Reconstructing the evolutionary history of desert-adapted *Cerastes* vipers in North Africa and the Arabian Peninsula
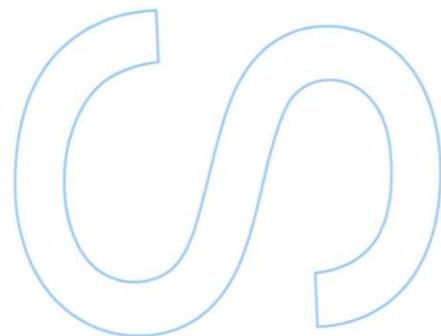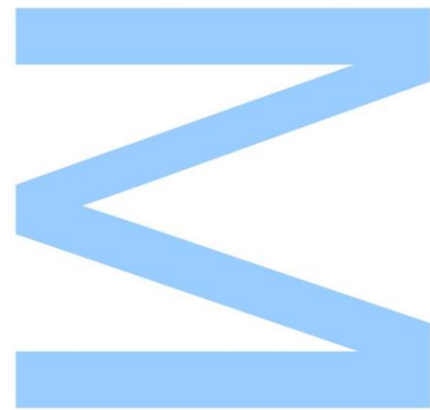
Margarida Isabel Oliveira Barros
Mestrado em Biodiversidade, Genética e Evolução
Departamento de Biologia
2019

Orientador
Fernando Martínez-Freiría, Researcher, CIBIO/InBIO

Co-orientador
Guillermo Velo-Antón, Assistant Researcher, CIBIO/InBIO

U.PORTO

FC FACULDADE DE CIÊNCIAS
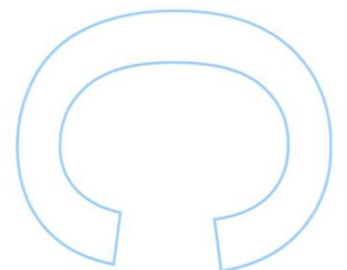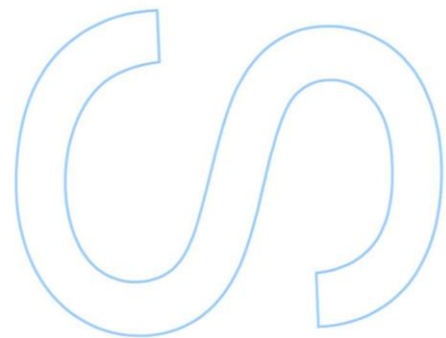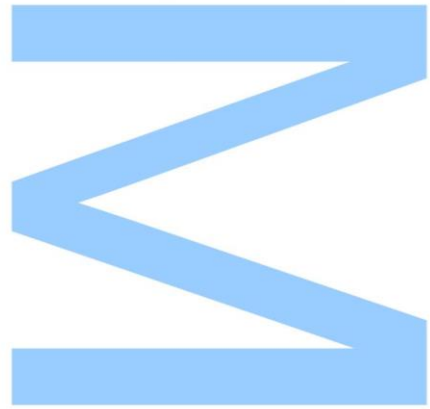UNIVERSIDADE DO PORTO

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____

Agradecimentos

Em primeiro lugar, quero agradecer aos meus orientadores, Doutor Fernando Martínez-Freiría e Doutor Guillermo Velo-Antón por tudo o que me ensinaram e pelo tempo que despenderam a orientar-me ao longo do decorrer desta tese. Quero ainda agradecer por toda a motivação, compreensão e apoio incansável que me deram desde o primeiro dia que nos conhecemos.

Quero agradecer a todos os membros do CTM, em especial à Patrícia, Sofia e Susana por toda a ajuda, disponibilidade e paciência que tiveram para comigo ao longo de todo o meu percurso no laboratório.

Também gostaria de agradecer ao Doutor José Carlos Brito e a todos os membros do grupo Biodeserts por me acolherem e pelos conhecimentos que partilharam comigo desde que faço parte deste grupo.

Agradeço a todos os que permitiram que este trabalho fosse possível através do seu esforço na recolha de amostras e dados de presença. Agradeço ainda a Jirí Moravec que permitiu a colheita de amostras de tecido de espécimes do Museu de História Nacional de Praga (República Checa).

Quero também agradecer à Inês Freitas por me ter acompanhado e ajudado ao longo de toda esta etapa e a todos os meus amigos e colegas que conheci no início deste mestrado por toda a companhia e amizade ao longo destes anos.

Um obrigado muito especial à minha família por me darem a oportunidade de ir mais longe na minha educação, por estarem sempre presentes e por toda a preocupação, conselhos e conforto que me deram não só ao longo desta etapa, mas também durante toda a minha vida.

Por fim quero agradecer aos meus amigos, em especial ao Francisco, à Catarina e à Marta por todo o apoio emocional e por terem estado sempre presentes ao longo dos últimos 12 anos.

Abstract

Pleistocene climatic oscillations have influenced biogeographical patterns of species worldwide, affecting their distributional ranges and shaping their genetic diversity. Desert environments have been mistakenly characterized as regions extremely poor in biodiversity and with low variability. However, the opposite has been proven, in major deserts such as the Sahara and Arabian deserts and its surrounding regions, which present several types of ecoregions. This great variety of ecoregions exist due to the wide range of climatic and topographical conditions that characterize North Africa and the Arabian Peninsula. These characteristics demonstrate the potential that these outstanding regions have for conducting scientific research. In addition, recent studies have shown the importance of the climate influence in the genetic structure and variability of species given their accentuated and dynamic climatic history, and diverse life history and habitat traits of taxa inhabiting such extreme regions. This study aims to address the role of Pleistocene climatic oscillations in the evolutionary history of the three *Cerastes* species (Viperinae), *C. cerastes* and *C. vipera* from the Sahara Desert, and *C. gasperettii* from the Arabian Peninsula deserts.

Phylogenetic structure and variability were inferred using Bayesian inference over sequences (68 samples) for one mtDNA (COI) and three nuDNA (PRLR, NT3, VIM) gene markers. Paleoclimatic models combined 318 occurrences and five climatic variables in Maxent to infer climatic suitability for current and past (mid Holocene, Last Glacial Maximum and Last Inter Glacial) events, and stability over time. Paleoclimatic models were conduct in two different scales, the first carried for all three species at a 10x10 kms scale, covering the total range of the genus, and the second at a 1x1kms scale conducted only in North West Africa.

Mitochondrial inferences show *C. cerastes* and *C. gasperettii* as sister taxa, while *C. vipera* is identified as a phylogenetically more distant species, which is concordant with previously recent studies. The three nuDNA genes analysed in this study (PRLR, NT3 and VIM) well differentiated *C. gasperettii* from the other two taxa. However, PRLR and NT3 showed extensive haplotype sharing between *C. cerastes* and *C. vipera*. The different results obtained using mitochondrial and nuclear markers raise questions on the true phylogenetic relationships between *Cerastes* vipers. Further levels of mtDNA structure within the three species originated along the middle and late Pleistocene. A clear division between *C. cerastes* and *C. vipera* populations was found between eastern and western areas of North Africa. In the phylogenetic reconstruction was also possible

to observe structuration within each lineage of *C. cerastes* and *C. vipera*. Within the western group, two distinct population were recovered for *C. cerastes* (West-North and West-South) and for *C. vipera* (West-west and West-Central). *C. gasperettii* shown to be the species with the most recent divergence and for which two lineages (North and South) were recovered.

Due to the vast intraspecific diversity found for *C. cerastes* in western North Africa, and since this region has been hypothesized to have acted as a refuge and corridor for several other species with different ecological requirements through time, further paleoclimatic modelling was conducted for Western populations of this species.

Paleoclimatic models identified Inter Glacial events as major drivers of range reduction and isolation in the three species. Distinct areas of high climatic stability across the Sahara and Arabian deserts fit spatial patterns of genetic structure and likely acted as Pleistocene climatic refugia for species and lineages. Mito-nuclear discordances are discussed in the light of morphological and ecological traits of species. This multidisciplinary approach allows to propose biogeographic scenarios for the evolution of these desert-adapted species.

Resumo

Oscilações climáticas do Pleistoceno têm influenciado os padrões biogeográficos de espécies em todo o mundo, afetando a sua distribuição e moldando a sua diversidade genética. Os ambientes desérticos têm sido erradamente caracterizados como regiões extremamente pobres em biodiversidade e com baixa variabilidade. No entanto, o oposto tem sido comprovado em grandes desertos, tais como o deserto do Sahara r os desertos da Península Arábica e as suas regiões vizinhas, que apresentam vários tipos de ecorregiões. Esta grande variedade de ecorregiões existe devido à ampla gama de condições climáticas e topográficas que caracterizam o Norte de África e a Península Arábica. Estas características demonstram o potencial que estas regiões excecionais têm para realizar investigação científica. Além disso, estudos recentes têm mostrado a importância da influência climática na estrutura genética e na variabilidade das espécies, devido à sua história climática acentuada e dinâmica, e à história de vida diversificada e às características de habitat de taxa que habitam nestas regiões extremas. Este estudo tem como objetivo compreender o papel das oscilações climáticas do Pleistoceno na história evolutiva das três espécies de *Cerastes* (Viperinae), C. *cerastes* e C*. vipera* do deserto do Saara, e C. *gasperettii* dos desertos da Península Arábica.

A estrutura e variabilidade filogenética foram inferidas através de inferência Bayesiana sobre sequências (68 amostras) para um marcador mitocondrial (COI) e três marcadores nucleares (PRLR, NT3, VIM). Modelos paleoclimáticos combinaram 318 ocorrências e cinco variáveis climáticas no software Maxent para inferir a adequação climática durante os períodos presente e do passado (Holoceno Médio, Último Máximo Glaciar e Último InterGlaciar), assim como estabilidade ao longo do tempo. Modelos paleoclimáticos foram conduzidos em duas escalas diferentes, a primeira realizada para todas as três espécies numa escala 10x10 kms, cobrindo a área de distribuição total do género, e a segunda a uma escala 1x1kms conduzida apenas no Norte de África Ocidental.

Inferências mitocondriais mostram C. *cerastes* e C. *gasperettii* como espécies irmãs, enquanto que C. *vipera* é identificada como uma espécie filogeneticamente mais distante, o que é concordante com estudos recentes. Os três genes nucleares analisados neste estudo (PRLR, NT3 e VIM) diferenciaram bastante a C. *gasperettii* dos outros dois taxa. No entanto, PRLR e NT3 mostraram uma partilha extensa de haplótipos entre C. *cerasta* e C. *vipera*. Os diferentes resultados obtidos utilizando marcadores mitocondriais e nucleares levantam questões sobre as verdadeiras relações

filogenéticas entre as víboras *Cerastes*. Outros níveis de estruturação mitocondrial dentro das três espécies foram originados ao longo do Pleistoceno Médio e Superior. A divisão clara entre as populações de *C. cerastes* e C. *vipera* foi encontrada entre as áreas oriental e ocidental do Norte de África. Na reconstrução filogenética foi também possível observar estruturação dentro de cada linhagem de *C. cerastes* e *C. vipera*. Dentro do grupo ocidental, duas populações distintas foram recuperadas para *C. cerastes* (West-North e West-South) e para *C. vipera* (West-West e West-Central). *C. gasperettii* mostrou ser a espécie que divergiu mais recentemente e para a qual duas linhagens (North e South) foram recuperadas.

Devido à grande diversidade intraespecífica encontrada para *C. cerastes* no oeste do Norte da África, e uma vez que tem sido levantada a hipótese desta região ter agido como um refúgio e corredor para várias outras espécies com diferentes requisitos ecológicos ao longo do tempo, foram conduzidos mais estudos de modelação paleoclimáticos para as populações ocidentais desta espécie.

Os modelos paleoclimáticos identificaram os eventos interglaciais como principais motores de redução da distribuição e isolamento nas três espécies. Áreas distintas de alta estabilidade climática ao longo dos desertos Sahara e Arábicos correspondem aos padrões espaciais de estrutura genética e provavelmente agiram como refúgio climáticos para espécies e linhagens. Durante o Pleistoceno. As discordâncias mito-nucleares são discutidas à luz de características morfológicas e ecológicas das espécies. Esta abordagem multidisciplinar permite propor cenários biogeográficos para a evolução destas espécies adaptadas ao deserto.

Palavras chave: Biogeografia, filogeografia, modelos baseados em nichos ecológicos, Pleistoceno, áreas climaticamente estáveis, répteis, filogenia multilocus, abordagem integrativa

Index

Table Index

FCUP
Reconstructing the evolutionary history of desert-adapted Cerastes vipers in North Africa and the
Arabian Peninsula

x

Figure Index

Abreviation Index

min - minutes

sec - seconds

Kya – Kilo years agora

Mya - Million years ago

LGM – Last Glacial Maximum

LIG – Last inter-glacial

1. Introduction

1.1 - Climatic variability

Climate is one of the strongest factors affecting species diversification and extinction processes (Peterson, 2009; Quante, 2010). Current species biogeographical patterns and genetic structure were largely shaped by climatic spatial variability, particularly during the Pleistocene, and landscape features acting as barriers to gene flow (Hewitt, 2000).

In the Early Pleistocene, ca 2.6 Mya, the Arctic ice cap started to grow, marking the beginning of the strong climatic oscillations characterized by glacial-interglacial periods. During glacial periods, the increasing colder climate lead to the expansion of the existing ice caps, occupying most of the Northern hemisphere, compressing the tropical and temperate regions to the equator (Hewitt, 2000; 2004). During the inter-glacial periods, the increasing of Earth's temperature and consequently of the aridity produced the melting of the ice caps, restricting them to the northernmost areas of the globe, leading to the expansion of desert areas and reducing tropical forests (Hewitt 2000; 2004). At the end of the last glacial period (Last Glacial Maximum, ca 21 Kya), ice sheets reached maximum extents, provoking the drop of sea levels and consequently droughts and desertification (Mithen et al., 2004). Around 12 Kya, ice caps started to melt rapidly, which led to an abrupt rise of the sea level, marking the transition between the ending of the Last Glacial Maximum and the beginning of the Holocene period (Clark et al., 2009; Walker et al., 2009). During this period, coastal lines areas decreased, and water subsequent from the melting ice caps filled lakes and created continental islands (Quante, 2010). Posteriorly, in the Middle Holocene, ca 6 Kya, the amount of solar radiation changed due to Earth's orbital position, leading to warmer summers and colder winters in the Northern Hemisphere. This period was also marked by an increase in ocean evaporation and consequently higher continental precipitation (Quante, 2010).

During these warming-cooling cycles, biodiversity was forced to retract or spread their distributional ranges, accordingly to their ecological requirements and life history-traits, in order to survive (Hewitt, 2000). Overall, during favourable conditions, species expanded their distribution and thrived, with potential reconnections of the existent populations, while during unfavourable conditions, species likely reduced distributional ranges and/or became extinct (e.g. Bannikova et al., 2010; Carranza *et al.,* 2008; Dobigny et al., 2005; Velo-Antón et al., 2013, 2018; Martínez-Freiría et al., 2015, 2017; Gonçalves

et al., 2018a). However, some species were also able to adapt to the new climatic conditions.

Although the effects of climatic oscillations on biodiversity patterns have been extensively studied in many regions of the world, specifically in Europe and North America (see Beddek et al., 2018; Habel et al., 2010; Hewitt, 2000, 2004; Husemann et al., 2014; Murphy and Weiss, 1992; Sommer and Zachos, 2009; Weiss and Ferrand, 2007) further studies on less studied continents and remote regions are needed (Brito et al., 2014).

1.2 -The deserts

The scientific community has tried to define desert environments using factors such as climate (precipitation, evaporation and temperature) as well as geomorphic features, fauna and flora (Laity, 2008). Characterizing a desert by its high temperatures, lack of precipitation, low humidity or scarce presence of plant communities can be a definition that could fit some desert areas but could completely fail others. So, the most fitting definition is the one that characterizes deserts by their extreme arid conditions, which may vary in a large range of abiotic characteristics (Laity, 2008; Ward, 2016).

Due to the general misconception that deserts are homogeneous environments, with little variability and due to the remoteness of some areas, few scientific research has been done in these areas. In addition, it is believed that these areas present scarce biodiversity due to its extreme environmental conditions, and so very little is known about the species biodiversity and history, as well as the role of climate and landscape on shaping species patterns of distribution and biogeography (Brito et al., 2014). However, deserts represent 18% of land (Trabucco and Zomer, 2009) and include 25% of terrestrial forms (Mace et al., 2005) which are highly adapted to the desert's harsh climatic conditions, many of which being endemic to these regions (Brito et al., 2016).

As other regions of the world, deserts also experienced climatic oscillations during the Pleistocene (Laity; 2008). The warm-cooling cycles that marked this period, characterized by humid or dry conditions, would cause the expansion and retraction of deserts. However, not all Earth's deserts responded to these climatic changes in the same way. During the Last Glacial Maximum, when the highest rates of aridity were felt in Africa, Australia and Asia, deserts of these regions have undergone by a massive expansion, which made them reach a size of about 5 times higher than their current area (Sarnthein, 1978; Stokes et al., 1997). On the other hand, the presence of the Laurentide

Ice Sheet in North-America caused wetter and cooler conditions across the deserts south of this region. An increase in effective moisture was verified in the region due to the winds coming from the west and the jet stream, which propitiated the formation of lakes in southwestern desert regions and the expansion of woodlands (Benson et al., 1990; Benson et al., 1995; Orme, 2002).

Scientific research is needed on deserts since they are known to be the most climate change impacted areas of the globe, therefore attention needs to be brought to these areas in order to apply effective conservation measures (Brito et al., 2014; Durant et al., 2012).

1.2.1 - The Sahara and Arabian deserts

The Sahara desert is the biggest and warmest desert on Earth, spreading throughout 11 countries and occupying the majority of North Africa's area (Olson et al., 2001; Harris 2003). The Sahara is one of the major ecoregions in Africa, marking the transition between the Palearctic and Afro-Tropical biogeographic realms (Olson et al., 2001). This ecoregion is not topographically homogeneous, presenting a high variability of features such as mountains, oasis, seasonal rivers, rocky areas and sand dunes and also a very heterogenous climate, with high variability of temperature (9.4 to 30.8 ºC) and precipitation (up to 981 mm) (Anthelme et al., 2008; Dinerstein et al., 2017; Le Houérou, 1997; www.wordclim.org).

The Sahara was not always a desert, its desertification process only begun around 7 million years ago (Schuster et al., 2006; Zhang et al., 2014). From the early Pliocene (5.3 Mya) to the late Pleistocene, the Sahara experienced strong climatic oscillations, characterized by multiple dry-wet cycles (Brito et al., 2014). As in the rest of the planet, these cycles were mostly related to the Earth's orbital cycles (i.e. Milankovitch cycles), but also to the variability of the Western African monsoons and the vegetation feedbacks in regional climates (Armitage et al., 2015; Maley, 2010). During the wet periods, the arid areas shrink due to the expansion of the afro-tropical environments towards the north, giving place to more humid areas and refilling the hydrographic networks. During these phases, Sahara was characterized by steppe habitats, prevailing in the north by Mediterranean vegetation, some temperate vegetation and with Sahelian and Sudanian elements in the driest places. During the dry phases, climate was hotter and drier than the present time due to an increase in aridity, which led to the expansion of the Sahara region towards the south (Le Houérou, 1997). Since the last wet phase,

around 6 Kya, until today, the Sahara lost most of the vegetation communities and river networks due to the increment in aridity (Foley et al., 2003; Holmes, 2008). These climatic events likely affected the patterns of distribution of North African taxa, allowing species to retract or spread their ranges according to their ecological requirements (e.g. Brito et al., 2011; Carranza et al., 2008; Douady et al., 2003; Tamar et al., 2016; Martínez-Freiría et al., 2017; Gonçalves et al., 2018 a,b; Velo-Antón et al., 2012; 2018).

