	na

	matrix	11b	beads+1ml trizol O/N	6 m/s 60''	Zymo	yes	yes	kit	650ul trizol+200ul chlorf+400ul abs EtOH	30ul	198.71	1.96	1.96	na
	ventral	12a	beads+1ml trizol O/N	6 m/s 60''	Zymo	yes	yes	kit	350ul trizol+350ul abs EtOH	30ul	112.91	2.00	2.20	na
	ventral	12b	beads+1ml trizol O/N	6 m/s 60''	Zymo	yes	yes	kit	650ul trizol+200ul chlorf+400ul abs EtOH	30ul	213.50	2.02	2.23	na
RAJ54 Leucoraja melitensis	eyespot	13	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit	200ul chlorf+ 400ul abs EtOH	30ul	298.30	1.90	1.57	na
	matrix	14	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit	200ul chlorf+ 400ul abs EtOH	30ul	319.60	1.94	1.91	na
	ventral	15	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit	200ul chlorf+ 400ul abs EtOH	30ul	368.20	2.03	2.25	na
RAJ02 Raja asterias	pseudo eyespot	16	beads+750ul trizol	4 m/s 50''	Zymo	yes	yes	kit	200ul chlorf+ 400ul abs EtOH	30ul	120.50	1.64	1.15	5.50
	matrix	17	beads+750ul trizol	4 m/s 50''	Zymo	yes	yes	kit	200ul chlorf+ 400ul abs EtOH	30ul	95.50	1.71	1.35	3.40
	ventral	18	beads+750ul trizol	4 m/s 50''	Zymo	yes	yes	kit	200ul chlorf+ 400ul abs EtOH	30ul	296.90	2.02	2.25	N/A
RAJ03 Raja asterias	pseudo eyespot	19	beads+750ul trizol	4 m/s 50''	Zymo	yes	yes	kit	200ul chlorf+ 400ul abs EtOH	30ul	84.70	1.79	1.54	na
	matrix	20	beads+750ul trizol	4 m/s 50''	Zymo	yes	yes	kit	200ul chlorf+ 400ul abs EtOH	30ul	99.20	1.80	1.52	na

	ventral	21	beads+750ul trizol	4 m/s 50''	Zymo	yes	yes	kit	200ul chlorf+ 400ul abs EtOH	30ul	147.00	2.04	2.27	na
RAJ04 Raja asterias	pseudo eyespot	22	beads+750ul trizol	4 m/s 50''	Zymo	yes	yes	kit	200ul chlorf+ 400ul abs EtOH	30ul	77.20	1.80	1.56	6.20
	matrix	23	beads+750ul trizol	4 m/s 50''	Zymo	yes	yes	kit	200ul chlorf+ 400ul abs EtOH	30ul	232.90	1.72	1.40	na
	ventral	24	beads+750ul trizol	4 m/s 50''	Zymo	yes	yes	kit	200ul chlorf+ 400ul abs EtOH	30ul	271.20	2.03	2.24	na
RAJ05 Raja asterias	pseudo eyespot	25	beads+750ul trizol	4 m/s 50''x2 20''	Zymo	yes	yes	kit	200ul chlorf+ 400ul abs EtOH	30ul	178.00	1.77	1.47	na
	matrix	26	beads+750ul trizol	4 m/s 50''x2 20''	Zymo	yes	yes	kit	200ul chlorf+ 400ul abs EtOH	30ul	263.10	1.71	1.40	na
	ventral	27	beads+750ul trizol	4 m/s 50''x2 20''	Zymo	yes	yes	kit	200ul chlorf+ 400ul abs EtOH	30ul	264.20	2.00	2.24	na
RAJ10 Raja asterias	pseudo eyespot	28	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 400ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	210.50	1.88	1.74	N/A
	matrix	29	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 400ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	311.30	1.91	1.80	7.20
	ventral	30	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 400ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	100.70	1.98	2.02	na
	pseudo eyespot	31	beads+750ul trizol	6 m/s 60''x2	Zymo	yes	yes	kit + 400ul WB (4	200ul chlorf+ 400ul abs EtOH	30ul	465.00	2.05	2.08	7.30

							washes tot)						
RAJ21 Raja clavata	matrix	32	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 400ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	244.20	2.00	2.13	N/A
	ventral	33	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 400ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	267.70	2.05	2.25	7.10
RAJ22 Raja clavata	pseudo eyespot	34	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 400ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	302.90	1.94	1.98	N/A
	matrix	35	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 400ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	287.70	2.00	2.08	6.90
	ventral	36	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 400ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	485.40	2.03	2.25	8.00
RAJ23 Raja clavata	pseudo eyespot	37	beads+750ul trizol	6 m/s 60''x2	Zymo	yes	yes	kit + 400ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	293.60	1.98	1.97	N/A
	matrix	38	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 400ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	119.70	1.98	1.84	N/A
	ventral	39	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 400ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	118.20	2.06	2.24	N/A

RAJ34 Raja clavata	pseudo eyespot	40	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 500ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	459.10	1.98	1.98	2.70
	matrix	41	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 500ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	572.40	1.99	1.94	N/A
	ventral	42	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 500ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	241.90	2.06	2.26	N/A
RAJ35 Raja clavata	pseudo eyespot	43	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 500ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	357.80	1.96	1.74	na
	matrix	44	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 500ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	459.10	1.93	1.79	na
	ventral	45	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 500ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	434.40	2.03	2.17	na
RAJ24 Raja miraletus	eyespot	46	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 500ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	216.60	1.89	1.59	na
	matrix	47	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 500ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	343.70	1.85	1.50	na
	ventral	48	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 500ul WB (4	200ul chlorf+ 400ul abs EtOH	30ul	361.20	2.04	2.23	na

								washes tot)						
RAJ26 Raja miraletus	eyespot	49	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit + 500ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	138.00	1.81	1.41	na
	matrix	50	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 500ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	201.60	1.66	1.10	na
	ventral	51	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit + 500ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	178.70	2.05	2.14	na
RAJ24 Raja miraletus	eyespot	52	beads+1000ul trizol	4 m/s 60''	Zymo	yes	yes	3 prewash+2 wash	200ul chlorf+ 400ul abs EtOH	30ul	42.80	2.04	2.06	N/A
	matrix	53	beads+1000ul trizol	4 m/s 60''	Zymo	yes	yes	3 prewash+2 wash	200ul chlorf+ 400ul abs EtOH	30ul	136.30	2.02	1.98	N/A
	ventral	54	beads+1000ul trizol	4 m/s 60''	Zymo	yes	yes	3 prewash+2 wash	200ul chlorf+ 400ul abs EtOH	30ul	123.10	2.09	2.21	2.40
RAJ26 Raja miraletus	eyespot	55	beads+1000ul trizol	4 m/s 60''	Zymo	yes	yes	3 prewash+2 wash	200ul chlorf+ 400ul abs EtOH	30ul	23.70	1.80	1.48	na
	matrix	56	beads+1000ul trizol	4 m/s 60''	Zymo	yes	yes	3 prewash+2 wash	200ul chlorf+ 400ul abs EtOH	30ul	46.40	1.69	1.13	na
	ventral	57	beads+1000ul trizol	4 m/s 60''	Zymo	yes	yes	3 prewash+2 wash	200ul chlorf+ 400ul abs EtOH	30ul	63.30	2.00	1.92	na
RAJ27 Raja miraletus	eyespot	58	beads+1000ul trizol	4 m/s 60''	Zymo	yes	yes	3 prewash+2 wash	200ul chlorf+ 400ul abs EtOH	30ul	49.30	1.77	1.44	na

	matrix	59	beads+1000ul trizol	4 m/s 60''	Zymo	yes	yes	3 prewash+2 wash	200ul chlorf+ 400ul abs EtOH	30ul	23.50	1.69	1.40	na
	ventral	60	beads+1000ul trizol	4 m/s 60''	Zymo	yes	yes	3 prewash+2 wash	200ul chlorf+ 400ul abs EtOH	30ul	72.50	1.98	2.18	na
RAJ28 Raja miraletus	eyespot	61	beads+1000ul trizol	4 m/s 60''	Zymo	yes	yes	3 prewash+2 wash	200ul chlorf+ 400ul abs EtOH	30ul	31.10	1.73	1.32	na
	matrix	62	beads+1000ul trizol	4 m/s 60''	Zymo	yes	yes	3 prewash+2 wash	200ul chlorf+ 400ul abs EtOH	30ul	26.99	1.69	1.38	na
	ventral	63	beads+1000ul trizol	4 m/s 60''	Zymo	yes	yes	3 prewash+2 wash	200ul chlorf+ 400ul abs EtOH	30ul	174.10	2.05	2.25	na
RAJ40 Raja miraletus	eyespot	64	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	38.70	1.82	1.50	6.80
	matrix	65	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	130.50	1.80	1.39	N/A
	ventral	66	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	140.70	2.03	2.16	9.30
RAJ44 Raja miraletus	eyespot	67	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	100.50	1.86	1.48	7.70
	matrix	68	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	110.50	1.86	1.44	8.80

	ventral	69	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	168.80	2.04	2.19	9.10
RAJ45 Raja miraletus	eyespot	70	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	197.00	1.83	1.36	8.50
	matrix	71	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	137.80	1.75	1.22	7.50
	ventral	72	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	230.70	2.05	2.19	N/A
RAJ47 Raja miraletus	eyespot	73	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	95.50	1.89	1.52	N/A
	matrix	74	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	106.40	1.87	1.47	7.80
	ventral	75	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	204.10	2.05	2.20	9.30
RAJ61 Raja ocellifera	eyespot	76	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	78.30	1.99	2.06	na
	matrix	77	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4	200ul chlorf+ 400ul abs EtOH	30ul	100.70	2.00	2.01	na

								washes tot)					
	ventral	78	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	268.90	2.01	2.28	na
RAJ62 Raja ocellifera	eyespot	79	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	127.40	2.00	2.10	na
	matrix	80	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	62.20	1.96	2.10	na
	ventral	81	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	260.80	2.05	2.27	na
RAJ63 Raja ocellifera	eyespot	82	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	41.90	1.97	1.63	na
	matrix	83	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	29.10	1.86	1.61	na
	ventral	84	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	167.30	2.06	2.25	na
RAJ64 Raja ocellifera	eyespot	85	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	16.90	1.96	1.36	na

	matrix	86	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	8.30	2.44	1.13	na
	ventral	87	beads+750ul trizol	4 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	78.40	2.11	2.24	na
RAJ73 Raja straeleni	eyespot	88	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	322.70	1.97	1.91	7.30
	matrix	89	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	210.10	1.95	1.88	6.60
	ventral	90	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	215.10	2.05	2.21	7.70
RAJ74 Raja straeleni	eyespot	91	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	228.70	1.92	1.79	8.10
	matrix	92	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	187.60	1.95	1.76	6.80
	ventral	93	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	371.69	2.03	2.29	6.70
	eyespot	94	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4	200ul chlorf+ 400ul abs EtOH	30ul	295.30	1.95	1.95	8.30

							washes tot)						
RAJ75 Raja straeleni	matrix	95	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	350.70	1.88	1.58	8.60
	ventral	96	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	340.90	2.03	2.26	6.40
RAJ76 Raja straeleni	eyespot	97	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	203.30	1.97	2.04	7.50
	matrix	98	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	377.60	1.96	1.99	8.00
	ventral	99	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	456.50	2.04	2.26	8.50
RAJ65 Raja ocellifera	eyespot	100	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	186.10	1.91	1.76	6.40
	matrix	101	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	212.00	1.85	1.53	6.50
	ventral	102	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	175.30	2.04	2.24	8.10

RAJ77 Raja straeleni	eyespot	103	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	178.20	1.97	1.92	6.60
	matrix	104	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	371.30	1.91	1.64	8.00
	ventral	105	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	285.10	1.99	2.00	6.80
RAJ19 Raja asterias	pseudo eyespot	106	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	454.60	1.95	1.89	7.10
	matrix	107	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	500.80	1.98	1.87	6.50
	ventral	108	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	382.30	2.06	2.30	6.10
RAJ20 Raja asterias	pseudo eyespot	109	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	636.00	1.94	1.80	7.80
	matrix	110	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	554.40	1.91	1.72	8.10
	ventral	111	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4	200ul chlorf+ 400ul abs EtOH	30ul	291.50	1.99	1.91	7.20

								washes tot)					
RAJ66 Raja ocellifera	eyespot	112	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	199.40	1.99	2.07	7.20
	matrix	113	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	244.00	1.94	1.84	7.80
	ventral	114	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	230.30	2.04	2.26	6.80
RAJ67 Raja ocellifera	eyespot	115	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	167.60	1.86	1.62	6.20
	matrix	116	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	169.70	1.84	1.47	6.20
	ventral	117	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	265.80	2.05	2.32	7.40
RAJ68 Raja ocellifera	eyespot	118	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	208.70	1.85	1.45	6.60
	matrix	119	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	220.90	1.82	1.36	5.40

	ventral	120	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	166.60	1.99	2.26	6.20
RAJ69 Raja ocellifera	eyespot	121	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	184.60	1.86	1.57	6.80
	matrix	122	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	376.50	1.88	1.65	7.10
	ventral	123	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	226.80	1.95	1.74	6.90
RAJ11 Raja asterias	pseudo eyespot	124	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	305.60	1.91	1.73	6.30
	matrix	125	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	286.20	1.77	1.81	5.90
	ventral	126	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	277.10	2.05	2.23	6.00
RAJ14 Raja asterias	pseudo eyespot	127	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	481.50	1.94	1.67	7.90
	matrix	128	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4	200ul chlorf+ 400ul abs EtOH	30ul	434.70	1.89	1.67	6.80

								washes tot)					
	ventral	129	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	404.40	2.03	2.17	8.50
RAJ17 Raja asterias	pseudo eyespot	130	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	482.00	1.99	1.86	6.80
	matrix	131	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	483.60	1.97	1.81	6.40
	ventral	132	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	380.20	2.04	2.26	6.90
RAJ46 Raja clavata	pseudo eyespot	133	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	198.10	2.01	2.08	8.90
	matrix	134	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	363.00	2.00	2.06	8.80
	ventral	135	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	339.50	2.06	2.28	8.80
RAJ55 Raja clavata	pseudo eyespot	136	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	169.00	1.94	2.11	7.10

	matrix	137	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	236.40	1.93	2.05	7.10
	ventral	138	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	144.20	1.98	2.18	7.80
RAJ57 Raja clavata	pseudo eyespot	139	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	178.80	1.94	2.15	N/A
	matrix	140	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	332.00	1.92	1.97	6.80
	ventral	141	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	201.70	2.00	2.20	7.70
RAJ56 Raja miraletus	eyespot	142	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	382.20	1.85	1.64	N/A
	matrix	143	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	332.10	1.86	1.75	6.60
	ventral	144	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	312.50	1.92	2.19	5.40
	pseudo eyespot	145	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4	200ul chlorf+ 400ul abs EtOH	30ul	370.70	1.90	1.92	N/A

								washes tot)					
RAJ18 Raja asterias	matrix	146	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	384.00	1.88	1.76	6.90
	ventral	147	beads+750ul trizol	6 m/s 60''	Zymo	yes	yes	kit +700ul WB (4 washes tot)	200ul chlorf+ 400ul abs EtOH	30ul	189.00	2.03	2.21	6.90

Table S11 Code conversions and well position of samples in the Illumina sequencing plate