The Arabian Peninsula is the largest peninsula of the world, located eastwards to North-Africa and spreading throughout nine different countries (Antonsich, 2004). It is divided in two major geologic provinces, the Arabian Shield and the Arabian Shelf (Al-Juaidi et al., 2003) and constituted mainly by desert and shrubland areas with only savannah and temperate grasslands in the north (Dinerstein et al., 2017). Overall this region is arid or extremely arid, with less than 100 mm of annual precipitation (Glennie, 1998) and high temperatures that can reach 50°C. The Arabian desert covers almost the entirety of the Arabian Peninsula area (Harris, 2004). This high-pressure climate desert is considered an extension of the Sahara Desert thus presenting a very similar climate. However, few areas of the Arabian desert are considered hyper-arid, with less than 50 mm of annual rainfall, in comparison to the large extent observed in Sahara (Edgell, 2006; Harris, 2004). The An Nafud (or Great Nafud) and the Rub' al-Khali ("Empty Quarter") are the two major sand regions in the north and south areas respectively, that comprise around a third of the total Arabian desert region (Harris, 2004).

As other areas of the globe, Arabian Peninsula climate and environment has been influenced by the climatic oscillations of the Quaternary period (Edgell, 1989; Glennie, 1998; Al-Farraj and Harvey, 2004), marked by a change from temperate to arid conditions during the Pleistocene/Holocene boundary (Kotwicki and al Sulaimani, 2009). As described for the Sahara Desert, the Arabian Peninsula was subject to glacial and interglacial periods, characterized by dry and humid phases, respectively. During the Last Inter-glacial period (LIG) ca 130 Kya, due to the change on Earth's rotational axis, North hemisphere annual temperature increased, and glaciers started to melt. This led to the rising of sea levels, which caused wetter conditions to be felt on the Arabian Peninsula, allowing the filling of lakes in the region. Posteriorly, during the Last Glacial Maximum (~21 Kya), extreme aridification hit the region, leading to an increase of arid zones, a decrease in the sea levels and the desiccation of water courses. The most recent humid period occurred around ~10 - 6 Kya, during the Holocene, when an extreme increase of precipitation was felt, especially in the northern areas (around 300% increment). In the last 6 Kya, Arabian Peninsula has been through an intense aridification

period. Although this pattern is very similar to the one verified in Sahara Desert, it is believed that the Arabian Peninsula had short dry and humid phases during the major wet and dry periods, however these phases are still not very well understood.

Species inhabiting the Sahara and Arabian deserts can be broadly characterized by their ecological requirements in: i) xeric (i.e. adapted to arid conditions); ii) mesic (i.e. not fully adapted to arid conditions and dependent of moist habitats) and iii) humid (i.e. water dependent) species (Brito et al., 2014). Currently, xeric species have widespread distributions throughout the desert, while mesic and humid species are present in the peripheral parts or in mountain refuges throughout the desert (Brito et al., 2014). Due to their different affinities and ecological requirements, each species likely reacted differently to the Pleistocene climatic oscillations, adapting to extreme environmental conditions or shifting to more suitable areas (Brito et al., 2014). In general, the distribution of xeric species would have contracted during wet periods, leading to the isolation and diversification in more suitable areas. However, during the dry phases, since the climate was warmer, optimum conditions were reunited for the expansion of these species. Species with other requirements were likely affected in a different way. For instance, mesic species probably expanded their distributional ranges in humid periods, while contracted and suffered population isolation when warmer climate was present. Overall, the emergence of Sahara acted as a barrier to the dispersion of species not adapted to the arid conditions of the region, creating a latitudinal vicariance in a North-South axis, affecting the diversification processes of many species (Carranza et al., 2008). Recent phylogeographic studies revealed that diversification and speciation in some species are mostly related to Sahara's spatial and temporal extension (e.g. Gonçalves et al., 2018a; Tamar et al., 2016). A similar process might be expected for the Arabian Peninsula, due to the formation of the Rub' al Khali and Nefud deserts during the onset of arid conditions. The Arabian deserts extend across most of the territory, which have made it difficult for species to disperse and functioned as a driver of diversification or extinction of many species (e.g. Machado et al., 2019; Šmíd et al., 2013; Tamar et al., 2016).

1.2.2 - Important geographic features across the Sahara and Arabian Peninsula

As previously mentioned, both Sahara and Arabian deserts are climatically and topographically heterogenous regions. Although large parts are constituted by extensive sand areas, important geographic features are present and comprise important biodiversity hotspots.

The Sahara and the Arabian deserts have different mountain complexes scattered along their ranges. On the African part, five mountain complexes can be found: the Aïr Massif (northern Niger), Hoggar Mountains (southern Algeria), Atlas Mountains (Morocco, Algeria and Tunisia), the Tibesti Mountains (north of Chad and to southern Libya) and the Adrar des Iforas massif (northeast Mali to Algeria). In the Arabian Peninsula can be found the Western Mountain Complex constituted by the Hijaz and Asir mountain ranges (along the Red Sea rift), the Hadramaut Mountains (Yemen), the Dhofar Mountains (southern Oman) and the Al Hajar Mountains (north-eastern Oman to eastern United Arab Emirates) (Fig. 1). Mountains have played an important role in the specie's diversification processes by acting as refugia and biodiversity corridors through time (Brito et al., 2014). For instance, Gonçalves et al., 2018b found that central Saharan mountains could have acted as refugia for the mesic species *Psammophis schokari* during unfavourable periods. This species would have reached mountain areas due to the presence of Trans-Sahara corridors, which connected suitable areas for the species along the Sahara Desert (Gonçalves et al., 2018b). It is believed that in the past humid periods, mountains where surrounded by savanna-like habitats which made possible the connection between the different complexes. Today they harbour species with different ecological requirements (e.g. Mediterranean and Sahelian species) in arid regions where such species could not survive if they did not exist (Brito et al., 2014).
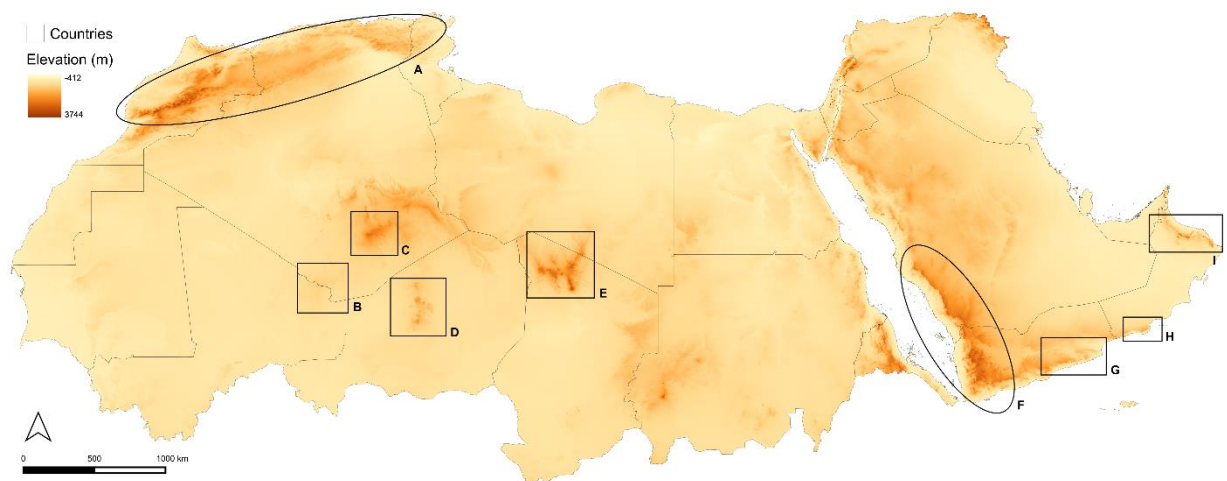
Fig. 1 - Location of mountain ranges in North Africa and Arabian Peninsula. A - Atlas Mountains, B- Adrar des Iforas massf , C- Hoggar Mountains, D- Aïr Massif , E- Tibesti Mountains, F- Western Mountain Complex, G- Hadramaut Mountains, H - Dhofar Mountains and I- Al Hajar Mountains.

In addition to mountains, other biodiversity corridors might have persisted to the present. The Atlantic and Red Sea coastal areas although still hot and dry, have less harsh conditions comparatively to the surrounding regions since mists coming from the

Atlantic Ocean and the Red Sea bring humidity to the area (Olson et al., 2001). Not all biodiversity corridors that existed during favourable periods were maintained when less favourable conditions appeared. However, it is believed that both Western Sahara and Red Sea coastal areas and the Nile river prevailed during these periods, acting as climatic refugia (Brito et al., 2014). Accordingly, Velo-Antón et al., 2018 demonstrated, for the endemic spiny-footed lizards *Acanthodactylus aureus,* the importance of the Atlantic coastal area as a dispersal corridor for biodiversity through the Sahara Desert during climatic fluctuations, which also acted as a centre of lineage diversification. The Nile River and the Sahara dune massifs and empty quarters also acted as corridors that allowed the dispersion of species throughout the Sahara Desert to reach more suitable areas (Drake et al., 2011; Dumont, 1982).

The role and influence of climatic oscillations and geographic features on species´ evolution and genetic structure have been studied across the globe. Several works suggest that mesic and humid species have dispersed through biodiversity corridors during wet periods in the past (e.g. Gonçalves et al., 2018b; Martínez-Freiría et al., 2017; Velo-Antón et al., 2018). Xeric species, although have been less studied (but see Tamar et al., 2016), are expected to have retracted their ranges during wet periods (Gonçalves et al., 2018a, Hoelzmann et al., 1998). Nevertheless, further studies are needed to assess species responses to climatic oscillations in the Sahara Desert and Arabian Peninsula.

1.3 - The Species

Reptiles are one of the most diversified groups of vertebrates in the Sahara and Arabian deserts, many of them presenting important adaptations to arid conditions (Brito et al., 2014, 2016; Carranza et al., 2018 or other works). Their ectothermic physiology and frequent low dispersal ability make them highly dependent to environmental factors and thus, an excellent model to analyse the link between climate and their current distributions (Kearney et al., 2009; Pough, 1980; Sinervo et al., 2010). Within this group, vipers (Serpentes, Viperidae) have particular life-history traits such as low growth rates and frequency of reproduction that likely accentuate responses to climatic oscillations (Martinez-Freiria et al., 2015, 2017; Santos et al., 2006).

The genus *Cerastes* is composed by three venomous species of small-medium body size, which are adapted to arid environments, commonly known as Horned vipers

(Phelps, 2010). The Sahara Horned viper, *Cerastes cerastes* (Linnaeus, 1758) is a Saharan species, being found in North Africa (across the whole Sahara desert) and north-western (i.e. Sinai and the northern Negev deserts) and south-western Arabian Peninsula (i.e. Hadramaut Mountains in Yemen; Fig. 2), inhabiting sandy and rocky desert areas (Bons and Montori, 1996; Phelps, 2010; Werner et al., 1999). The Sahara Sand viper, *Cerastes vipera* (Linnaeus 1758) inhabits the arid regions of North Africa and NW Arabian Peninsula (i.e. Sinai Peninsula; Fig. 2), being found in sandy areas, such as wind-blown dunes (Bons and Montori, 1996; Le Berre, 1989; Phelps, 2010; Welch, 1982). The Arabian Horned viper, *Cerastes gasperettii* Leviton and Anderson, 1967, is found throughout the Arabian Peninsula in sparsely vegetated wind-blown dunes and other sandy areas, being absent in the mountain regions (Egan, 2007; Phelps, 2010; Fig. 2).

The three species are morphologically similar although exhibit some differences: *C. cerastes* and *C. gasperettii* are more similar to each other, both presenting large eyes and similar body sizes (avg. 60 cm, up to 85) (Phelps, 2010). Both species are constituted by individuals with and without horns, characteristic that varies regionally, and can be distinguished by the number of ventral scales (Sindaco et al., 2013). Contrarily, *C. vipera* is the smallest species of this genus (avg. 35 cm, up to 50), has small eyes upwards and lacks horns (Phelps, 2010).

A fourth species, *Cerastes boehmei*, was described by Wagner and Wilms 2010 using only morphological characteristics of one specimen. This specimen was found in Tunisia between Bani Kheddache and Ksar el Hallouf (Fig. 2), and differently to all three species described previously, has crown structures above the eyes formed by erected supraocular scales (Wagner and Wilms, 2010). This species was described as being more closely related to the North African hornless viper *C. vipera* than to *C. cerastes,* but due to the lack of information, doubts on the veracity of the existence of this species have arisen.

Phylogenetic studies, including many other species of viper taxa, revealed that *Cerastes* is a sister taxon to the *Echis* genus, both included in the subfamily Viperinae (Pook et al., 2009; Alencar et al., 2016; Zheng and Wiens, 2016; Šmíd and Tolley, 2019). According to these works, which are based on mitochondrial DNA (mtDNA), *Cerastes cerastes* is phylogenetically closer to *C. gasperettii* than to *C. vipera*. Nevertheless, the evolutionary history of these species and the role of climatic oscillations as driver of their potential intraspecific diversification remains unknown.
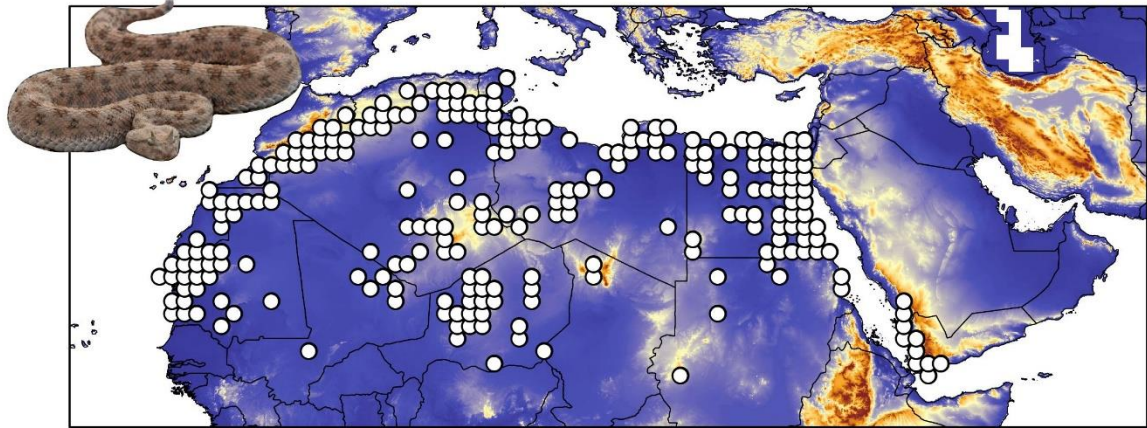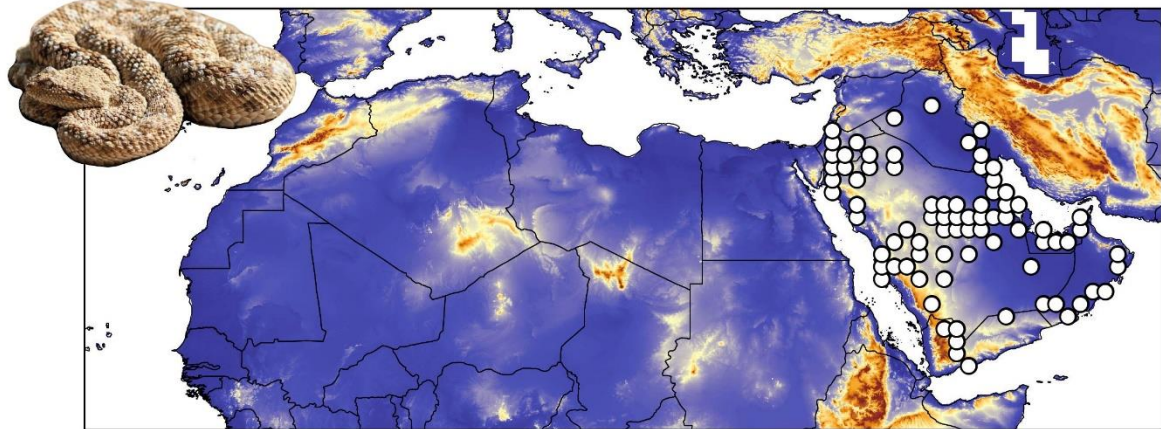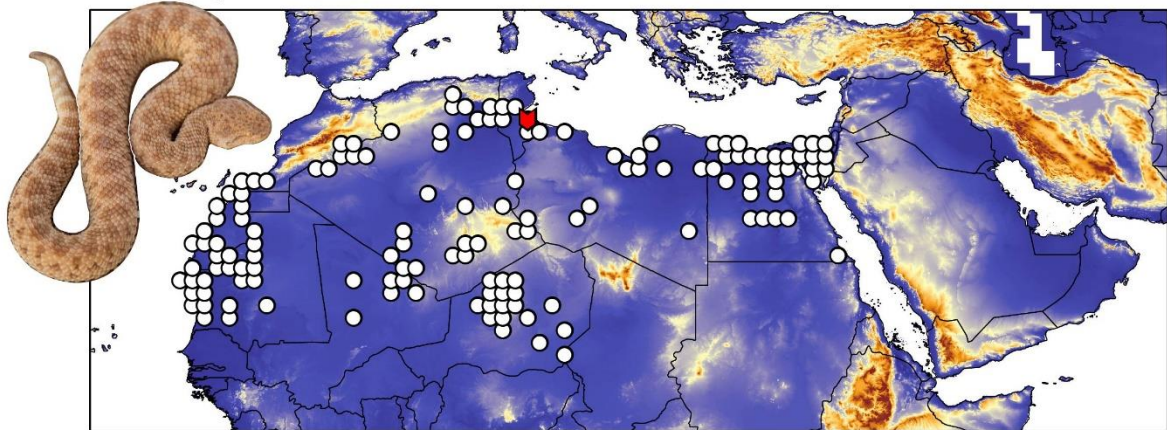
Fig. 2 - Distribution of occurrences for the *Cerastes* genus according to Sindaco et. al, 2013. Representative photos of one specimen of each species are shown in the top left corner of each map. All currently recognized species are represented by white circles, and the only specimen of *C. bohemei* found is represented by a red arrow on the same map of *C. vipera*.

1.4 - Approaches

Phylogeography is a multidisciplinary area that lays between phylogenetics and population genetics. It is used to understand the evolutionary history of species in a biogeographic context (Hickerson et al., 2010). In phylogeographic studies, mitochondrial markers are commonly used to infer intra-specific genetic structure. The choice of a mitochondrial marker facilitates PCR amplifications due to the existence of multiple molecules in each cell, contrarily to nuclear DNA. However, mtDNA only represents a fourth of the effective population size since it is only maternally inherited. By using only the matrilineal evolution history of a species we might fail to infer the true evolutionary history of species (Avise 2000, 2009; Ballard and Whitlock, 2004). Nuclear DNA (nuDNA) is inherited from both parental sides and they tend to show high conservative evolutionary rates (Wan et al., 2004).

Ecological Niche based models (ENMs) use species presence and/or absence data to predict the ecological niche of a species using a set of environmental variables (Franklin and Miller, 2010). These methods allow to map the suitable areas where species are theoretically able to thrive, that is, the geographical areas environmentally similar to the ones that the species prefer, as well as to infer them in other time periods, assuming that the niche was the same over time (Maiorano et al., 2013; Nogués-Bravo, 2009).