Sample ID	Extraction tube ID	4 digit code	Well position
RAJ21 Raja clavata	31	R21S	B7
	32	R21D	E1
	33	R21W	H7
RAJ22 Raja clavata	34	R22S	F1
	35	R22D	C6
	36	R22W	C2
RAJ40 Raja miraletus	64	R40S	B3
	65	R40D	G2
	66	R40W	B4
RAJ44 Raja miraletus	67	R44S	A2
	68	R44D	C3
	69	R44W	D8
RAJ45 Raja miraletus	70	R45S	D4
	71	R45D	D1
	72	R45W	E5
RAJ47 Raja miraletus	73	R47S	F6
	74	R47D	C7
	75	R47W	A6
RAJ73 Raja straeleni	88	R73S	F8
	89	R73D	D9
	90	R73W	E9
RAJ74 Raja straeleni	91	R74S	G4
	92	R74D	A3
	93	R74W	A10
RAJ75 Raja straeleni	94	R75S	A5
	95	R75D	C4
	96	R75W	B9
RAJ76 Raja straeleni	97	R76S	C8
	98	R76D	E8
	99	R76W	B10
RAJ65 Raja ocellifera	100	R65S	F2
	101	R65D	G8
	102	R65W	D6
RAJ77 Raja straeleni	103	R77S	H6
	104	R77D	H3
	105	R77W	F3
RAJ19 Raja asterias	106	R19S	C5
	107	R19D	D5
	108	R19W	C9

Sample ID	Extraction tube ID	4 digit code	Well position
RAJ20 Raja asterias	109	R20S	A7
	110	R20D	E4
	111	R20W	C1
RAJ66 Raja ocellifera	112	R66S	F5
	113	R66D	G1
	114	R66W	A9
RAJ67 Raja ocellifera	115	R67S	C10
	116	R67D	G7
	117	R67W	B1
RAJ68 Raja ocellifera	118	R68S	E3
	119	R68D	G3
	120	R68W	G5
RAJ69 Raja ocellifera	121	R69S	A1
	122	R69D	H2
	123	R69W	H9
RAJ11 Raja asterias	124	R11S	E7
	125	R11D	B8
	126	R11W	A8
RAJ14 Raja asterias	127	R14S	A4
	128	R14D	D2
	129	R14W	H4
RAJ17 Raja asterias	130	R17S	G9
	131	R17D	H5
	132	R17W	B6
RAJ46 Raja clavata	133	R46S	H1
	134	R46D	D3
	135	R46W	H8
RAJ55 Raja clavata	136	R55S	G6
	137	R55D	B2
	138	R55W	D7
RAJ57 Raja clavata	139	R57S	F4
	140	R57D	F7
	141	R57W	E2
RAJ56 Raja miraletus	142	R56S	B5
	143	R56D	F9
	144	R56W	E6

Table S12 Total number of survived and dropped sequences per sample after clipping adapters with Trimmomatic v.0.36.

	SAMPLE ID	INPUT	SURVIVING	DROPPED	%
Raja asterias	R11D	13838380	13766724	71656	-0.52%
	R11S	12820964	12692278	128686	-1.00%
	R11W	12938667	12884993	53674	-0.41%
	R14D	12308968	12247765	61203	-0.50%
	R14S	13273379	13212389	60990	-0.46%
	R14W	14221010	14181371	39639	-0.28%
	R17D	12650280	12595214	55066	-0.44%
	R17S	14036178	13965105	71073	-0.51%
	R17W	12521340	12452130	69210	-0.55%
	R19D	13685146	13626808	58338	-0.43%
	R19S	15457286	15409839	47447	-0.31%
	R19W	10825389	10717434	107955	-1.00%
	R20D	11675374	11636017	39357	-0.34%
	R20S	15022304	14965929	56375	-0.38%
	R20W	10884592	10826177	58415	-0.54%
Raja clavata	R21D	11559056	11491959	67097	-0.58%
	R21S	10663782	10621709	42073	-0.39%
	R21W	12788412	12661664	126748	-0.99%
	R22D	12423020	12368580	54440	-0.44%
	R22S	12637239	12555814	81425	-0.64%
	R22W	11287016	11233390	53626	-0.48%
	R46D	11524795	11452729	72066	-0.63%
	R46S	13567118	12895465	671653	-4.95%
	R46W	17605423	17492773	112650	-0.64%
	R55D	10773215	10707136	66079	-0.61%
	R55S	12175734	5551256	6624478	-54.4%
	R55W	11823530	11733759	89771	-0.76%
	R57D	12702711	12653209	49502	-0.39%
	R57S	12162639	12119674	42965	-0.35%
	R57W	10269441	10220217	49224	-0.48%
Raja straeleni	R73D	12671021	12586713	84308	-0.67%
	R73S	13953473	13911868	41605	-0.30%
	R73W	14302922	14216173	86749	-0.61%
	R74D	10177884	10078912	98972	-0.97%
	R74S	13285957	13237445	48512	-0.37%
	R74W	14348183	14297047	51136	-0.36%
	R75D	13199501	13151270	48231	-0.37%
	R75S	14730531	14674842	55689	-0.38%
	R75W	12438760	12377423	61337	-0.49%
	R76D	12522187	12481762	40425	-0.32%
	R76S	13161702	13113927	47775	-0.36%
	R76W	13810068	13769954	40114	-0.29%

	R77D	10611149	10538983	72166	-0.68%
	R77S	13193103	13089781	103322	-0.78%
	R77W	9349305	9279918	69387	-0.74%
Raja miraletus	R40D	11712421	11617861	94560	-0.81%
	R40S	10321698	10249532	72166	-0.70%
	R40W	13110491	13042902	67589	-0.52%
	R44D	10682845	10612763	70082	-0.66%
	R44S	12030188	11953121	77067	-0.64%
	R44W	12935873	12884286	51587	-0.40%
	R45D	10506216	10439014	67202	-0.64%
	R45S	11202695	11157757	44938	-0.40%
	R45W	13389969	13342995	46974	-0.35%
	R47D	13023376	12974573	48803	-0.37%
	R47S	13336447	10441862	2894585	-21.7%
	R47W	14244306	14186086	58220	-0.41%
	R56D	12416633	12355800	60833	-0.49%
	R56S	14400342	14340160	60182	-0.42%
	R56W	9285364	9186428	98936	-1.07%
Raja ocellifera	R65D	11660428	11261008	399420	-3.43%
	R65S	10674548	10614074	60474	-0.57%
	R65W	14072998	14020251	52747	-0.37%
	R66D	14142710	14054206	88504	-0.63%
	R66S	12386643	12338122	48521	-0.39%
	R66W	11899227	11733109	166118	-1.40%
	R67D	11647995	11555265	92730	-0.80%
	R67S	12664324	12624869	39455	-0.31%
	R67W	10693108	9241174	1451934	-13.6%
	R68D	9325192	9241174	84018	-0.90%
	R68S	10596473	10504819	91654	-0.86%
	R68W	12372117	12309709	62408	-0.50%
	R69D	13609891	13533207	76684	-0.56%
	R69S	9263906	9188707	75199	-0.81%
	R69W	14823518	14656260	167258	-1.13%

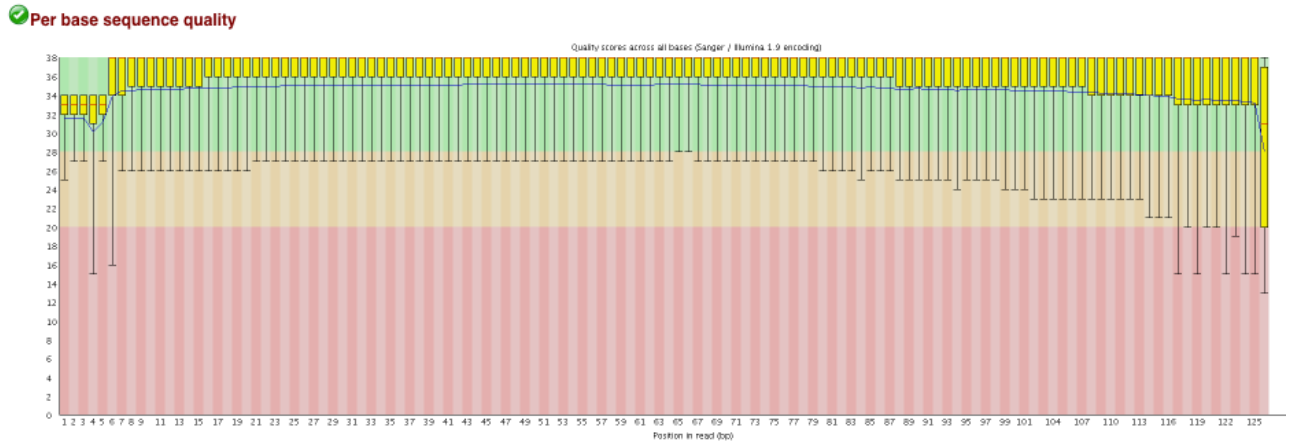


Figure S3 Plot illustrating the 'Per base sequence quality' of raw sequences generated from sample R55S, corresponding to the pseudo-eyespot ornament of *Raja clavata*.

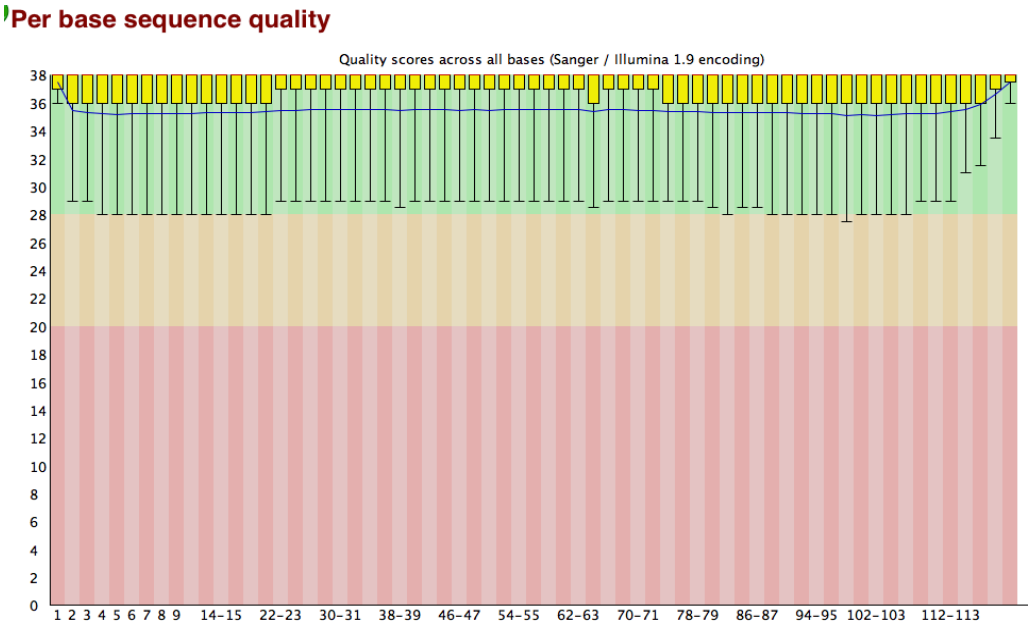


Figure S4 Plot illustrating the 'Per base sequence quality' of trimmed and clean sequences characterising sample R55S, corresponding to the pseudo-eyespot ornament of *Raja clavata*.

Appendix III. Chapter 4

Table S13 Sampling details for each individual, including the tissues collected.

Sample_ID	Site	Date	Haul	Lat_D_S	Lon_D_S	Depth	*Species*	Weight	Length (cm)	Disc length (cm)	Disc Width (cm)	Sex	Tissue
RAJG133SD	GSA17	18/11/15	32	44°00'73"N	14°32'35"E	72.5	*Raja miraletus*	250	39	N.a.	N.a.	M/adult	SKIN1
RAJG133SS													SKIN2
RAJC133SW													SKIN3
RAJC133M													MUSCLE
RAJF133T													GONAD
RAJF133L													LIVER
RAJG133H													HEART
RAJG133B													BRAIN
RAJC133D													DORSAL FIN
RAJC133E													EYE
RAJF133J													JAW BONE
RAJF133G													GILL
RAJG134SD	GSA17	18/11/15	32	44°00'73"N	14°32'35"E	72.5	*Raja miraletus*	245	38	N.a.	N.a.	F/adult	SKIN1
RAJG134SS													SKIN2
RAJC134SW													SKIN3
RAJC134M													MUSCLE
RAJF134T													GONAD
RAJF134L													LIVER
RAJG134H													HEART
RAJG134B													BRAIN
RAJC134D													DORSAL FIN
RAJC134E													EYE
RAJF134J													JAW BONE
RAJF134G													GILL

Table S14 List of RNA extractions, purity and integrity.

	Raja miraletus			NANODROP			BIOANALYZER	
	TUBE CODE	Sample ID	Tissue	ng/ul	260/280	260/230	ng/ul	RIN
Male	RAJG133SD	18T	SKIN1 D	110.60	1.71	1.43	93.00	N/A
	RAJG133SS	19T	SKIN2 S	78.80	1.74	1.34	66.00	N/A
	RAJC133SW	20T	SKIN3 W	90.50	1.95	1.94	72.00	8.40
	RAJC133M	16T	MUSCLE	84.20	1.73	1.17	44.00	8.40
	RAJF133T	13T	TESTICLE	1021.60	1.99	2.32	1428.00	8.00
	RAJF133L	15T	LIVER	309.30	1.89	1.82	221.00	9.00
	RAJG133H	24T	HEART	161.40	1.90	2.01	174.00	8.80
	RAJG133B	14T	BRAIN	163.30	1.94	2.09	175.00	9.30
	RAJC133D	17T	DORSAL FIN	140.00	1.65	1.39	157.00	7.20
	RAJC133E	21T	EYE	117.60	1.76	1.36	197.00	6.50
	RAJF133J	22T	JAW BONE	498.70	2.01	2.21	779.00	9.00
	RAJF133G	23T	GILL	399.40	1.80	1.75	505.00	8.70
Female	RAJG134SD	25T	SKIN1 D	77.40	1.75	1.38	88.00	6.70
	RAJG134SS	26T	SKIN2 S	26.70	1.61	0.96	27.00	6.00
	RAJC134SW	27T	SKIN3 W	82.30	1.95	1.91	103.00	7.30
	RAJC134M	28T	MUSCLE	239.80	1.95	1.74	201.00	8.60
	RAJF134T	29T	OVARY	733.50	1.99	2.12	640.00	5.70
	RAJF134L	30T	LIVER	676.40	2.06	2.08	683.00	6.80
	RAJG134H	31T	HEART	348.20	2.01	2.14	424.00	6.50
	RAJG134B	32T	BRAIN	317.80	2.01	2.08	379.00	8.30
	RAJC134D	33T	DORSAL FIN	226.90	1.73	1.42	237.00	5.90
	RAJC134E	34T	EYE	241.40	2.00	2.03	355.00	5.50
	RAJF134J	35T	JAW BONE	275.50	2.02	2.20	331.00	7.40

Table S15 List of Blastx matchings between the assembled transcripts and putative genes related to pigmentation. Gene Ontology (GO) is also reported.