Ecological models can be divided in two types of models: correlative and mechanistic (Peterson et al., 2011; Guisan et al., 2017). Correlative models relate the distribution of species with environmental variables, aiming to determine the ecological niche of a species. Mechanistic models combine species functional traits with environmental conditions in order to describe the processes that potentially constrained a species distribution and to map the different aspects of the fundamental ecological niche of a species (Alvarado-Serrano and Knowles, 2014; Wiens et al., 2009). Since it is necessary a vast knowledge on species physiological traits to perform mechanistic models and as most of the times this knowledge doesn't exist, correlative models are the most used in studies(Alvarado-Serrano and Knowles, 2014; Elith et al., 2010). Also, the use of correlative models has the advantage that many presence and/or absence data are available. Presence data is usually more used than absence data because it is largely more available and because it is very difficult to make sure that a species is effectively absent at a certain location (e.g. it may be considered that a species is absent

in a certain region when in reality it has not been detected because it is rare in that region or present a low detectability rate).

ENMs also present some constrains and problems. To truly predict the ecological niche of a species it would be necessary that the total species distribution was assessed, which is most of the time difficult and/or impossible (e.g. Martínez-Freiría et al., 2016). Also, even if we have a lot of knowledge about the species we choose to study, it is very likely that we will not take into consideration all variables that influence its ecological niche or that these variables are not available to use (Wiens et al., 2009). Another problem is the assumption that the ecological niche of a species is an independent unit, not being affected by other biotic interactions (e.g. biological constrains of a species and interaction between different species) (Wiens et al., 2009). So, this gap of information affects the veracity of the model since without it, it will never be possible to predict the true ecological niche of a species.

The use of ENMs allow to respond to evaluate or develop phylogeographic hypothesis and this combined methodology (phylogeography + ENMs) allows a better understanding of species biogeographic and evolutionary histories, since they complement each other very well (Alvarado-Serrano and Knowles, 2014). For instance, this combined methodology was successfully used recently in different works on North African reptiles (e.g. Gonçalves et al., 2018a; Martínez-Freiría et al., 2017; Velo-Antón et al., 2018) to unveil species genetic struture and infer biogeographic scenarios for their evolution. Here, we apply this approach to infer the genetic structure and evolutionary scenarios for the *Cerastes* vipers in North Africa and Arabia.

1.5 - Aims

Despite the available phylogenies including the genus *Cerastes* (Alencar et al., 2016; Šmíd and Tolley, 2019; Zheng and Wiens, 2016), little is known regarding the phylogeographic patterns of variation and the intraspecific relationships within each species, as well as, how climate has influenced these patterns. Similarly, knowledge on biogeographic patterns and ecological requirements for these species is limited to the western populations of Saharan species (e.g. Brito et al., 2011).

The aim of this study is to address the role of climatic oscillations in the evolutionary histories of the three *Cerastes* vipers (*Cerastes cerastes, Cerastes vipera and Cerastes gasperettii*) with special emphasis on the western Sahara region, as it includes the Atlantic Sahara region that likely promoted diversification in the first two species, and for which our sampling is particularly detailed. To do so, currently available georeferenced tissue samples for the three species will be analysed using an integrative approach combining phylogeographic analysis for both mitochondrial and nuclear DNA and ecological niche-based modelling for present and past times.

Phylogeographic analysis are expected to answer the following questions:

a) Are species genetically structured across the region?
b) How are the different lineages geographically distributed?
c) How old is the divergence of major lineages?

Ecological Niche-based Models are expected to answer the questions:

d) Which are the climatic requirements of each species?
e) How suitable is the region for each species/clade over time?
f) Do climatically stable areas match with suitably stable areas for species/clade?

2. Materials and Methods

2.1 - Phylogenetic analyses

2.1.1 - Sampling

Samples from 77 specimens of the three species of *Cerastes*, i.e. *C. cerastes* (n = 51), *C. vipera* (n = 19) and *C. gasperettii* (n = 7), were obtained for phylogenetic analyses (Appendix 3; Fig. 3). Samples consisted of tail-tips, bones and shredded skins, from alive, road-killed or ethanol stored specimens. These samples were collected across North Africa and the Arabian Peninsula by members of BIODESERTS research group and collaborators during field-work campaigns for more than a decade, and visits to three museum collections (CEFE-CNRS of Montpellier, France; Národní Muzeum of Prague, Czech Republic; Natural History Museum of Vienna, Austria; Appendix 3).
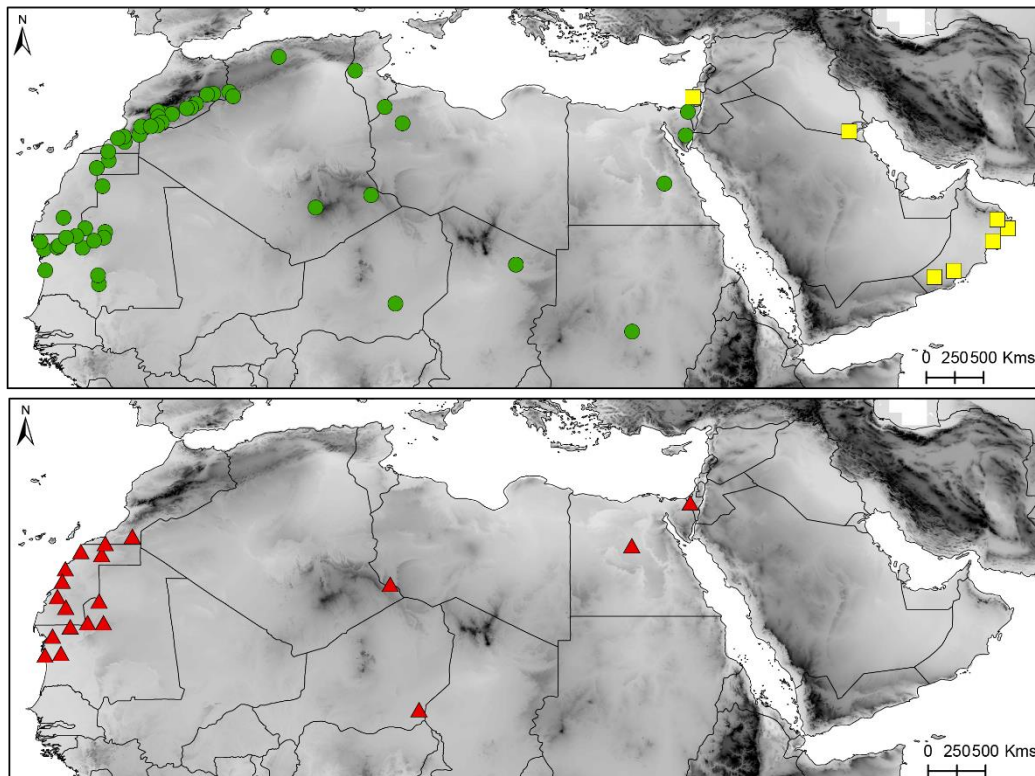
Fig. 4 - Spatial distribution of samples considered in this study. The distributions of *C. cerastes*, *C. gasperettii* and *C. vipera* are represented with green dots, yellow squares and red triangles, respectively.

2.1.2 -DNA extraction

Samples were cut in small pieces and placed in Phosphate buffered saline solution (PBS) over-night at room temperature to remove the ethanol solution and possible impurities present that could act as inhibitors. Posteriorly, total Genomic DNA was extracted using different protocols depending on the type and quality of the samples: QIAGEN's EasySpin protocol for samples collected from alive and road-killed specimens, and QIAGEN 's QIAmp® DNA MicroKit for museum samples, bones, shed skins, and samples with limited amount of tissue. For the samples in which the QIAGEN's QIAmp® DNA MicroKit failed, a "Non-Invasive protocol" optimized for the manipulation of low-quality DNA was performed by a technician from CTM laboratory in isolated and sterile conditions (Appendix 1).

The extraction success and DNA quantity and quality were then verified by electrophoresis on a 0.8% agarose gel died with GelRed™ (Biotium), once total Genomic DNA was extracted. Samples with high DNA quantity and/or degraded DNA where diluted accordingly with ultra-pure water to increase the success of amplification.

2.1.3 - Amplification and sequencing

A fragment of one mitochondrial gene, Cytochrome C oxidase subunit I (hereafter COI), two nuclear coding gene fragments, 3'-Nucleotidase (NT3), Prolactin Receptor (PRLR), and one non-coding nuclear gene, Vimentin (VIM), were amplified by Polimerase Chain reaction (PCR) (Table 1).

Table 1 - List of primers and respective sequences used to conduct the amplification.

Gene	Primer	Sequence	Source
COI	Rep-COI-F	5'-TNTTMTCAACNAACCACAAAGA-3'	Nagy et al., 2012
	Rep-COI-R	5'-ACTTCTGGRTGKCCAAARAATCA-3'	Nagy et al., 2012
NT3	NT3-F3	5'-ATATTTCTGGCTTTTCTCTGTGGC-3'	Noonan and Chippindale, 2006
	NT3-R4	5'-GCGTTTCATAAAAATATTGTTTGACCGG-3'	Noonan and Chippindale, 2006
PRLR	PRLR_F1	5'-GACARYGARGACCAGCAACTRATGCC-3'	Towsend et al., 2008
	PRLR_R3	5'-GACYTTGTGRACTTCYACRTAATCCAT-3'	Towsend et al., 2008
VIM	VIM_Ex5_F2	5'-AACAATGATGCCCTGCGCCA-3'	Pyron and Burbrink, 2009
	VIM_Ex6_R2	5'-CAATATCAAGAGCCATCTTTACATT- 3'	Pyron and Burbrink, 2009

PCRs were performed in a total volume of 10µl, which contained 5 µL of QIAGEN PCR MasterMix (for COI and PRLP amplification) or MyTaq (MyTaqTM Mix, Bioline) (for

NT3 and VIM amplification), 3.2 µl of ultrapure water, 0.4 µl of both reverse and forward primers at a concentration of 10 µM, and 1-3 µl of DNA (approximately 50 ng/µl) depending on the DNA concentration in each sample. PCR conditions were optimized for each gene for the species under study (see PCR conditions on Appendix 2). In all PCRs a negative control was used, which included every reagent except the DNA, to assure possible contamination weren't due to the reagents used. Quality and quantity of PCR products were checked by visual examination in electrophoresis on a 2% agarose gel died with GelRed™ (Biotium), and PCR products were properly diluted when high DNA concentration was observed. Purification and sequencing of PCR products were outsourced to GENEWIZ company (United Kingdom), using ExoSAP and Sanger methods, respectively.

The obtained sequences (see Appendix 3) were manually inspected, edited and aligned using Geneious Pro v.4.8.5 (Biomatters Ltd.). When heterozygous sites were identified in the chromatograms, IUPAC nucleotide ambiguity codes were used. Finally, sequences were translated into amino acids and subjected to a search for STOP codons to ensure there were no pseudogenes.

2.1.4 - Phylogenetic tree analysis

Phylogenetic reconstruction was done primarily for the mtDNA dataset only, as the nuclear markers resulted largely uninformative at intra-specific level and some at inter-specific level (see results). COI most appropriate substitution model was determined in jModeltest (Posada, 2008) using the Bayesian Information Criterion (BIC). The best inferred model was HKI.

The software Beast v1.10.0 (Drummond et al., 2012; Drummond and Rambaut, 2007) was used to perform coalescent-based Bayesian phylogenetic inference on *Cerastes* sequences. Analyses were run using a lognormal relaxed clock, with a constant population size as the coalescent tree prior. Three independent runs were performed and combined with a total of 100 million generations with sampling trees and parameter estimates every 10,000 generations, 10% of the trees discarded as burn-in. The quality of the runs (i.e. parameter convergence) was evaluated by observation of the posterior trace plots and effective sample sizes of all parameters using Tracer v.1.7 (ESS>300). A maximum clade credibility summary tree with Bayesian posterior probabilities (BPP) for each node was obtained using TreeAnnotator v1.7.1 (available in the BEAST package). The resulting phylogenetic tree was visualized and edited with FigTree v1.4.3

(Rambaut, 2016). BEAST analyses were run in the Cipres Science Gateway (Miller et al., 2010).

Time to most recent common ancestor (TMRCA) of main intra-specific lineages within each species were estimated in Beast. Two calibration points based on the split between *C. vipera* and *C. cerastes / C. gasperettii* and between the two later were applied (Zheng and Wiens, 2016). A prior distribution for each TMRCA following a lognormal distribution with a mean of 18.16 (SD= 0.001) and 2.94 (SD= 0.001) was used, respectively. Clock, tree models and running parameters were the same as above.

2.1.5 - Haplotype network analysis

Haplotype networks were constructed for each nuclear fragment. The haplotype phases of the nuclear fragments were determined using a coalescent-based Bayesian method using the PHASE algorithm (Stephens et al., 2001; 2005) implemented in DNAsp v. 5 (Librado and Rozas, 2009). The phases were estimated with 100 iterations, 1 as thinning interval and 100 burn-in iterations. This approach has been proved to have a good performance in haplotype reconstruction (Garrick et al., 2010). The phased datasets of each nuclear gene were used as input data in TCS v1.21 (Clement et al., 2000) and the construction of the three haplotype networks was performed using statistical parsimony implemented in the software. Finally, TCSBU was used to visualize the haplotypes relationship among the three *Cerastes* species (Múrias Dos Santos et al., 2015).

Due to the incoherent results obtained with mitochondrial and nuclear markers, two additional phylogenetic trees where reconstructed using a concatenated nuclear dataset containing all three nuclear markers and another one using a concatenated dataset of both mitochondrial and nuclear markers combined. The same previously described methodology was used, and the most appropriate substitution model obtained was HKI for both phylogenetic reconstructions.

2.2 - Paleoclimatic modelling

In this study, we aimed to determine the stable and suitable climatic areas for each species and respective lineages through time, and thus, our modelling strategy was divided in two stages. The first stage consisted of modelling the paleoclimatic scenarios for the three species regarding the entire distribution of the genus in order to determine which climatic variables are relevant to the distribution of each species and which areas

likely allowed suitable climatic conditions in past and present scenarios. In the second stage, paleoclimatic modelling was conducted at a smaller scale taking into account only the West Sahara. This region is known as an important biodiversity corridor and is thought to have acted as a stable area (refugia) and has a promoter of diversification processes for several species in North Africa (Gonçalves et al., 2018a). Paleoclimatic modelling at the scale of the West Sahara was only performed for *C. cerastes* lineages as this species in the only one showing reasonable levels of structure in this region and also enough number of occurrences to be attributed to the different levels of intraspecific structure (see results).

2.2.1 - Paleoclimatic modelling for the three *Cerastes* species

Sampling

A dataset with a total of 318 occurrences at 10 x 10 km (WGS 1984 datum) was built for the distributional range of *C. cerastes* (n = 180), *C. vipera* (n = 74) and *C. gasperettii* (n = 64), covering all North Africa and the Arabian Peninsula (Fig. 4). Records were collected from fieldwork expeditions conducted by BIODESERTS team members and collaborators, from the Global Biodiversity Information Facility (GBIF; https://www.gbif.org/), scientific articles and from seven museum collections (CEFE-CNRS of Montpellier, France; Doñana Biological Station – CSIC, Seville, Spain; NM of Prague, Check Republic; NHM of Vienna, Austria; National Museum of Natural History, Paris, France; Natural History Museum of London, UK; ZMFK of Bonn, Germany). Some occurrences had the same geographical coordinates thus the "Remove duplicates" function in ArcGIS was used to eliminate duplicates.
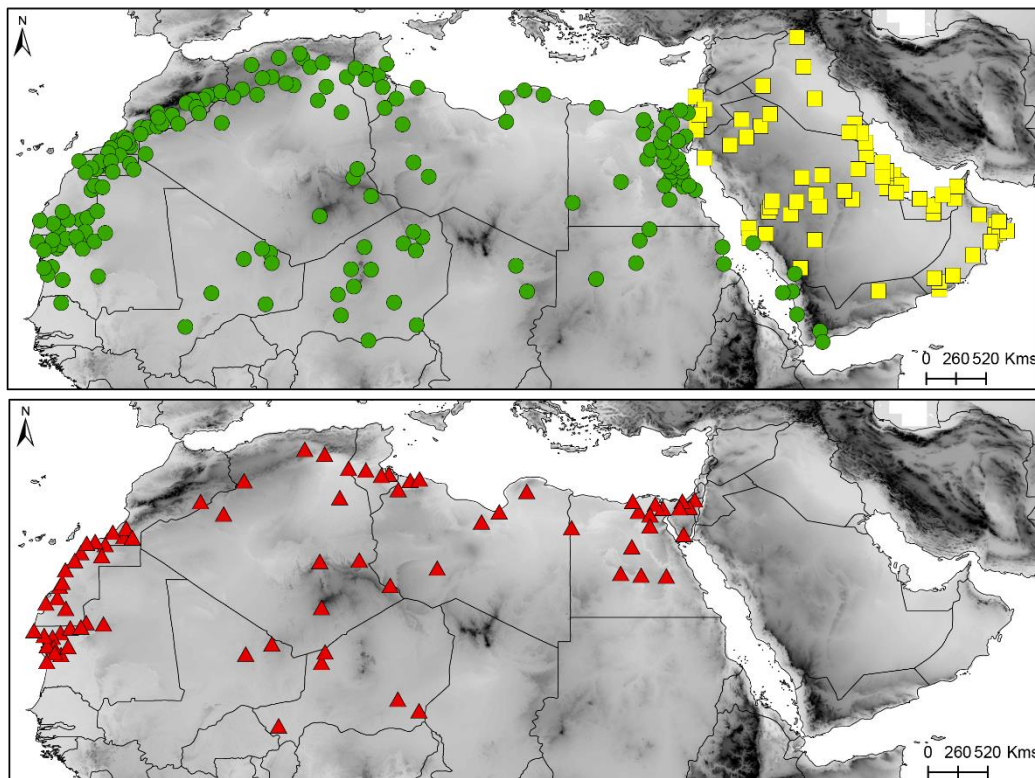
Fig. 5 - Distribution of *Cerastes* genus occurrences used to build the ENMs. Occurrences are represented by green dots for C. *cerastes*, yellow squares for C. *gasperettii* and red triangles for C. *vipera.*

Environmental factors

19 bioclimatic, temperature and precipitation related variables for the current conditions were download from WorldClim version 1.4 (Hijmans et al., 2005) at ca. 10km^2 of resolution (5 minutes).

Mean Temperature of Wettest Quarter (bio 8) and Mean Temperature of Driest Quarter (bio 9) were excluded due to the presence of spatial artefacts (e.g. Martínez-Freiría et al., 2017). Precipitation of Driest Month and Precipitation Seasonality (bio 15) were also excluded due to the high levels of discrepancy between Global circulation models (GCMs) as presented in Varela et al., (2015). Afterwards, the correlation between variables was computed using the "Band collection statistics" tool, available on ArcGis v10.1 (ESRI, 2010). The final set included four low-correlated variables (R < 0.7), with potential biological meaning for the species' biology (Table 2).

Climatic variables available for past conditions were downloaded from WorldCim for three time periods: Last Interglacial (LIG ca ~120,000 – 140,000 years BP; Otto-

Bliesner et al., 2006), Last Glacial Maximum (LGM; ~21,000 years BP, Paleoclimate Modelling Intercomparison Project Phase II) and Middle Holocene (MidHol; ~6,000 years BP, Coupled Model Intercomparison Project #5). Variables were at 30 arc minutes (~1x1 km) for LIG and later upscaled at 5 arc minutes (~10x10 km). Variables were at 5 arc minutes (~10x10 km) for LGM and Mid-Hol. Distinct Global Circulation Models (GCMs) were used for the three time periods. For LIG the NCAR-CCSM (Community Climate System Model; Otto-Bliesner et al., 2006) was used. For LGM and Mid-Hol three different GCMs were available and used: CCSM4 (Community Climate System Model, ver. 4, Collins *et al.,* 2006), MIROC-ESM (Model for Interdisciplinary Research on Climate, ver. 3.2, Hasumi and Emori, 2004) and MPI-ESM-P (Max Planck Institute, Giorgetta et al., 2013).