Transccript_ID	UniProt accession ID	GENE_ORGANISM	%Identity	E-value	PFAM accession ID	PFAM domain	GO terms
TRINITY_GG_48136_c0_g1_i1	O54967	ACK1_MOUSE	52.63	E:6e-42	PF11555.6	EGFR receptor inhibitor Mig-6	
TRINITY_GG_48136_c0_g2_i1	O54967	ACK1_MOUSE	57.24	E:6e-39	PF11555.6	EGFR receptor inhibitor Mig-6	
TRINITY_GG_1530_c0_g1_i1	P78536	ADA17_HUMAN	72.58	E:2e-58	PF16698.3	Membrane-proximal domain, switch, for ADAM17	
TRINITY_GG_1951_c0_g1_i1	O42574	ADRB1_XENLA	60.67	E:5e-63	.		
TRINITY_GG_12802_c0_g1_i1	Q9Y4K1	AIM1_HUMAN	50.88	E:2e-26	PF00030.17	Beta/Gamma crystallin	
TRINITY_GG_22903_c0_g1_i1	Q9Y4K1	AIM1_HUMAN	49.23	E:7e-37	PF00030.17	Beta/Gamma crystallin	
TRINITY_GG_31829_c0_g1_i1	Q9Y4K1	AIM1_HUMAN	57.14	E:1e-57	PF00030.17	Beta/Gamma crystallin	
TRINITY_GG_6872_c0_g3_i1	Q9Y4K1	AIM1_HUMAN	45.08	E:3e-42	.		
TRINITY_GG_6872_c0_g1_i1	Q9Y4K1	AIM1_HUMAN	52.73	E:4e-33	.		
TRINITY_GG_8356_c0_g1_i1	Q9Y4K1	AIM1_HUMAN	50	E:1e-21	.		
TRINITY_GG_41838_c0_g1_i1	Q9UJX6	ANC2_HUMAN	77.78	E:7e-45	PF08672.9	Anaphase promoting complex (APC) subunit 2	
TRINITY_GG_29528_c0_g2_i1	O14617	AP3D1_HUMAN	75	E:8e-59	PF06375.9	AP-3 complex subunit delta-1	GO:0015031: biological_process: protein transport GO:0030123: cellular_component: AP-3 adaptor complex
TRINITY_GG_29528_c0_g1_i1	O54774	AP3D1_MOUSE	76.82	E:8e-64	PF06375.9	AP-3 complex subunit delta-1	GO:0015031: biological_process: protein transport GO:0030123: cellular_component: AP-3 adaptor complex
TRINITY_GG_29528_c0_g3_i1	O54774	AP3D1_MOUSE	77.48	E:3e-64	PF06375.9	AP-3 complex subunit delta-1	GO:0015031: biological_process: protein transport GO:0030123: cellular_component: AP-3 adaptor complex

TRINITY_GG_64983_c0_g2_i1	O54774	AP3D1_MOUSE	74.58	E:9e-18	PF06375.9	AP-3 complex subunit delta-1	GO:0015031: biological_process: protein transport GO:0030123: cellular_component: AP-3 adaptor complex
TRINITY_GG_7_c1_g1_i1	P25054	APC_HUMAN	66.76	E:0.00	PF05956.9	APC basic domain	GO:0008017: molecular_function: microtubule binding GO:0016055: biological_process: Wnt signalling pathway GO:0008013: molecular_function: beta-catenin binding
TRINITY_GG_7_c1_g2_i1	P25054	APC_HUMAN	66.76	E:0.00	PF05956.9	APC basic domain	GO:0008017: molecular_function: microtubule binding GO:0016055: biological_process: Wnt signalling pathway GO:0008013: molecular_function: beta-catenin binding
TRINITY_GG_7_c3_g1_i1	P25054	APC_HUMAN	66.24	E:5e-99	PF16629.3	Armadillo-associated region on APC	GO:0008013: molecular_function: beta-catenin binding GO:0016055: biological_process: Wnt signalling pathway
TRINITY_GG_7_c5_g1_i1	P25054	APC_HUMAN	68	E:2e-19	PF05923.10	APC repeat	GO:0016055: biological_process: Wnt signalling pathway GO:0008013: molecular_function: beta-catenin binding

TRINITY_GG_37150_c0_g1_i1	Q61315	APC_MOUSE	61.9	E:2e-68	PF00514.21	Armadillo/beta-catenin-like repeat	GO:0005515: molecular_function: protein binding
TRINITY_GG_7_c0_g2_i1	Q61315	APC_MOUSE	60	E:1e-32	PF16634.3	Unstructured region on APC between APC_crr a SAMP	GO:0008013: molecular_function: beta-catenin binding GO:0016055: biological_process: Wnt signalling pathway
TRINITY_GG_7_c4_g1_i1	Q61315	APC_MOUSE	64.71	E:9e-09	PF05972.9	APC 15 residue motif	GO:0008013: molecular_function: beta-catenin binding GO:0016055: biological_process: Wnt signalling pathway
TRINITY_GG_7_c0_g1_i1	P70039	APC_XENLA	53.14	E:7e-85	PF16634.3	Unstructured region on APC between APC_crr a SAMP	GO:0008013: molecular_function: beta-catenin binding GO:0016055: biological_process: Wnt signalling pathway
TRINITY_GG_63968_c0_g1_i1	A8CEM5	ASIP_MACHE	37.5	E:6e-12	PF05039	Agouti protein	GO:0009755: biological_process: hormone-mediated signalling pathway GO:0005576: cellular_component: extracellular region
TRINITY_GG_20560_c0_g1_i1	A5PJP1	BL1S3_BOVIN	41.98	E:2e-13	PF15753.3	Biogenesis of lysosome-related organelles complex 1 subunit 3	
TRINITY_GG_21185_c0_g1_i1	Q99MP8	BRAP_MOUSE	88.73	E:3e-91	PF07576.10	BRCA1-associated protein 2	GO:0005515: molecular_function: protein binding GO:0008270: molecular_function: zinc ion binding GO:0046872:

							molecular_function: metal ion binding
TRINITY_GG_32340_c0_g1_i1	Q5F3N9	CF106_CHICK	82.67	E:1e-121	PF16158.3	Ig-like domain from next to BRCA1 gene	
TRINITY_GG_32340_c0_g1_i2	Q5F3N9	CF106_CHICK	83.57	E:6e-131	PF16158.3	Ig-like domain from next to BRCA1 gene	
TRINITY_GG_44132_c0_g1_i1	P27925	CREB1_BOVIN	75	E:1e-57	PF00170.19	bZIP transcription factor	GO:0003700: molecular_function: transcription factor activity, sequence-specific DNA binding GO:0043565: molecular_function: sequence-specific DNA binding GO:0006355: biological_process: regulation of transcription, DNA-templated
TRINITY_GG_44132_c0_g2_i1	P27925	CREB1_BOVIN	66.67	E:8e-44	PF00170.19	bZIP transcription factor	GO:0003700: molecular_function: transcription factor activity, sequence-specific DNA binding GO:0043565: molecular_function: sequence-specific DNA binding GO:0006355: biological_process: regulation of transcription, DNA-templated
TRINITY_GG_23196_c0_g1_i1	Q9I8N6	CSF1R_DANRE	56.9	E:6e-44	PF07714.15	Protein tyrosine kinase	GO:0004672: molecular_function: protein kinase activity GO:0006468: biological_process: protein phosphorylation GO:0005524: molecular_function: ATP binding

TRINITY_GG_7012_c0_g1_i1	P11348	DHPR_RAT	68.06	E:1e-30	.		
TRINITY_GG_57720_c0_g1_i1	P97386	DNLI3_MOUSE	81.76	E:3e-29	PF16759.3	DNA ligase 3 BRCT domain	
TRINITY_GG_52960_c0_g1_i1	Q90YB1	DNLI4_CHICK	82.41		PF01068.19	ATP dependent DNA ligase domain	GO:0003910: molecular_function: DNA ligase (ATP) activity GO:0005524: molecular_function: ATP binding GO:0006281: biological_process: DNA repair GO:0006310: biological_process: DNA recombination
TRINITY_GG_26903_c0_g1_i1	P56497	EDNRB_CANLF	45.11	E:2e-44	.		
TRINITY_GG_34515_c0_g1_i1	Q90328	EDNRB_COTJA	68.95	E:1e-98	PF00001.19	7 transmembrane receptor (rhodopsin family)	GO:0004930: molecular_function: G-protein coupled receptor activity GO:0007186: biological_process: G-protein coupled receptor signalling pathway GO:0016021: cellular_component: integral component of membrane
TRINITY_GG_9663_c0_g1_i1	Q90328	EDNRB_COTJA	86.81	E:6e-133	PF00001.19	7 transmembrane receptor (rhodopsin family)	GO:0004930: molecular_function: G-protein coupled receptor activity GO:0007186: biological_process: G-protein coupled receptor signalling pathway GO:0016021: cellular_component: integral component of membrane
TRINITY_GG_50561_c0_g1_i1	Q90328	EDNRB_COTJA	72.93	E:1e-21	.		
TRINITY_GG_50561_c0_g2_i1	Q90328	EDNRB_COTJA	70.15	E:7e-21	.		