All variables were clipped to the study area using the Raster Analysis tool in ArcGIS. Afterwards, the raster files were exported in ascii format using the "Raster to ascii" tool, to later use in MaxEnt for the construction of models.

Table 2 - Climatic variables used for building ENMs of *Cerastes*, with respective range variation and units.

EGVS	Range (units)
Temperature Seasonality (Bio4)	745 - 10202 (cv)
Mean Temperature of Warmest Quarter (Bio10)	80 - 381 (10 x ºC)
Precipitation of Wettest Quarter (Bio16)	0 - 2079(mm)
Precipitation of Coldest Quarter (Bio19)	0 - 2077(mm)

Study area and occurrence rarefication

A study area consisting of 300 km buffer around a Minimum Convex Polygon (created with the "Minimum bounding geometry" function) including all occurrences was considered for developing species paleoclimatic models.

Due to the use of different sources for the dataset construction, it was observed occurrence over-representation in the more accessible and surveyed areas of the study area. Sampling bias issues, could affect the quality and reliability of the models (Merow et al., 2013; Yackulic et al., 2013), and thus, species occurrences were subject to a process of spatial rarefication performed using the climatic heterogeneity of the study area as surrogate. The "Spatially Rarefy Occurrence Data for SDM" function from the SDMToolBox package (Brown et al., 2014) for ArcGis 10.1 was used for this purpose. First, a Principal Component Analyses of the 19 bioclimatic variables was derived and then, a climatic heterogeneity was created using the first three PCs of this PCA. For each

species, locations from 30 to 50 km of distance each to the other were removed according to climatic heterogeneity of the study area. The final dataset included 180 records for *C. cerastes*, 74 for *C. vipera* and 64 for *C. gasperettii*.

Species Modelling

Paleoclimatic Models were developed using the Maximum Entropy approach in the MaxEnt ver. 3.4.1 software (Phillips et al., 2006). This method has a good performance comparatively to other methods (Elith et al., 2006), it allows the construction of models using only presence data and a low sample size (Elith et al., 2006; Hernandez et al., 2006; Wisz et al., 2008) and has been successfully used in the construction of models for many snake species, vipers included (Brito et al., 2011; Martínez-Freiría et al., 2015).

In a first stage, model calibration was performed by conducting modelling essays with different values for the regularization parameter and using distinct feature combinations. Best performance options included regularization parameter = 0.5 and linear + quadratic + product + threshold + hinge features for the three species (Supplementary material). Once the options that demonstrated the best performance for our data were chosen, final models were built for the present for each species, with a total of 50 replicates and combining random seed, 80%/20% training/test partition and bootstrap replacement. Evaluation of individual model fit was assessed through the area-under-the-curve (AUC) of the receiver-operating-characteristics (ROC) plot (Fielding and Bell, 1997). Posteriorly, individual replicates were projected to current and past conditions (Mid-Holocene, Last Glacial Maximum and Last Interglacial) using a larger area than was used for training models. Clamping masks and the "fade by clamping" option were used to reduce the probability of inflation of the predictable suitable areas that occur when past climatic conditions fall outside the physiological tolerance of a species. These allows that the values of the projected areas fall in the range of values of the study area (Alvarado-Serrano and Knowles, 2014; Elith et al., 2011).

Average predictions were built for the three species for each scenario in the ArcGIS v. 10.1(Appendix 5, 6 and 7). Prediction uncertainty was assessed by using the Standard Deviation from model replicates for current and past projections (e.g. Brito et al., 2011; Martínez-Freiría et al., 2015). The importance of climatic variables to explain the distribution of each species was determined by the mean percentage contribution to

the models (e.g. Gonçalves et al., 2018a). Response curves of the most important variables for each species were graphically represented (Phillips et al., 2006).

Potential areas that could act as refugia through time for each species were determined by calculating stable areas of climatic suitability (i.e. stable areas) for each species (Carnaval et al., 2009). These areas were obtained by merging the average projections for each scenario obtained with MaxEnt into a single projection, applying the "AND" overlay type in the "Fuzzy overlay" available in ArcGIS (e.g. Present + CCSM4(Mid-Hol) + CCSM4(LGM) + NCAR-CCSM(LIG)). This function allows to calculate the intersection between the model predictions, which represent the maintenance of favourable conditions for each species through time. Posteriorly stable areas for each climatic scenario for each species were displayed in ArcMap.

2.2.2 - Paleoclimatic modelling of *C. cerastes* lineages

To geographically delimit the range of each *Cerastes cerastes* haplogroup, a full genetical analysis of all the records would be needed, requiring tissue samples for all occurrences and financial resources. Since this approach is not realistic, the occurrences not represented in the genetic analysis were assigned to the mitochondrial haplogroups recovered in the spatial interpolation of mtDNA data (e.g. Martínez-Freiría et al., 2017). We used the *phylin* package in R package for this purpose (Tarroso et al., 2015). Firstly, we used the *ape* package in R to extract the branch where *C. cerastes* was located from the complete phylogenetic tree containing *C. cerastes*, *C. vipera* and *C. gasperettii*. Secondly, we used the *phylin* package to create a distance matrix between samples and correlate the geographic distance between points and the semi-variance of genetic distances (Tarroso et al., 2015). Using a modified method of kriging interpolation based in a spherical model, the distances between each sub-lineage (West-N, West-S, NW-SE and SW-NE) were interpolated against the other lineages and sub-lineages present in the study area (e.g. sub-lineage West-N against sub-lineages West-S + lineage East). Rasters depicting the probability of the occurrence of each genetic group (lineage or sublineage) were created. Posteriorly the results were imported to ArcGIS and the values of probability of sub-lineages occurrence were extracted for each occurrence using the "Extract multi-values to points" function. A minimum value of probability of occurrence (0.75) common to all sub-lineages was determined and used to attribute the occurrences to each sub-lineage.

Study area

A study area consisting of 300 km buffer around a Minimum Convex Polygon (created with the "Minimum bounding geometry" function) including all occurrences of lineage West was created for developing species paleoclimatic models for NW haplogroups of *C. cerastes*.

Environmental factors

The four climatic variables for the present and for each past scenario, previously used on Paleoclimatic modelling of the *Cerastes* species distributions, were clipped to the study area using the "extract by mask" function in ArcGIS (ESRI, 2006).

Species Modelling

Paleoclimatic Models were developed with MaxEnt, for the whole distribution of lineage West and sub-lineages West-N and West-S. Models for the present were built using the same settings as in the previous models (see A. Paleoclimatic modelling of *Cerastes spp* distribution), and posteriorly projected to past conditions (Mid-Holocene, Last Glacial Maximum and Last Interglacial). Evaluation of individual model fit was assessed through the AUC metric from the ROC plots (Fielding and Bell, 1997).

Average models/projections and standard deviation plots were built for lineage West and each sub-lineage (NW-N and NW-S) for each scenario in ArcGIS v. 10.1 (ESRI, 2006) (Appendix 8,9 and 10). The importance of climatic variables in lineage and haplogroups distributions, response curves for the most important variables and stable climatic areas were determined using the same methods previously described.

3. Results

3.1 - Genetic analyses

3.1.1 - Laboratory overview

Overall amplification was highly successful for the mitochondrial region (COI) and for two of the three nuclear fragments (NT3 and PRLP), with 82% of success for COI (64 out of 78 samples), 96% for NT3 (24 out of 25 samples) and 100% for PRLR (25 out of 25 samples). VIM nuclear fragment amplification was less successful, with only 68% of the initial samples amplified (17 out of 25 samples). The final sequence alignments had 629 base pairs (bp) for COI, 480 bp for NT3, 523 bp for PRLR and 606 bp for VIM.

3.1.2 - mtDNA phylogenetic reconstruction

The Bayesian mtDNA phylogenetic tree revealed two well supported clades (Bayesian Posterior Probability, BPP > 0.98; Fig. 5): samples of *C. vipera* group in one lineage (in red), which is sister to a clade formed by the sister species *C. cerastes* (green) and *C. gasperettii* (yellow). TMRCA for *C. vipera* from *C. cerastes* and *C. gasperettii* set around 18 Mya, resulted in TMRCA for *C. cerastes* and *C. gasperettii* around 3 million years ago.

Further intraspecific substructure is observed within each species. *Cerates vipera* is structured in two lineages (BPP > 0.98, Fig. 5): lineage CV-West (W) and lineage CV-East (E), which diverged around 2 million years ago. Lineage East is only distributed in the half part of North Africa closest to the Arabian Peninsula, with specimens from Egypt and Israel, while lineage West occupies the half of North Africa closest to the Atlantic Ocean (Fig. 5). The West lineage is further subdivided in two sub-lineages: sub-lineage West-Central (W-C) constituted by specimens from South Algeria and Niger and sub-lineage West-West (W-W) with specimens from Morocco, Western Sahara and Mauritania. The divergence of sub-lineages Central-West and West-West occurred within the last 1 million year (Fig. 5).

Within *C. cerastes*, two well supported main lineages were found (BPP > 0.98; Fig. 5): lineage West (W), distributed in Morocco, Mauritania, Western Sahara, North Algeria, Tunisia and northwest Libya and lineage East (E), distributed in South Algeria (Hoggar Mountains), Niger, Chad, Egypt, and Central Sudan. The divergence of the West and East lineages in *C. cerastes* likely occurred around 2 million years ago.

Lineage West is subdivided in two well supported sub-lineages (Fig. 1): West-North (NW-N), occurring across Morocco, North Algeria, Tunisia and northwest Libya, and West-South (W-NW) distributed across Western Sahara and Mauritania. Lineage East is also subdivided in two sub-lineages: East-SouthEast (E-SE) which includes specimens from Niger, Chad and Egypt and East-Central (E-C) with specimens from South Algeria (Hoggar Mountains) and Central Sudan (Fig. 5). TMRCA between West and East sublineages is dated for less than 1.5 Mya.

Cerastes gasperettii is also divided in two well supported lineages (BPP > 0.95; Fig. 1): lineage North (N) and lineage South (S). Specimens from lineage North are found in Israel, Kuwait and near Oman Gulf coastal area, while lineage South only occurs in the Oman Pacific Coastal area. *C.gasperettii* lineage diversification occurred around 0.5 million years ago.

Haplotype networks

The obtained haplotype networks for the mitochondrial marker are in concordance with the phylogenetic results (Fig. 5). For *C. vipera* three haplogroups were found corresponding to the lineages recovered in the phylogenetic tree. The haplogroup corresponding to the West-West lineage consists of a main haplotype (N = 8), with individuals from Mauritania and Western Sahara, separated by one and four mutations from the other haplotypes. The haplotype, which only differs by one step, is constituted by a single individual from Morocco. The remaining haplotypes are composed by individuals located on the border between Morocco and Western Sahara, and the others branch out of the main haplotype and consist of one haplotype with three individuals of south Western Sahara and west and northwest Mauritania which diverges from the main haplotype by four mutations and another one with one individual of northwest Western Sahara, which diverges from the previous described haplotype by seven mutations. Another haplogroup represents the West-Central lineage and consists of two haplotypes (n = 1) that diverge between themselves in seven mutations and are from Algeria and Niger. Finally, the representative haplogroup of the East lineage, with only two haplotypes consisting of one individual from Egypt and the other by one from Israel, diverging only one mutation.

For *C. cerastes*, there are four distinct haplogroups that are not connected to each other and correspond to the sub-lineages found in the West and East lineages in the phylogenetic tree (Fig. 5). The haplogroup corresponding to the West-North sub-

lineage presents three branches each consisting of three different haplotypes. One of these branches constituted by individuals from Morocco, Algeria, Libya and Tunisia, differing from each other by only one mutation and differing from the remaining branches of haplotypes by four steps. The second branch consists of a main haplotype (n = 6) consisting of individuals from the eastern zone of Morocco, and from which branch out two haplotypes composed by individuals from Morocco. This second and third branches diverge among themselves by three mutations, the latter being constituted by one haplotype with individuals from Morocco, diverging by three other mutations from another haplotype with individuals from Morocco, from which branch out another haplotype with only one mutation of difference, from Tata, Morocco. The haplogroup representing the West-South lineage does not present any structuring, and the haplotypes contained in it are interconnected with each other. This haplogroup is composed by individuals from the Drâa river Valley (Morocco and Western Sahara) to the southwest of Mauritania. The haplogroup constituted by individuals of the East-Central lineage, has two different haplotypes composed by individuals from Sudan and Algeria, respectively. Contrary to what would be expected, the two individuals from Algeria find themselves in separate haplotypes. Finally, the haplogroup corresponding to the East-South East lineage consists of 4 different haplotypes, two of which are formed by individuals from Chad, Niger and Egypt differing 4 mutations from the remaining two haplotypes composed by individuals from Egypt (Sinai) and Israel.

C. gasperettii presents two haplogroups, one representative of the North lineage, consisting of only one haplotype with individuals from Israel, Kuwait and Oman. The second haplogroup, representative of the South lineage, presents three different haplotypes, one haplotype with a single individual of Oman that differs from the second haplotype with individuals of Oman of which branch out a third haplotype with a single individual also of Oman.
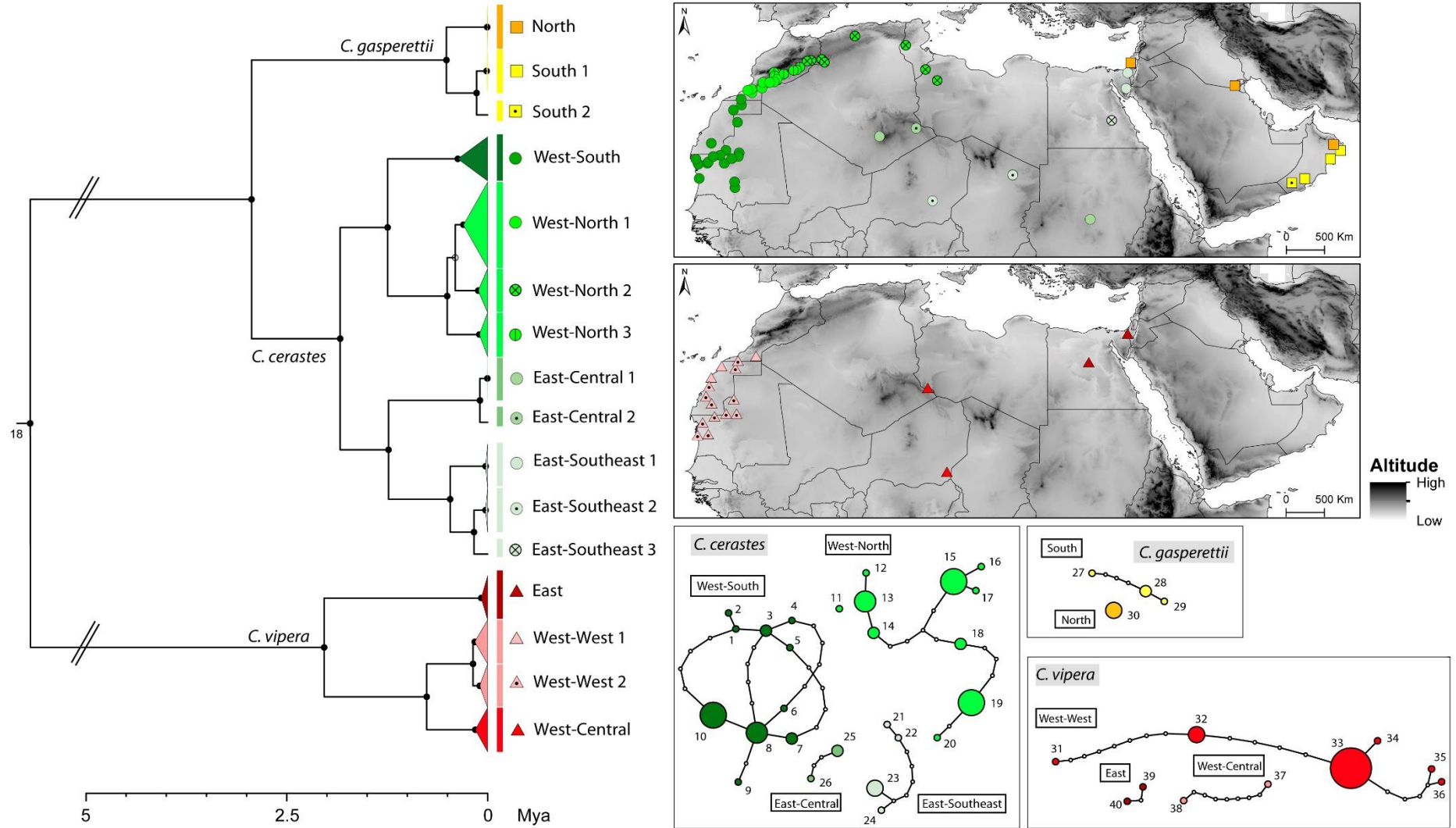
Fig. 6- Time-calibrated Bayesian phylogenetic tree for the three species of *Cerastes* using one mitochondrial marker (COI). Nodes with Bayesian Posterior Probability (BPP) over 90% and 95% are represented with white and black dots respectively. Geographic distribution of the lineages recovered are represented for each species in the top right corner, along with TCS haplotype networks on the bottom right corner.

3.1.3 - nuDNA phylogenetic reconstruction

Haplotype networks

For the nuclear markers, a total of 19, 16 and 25 different haplotypes where found for NT3, PRLR and VIM, respectively (Fig. 6). In PRLR, 12 out of the 25 analysed samples corresponded to heterozygotic individuals, while all the samples were identified as heterozygous for NT3 and VIM. *Cerastes gasperettii* is well differentiated from *C. vipera* and *C. cerastes* in all the three nuclear markers, and it is also represented as an independent haplogroup in VIM. Extensive haplotype sharing is observed for PRLR and NT3 between *C. vipera* and *C. cerastes*, but not for VIM that differentiate all the species (Fig. 6). VIM haplotype network also recovered one intraspecific lineage for *C. vipera*, formed by one single sample from Israel, corresponding to the previously described lineage East found with COI. The remaining intraspecific diversity found with mtDNA was not recovered since different samples of different lineages obtained in the phylogenetic tree and haplotypic network for COI share the same haplotypes in VIM's haplotype network.