TRINITY_GG_52270_c0_g1_i1	Q90328	EDNRB_COTJA	70.15	E:3e-42	.		
TRINITY_GG_56604_c0_g2_i1	Q90328	EDNRB_COTJA	80.25	E:1e-38	.		
TRINITY_GG_26903_c0_g3_i1	P24530	EDNRB_HUMAN	75	E:1e-58	PF00001.19	7 transmembrane receptor (rhodopsin family)	GO:0004930: molecular_function: G-protein coupled receptor activity GO:0007186: biological_process: G-protein coupled receptor signalling pathway GO:0016021: cellular_component: integral component of membrane
TRINITY_GG_26903_c0_g2_i1	P24530	EDNRB_HUMAN	83.1	E:3e-43	.		
TRINITY_GG_28236_c0_g1_i1	P24530	EDNRB_HUMAN	86.44	E:2e-29	.		
TRINITY_GG_56604_c0_g1_i1	P24530	EDNRB_HUMAN	90.2	E:1e-30	.		
TRINITY_GG_1393_c0_g1_i1	Q9N0W7	EDNRB_RABIT	58.51	E:5e-53	PF10320.7	Serpentine type 7TM GPCR chemoreceptor Srsx	GO:0004930: molecular_function: G-protein coupled receptor activity GO:0007186: biological_process: G-protein coupled receptor signalling pathway GO:0016021: cellular_component: integral component of membrane
TRINITY_GG_56502_c0_g1_i1	Q99JZ7	ERRFI_MOUSE	86.15	E:4e-74	PF11555.6	EGFR receptor inhibitor Mig-6	
TRINITY_GG_56502_c0_g2_i1	Q99JZ7	ERRFI_MOUSE	62.41	E:4e-74	PF11555.6	EGFR receptor inhibitor Mig-6	
TRINITY_GG_38169_c0_g1_i1	P23610	F8I2_HUMAN	86.08	E:4e-17	PF14938.4	Soluble NSF attachment protein, SNAP	GO:0005515: molecular_function: protein binding
TRINITY_GG_18212_c0_g1_i1	Q01721	GAS1_MOUSE	75.14	E:5e-68	PF02351.14	GDNF/GAS1 domain	
TRINITY_GG_8121_c0_g1_i1	Q9I9R0	HABP4_CHICK	75.14	E:1e-38	PF04774.13	Hyaluronan / mRNA binding family	
TRINITY_GG_14437_c0_g1_i1	Q32PJ8	HDAC1_BOVIN	38.36	E:1e-95	PF00850.17	Histone deacetylase domain	

TRINITY_GG_29429_c0_g2_i1	Q32PJ8	HDAC1_BOVIN	57.84	E:4e-86	PF00850.17	Histone deacetylase domain	
TRINITY_GG_14828_c0_g1_i1	Q5RAG0	HDAC1_PONAB	61.64	E:3e-82	PF00850.17	Histone deacetylase domain	
TRINITY_GG_1516_c0_g1_i1	P56518	HDAC1_STRPU	98.33	E:2e-67	PF00850.17	Histone deacetylase domain	
TRINITY_GG_29429_c0_g1_i1	P56518	HDAC1_STRPU	96.83	E:7e-66	.		
TRINITY_GG_9432_c0_g1_i1	Q8BLY7	HPS6_MOUSE	99.17	E:5e-11	PF15702.3	Hermansky-Pudlak syndrome 6 protein	
TRINITY_GG_13966_c1_g1_i1	P09914	IFIT1_HUMAN	93.94	E:2e-27	PF02071.18	Aromatic-di-Alanine (AdAR) repeat	GO:0005515: molecular_function: protein binding
TRINITY_GG_68916_c0_g1_i1	Q13325	IFIT5_HUMAN	96.23	E:1e-25	PF02071.18	Aromatic-di-Alanine (AdAR) repeat	GO:0005515: molecular_function: protein binding
TRINITY_GG_11160_c0_g1_i1	Q08156	KIT_CHICK	91.25	E:9e-111	PF07714.15	Protein tyrosine kinase	GO:0004672: molecular_function: protein kinase activity GO:0006468: biological_process: protein phosphorylation GO:0005524: molecular_function: ATP binding
TRINITY_GG_1988_c1_g1_i1	Q5RES4	MED1_PONAB	60	E:7E-21			
TRINITY_GG_1990_c0_g1_i1	Q5RES4	MED1_PONAB	69.06	E:7E-61			
TRINITY_GG_3059_c0_g1_i1	Q2QCI8	MED12_DANRE	100	E:3e-115	PF12145.6	Eukaryotic Mediator 12 subunit domain	
TRINITY_GG_33816_c0_g1_i1	Q2QCI8	MED12_DANRE	36.36	E:2e-67	.		
TRINITY_GG_33816_c0_g2_i1	Q2QCI8	MED12_DANRE	47.47	E:1e-46	.		
TRINITY_GG_53292_c0_g1_i1	Q2QCI8	MED12_DANRE	42.65	E:7e-41	PF12144.6	Eukaryotic Mediator 12 catenin- binding domain	GO:0008013: molecular_function: beta-catenin binding GO:0016592: cellular_component: mediator complex
TRINITY_GG_1522_c0_g1_i1	Q7YQK8	MED12_PANTR	66.81	E:3e-79	.		
TRINITY_GG_58464_c0_g1_i1	Q7YQK8	MED12_PANTR	84.4	E:5e-62	PF12144.6	Eukaryotic Mediator 12 catenin-binding domain	GO:0008013: molecular_function: beta-

							catenin binding GO:0016592: cellular_component: mediator complex
TRINITY_GG_14816_c0_g1_i1	Q7ZUL9	MGRN1_DANRE	98.21	E:3e-40	.		
TRINITY_GG_28473_c0_g1_i1	Q7ZUL9	MGRN1_DANRE	98.77	E:2e-30	.		
TRINITY_GG_43917_c0_g1_i1	Q7ZUL9	MGRN1_DANRE	58.33	E:3e-46	.		
TRINITY_GG_43100_c0_g1_i1	O60291	MGRN1_HUMAN	90.09	E:2e-63	PF13920.4	Zinc finger, C3HC4 type (RING finger)	GO:0005515: molecular_function: protein binding GO:0008270: molecular_function: zinc ion binding
TRINITY_GG_31461_c0_g1_i1	O60291	MGRN1_HUMAN	69.39	E:8e-30	.		
TRINITY_GG_24228_c0_g1_i1	Q5XIQ4	MGRN1_RAT	92.65	E:2e-67	.		
TRINITY_GG_65360_c0_g1_i1	O75030	MITF_HUMAN	86.75	E:1e-44	PF15951.3	MITF/TFEB/TFEC/TFE3 N-terminus	
TRINITY_GG_36802_c0_g1_i1	Q08874	MITF_MOUSE	75.86	E:5e-42	PF11851.6	Domain of unknown function (DUF3371)	GO:0006355: biological_process: regulation of transcription, DNA-templated
TRINITY_GG_42845_c0_g1_i1	P97432	NBR1_MOUSE	67.11	E:2e-38	PF16158.3	Ig-like domain from next to BRCA1 gene	
TRINITY_GG_27742_c0_g1_i1	Q8WX92	NELFB_HUMAN	78.89	E:1e-60	PF06209.11	Cofactor of BRCA1 (COBRA1)	GO:0045892: biological_process: negative regulation of transcription, DNA-templated GO:0005634: cellular_component: nucleus
TRINITY_GG_3724_c0_g1_i1	Q8WX92	NELFB_HUMAN	88.33	E:4e-54	PF06209.11	Cofactor of BRCA1 (COBRA1)	GO:0045892: biological_process: negative regulation of transcription, DNA-templated GO:0005634: cellular_component: nucleus