Phylogenetic analyses

The results obtained with the nuclear dataset, consisting of the combination of the NT3, PRLR and VIM markers, proved to be very different from the results obtained with the mitochondrial dataset, as was already expected since the nuclear haplotype networks did not resemble the mitochondrial haplotype network. Accordingly, the nuDNA phylogenetic tree showed that *C. gasperettii* was the first species of the *Cerastes* genus to diverge from the remaining species (Fig. 6). Afterwards, there is no divergence between *C. cerastes* and *C. vipera*, but rather the division of the individuals of both species into two groups without support (Fig. 6). Within the groups composed by *C. cerastes* and *C. vipera* individuals, in addition to the relationships not being supported, geographic structuring with some meaning was not found.
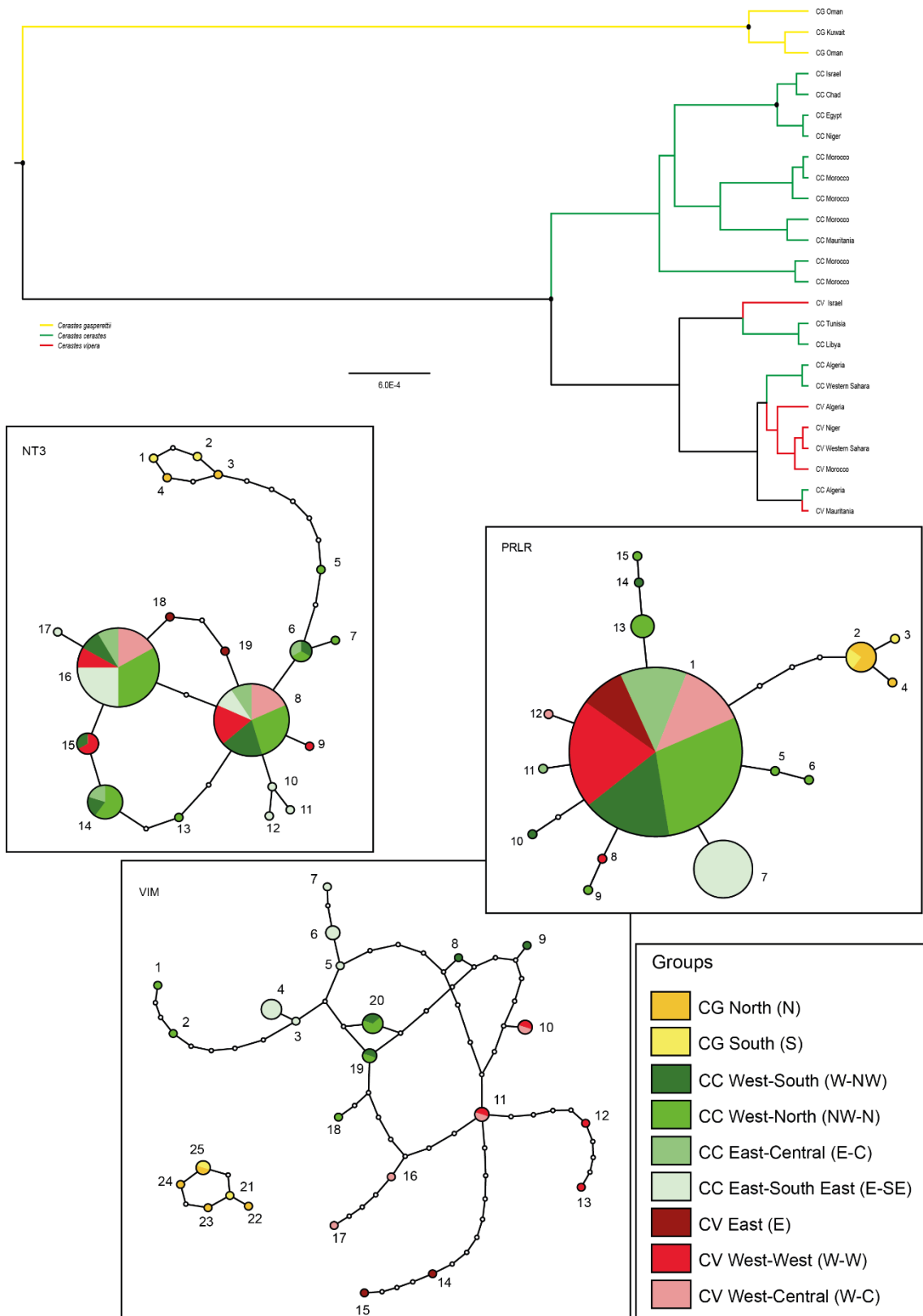
Fig. 7 - Bayesian phylogenetic tree for the three species of *Cerastes* using the combination of three nuclear markers (NT3, PRLR and VIM). Nodes with Bayesian Posterior Probability (BPP) over 90% and 95% are represented with white and black dots respectively. On the bottom, TCS haplotype networks are represented for each nuclear marker, colored according to the obtained groups with the mitochondrial marker.

3.1.4 - Phylogenetic reconstruction using both mitochondrial and nuclear markers

Once the analysis with the different markers obtained incongruent results, a new analysis was done combining the previously used datasets. The phylogenetic reconstruction obtained with the concatenated mtDNA and nuDNA markers (COI + NT3 + PRLR + VIM) is in agreement with the results obtained with the mitochondrial marker COI. *C. vipera* is the first species to diverge and *C. gasperettii* and *C. cerastes* remain together until they diverge later (Fig. 6). However, the sister relationship between *C. cerastes* and *C. gasperettii* is not supported. As for the intraspecific diversity, the same lineages and sub-lineages that were obtained in the analysis with the mitochondrial dataset were recovered, all of which are well supported.
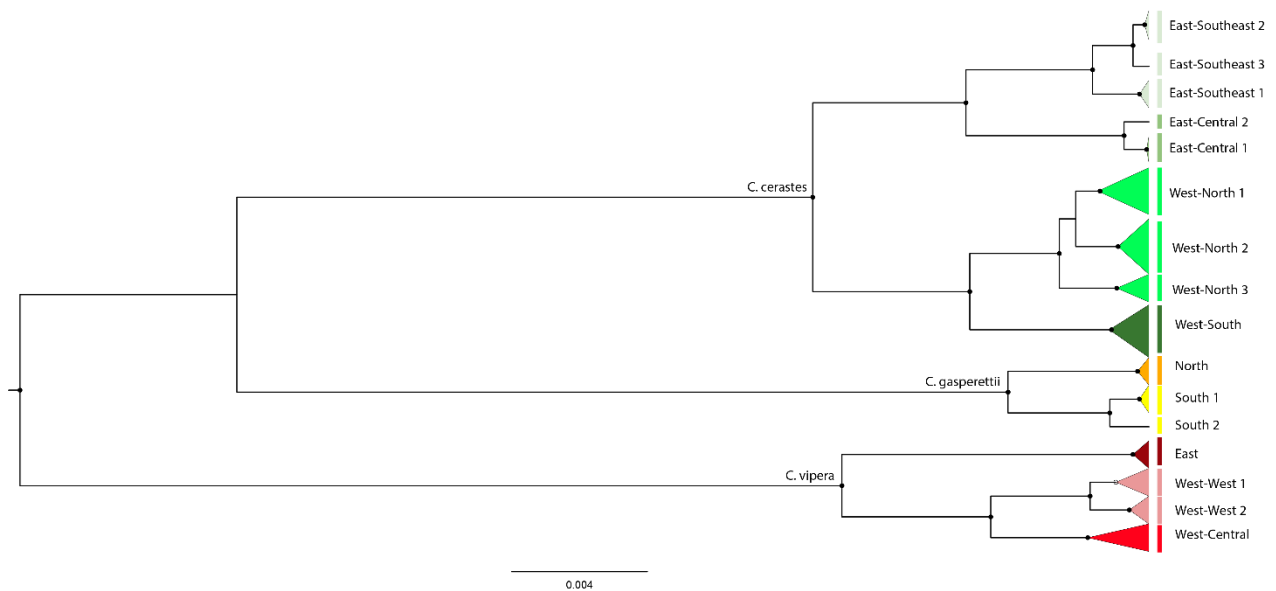
Fig. 8 - Bayesian phylogenetic tree for the three species of *Cerastes* spp using the combination of both mitochondrial (COI) and nuclear markers (NT3, PRLR and VIM). Nodes with Bayesian Posterior Probability (BPP) over 90% and 95% are represented with white and black dots respectively.

3.2 - Modelling

3.2.1 -. Paleoclimatic modelling of the three *Cerastes* species distributions

Models evaluation

Models were developed for the full extent of North Africa and Arabian Peninsula, covering the entire geographical distribution of *Cerastes cerastes*, *C. vipera* and *C. gasperettii*. All AUC values were high (> 0.75) and with low standard deviations in both training and test datasets. The lowest training AUC values were obtained for *C. vipera* (0.855 ±0.023) and the highest for *C. gasperettii (*0.929 ± 0.012), while the lowest test AUC values where obtained for *C. cerastes* (0.788 ± 0.036), verifying again the highest values for *C. gasperettii* (0.895 ± 0.038) (Table 3).

Table 3 - Detailed information regarding the 50 model replicates developed for *Cerastes* spp., including the number of records to train and test the models, average (standard deviation) training and test AUC (area under ROC curve), Minimum training presence Logistic threshold (MTL thr) and Maximum training sensitivity plus specificity Logistic threshold (MTSPSLT).

metrics	*C. cerastes*	*C. vipera*	*C. gasperettii*
N Training samples	144	60	52
N Test samples	36	14	12
Training AUC	0.877 (±0.013)	0.855 (±0.023)	0.929 (±0.012)
Test AUC	0.788 (±0.036)	0.820 (±0.063)	0.895 (±0.038)
MTL thr	0.060 (±0.026)	0.063 (±0.051)	0.073 (±0.039)
MTSPSLT	0.419 (±0.039)	0.350 (±0.078)	0.300 (±0.077)

Eco-geographical correlates

The most important variables related to the species distributions were: Mean Temperature of Warmest Quarter (bio 10) for *C. cerastes*, Precipitation of Wettest Quarter (bio 16) for *C. vipera* and Precipitation of Coldest Quarter (bio 19) for *C. gasperettii*. Temperature Seasonality (Bio 4) and Precipitation of Wettest Quarter also showed a high percentage contribution for the distribution of *C. cerastes*, while bio 19 had the lowest percentage contribution. For *C. vipera* distribution, the remaining three variables (bio 4, bio 10 and bio 19) showed similar percentage contributions. Bio 4, bio 10 and bio 16 clearly showed less percentage contribution for the total distribution of *C. gasperettii* comparatively to bio 19.

Table 4 - Average (standard deviation) percent contribution of each variable to the model replicates for each
Cerastes specie.

Variables	*C. cerastes*	*C. vipera*	*C. gasperettii*
Bio 4	28.7 (±5.83)	21.9(±6.40)	10.2(±3.90)
Bio 10	32.6 (±5.53)	26.2 (±8.09)	16.2 (±3.45)
Bio 16	27.4 (±5.46)	29.4 (±6.15)	11.2 (±6.83)
Bio 19	11.3 (±4.94)	22.5 (±7.45)	62.3 (±4.92)

The most important common EGVs were Mean Temperature of Warmest Quarter (bio10) and Precipitation of Wettest Quarter (bio16). For the first (Fig. 8), response curves revealed differences between all three species: *C. vipera* appears to be more frequently distributed in areas with temperatures between 20 - 25ºC, whereas *C. gasperettii* seems to prefer temperatures between 30 - 35ºC. Curiously, the results suggest that *C. cerastes* selects the same areas as *C. vipera* and *C. gasperettii*, with a decrease in its distribution at intermediate temperatures. As for Precipitation of Wettest Quarter (bio16), both *C. cerastes* and *C. vipera* are restricted to areas with less than 400 mm of precipitation, being more frequently distributed in areas with less than 200 mm of precipitation. Regarding Precipitation of Coldest Quarter, *C. gasperettii* appears to be more frequently distributed in areas with 50 mm or 400 mm of precipitation.
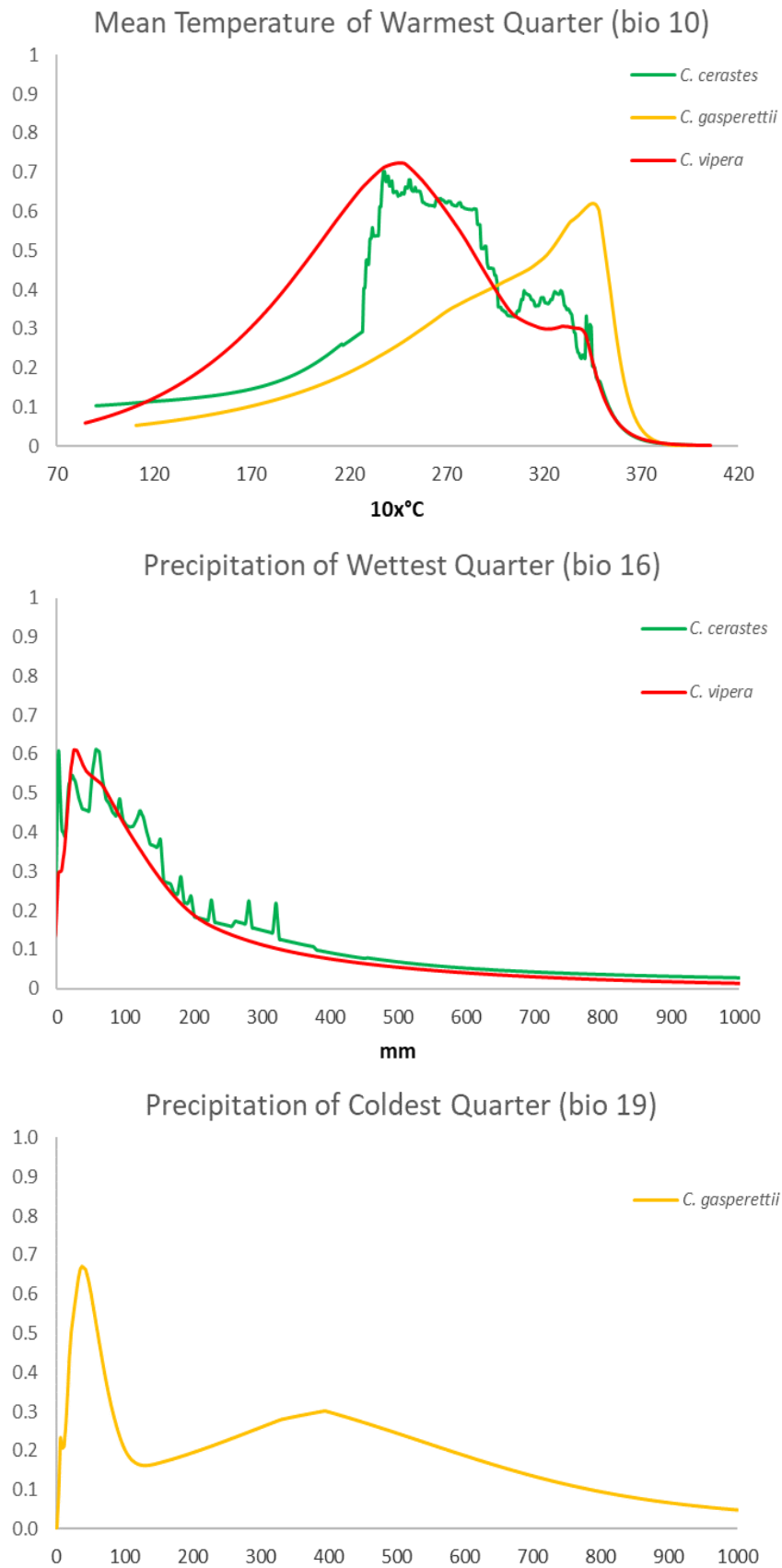
Fig. 9 - Response curves for the bioclimatic variables most related to the distribution of *Cerastes* species.

Predicted suitable areas

The models obtained for *C. cerastes*, *C. vipera* and *C. gasperettii* for the current conditions predicted potential areas of occurrence that overall fit the distribution of the occurrence dataset (Appendix 6, 7 and 8). For *C. cerastes*, the models identified areas with the highest climatic suitability along the Atlantic, Mediterranean, Red Sea, Yemen and Oman coasts. *C. vipera* has high probability of occurrence in the Western Sahara and northern areas of Mauritania, south of Morocco, in Libya and Egypt Mediterranean Coasts, as well as in the Yemen and Oman Pacific Coasts. The climatic suitability of *C. gasperettii* predicted by the model has its highest values in the United Arab Emirates, Saudi Arabian Persian Gulf Coast and Red Sea Coast, as well as some areas in the African continent such as the north Eritrea and northeast Sudan coastal areas and near Marrakech (Morocco), although this is an Arab species. All three models appear to have some areas in which the predicted climatic suitability is low even though occurrences are present, more specifically the southern part of North Africa for *C. cerastes* and *C. vipera* and in the northwest area of the Arabian Peninsula for *C. gasperettii* models.

Concerning projections to the past conditions, all three species mostly showed an identical pattern of expansion/retraction of suitable areas over time (Appendix 5,6 and 7). In the Last Interglacial, suitable areas for *Cerastes cerastes* and *Cerastes vipera* are restricted to the Mauritanian and Western Saharan coastal area, while for *Cerastes gasperettii* they are located in central Saudi Arabia. Subsequently, the suitable areas for *C. cerastes* and *C. vipera* increased greatly, being dispersed throughout North Africa, with the exception of the areas at the south of North Africa, and a large area constituted by todays north-eastern Mauritania, south-western Algeria and northern Mali, during the Last Glacial Maximum period. Expansion of suitable areas for *C. gasperettii* were restricted to the eastern areas of Saudi Arabia. During the Mid-Holocene period, the suitable areas for the three species were contracted, in a similar way to the species' models for current conditions.

Predicted stable areas

The overlapping of models and projections with the fuzzy functions successfully identified stable areas for the three *Cerastes* species (Fig. 9). Regarding *Cerastes cerastes*, high probability of occurrence is seen along the Atlantic coast of North-Western Africa, in the Mediterranean Coast of Libya and Egypt and in the Arabian Coast of Yemen

and Oman. For *Cerastes vipera* stable areas are located along the Atlantic coast of North-Western Africa, not occurring in any other place where the species is currently distributed. Stable areas for *Cerastes gasperettii* occur only along the Red Sea coast of western Arabia between Mekka and Abha, Port Sudan to Eritrea border and between the border of Eritrea and Djibouti and also in north Oman, United Arab Emirates and Qatar. Stable areas for each climatic scenario (CCSM, MPI-ESM-P, MIROC) were very similar concerning each species.
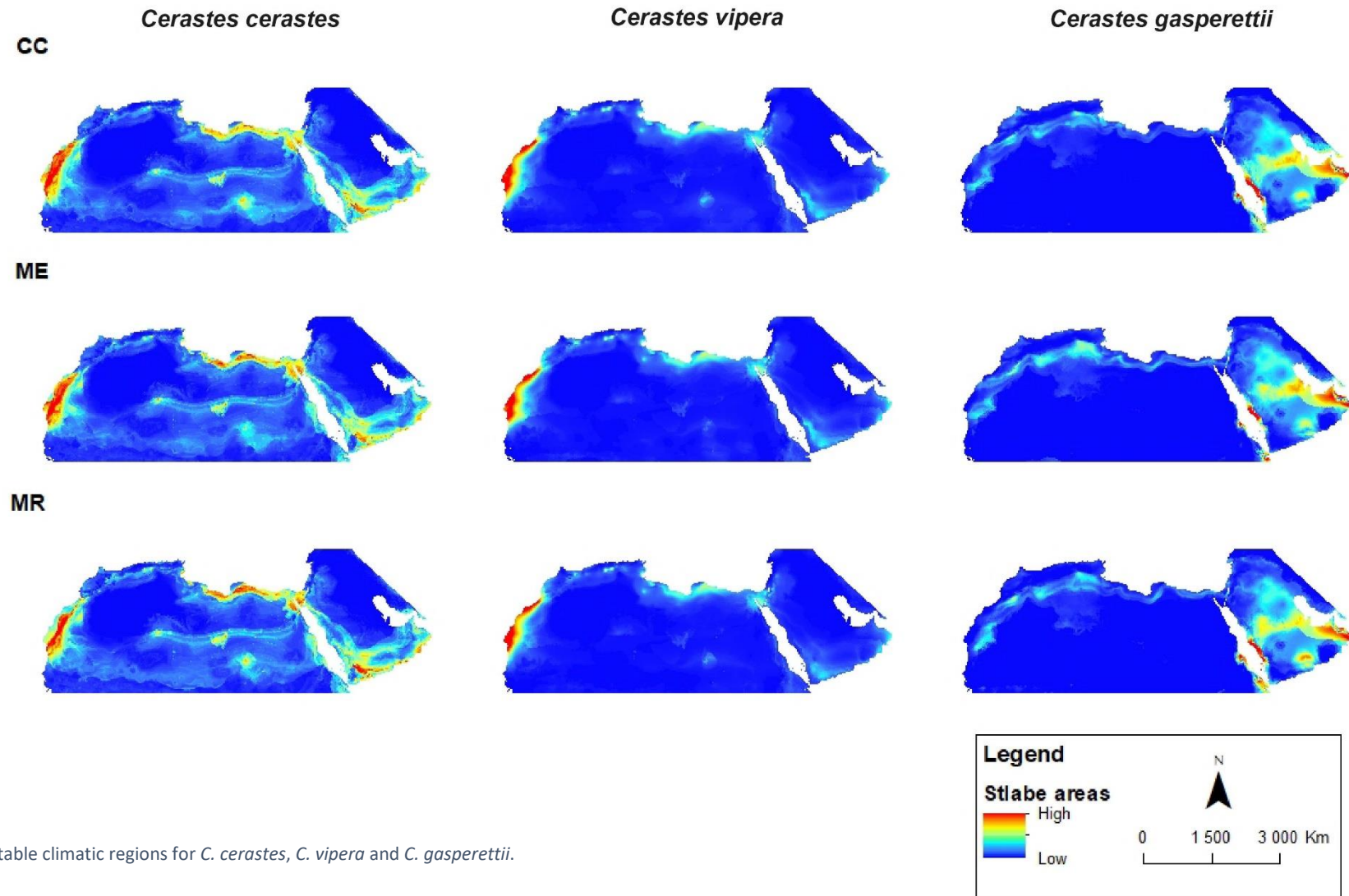
Fig. 10 – Stable climatic regions for *C. cerastes*, *C. vipera* and *C. gasperettii*.