TRINITY_GG_68879_c0_g1_i1	Q8WX92	NELFB_HUMAN	90.38	E:1e-76	PF06209.11	Cofactor of BRCA1 (COBRA1)	GO:0045892: biological_process: negative regulation of transcription, DNA-templated GO:0005634: cellular_component: nucleus
TRINITY_GG_32932_c0_g1_i1	Q8C4Y3	NELFB_MOUSE	83.67	E:4e-51	PF06209.11	Cofactor of BRCA1 (COBRA1)	GO:0045892: biological_process: negative regulation of transcription, DNA-templated GO:0005634: cellular_component: nucleus
TRINITY_GG_32932_c0_g2_i1	Q8C4Y3	NELFB_MOUSE	74.48	E:3e-82	PF06209.11	Cofactor of BRCA1 (COBRA1)	GO:0045892: biological_process: negative regulation of transcription, DNA-templated GO:0005634: cellular_component: nucleus
TRINITY_GG_68799_c0_g1_i1	Q8C4Y3	NELFB_MOUSE	79.31	E:2e-71	PF06209.11	Cofactor of BRCA1 (COBRA1)	GO:0045892: biological_process: negative regulation of transcription, DNA-templated GO:0005634: cellular_component: nucleus
TRINITY_GG_45475_c0_g1_i1	P18708	NSF_CRIGR	65.87	E:2e-90	PF07724.12	AAA domain (Cdc48 subfamily)	GO:0005524: molecular_function: ATP binding GO:0016887: molecular_function: ATPase activity
TRINITY_GG_10296_c0_g1_i1	P46459	NSF_HUMAN	79.31	E:3e-76	PF00004.27	ATPase family associated with various cellular activities (AAA)	GO:0005524: molecular_function: ATP binding

TRINITY_GG_18026_c0_g2_i1	P46459	NSF_HUMAN	68.13	E:4e-75	PF00004.27	ATPase family associated with various cellular activities (AAA)	GO:0005524: molecular_function: ATP binding
TRINITY_GG_41293_c0_g1_i1	P46459	NSF_HUMAN	76.42	E:2e-42	.		
TRINITY_GG_33608_c0_g1_i1	P46460	NSF_MOUSE	69.17	E:4e-72	PF02359.16	Cell division protein 48 (CDC48), N-terminal domain	
TRINITY_GG_33608_c0_g2_i1	P46460	NSF_MOUSE	90.37	E:9e-62	PF02359.16	Cell division protein 48 (CDC48), N-terminal domain	
TRINITY_GG_2443_c0_g1_i1	Q5R410	NSF_PONAB	77.78	E:3e-44	PF00004.27	ATPase family associated with various cellular activities (AAA)	GO:0005524: molecular_function: ATP binding
TRINITY_GG_32883_c0_g1_i1	Q5R410	NSF_PONAB	89.78	E:1e-72	PF00004.27	ATPase family associated with various cellular activities (AAA)	GO:0005524: molecular_function: ATP binding
TRINITY_GG_33579_c0_g1_i1	Q5R410	NSF_PONAB	92.11	E:4e-81	PF00004.27	ATPase family associated with various cellular activities (AAA)	GO:0005524: molecular_function: ATP binding
TRINITY_GG_66607_c0_g1_i1	Q9QUL6	NSF_RAT	94.29	1.00E-27		Vesicle-fusing ATPase	
TRINITY_GG_18026_c0_g1_i1	Q9QUL6	NSF_RAT	90.77	E:4e-23	.		
TRINITY_GG_32418_c0_g2_i1	Q9CY58	PAIRB_MOUSE	94.4	E:3e-53	PF04774.13	Hyaluronan / mRNA binding family	
TRINITY_GG_5848_c0_g1_i1	Q9CY58	PAIRB_MOUSE	94.32	E:2e-34	PF04774.13	Hyaluronan / mRNA binding family	
TRINITY_GG_5848_c0_g2_i1	Q9CY58	PAIRB_MOUSE	82.31	E:6e-34	PF04774.13	Hyaluronan / mRNA binding family	
TRINITY_GG_60948_c0_g1_i1	Q6ZW49	PAXI1_HUMAN	81.74	E:1e-91	PF12738.5	twin BRCT domain	
TRINITY_GG_6251_c0_g1_i1	Q6ZW49	PAXI1_HUMAN	83.78	E:4e-57	PF16770.3	Regulator of Ty1 transposition protein 107 BRCT domain	
TRINITY_GG_64272_c0_g1_i1	B5X171	PESC_SALSA	84.62	E:3e-67	PF16589.3	BRCT domain, a BRCA1 C-terminus domain	
TRINITY_GG_26850_c0_g1_i1	Q98917	PMEL_CHICK	90.14	E:9e-41	PF00801.18	PKD domain	
TRINITY_GG_3716_c0_g1_i1	Q98917	PMEL_CHICK	97.83	E:2e-39	.		
TRINITY_GG_47250_c0_g1_i1	Q98917	PMEL_CHICK	71.43	E:8e-26	.		
TRINITY_GG_47250_c0_g2_i1	Q98917	PMEL_CHICK	66.46	E:2e-32	.		
TRINITY_GG_9521_c0_g1_i1	Q98917	PMEL_CHICK	65.22	E:3e-34	.		

TRINITY_GG_9314_c0_g1_i1	Q96D21	RHES_HUMAN	43.44	E:5e-21	PF00071.20	RAS family	GO:0005525: molecular_function: GTP binding GO:0007264: biological_process: small GTPase mediated signal transduction
TRINITY_GG_9314_c0_g2_i1	Q96D21	RHES_HUMAN	43.44	E:5e-21	PF00071.20	RAS family	GO:0005525: molecular_function: GTP binding GO:0007264: biological_process: small GTPase mediated signal transduction
TRINITY_GG_31390_c0_g1_i1	Q60520	SIN3A_MOUSE	66.29	E:2e-42	PF02671.19	Paired amphipathic helix repeat	GO:0006355: biological_process: regulation of transcription, DNA-templated
TRINITY_GG_31390_c0_g1_i1	Q60520	SIN3A_MOUSE	66.29	E:2e-42	PF02671.19	Paired amphipathic helix repeat	GO:0006355: biological_process: regulation of transcription, DNA-templated
TRINITY_GG_57177_c0_g1_i1	Q60520	SIN3A_MOUSE	79.13	E:3e-44	PF02671.19	Paired amphipathic helix repeat	GO:0006355: biological_process: regulation of transcription, DNA-templated
TRINITY_GG_5915_c0_g1_i1	Q60520	SIN3A_MOUSE	64.9	E:1e-60	PF02671.19	Paired amphipathic helix repeat	GO:0006355: biological_process: regulation of transcription, DNA-templated
TRINITY_GG_51986_c0_g1_i1	O75182	SIN3B_HUMAN	67.72	E:1e-27	PF02671.19	Paired amphipathic helix repeat	GO:0006355: biological_process: regulation of transcription, DNA-templated
TRINITY_GG_22024_c0_g1_i1	Q62141	SIN3B_MOUSE	53.78	E:1e-82	PF02671.19	Paired amphipathic helix repeat	GO:0006355: biological_process: regulation of transcription, DNA-templated