3.2.2 - Paleoclimatic modelling of *C. Cerastes* lineages

Models evaluation

Models were developed for the North-West African area concerning only *C. cerastes lineages*, since it was of our interest to understand the phylogeographic history of the lineage and sub lineages present in this area. Models presented high AUC values (>0.90) with low standard deviations in both training and test datasets. The lowest training AUC values were obtained for sub lineage West-N (0.953 ± 0.008) and the highest for sub lineage West-S *(*0.980 ± 0.003) while the lowest test AUC values where obtained for lineage West (0.925 ± 0.018) and the highest values for lineage West-N (0.970 ± 0.008) (Table 4).

Table 5 - Detailed information regarding the 50 model replicates developed for C. *cerastes* western lineages, including the number of records to train and test the models, average (standard deviation) training and test AUC (area under ROC curve), Minimum training presence Logistic threshold (MTL thr) and Maximum training sensitivity plus specificity Logistic threshold (MTSPSLT).

metrics	West	West-N	West-S
N Tr samples	80	44	32
N Test samples	19	10	8
Tr AUC	0.961 (±0.004)	0.953 (±0.008)	0.980 (±0.003)
Test AUC	0.925 (±0.018)	0.939 (±0.018)	0.970 (±0.008)
MTL thr	0.078 (±0.051)	0.092 (±0.056)	0.163 (±0.086)
MTSPSLT	0.265 (±0.072)	0.218 (±0.065)	0.187 (±0.086)

Eco-geographical correlates

Concerning *C. cerastes* western North African lineages, three of the four variables showed high percentage of contribution in the different ENMs. Precipitation of Coldest Quarter (bio 19) presented the highest values of percentage of contribution for lineage West and sub lineage West-N, and second highest value for lineage West-S. Contrarily, Temperature Seasonality (bio 4) as the highest values of percentage contribution only for lineage West-S, having the lowest values for lineage West and West-N. Precipitation of Wettest Quarter (bio 16) also showed high percentage of contribution for the total distribution of lineage West.

Table 6 - Average (standard deviation) percent contribution of each variable to the model replicates of C. *cerastes* western lineages.

Variables	West	West-N	West-S
Bio 4	13.4 (±3.42)	6.30 (±3.67)	43.2 (±3.03)
Bio 10	11.4 (±5.36)	2.50 (±2.31)	2.56 (±2.01)
Bio 16	30.4 (±3.84)	15.8 (±5.93)	20.8 (±2.65)
Bio 19	44.8 (±4.83)	75.4 (±5.94)	33.4 (±3.56)

For these models, Average profiles of response curves where created for Temperature Seasonality (bio 4), Precipitation of Wettest Quarter (bio 16) and Precipitation of Coldest Quarter (bio19). Response curve for bio 4 suggests that sub lineage West-N occurs in areas with moderate levels of temperature seasonality (3500 cv.). Both lineage and sub lineages occur more frequently in areas with precipitation below 100 mm, observing that sub-lineage West-N occurs in the driest areas (20 mm).
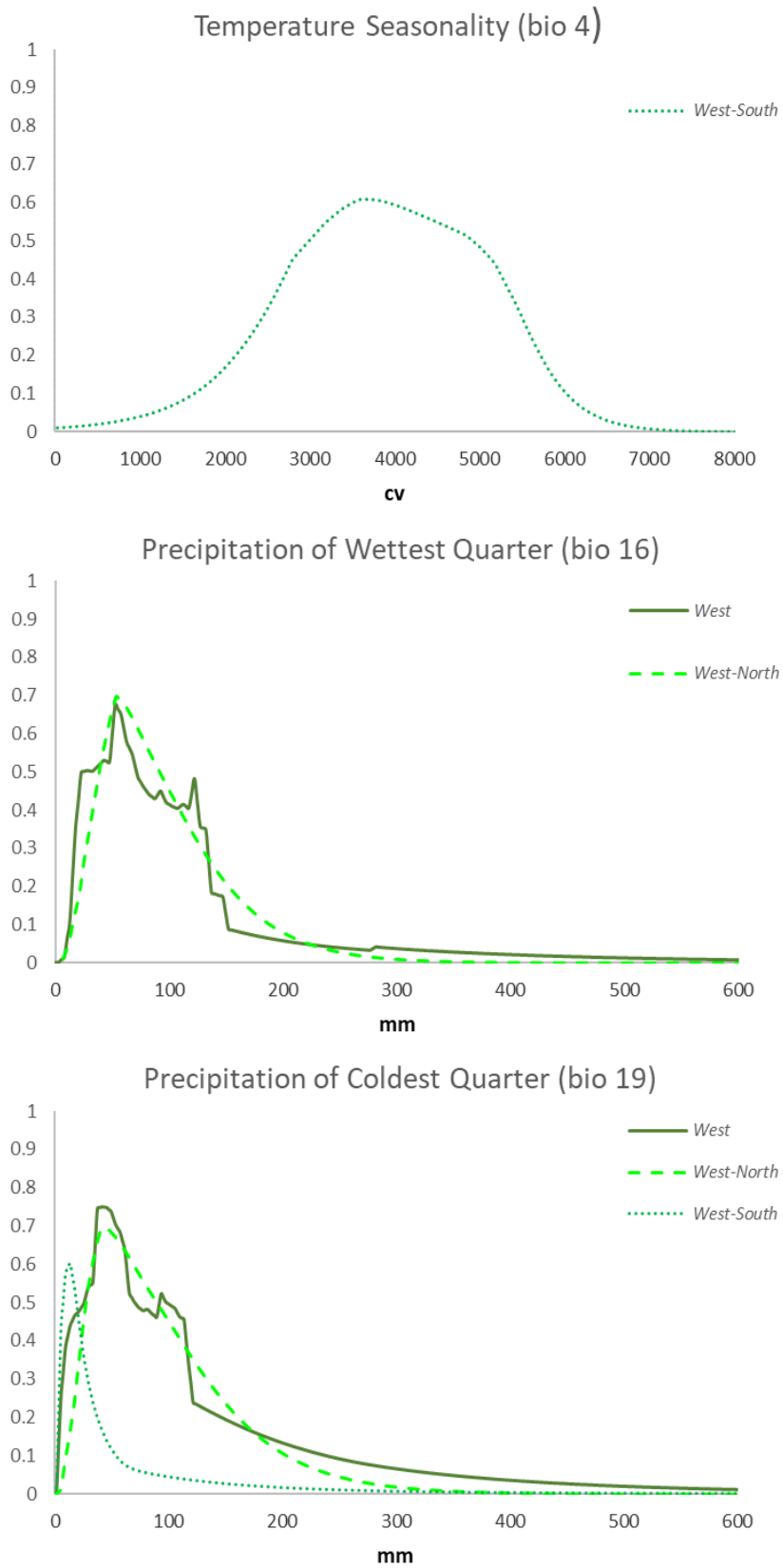
Fig. 11 - Response curves for the bioclimatic variables most related to the distribution of *C. cerastes*
Western lineages

Predicted suitable areas

Overall, models for *Cerastes cerastes* western lineages predicted potential areas of occurrence that fit the distributional data (Appendix 8, 9,10). The predicted models for the past conditions for lineage West show that suitable areas were restricted to the north west coast of Mauritania and Western Sahara during the Last Interglacial period. During the Last Glacial Maximum suitable areas were located along the Mediterranean coast and posteriorly, during the Mid-Holocene period, they were contracted. A similar pattern can be found in the predicted models for the past conditions for sub-lineages West-N and West-S (Appendix. 9 and 10), where it should be noted that in the model of sub-lineage NW-N it is possible to see that the same suitable area predicted for the Last Glacial Maximum period for lineage West and sub-lineage W-NW was also predicted for lineage NW-N, although none of the occurrences used to build these model where located in this area.

Predicted stable areas

Stable areas for lineage West occur approximately between Tan Tan (Morocco) and Saint Louis (Senegal), extending to the interior of Morocco, Western Sahara and Mauritania (Fig. 11). Also, some areas are spotted in the Mediterranean Coast of Tunisia and Libya.

Overall, the models for sub-lineage West-N identified stable areas between Agadir and Ourzazate in Morocco, near Atar, Tergit and Aoujeft in Mauritania and in the Mediterranean Coast of Tunisia and Libya, but nearest the coast than the areas identified for lineage West (Fig. 11). Regarding sub-lineage West-S, stable areas are mostly located in south Morocco, Western Sahara and Mauritania, being found in the same regions as the ones found for lineage West, with the exception for the ones in Tunisia and Libya (Fig. 11).
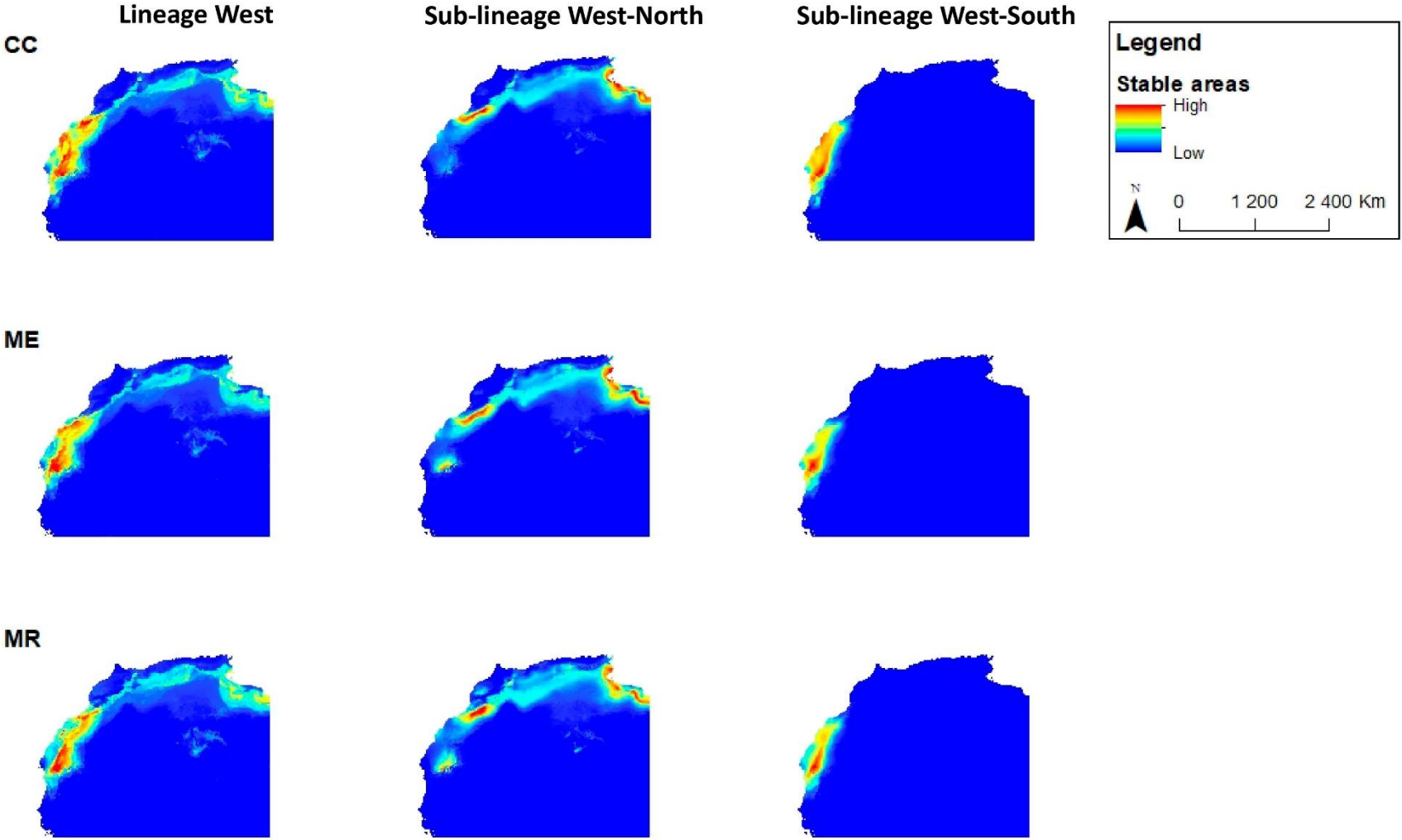
Fig. 12 - Stable climatic regions for C. *cerastes* Western lineages

4. Discussion

This study integrates genetic and ecological analyses to reconstruct the biogeographical history of the species composing the *Cerastes* genus, unveiling both inter- and intraspecific genetic relationships and variability, and signalling areas likely acting as climatic refugia during Pleistocene climatic oscillations. This is the first phylogeographic study in which mitochondrial and nuclear markers for the entire genus are combined, providing insights into the evolutionary history of this group of desert vipers. Furthermore, it is also the first study using ecological models to identify the climatic correlates for the whole distributional ranges of these species (but see Brito et al., 2011) in North Africa and in the Arabian Peninsula.

4.1 - Phylogenetic inferences

4.1.1 Interspecific relationships and mito-nuclear discordances

Similarly to previous studies (e.g. Alencar et al., 2016; Šmíd and Tolley, 2019; Zheng and Wiens, 2016), our mtDNA inferences recovered two main monophyletic clades within the genus, one represented by *C. vipera* and another constituted by two sister taxa, *C. cerastes* and *C. gasperettii* (Fig. 5). In contrast, the results obtained from nuDNA were not concordant to those obtained with mtDNA (Fig. 6): *C. vipera* and *C. cerastes* are included in the same clade, while *C. gasperettii* is in another, being this species the first to diverge. Extensive haplotype sharing was verified for two of the three nuclear markers (NT3, PRLR; Fig. 6), while *C. cerastes* and *C. vipera* are separated but interconnected in the same network obtained with VIM (Fig. 6). Accordingly, the combination of both mitochondrial and nuclear genes (Fig. 7) was concordant with the interspecific diversity found with the mitochondrial marker, however, the relation between *C. gasperettii* and *C. cerastes* no longer is supported. This pattern was never recovered in any previous phylogenetic study on *Cerastes* (Pook et al., 2009)(e.g. Alencar et al., 2016; Šmíd and Tolley, 2019; Wüster et al., 2008; Zheng and Wiens, 2016) and raises doubt about the true phylogenetic relationships and evolutionary history of the genus.

Mitochondrial discordances were also found in other taxa with Afro-Arabian distributions (e.g. African wolves, Koepfli et al., 2015; North-African foxes, Leite et al., 2015). Two possible hypothesis might explain the relation between *C. cerastes* and *C. vipera.* The first scenario is the possible introgression of the nuclear genome of *C. cerastes* in *C. vipera*, due to a past contact between species. Introgression occurs when

a portion of a gene of one entity (species) is incorporated in the gene pool of another entity (divergent species). This process usually happens by hybridization and posterior backcrossing (between the hybrid and one of the parent populations). For introgression to occur it is necessary the existence of gene flow through successive contacts between both populations/species, which are possible due to the existence of contact zones, where species can overlap and exchange gene flow (Harrison and Larson, 2014). *Cerastes cerastes* and *C. vipera* are co-distributed across North-Africa (Phelps, 2010) and although they differ in morphology and ecological requirements, sympatric areas where species can meet and hybridize have been recently identified (e.g. del Marmol et al., 2019; Brito et al., 2011; García-Cardenete et al., 2017). A case of nuclear introgression with absence of mitochondrial introgression has already been described in *Graptemys* turtles (*Graptemys pseudogeographica* into *G. geographica*). Although this study demonstrated that it is possible for introgression to occur only at a nuclear level in reptiles (Mitchell et al., 2016), the introgression between *C. cerastes* and *C. vipera* seems unlikely due to the morphological and ecophysiological differences mentioned previously.

We also consider the hypothesis that the nuclear data used in this study does not have enough power or resolution to delimit taxa due to ancestral polymorphism (occurrence of genetic variation prior to the speciation process). Lack of power and/or resolution in nuclear data it is not unusual in vipers since it has already been documented by several authors in different species of vipers (e.g. *Vipera latastei/monticola* (Freitas et al., 2018; Velo-Antón et al., 2012), *Daboia* (Martínez-Freiría et al., 2017), *Montivipera* (Stümpel et al., 2016)). For this reason, further works should increase nuDNA inferences with more markers (e.g. SNPs, (e.g. SNPs, Schield et al., 2017).

4.1.2 Intraspecific diversity

In this study, the phylogenetic analysis obtained with mtDNA demonstrated a complex intraspecific structure relatively to *C. vipera* and *C. cerastes*, with deep lineages and high levels of genetic structure (FIG 5). *C. gasperettii* presented a relatively simplified intraspecific structure compared to *C. vipera* and *C. cerastes* with the remaining species. Only few specimens of *C. gasperettii* were available to conduct this study, with large sampling gaps, and thus, intraspecific structure within this species must be further studied. All previous studies (e.g. Alencar et al., 2016; Zheng et al., 2016) only

determined the interspecific structure of the genus, never unveiling the intraspecific diversity within each species, making this work the first assessment of the intraspecific diversity for the *Cerastes* genus.

The genetic diversity of the genus regarding North Africa is divided into two distinct regions, the West, ranging from the Atlantic coast to the Hoggar mountains, and the East, ranging from the Hoggar Mountains to the coast of the Red Sea and the Sinai Peninsula. However, populations of *C. vipera* from the Hoggar Mountains belong to the western clade while the populations of *C. cerastes* are genetically closer to eastern populations. The existence of different spatial patterns of genetic structure within these species could indicate that different events may be responsible for their current genetic diversity. However, both divergence events took place during the early Pleistocene about 2-1.7 Mya, pointing to the Pleistocene climatic oscillations as the main driver of this separation between eastern and western populations in both species. Other studies concerning different species inhabiting North Africa present similar genetic patterns and diversification processes occurring in the Pleistocene (e.g. *Agama* lizards, Gonçalves et al., 2018a).

Further lineage diversification occurred during the mid-Pleistocene, starting first for eastern and western lineages of *C. cerastes* (ca. 1.25 Mya), and later in *C. vipera* western lineage (ca. 1 Mya). Similarly to past intraspecific diversification events, allopatric diversification induced by the climatic oscillations during dry-wet periods of the Pleistocene are likely the responsible of such diversification. This patter was already reported for other species inhabiting this region such as *Ptyodactylus* geckos (Metallinou et al., 2015), *Agama* (Gonçalves et al., 2012) and *Uromastyx* (Tamar et al., 2018) lizards or *Echis* vipers (Pook et al., 2009; Robinson et al., 2009).

Given the results obtained in the mtDNA phylogenetic analysis and haplotype networks for *C. cerastes*, further levels of structure appear to exist in central Saharan mountains. The presence of one specimen of *C. cerastes* in Sudan belonging to the same lineage as the specimens of the Hoggar mountains suggest that populations from these two regions were likely connected in the past.

The genetic diversity of the genus in North Africa is mostly concentrated along the western Atlantic coast, where two and three sub-lineages of *C. vipera* and *C. cerastes* were found, respectively. In this region suitable climatic conditions persisted through time, acting not only as a refuge for species during unfavorable periods but also as a biodiversity corridor which will have allowed the connection between the regions at north and south of the Sahara Desert (Gonçalves et al., 2018a; Velo-Antón et al., 2018).

However, during the mid-Pleistocene ca. 1.25 Mya, a geographical or climatic barrier may have existed, preventing the contact between the populations that exist in the north and south of this region since it wasn't found shared haplotypes between these populations. The most likely geographical barrier in this region is the Drâa valley, which during wet and dry periods will have had its river filled or formed a sand body respectively. Since *C. cerastes* is a xeric species that does not inhabit extensive sandy soils, the Drâa valley probably has prevented the passage of populations in both periods, forming a barrier to species dispersal. However, more genetic information is needed in order to determine if gene flow between the northern and southern populations does not occur.

In this study, eastern North African populations were not well represented. Even so, high genetic diversity was found for *C. vipera* and *C. cerastes*, recovering two different populations for the latter: one found only in the Sinai Peninsula and the other one across Egypt, Chad and Niger. It is difficult to identify the possible cause for the diversification of these populations based on these results, although the existence of the Nile river seems to be the only geographical barrier that may have led to this separation. Moreover, the small number of samples for this region may not be representative of the true genetic diversity of the region and therefore future studies should be done for this region with a larger sample size, not only for *C. cerastes* but also for *C. vipera*.

Genetic diversity across the Arabian Peninsula is divided in northern and southern groups, being the Hajar mountains in northeast Oman the only area where both groups are present. The estimated divergence time for these two groups was about 0.5 Mya, during the late Pleistocene. The Hajar mountains could have acted has a stable climatic area for both groups during unfavorable periods, from which they could later have expanded their range and diverged. Another possible scenario is the use of both Dhofar and Hajar mountain ranges as a refuge by the North and South lineages, respectively, thus giving the allopatric isolation of both populations, which may later have expanded their range to the north and west, respectively. As was also previously mention for the eastern region of North Africa, further sample size is needed for *C. gasperettii* in order to understand the intraspecific diversity and evolutionary history of this species.

4.2 – Ecological modelling inferences

4.2.1 - Climatic correlates of species distributions

Despite the widespread distribution of the species that conform this genus, studies addressing environmental correlates related to their distribution are limited to the western region and were developed at regional scale (e.g. Brito et al., 2011; Sow et al., 2014). These studies have shown that the distributional ranges of both *C. cerastes* and *C. vipera* are slightly related to climatic variables, while strongly influenced by habitat variables such as distance to rocky and bare areas or croplands (Brito et al., 2011). Also, it was verified that *C. cerastes* probability of occurrence decreased in regions very close to sandy areas, preferring more distant areas (Sow et al., 2014). In contrast, *C. vipera* presented a high probability of occurrence close to sandy areas, being more frequently found in south plains and sebkhas (smooth flat plains formed by the desiccation of saline or salt lakes) (Sow et al., 2014). Our models, although only considered climatic variables since other types of variables were not available for the entire range for the past, reflected the xeric requirements of these species (see Brito et al., 2014), while showed similarities and particularities in the climatic correlates for the three species.

The distribution of *C. cerastes* was mostly affected by the Mean temperature of the warmest quarter, with the species occurring at the higher ranges (FIG 8). Precipitation also play an important role in the distribution of this species, particularly for the lineages and sub-lineages that are located in West Africa (Fig. 8 and 10), limiting species occurrence to the driest areas (Fig. 4). Although *C. vipera* also selects dry environments, it is mainly distributed in regions with lower temperatures than *C. cerastes.* Even though both species are sympatric across North Africa, distinct climatic, and habitat correlates (Brito et al., 2011), thus suggest that they must segregate across their ranges. The distribution of *C. gasperettii*, for which no information on climatic correlates was available before this work, is strongly affected by Precipitation of Coldest Quarter, preferring sites with approx. 100 mm (Fig. 8). In relation to temperature (i.e. Mean Temperature of Warmest Quarter), the species preferentially occur in areas with high temperature (Fig. 8). Therefore, *C. gasperettii* seems to replicate the pattern of *C. cerastes*, suggesting that the geographical divergence between these two species was led in absence of niche divergence (Peterson et al., 1999 (Peterson et al., 1999; Wiens et al., 2009; Wiens and Graham, 2005).

4.2.2 - Responses to climatic oscillations

Pleistocene cycles had different effects on the distribution and genetic structure of species with different ecological requirements, and in recent years these effects have

been studied for species that inhabit North Africa. According to Brito et al., 2014, during wet periods, xeric species inhabiting this region will have experienced diversification processes due to the decrease of hyper-arid zones (Sahara), which led to population contraction, inciting the divergence of lineages due to allopatric or vicariant events. On the contrary, during the dry periods, the expansion of arid zones would have led to the expansion of the range of these species, allowing the contact between individuals of different populations (Brito et al., 2014). Overall, the results of our models agree with this pattern, which was already expected since the species under study are xeric, and resemble studies done for species with the same ecological requirements in the area.

Our results show that *C. cerastes* and *C. vipera* responded similarly to the different climatic scenarios tested in this study (LIG, LGM, mid-Holocene and present) (Appendix 5 and 6). During the last interglacial period, the reduction of desert like areas in North Africa, caused by an increase in humidity, created the decrease of suitable climatic regions for both species. This loss in xeric environments likely drove to the diminishment of specie's range, which consequently made species seek refuge in more suitable climatic regions (Fig. 9). However, due to the ecological requirements of *C. cerastes* and *C. vipera*, the onset of milder climates, especially in coastal areas, would not have led these xeric species to take refuge in these areas, as it was recovered by our models. Recent paleoclimatic studies conducted for other species in North Africa (e.g. *Psammophis*, Gonçalves et al., 2018b; *Daboia*, Martínez-Freiría et al., 2017; *Agama*, Gonçalves et al., 2018a and *Acanthodactylus,* Velo-Antón et al., 2018) and Iberian Peninsula (e.g. *Vipera seoanei*, Martínez-Freiría et al., 2015) have shown that the Last Interglacial period would have been a very warm period rather than humid as previously though. Considering this new scenario, the coastal regions recovered as suitable climatic areas for both species would thus have gathered the favourable conditions for the specie's existence when no favourable conditions existed in other regions. The reduction in species range and the consequent isolation in more suitable areas will have enhanced allopatric diversification events, that possibly led to the separation of the main lineages.

During glacial periods (LGM), due to the vast increase in warmer and drier conditions, and therefore in xeric environments, a major expansion of species ranges occurred. This range was possibly larger than the one found today, making this period the most favourable for North African *Cerastes* of all the analysed climatic scenarios in this study (Appendix 5 and 6). The climatic conditions that characterized this period will have allowed the contact between previously isolated populations, possibly leading to

interbreeding and genetic admixture. From the mid-Holocene forward, due to the cooling of the temperature and the increase in humidity, similar events as those described for the LIG happened, although more mildly. During this period, it is identified a contraction in specie's ranges due to the loss of suitable climatic areas, inducing population isolation and further genetic structuration.

Although similar climatic events were registered in the Arabian Peninsula, climatic suitability for *C. gasperettii* did not follow the same pattern found for *C. cerastes* and *C. vipera* during the past periods analysed (Appendix 7). During LIG and LGM suitable climatic regions were reduced to small portions of the Arabian Peninsula coast, especially near the mountain regions of Dhofar and Hajar and in the Yemen Highlands. Specie's range expansion only seems to have occurred during the mid-Holocene and the present, due to the increase of suitable climatic areas in central regions that seem to connect the previous suitable climatic areas described for the LIG and LGM periods. This pattern could be explained by the increment of arid conditions during the mid-Holocene period, which led to the dry out of the river network in the southern areas of the Nefud desert. Consequently, the lack of a major water body in the region would have led to the decline and/or disappearance of vegetation and water-dependent and mesic species. The possible disappearance of this river network would not only have led to the installation of more arid conditions in the region, but also would have left the riverbed sediments uncovered, allowing species to move along it, thus enabling the expansion of the *C. gasperettii* range along the interior of the Peninsula (Stimpson et al., 2016). Interestingly no suitable climatic areas were found for *C. gasperettii* in the regions where populations of *C. cerastes* exist in western Yemen.

4.2.3 - Stability in North Africa and Arabian Peninsula

In this study we identified several stable climatic areas for each species using the combination of the climatic scenarios for each Pleistocene period (LIG, LGM, mid-Holocene and present; Fig. 9). Due to the climatic oscillations of the Pleistocene, characterized by dry-wet cycles, these regions were crucial for the persistence of the species through time, since they always sustained suitable climatic conditions, thus acting as refugia when less favourable conditions where present in other regions (Waltari et al., 2007). Overall the identified areas correspond to the suitable climatic regions found in the models for the LIG period (Fig. 9, Appendix 5,6 and 7). This result was already expected as it was during this period that was observed the largest reduction in suitable

climatic areas of all analysed periods. Thus, stable climatic areas were overall found near coastal or mountainous regions in North Africa and in the Arabian Peninsula. During unfavourable periods, mountain areas, characterized by various types of habitat due to their altitude ranges, provide climatic oasis for xeric and mesic species through time.

For both *C. cerastes* and *C. vipera*, similar stable climatic areas were found in the West. These areas are located along the Atlantic coast of Mauritania and Western Sahara, being similar to the current distribution of the West-South and West-West sub-lineages of *C. cerastes* and *C. vipera* respectively (Fig. 9). The sharing of the same stable climate zone may have allowed genetic introgression between these two species and may explain the haplotype sharing found in nuDNA haplotype networks (Fig. 6). However, as it was previously mentioned, the difference in morphology and ecological requirements make this hypothesis quite unlikely The suitable areas found in our models for the LIG and mid-Holocene periods suggest that this area functioned as a refugia during these unfavourable periods, not only for both *Cerastes* species but also for several other taxa (e.g. *Daboia mauritanica*, Martínez-Freiría et al., 2017; *Mauremys leprosa*, Verissimo et al., 2016). In recent studies, Velo-Antón et al., 2018 and Gonçalves et al., 2018a demonstrated that this region is very important for species as it functions as a biogeographic corridor that interconnects the regions above and below the Sahara, as well as a promotor and center of lineage diversification.

Three different stable climatic regions were also found for both *C. cerastes* western lineages, West-North and West-South (Fig. 11). For the West-South sub-lineage, only one area seems to have had favourable conditions for the persistence of the species through time, corresponding mostly to the stable area previously described for *C. cerastes* and *C. vipera* in the West. In contrast to these results, two climatically stable areas were found for the West-North sub-lineage, one located above the West-South sub-lineage stable area and another along the Mediterranean coast of Tunisia and Libya (Fig.10). These regions could have acted has refugia during unfavourable conditions, promoting allopatric differentiation, possibly explaining the separation between the West-North 1 and West-North2+3 populations. Also, the presence of two distinct climatic stable areas for the West-North and West-South lineage between Morocco, Western Sahara and Mauritania support the role of the Drâa valley as a geographical barrier for species divergence, being the most likely responsible for the separation of both lineages.

Two other stable climatic regions were found for *C. cerastes* along the eastern Mediterranean coast of North Africa and southern areas of the Arabian Peninsula. Only a small fraction of the area of the first region, located between Libya and the Sinai Peninsula, corresponds to the Eastern lineage distribution area (Fig. 8). The remaining areas seem to have only been suitable for *C. cerastes* during the LGM and mid-Holocene, with a continuous decrease in suitability through time (Appendix 5). It is also relevant to pinpoint that despite this reduction in specie's suitable climatic areas, individuals were found distributed in the southernmost regions of North Africa for which no suitable areas were found at present time. A possible explanation for this could be the lack of use of other types of variables, such as habitat or soil type, other than climatic, in the construction of the models. These variables most likely play an important role in species distribution, given that, for example, this species does not occur in sand dune environments unlike *C. vipera*. These results indicate this stable climatic area served as a refugia for eastern populations, contributing to lineage diversification within the species.

Finally, the presence of a stable climatic region in southern regions of the Arabian Peninsula, especially near the mountain areas of south-west Yemen, are consistent with the presence of populations of *C. cerastes* in the Arabian Peninsula (Fig. 9) There is still a great lack of information about the lineages of this species in this region since in this study it was not possible to determine their phylogeographic patterns because we had no access to samples and no other studies were available. We can only hypothesize that these populations were confined to this region during the LIG, losing contact with the populations of the Sinai Peninsula (Fig. 9). Posteriorly, this populations might have expanded their range along the coast of Saudi Arabia and Yemen as these zones appear to have favourable conditions for the species to exist.

Stable climatic regions for *C. gasperettii* appear mostly near the mountain complexes of Dhofar and Hajjar mountains and in a small portion of the Yemen Highlands. The stable climatic area in the Hajjar mountain coincides with most of the current distribution of the species lineages. This region may have acted as a speciation center for *C. gasperettii* as both lineages are present in this region, and their current distribution may have resulted of expansion events in more recent periods from this region to the central regions of the Peninsula. The mountainous areas of the Arabian Peninsula are characterized by a diversity of habitats sustained by the milder conditions present in these areas, due to the proximity of the seacoast and by the different altitude

of its cliffs. These areas contrast with the flat inland areas constituted mostly by large sand masses and sustain biodiversity and high levels of endemicity.

4.3 – Biogeographical history of *Cerastes* spp during the Pleistocene

Pleistocene cycles have been responsible for multiple events of species expansion and contraction, leading to population separation due to persistence and isolation in climatic areas and later expansion to suitable climatic areas, promoting the diversification and formation of the current existing lineages.

Cerastes and *C. vipera* began to diversify around 2-1.7 Mya and thus, diversification of both species was likely under the influence of the same events. The Quaternary period is responsible for the intra-specific diversification of both *C. cerastes* and *C. vipera* and was characterized by multiple dry-wet cycles, which had led to multiple range expansions and contractions of both species. Since the *Cerastes* genus is only constituted by xeric species, unlike Mediterranean and Sahelian species inhabiting North Africa, warm and dry periods likely favoured the expansion of their range, while cold and humid periods lead to the contraction of the species in climatic refugia. Major radiation processes in North Africa during this period are already described in other species of reptiles, caused by the wet-dry cycles that characterized the Pleistocene. As previously described, these cycles were responsible for the expansion and contraction of the arid areas of the region, causing the fragmentation, expansion and contraction of populations over time.

Since our results do not allow us to determine the origin of the *Cerastes spp.* ancestor, two possible scenarios may be hypothesized for the evolutionary history of *C. vipera*. The first scenario considers a possible expansion of the species from East to the West along North Africa. Contrary, a second scenario would involve the expansion from the West to the East. The combination of our genetic and ecological results point to the first scenario as the most likely hypothesis, with the expansion beginning from the Sinai Peninsula or surrounding eastern areas, heading west through the Mediterranean coast until reaching the Atlantic coast, the only apparent climatic refuge during unfavourable periods throughout the Pleistocene. Although the scenarios tested here occurred posteriorly to the separation of the main lineages (*C. cerastes* and *C. vipera* West and East lineages), similar wet-dry events will have occurred during the last 5 Mya (Trauth et al., 2009). This scenario is also supported by other studies, such as Gonçalves et al.,

2018b, in which similar suitable climatic areas along the entire Mediterranean and Atlantic coast of North Africa were recovered for *Psammophis schokari* group. As in this study, climate seems to play an important role in the diversification of both species, which was expected given specie's xeric requirements.

The Atlas Mountains situated along Morocco and Tunisia appear to be a major barrier for the dispersion of *C. vipera*, as they mark the limit between Mediterranean and desert ecoregions. As described previously, in this region were found two different sub-lineages located in Mauritania, Western Sahara and south of Morocco (West-West 1 and West-West 2). The presence of rivers in the area during wet periods may have functioned as a geographical barrier to species dispersal, promoting the diversification of the West-West sub-lineages.

The current distribution and diversity of *C. cerastes* was likely a result of an evolutionary history similar to *C. vipera*, as both have a very similar geographic distribution (North Africa) combined with similar divergence times and a division in East-West lineages. The origin of this species most likely occurred in the east of North Africa or in the Arabian Peninsula, since the sister taxa *C. gasperettii* only occurs in the Arabian Peninsula. Subsequently during warm periods, the specie's range will have expanded allowing it to reach the western coast of North Africa. Posteriorly, due to the onset of unfavourable conditions, western and eastern populations split, taking refuge on the west Atlantic coast and along the Mediterranean coast between Libya and Egypt, respectively. These two areas presented favourable conditions for the persistence of the species throughout the different Pleistocene periods and would have functioned as refugia for the species.

Further genetic diversification will have resulted from successive climatic fluctuations along the Pleistocene. Regarding the Eastern lineages, the possible isolation of individuals in the Hoggar Mountains of Algeria and others in the eastern stable climatic areas will have led to the separation of the East-Central and East-Southeast sub-lineages. Subsequently, the East-Southeast lineage will have expanded its range southwest from Egypt to Chad where further diversification events will have taken place. The range and diversity of the East-Southeast lineage is in agreement with the results obtained in the study by Gonçalves et al., 2018b, which points to the possibility of a corridor during warm conditions such as the ones found during LGM period, which allowed the connection between the eastern stable climatic area and all areas between this and the mountainous areas of Chad.

As for the East-Central sub-lineage, the presence of one individual from Sudan genetically similar to one individual in southwest Hoggar mountains suggests a possible route between both regions through the corridor or via a southern route to reach their current location. In previous studies, it was found that a greater number of species colonized these mountains through the north (e.g. *Psammophis schokari* (Gonçalves et al., 2018b), *Pelophylax saharicus* (Nicolas et al., 2015), *Bufotes boulengeri* (Nicolas et al., 2018). However, Gonçalves et al., 2018a determined that *Agama* lizards possibly colonized the mountains through the south.

At the same time, similar climatic events will have occurred in western North Africa, separating the West-North and West-South sub-lineages. Our models suggest that in the past both lineages shared the same area during the Last Interglacial and Last Glacial Maximum, so posterior contact between both lineages could have happened. Pleistocene climatic oscillations also played a strong role in the process of lineage diversification in western North Africa, however a geographical barrier, the Drâa valley seems to have functioned as a major geographic barrier for the contact between both sub-lineages. A more recent diversification of the West-North sub-lineage took place, originating three different populations located along Atlas Mountains. This geographic feature seems to have acted as a topographical barrier to the dispersion of individuals between populations.

Cerastes gasperettii had a later diversification compared to the North African *Cerastes* sister species. Our results suggest that this species only begun to diversify at around 0.5 Mya, resulting on the separation of two lineages (North and South). Due to the lack of sedimentary deposits, little is known about the geological and climatic history of the Arabian Peninsula and for this reason very few assumptions can be drawn on the evolution of *C. gasperettii* (Pook et al., 2008). Similarly to North Africa, the Arabian Peninsula suffered from severe aridification 2 Mya, leading to the formation of the Rub Al-Khali desert. During the Quaternary, this area was characterized by periods of humidity and rainfall and dry and arid periods (Kotwicki and al Sulaimani, 2009). During humid periods, the more favourable conditions in the area made possible the migration between northern and southern areas throw the central regions (Presseur, 2009). However, although similar climatic events to the ones registered for North Africa took place in the Arabian Peninsula, *C. gasperettii* seems to have behaved differently through the same periods.