TRINITY_GG_67679_c0_g1_i1	P81125	SNAA_BOVIN	52.94	E:1e-69	PF14938.4	Soluble NSF attachment protein, SNAP	
TRINITY_GG_30401_c0_g1_i1	Q9DB05	SNAA_MOUSE	56.84	E:1e-97	PF14938.4	Soluble NSF attachment protein, SNAP	
TRINITY_GG_53852_c0_g1_i1	Q9DB05	SNAA_MOUSE	44.44	E:8e-66	PF14938.4	Soluble NSF attachment protein, SNAP	
TRINITY_GG_31582_c0_g1_i1	P54921	SNAA_RAT	50.6	E:2e-61	PF14938.4	Soluble NSF attachment protein, SNAP	
TRINITY_GG_43021_c0_g1_i1	P28663	SNAB_MOUSE	51.33	E:2e-84	PF14938.4	Soluble NSF attachment protein, SNAP	
TRINITY_GG_63608_c0_g1_i1	P81127	SNAG_BOVIN	98	E:3e-54	PF14938.4	Soluble NSF attachment protein, SNAP	
TRINITY_GG_63608_c0_g4_i1	P81127	SNAG_BOVIN	76.84	E:5e-56	PF14938.4	Soluble NSF attachment protein, SNAP	
TRINITY_GG_63608_c0_g6_i1	P81127	SNAG_BOVIN	80.65	E:3e-98	PF14938.4	Soluble NSF attachment protein, SNAP	
TRINITY_GG_27414_c0_g1_i1	Q8AXX8	SOX10_XENLA	73.24	E:6e-44	.		
TRINITY_GG_27321_c0_g1_i1	P48434	SOX9_CHICK	88.19	E:5e-43	PF12444.6	Sox developmental protein N terminal	
TRINITY_GG_32038_c0_g1_i1	Q92547	TOPB1_HUMAN	93.86	E:6e-59	PF00533.24	BRCA1 C Terminus (BRCT) domain	
TRINITY_GG_16794_c0_g1_i1	Q98949	TYRO3_CHICK	85.19	E:5e-63	PF00069.23	Protein kinase domain	GO:0004672: molecular_function: protein kinase activity GO:0005524: molecular_function: ATP binding GO:0006468: biological_process: protein phosphorylation
TRINITY_GG_22756_c0_g1_i1	Q98949	TYRO3_CHICK	83.48	E:3e-46	PF07714.15	Protein tyrosine kinase	GO:0004672: molecular_function: protein kinase activity GO:0006468: biological_process: protein phosphorylation GO:0005524:

							molecular_function: ATP binding
TRINITY_GG_37430_c0_g1_i1	Q98949	TYRO3_CHICK	90.2	E:1e-76	PF07714.15	Protein tyrosine kinase	GO:0004672: molecular_function: protein kinase activity GO:0006468: biological_process: protein phosphorylation GO:0005524: molecular_function: ATP binding
TRINITY_GG_22756_c0_g2_i1	Q98949	TYRO3_CHICK	93.89	E:3e-36	.		
TRINITY_GG_50835_c0_g1_i1	A0JM20	TYRO3_XENTR	81.13	E:8e-31	.		
TRINITY_GG_22144_c0_g1_i1	Q8WN57	TYRP1_BOVIN	82.24	E:2e-60	PF00264.18	Common central domain of tyrosinase	GO:0016491: molecular_function: oxidoreductase activity GO:0008152: biological_process: metabolic process
TRINITY_GG_3909_c0_g1_i1	Q8WN57	TYRP1_BOVIN	81.76	E:7e-35	.		
TRINITY_GG_53969_c0_g1_i1	O57405	TYRP1_CHICK	92.19	E:9e-50	.		
TRINITY_GG_57110_c0_g1_i1	O93505	TYRP2_CHICK	96.92	E:1e-77	PF00053.22	Laminin EGF domain	
TRINITY_GG_15257_c0_g1_i1	P29812	TYRP2_MOUSE	75.83	E:1e-75	PF00264.18	Common central domain of tyrosinase	GO:0016491: molecular_function: oxidoreductase activity GO:0008152: biological_process: metabolic process
TRINITY_GG_28604_c0_g1_i1	P29812	TYRP2_MOUSE	72.93	E:2e-61	PF00264.18	Common central domain of tyrosinase	GO:0016491: molecular_function: oxidoreductase activity GO:0008152: biological_process: metabolic process
TRINITY_GG_13519_c0_g1_i1	P47990	XDH_CHICK	53.9	2.00E-42		Xanthine dehydrogenase/oxidase	

TRINITY_GG_13519_c0_g2_i1	P47990	XDH_CHICK	84.38	1.00E-42		Xanthine dehydrogenase/oxidase	
TRINITY_GG_25378_c0_g1_i1	P47990	XDH_CHICK	78.32	2.00E-35		Xanthine dehydrogenase/oxidase	
TRINITY_GG_28224_c0_g1_i1	P47990	XDH_CHICK	73.13	2.00E-27		Xanthine dehydrogenase/oxidase	
TRINITY_GG_28224_c0_g2_i1	P47990	XDH_CHICK	80.91	1.00E-28		Xanthine dehydrogenase/oxidase	
TRINITY_GG_44636_c0_g1_i1	P47990	XDH_CHICK	80.28	E:3e-39	PF01315.20	Aldehyde oxidase a xanthine dehydrogenase, a/b hammerhead domain	
TRINITY_GG_68752_c0_g1_i1	P47990	XDH_CHICK	79.2	3.00E-23		Xanthine dehydrogenase/oxidase	
TRINITY_GG_71332_c0_g1_i1	P47990	XDH_CHICK	81.02	1.00E-38		Xanthine dehydrogenase/oxidase	
TRINITY_GG_71332_c0_g2_i1	P47990	XDH_CHICK	64.14	9.00E-49		Xanthine dehydrogenase/oxidase	
TRINITY_GG_71332_c0_g3_i1	P47990	XDH_CHICK	63.27	3.00E-49		Xanthine dehydrogenase/oxidase	
TRINITY_GG_23943_c0_g2_i1	P22811	XDH_DROPS	56.94	E:5e-23	PF02738.16	Molybdopterin-binding domain of aldehyde dehydrogenase	GO:0016491: molecular_function: oxidoreductase activity GO:0055114: biological_process: oxidation-reduction process
TRINITY_GG_4100_c0_g2_i1	Q9MYW6	XDH_FELCA	64.91	E:6e-46	PF02738.16	Molybdopterin-binding domain of aldehyde dehydrogenase	GO:0016491: molecular_function: oxidoreductase activity GO:0055114: biological_process: oxidation-reduction process
TRINITY_GG_44636_c0_g2_i1	P47990	XDH_HUMAN	67.5	5.00E-25		Xanthine dehydrogenase/oxidase	

TRINITY_GG_4100_c0_g1_i1	P22985	XDH_RAT	69.23	2.00E-40		Xanthine dehydrogenase/oxidase	
TRINITY_GG_23700_c2_g1_i1	O95409	ZIC2_HUMAN	74.16	E:0.00	PF00096.24	Zinc finger, C2H2 type	GO:0046872: molecular_function: metal ion binding

Appendix IV. Chapter 5

Table S16 Number of differentially expressed genes in different tissues in pairwise comparisons.

	Raja asterias	*Raja clavata*	*Raja miraletus*	*Raja ocellifera*	*Raja straeleni*	
Raja asterias	*					
Raja clavata	804	*				Dorsal matrix
Raja miraletus	3762	3246	*			
Raja ocellifera	2668	2288	1937	*		
Raja straeleni	760	221	3249	2068	*	

	Raja asterias	*Raja clavata*	*Raja miraletus*	*Raja ocellifera*	*Raja straeleni*	
Raja asterias	*					
Raja clavata	591	*				Spot
Raja miraletus	3145	1931	*			
Raja ocellifera	3164	2328	1807	*		
Raja straeleni	927	261	3050	2868	*	

	Raja asterias	*Raja clavata*	*Raja miraletus*	*Raja ocellifera*	*Raja straeleni*	
Raja asterias	*					
Raja clavata	804	*				Ventral portion
Raja miraletus	3762	3246	*			
Raja ocellifera	2668	2288	1937	*		
Raja straeleni	760	221	3249	2068	*	

Table S17 Number of differentially expressed genes between different tissues within species. DvsW dorsal matrix compared to white ventral portion, SvsD spot compared to dorsal matrix, SvsW spot compared to white ventral portion.

	Raja asterias		*Raja clavata*		*Raja straeleni*		*Raja miraletus*		*Raja ocellifera*	
Comparisons	Down	Up	Down	Up	Down	Up	Down	Up	Down	Up
DvsW	90	70	97	125	64	69	41	203	128	32
SvsD	0	0	0	0	0	0	61	0	0	0
SvsW	70	91	213	38	44	76	207	95	358	50