4.4 – Unresolved biogeographic questions

The combination of the genetic analysis and ecological niche-based models for the different Pleistocene periods has brought insight into the evolutionary history of the *Cerastes* genus. However, several biogeographical questions still remain that could not be answered based on our results.

(1) Where did the *Cerastes* genus originated?

The Viperinae subfamily, to which the *Cerastes* genus belongs likely originated in Africa, as suggested by Wuster et al., 2008. Even though a general lack of knowledge exists about the history of the genus, the ancestors of the genus diversified during the mid-Miocene, around 18 Mya. However, it is unknown in which part of the African continent the ancestor of the *Cerastes* genus emerged, but, the occurrence of the 3 species in the Sinai Peninsula pinpoint to an important role of this region in the diversification of the genus.

2) What is the relationship among the three species and when did *Cerastes vipera* diverged from the remaining *Cerastes*?

During the late Oligocene and Miocene periods, North of Africa and the Arabian Peninsula underwent through several geological events that turned this area very unstable. This period was characterized by the opening of the Red Sea, the formation of the Gulf of Aden and the African Rift System causing the separation of the Arabian Peninsula from the African Continent (Hughes et al., 1991; Chorowicz, 2005). Although being separated, it is believed that a passage connecting Africa to Eurasia existed approximately around 18 Mya., the Gomphotherium landbridge (Rogl et al., 1998, 1999; Chorowicz, 2005). It is hypothesized that this land bridge allowed the passage and exchange of species between Africa, Arabia and Eurasia, therefore promoting the range expansion of several taxa. Based on our mtDNA phylogenetic results, these geological events could have been responsible for the separation of *Cerastes* species: *C. vipera* would have remained in Africa and *C. cerastes* and *C. gasperettii* would have colonized the Arabian Peninsula. Alternatively, *C. vipera* alone could have diverged in Africa, leaving the remaining species in the Arabian Peninsula. Similar diversification events have been suggested for other genus with Afro-Arabian ranges (e.g. *Naja,* (Pook et al., 2009; Wüster et al., 2008); *Echis*, (Pook et al., 2009); *Mesalina*, (Kapli et al., 2015). However, based only on nuDNA phylogenetic tree and the geologic changes observed

in this area, it may have occurred first the separation of *C. gasperettii* from the remaining *Cerastes*, following the separation of the African continent and Arabian Peninsula, with posterior colonization's and contact zones along the Sinai Peninsula.

Posteriorly, both the African continent and the Arabian Peninsula went through aggravated desertification/aridification processes, which led to the formation of the Sahara and Arabian deserts around 7 Mya. In addition to the formation of deserts, due to the orogenic processes resulting from the collision between the tectonic plates, large mountain ranges were formed such as the Atlas in Morocco, the Zagros in Iran and the mountain ranges in Yemen (Bohannon et al., 1989; Rogl et al., 1999, Popov et al., 2004; Bosworth et al., 2005; Edgell, 2006; Jolivet et al., 2006; Mouthereau, 2011). The increasing formation of arid habitats in the African continent would have led the ancestral of *C. vipera* to adapt and develop the present characteristics that make it morphologically very different from the other species of the *Cerastes* genus.

The onset of arid conditions combined with the inconsistent results obtained in the mtDNA and nuDNA may raise another hypothesis for the possible events that led to the origin of *C. vipera* as an ecotype of *C. cerastes,* having diverged in ecological niche, adapting to sandy environments. An ecotype is a distinct form within a species that is genetically different due to an adaptation to a specific environment (Begon et al., 2016). The increase of arid conditions during the Miocene in North Africa and Arabia would have triggered this process and has also been associated with other splits between closely related species, as for example *Stenodactylus* and *Stenodactylus mauritanicus* (Metallinou et al., 2012) and North-African *V. vulpes* and *V. ruepellii* (Leite et al., 2015). However, due to the results obtained in the mtDNA phylogenetic tree, *C. vipera* and *C. cerastes* were recovered as sister taxa, pinpointing to a past divergence.

(3) <u>What are the most likely dispersal/colonization routes for *Cerastes?*</u>

The increase in arid conditions around 3 Mya likely lead to the divergence of the species currently recognized as *C. cerastes* and *C. gasperettii*. We hypothesize that *C. cerastes* could have colonized Africa through dispersion westwards, while *C. gasperettii* remained in the Arabian Peninsula. As already described for other Afro-Arabian species (e.g. *Naja haje*, Wüster et al., 2008), *C. cerastes* may have reached Africa by dispersion along the Red Sea coast from the Arabian Peninsula, reaching Sinai and through there colonizing Egypt. Another possible scenario is that *C. cerastes* may have crossed the Red Sea southwards, through the strait of Bab-el-Mandeb. During glacial periods, a land

bridge between Yemen and Ethiopia would have existed and connected the Arabian Peninsula and Africa. This land bridge would have allowed the migration of several species between these areas (e.g. *Bitis arietans*, Barlow et al., 2019), being present for the last time in the Pliocene around 3.6-2.6 Mya (Haq et al., 1987). Similar hypothesis of dispersion were suggested in other reptile species such as *Mesalina guttulata* (Kapli et al., 2015) and *Uromastyx ocellata*(Tamar et al., 2018), since these species occur in both sides of the Red Sea as *C. cerastes*.

With all the possible hypothesis previously mentioned and being unable to understand which of these hypothesis seems more likely to explain the evolutionary history of the *Cerastes* genus, we reinforce the need for a more complete phylogenetic analysis to fill the gaps that currently exist in areas for which samples were not available, combined with further ecological models including more variables other than climatic.

5. Conclusion

This study, using the combination of ecological and genetic approaches, brought some insights on the evolutionary history of the desert species *Cerastes* genus, as well as, on the diversification processes and the impact of past climatic events on the distribution and diversity patterns of the species. However unexpected results were also obtained, which made it very complex to understand the biogeographic history of the *Cerastes* genus. Overall, climatic scenarios determined by the ecological models were concordant with the mtDNA phylogenetic results. However, the incongruences between mitochondrial and nuclear phylogenetic reconstructions raises doubts about the true relationship between the species and the biogeographic history of the group. In order to unveil a more convincing biogeographical history of the group, a more complete phylogenetic analysis should be done, covering the current sampling gaps and with a larger number of nuclear markers or using microsatelites. A more complete analysis with a greater number of samples from unsampled regions in Africa such as Algeria, Libya, Niger, Sudan and Chad for *C. cerastes* and *C. vipera*, as well as the majority of *C. gasperettii* distribution and of the Arabian populations of *C. cerastes* would also benefit the understanding of the evolutionary history of the genus. Further phylogenetic assessments should include the existing samples characterized as *C. bohemei* to determine if these specimens really constitute a different species and determine its phylogenetic relationship in the genus or whether these individuals only represent specimens of *C. cerastes* or *C. vipera* with supraorbital malformations.

Due to the life history constrains of vipers, such as ectothermic physiology and low dispersal rates, and because deserts environments are expected to be the highly affected by climate change, *Cerastes* species are expected to be very vulnerable to range contractions in future times (Alencar et al., 2018). For this reason, it is necessary to unveil the biogeographic patterns of the species and to predict the impact that future climatic conditions will have on these desert vipers. On this note, the findings of this work combined with other studies of other species in North Africa and in the Arabian Peninsula can be used as a starting point to design conservation units based on evolutionary realms that should be considered for protection and future management.

References

Al-Juaidi, F., Millington, A.C., and McLaren, S.J. (2003). "Merged remotely sensed data for geomorphological investigations in deserts: examples from central Saudi Arabia." *The Geographical Journal,* 169(2), pp. 117-130.

Alencar, L.R.V., Quental, T.B., Grazziotin, F.G., Alfaro, M.L., Martins, M., Venzon, M., and Zaher, H. (2016.) "Diversification in vipers: Phylogenetic relationships, time of divergence and shifts in speciation rates." *Molecular Phylogenetics and Evolution*, 105, pp. 50-62.

Alencar, L.R.V., Martins, M., and Greene, H.W. (2018). "Evolutionary History of Vipers." eLS, pp. 1-10.

Al-Farraj, A. and Harvey, A.M. (2004) "Late Quaternary interactions between aeolian and fluvial processes: a case study in the northern UAE." *Journal of Arid Environments*, 56, pp. 235-48.

Alvarado-Serrano, D.F. and Knowles, L.L. (2014). "Ecological niche models in phylogeographic studies: applications, advances and precautions." *Molecular Ecology Resources*, 14(2), pp.233-248.

Anthelme, F., Mato, M.W., and Maley, J. (2008). "Elevation and local refuges ensure persistence of mountain specific vegetation in the Nigerien Sahara." *Journal of Arid Environments*, 72(12), pp. 2232-2242.

Antonsich, M. (2004). "Geopolitics of the World System." *Political Geography.*

Armitage, S.J., Bristow, C.S., and Drake, N.A. (2015). "West African monsoon dynamics inferred from abrupt fluctuations of Lake Mega-Chad." *Proceedings of the National Academy of Sciences*, 112(28), pp. 8543-8548.

Avise, J.C. (2000). "Phylogeography: the history and formation of species." *Harvard university press.*

Avise, J.C. (2009). "Phylogeography: retrospect and prospect." *Journal of Biogeography*, 36(1), pp. 3-15.

Ballard, J.W. O. and Whitlock, M.C. (2004). "The incomplete natural history of mitochondria." *Molecular Ecology*, 13(4), pp. 729-744.

Bannikova, A.A., Dokuchaev, N.E., Yudina, E.V., Bobretzov, A.V., Sheftel, B.I., and Lebedev, V.S. (2010). "Holarctic phylogeography of the tundra shrew (*Sorex tundrensis*) based on mitochondrial genes." *Biological Journal of the Linnean Society*, 101(3), pp. 721-746.

Barlow, A., Wüster, W., Kelly, C.M., Branch, W.R., Phelps, T., and Tolley, K.A. (2019). "Ancient habitat shifts and organismal diversification are decoupled in the African viper genus *Bitis* (Serpentes: Viperidae)." *Journal of Biogeography*, 46(6), pp. 1234-1248.

Benson, L.V.,Currey, D.R., Dorn, R.I., Lajoie, K.R., Oviatt, C.G., Robinson, S.W., Smith, G.I., and Stine, S. (1990). "Chronology of expansion and contraction of four Great Basin lake systems during the past 35,000 years." *Palaeogeography, Palaeoclimatology, Palaeoecology,* 78(3-4), pp. 241-286.

Benson, L., Kashgarian, M., and Rubin,M. (1995). "Carbonate deposition, Pyramid Lake subbasin, Nevada: 2. Lake levels and polar jet stream positions reconstructed from radiocarbon ages and elevations of carbonates (tufas) deposited in the Lahontan basin." *Palaeogeography, Palaeoclimatology, Palaeoecology*, 117(1-2), pp. 1-30.

Bons, J., Geniez, P. and Montori, A. (1996). "Amphibiens et reptiles du Maroc (Sahara occidental compris): atlas biogéographique." *Associón Herpetológica Española*.

Brito, J.C., Fahd, S., Geniez, P., Martínez-Freiría, F., Pleguezuelos, J.M., and Trape, J.F. (2011). "Biogeography and conservation of viperids from North-West africa: an application of ecological niche-based models and GIS." *Journal of Arid Environments*, 75(11), pp. 1029-1037.

Brito, J.C., Godinho, R., Martínez-Freiría, F., Pleguezuelos, J.M., Rebelo, H., Santos, X., Vale, C.G., Velo-Antón, G., Boratyński, Z., Carvalho, S.B., Ferreira, S., Gonçalves, D.V., Silva, T.L., Tarroso, P., Campos, J.C., Leite, J.V., Nogueira, J., Álvares, F., Sillero, N., Sow, A.S., Fahd, S., Crochet, P.A., and Carranza, S. (2014). "Unravelling biodiversity, evolution and threats to conservation in the Sahara-Sahel." *Biological Reviews,* 89(1), pp. 215-231.

Brito, J.C., Tarroso, P., Vale, C.G., Martínez-Freiría, F., Boratyński, Z., Campos, J.C., Ferreira, S., Godinho, R., Gonçalves, D.V., Leite, J.V., Lima, V.O., Pereira, P., Santos, X., Ferreira da Silva, M.J., Silva, T.L., Velo-Antón, G., Veríssimo, J., Crochet, P.A., Pleguezuelos, J.M., and Carvalho, S.B. (2016). "Conservation Biogeography of the Sahara-Sahel: additional protected areas are needed to secure unique biodiversity." *Diversity and Distributions*, 22(4), pp. 371-384.

Brown, J.L. (2014). "SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses." *Methods in Ecology and Evolution,* 5(7), pp. 694-700.

Carnaval, A.C., Hickerson, M.J., Haddad, C.F.B., Rodrigues, M.T., and Moritz, C. (2009). "Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot." *Science*, 323(5915), pp. 785-789.

Carranza, S., Arnold, E.N., Geniez, P, Roca, J., and Mateo, J.A. (2008). "Radiation, multiple dispersal and parallelism in the skinks, *Chalcides* and *Sphenops* (Squamata: Scincidae), with comments on *Scincus* and *Scincopus* and the age of the Sahara Desert." *Molecular Phylogenetics and Evolution*, 46(3), pp. 1071-1094.

Carranza, S., Xipell, M., Tarroso, P., Gardner, A., Arnold, E.N., Robinson, M.D., Simó-Riudalbas, M., Vasconcelos, R., de Pous, P., Amat, F., Šmíd, J., Sindaco, R., Metallinou, M., Els. J., Pleguezuelos, J.M., Machado, L., Donaire, D., Martínez, G., Garcia-Porta, J., Mazuch, T., Wilms, T., Gebhart, J., Aznar, J., Gallego, J., Zwanzig, B.M., Fernández-Guiberteau, D., Papenfuss, T., Al Saadi, S., Alghafri, A., Khalifa, S., Al Farqani, H., Bilal, S.B., Alzari, I.S., Al Adhoobi, A.S., Al Omairi, Z.S., Al Shariani, M., Al Kiyumi, A., Al Sariri, T., Al Shukaili, A.S. and Al Akhzami, S.N. (2018) "Diversity, distribution and conservation of the terrestrial reptiles of Oman (Sauropsida, Squamata)." *PLoS one*, 13(2), e0190389.

Clark, P.U., Dyke, A.S., Shakun, J.D., Carlson, A.E., Clark, J., Wohlfarth, B., Mitrovica, J.X., Hostetler, S.W. and McCabe, A.M. (2009). "The last glacial maximum." *Science*, 325(5941), pp. 710-714.

Clement, M., Posada, D., and Crandall, K.A. (2000). "TCS: a computer program to estimate gene genealogies." *Molecular Ecology*, 9(10), pp. 1657–59.

del Marmol, G.M., Fernandez, B.R., Clerc, S., Hakonsen, T., Acierno, S., and Sassoe, M. (2019). "Sympatry of vipers in southwestern Morocco." *Herpetology Notes*, 12, pp. 337-342.

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N.D., Wikramanayake, E., Hahn, N., Palminteri, S., Hedao, P., Noss, R., Hansen, M., Locke, H., Ellis, E.C., Jones, B., Barber, C.V., Hayes, R., Kormos, C., Martin, V., Crist, E., Sechrest, W., Price, L., Baillie, J.E.M., Weeden, D., Suckling, K., Davis, C., Sizer, N., Moore, R., Thau, D., Birch, T., Potapov, P., Turubanova, S., Tyukavina, A., Souza, N.D., Pintea, L., Brito, J.C., Llewellyn, O.A., Miller, A.G., Patzelt, A., Ghazanfar, S.A., Timberlake, J., Klöser, H., Shennan-Farpón, Y., Kindt, R., Lillesø, J.P.B., Breugel, P.V., Graudal, L., Voge, M., Al-Shammari, K.F., and Saleem, M. (2017). "An ecoregion-based approach to protecting half the terrestrial realm." *BioScience*, 67(6), pp. 534-545.

Dobigny, G., Aniskin, V., Granjon, L., Cornette, R. and Volobouev, V. (2005) "Recent radiation in West African *Taterillus* (Rodentia, Gerbillinae): the concerted role of chromosome and climatic changes." *Heredity*, 95(5), pp. 358-368.

Douady, C.J., Catzeflis, F., Raman, J., Springer, M.S. and Stanhope, M.J. (2003). "The Sahara as a vicariant agent, and the role of Miocene climatic events, in the diversification of the mammalian order Macroscelidea (elephant shrews)." *Proceedings of the National Academy of Sciences of the United States of America*, 100, pp. 8325-8330.

Drake, N. A., Blench, R.M., Armitage, S. J., Bristow, C.S. and White, K.H. (2011). "Ancient watercourses and biogeography of the Sahara explain the peopling of the desert." *Proceedings of the National Academy of Sciences*, 108(2), pp. 458-462.

Drummond, A.J. and Rambaut, A. (2007). "BEAST: Bayesian evolutionary analysis by sampling trees." *BMC Evolutionary Biology*, 7(1), p. 214.

Drummond, A.J., Suchard, M.A., Xie, D. and Rambaut, A. (2012). "Bayesian phylogenetics with BEAUti and the BEAST 1.7." *Molecular Biology and Evolution*, 29(8), pp. 1969-1973.

Dumont, H. J. (1982). "Relict distribution patterns of aquatic animals: another tool in evaluating Late Pleistocene climate changes in the Sahara and Sahel." *Palaeoecology of Africa*, 14, pp. 1-24.

Durant, S.M., Pettorelli, N., Bashir, S., Woodroffe, R., Wacher, T., Ornellas, P.D., Ransom, C., Abáigar, T., Abdelgadir, M., Alqamy, H.E., Beddiaf, M., Belbachir, F., Belbachir-Bazi, A., Berbash, A.A., Beudels-Jamar, R., Boitani, L., Breitenmoser, C., Cano, M., Chardonnet, P., Collen, B., Cornforth, W.A., Cuzin, F., Gerngross, P., Haddane, B., Hadjeloum, M., Jacobson, A., Jebali, A., Lamarque, F., Mallon, D., Minkowski, K., Monfort, S., Ndoassal, B., Newby, J., Ngakoutou, B.E., Niagate, B., Purchase, G., Samaïla, S., Samna, A.K., Sillero-Zubiri, C., Soultan, A. E., Price, M.R.S., and Baillie, J. E. M. (2012). "Forgotten biodiversity in desert ecosystems." *Science*, 336(6087), pp. 1379-1380.

Edgell, G. (1989) "Evolution of the Rub al Khali desert." *Earth Science*, 3, pp. 109-126.

Edgell, H.S. (2006). "Arabian deserts: nature, origin, and evolution." *Springer Science & Business.*

Egan, D. (2007). "Snakes of Arabia: a field guide to the snakes of the Arabian Peninsula and its shores." *Motivate Publishing Limited*.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C, Nakamura, M., Nakazawa, Y., Overton, J.M.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberón, J., Williams, S., Wisz, M.S. and Zimmermann, N.E. (2006). "Novel methods improve prediction of species' distributions from occurrence data." *Ecography*, 29(2), pp. 129-151.

Elith, J., Kearney, M., and Phillips, S. (2010). "The art of modelling range-shifting species." *Methods in Ecology and Evolution*, 1(4), pp. 330-342.

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E. and Yates, C.J. (2011). "a statistical explanation of MAXENT for ecologists." *Diversity and Distributions*, 17(1), pp. 43-57.

ESRI (2006) "9.2. Environmental Systems." *Research Institute*, Inc, USA

Fielding, A.H. and Bell, J.F. (1997). "A review of methods for the assessment of prediction errors in conservation presence/absence models." *Environmental Conservation,* 24(1), pp. 38-49.

Foley, J.A., Coe, M.T., Scheffer, M. and Wang, G. (2003). "Regime shifts in the Sahara and Sahel: interactions between ecological and climatic systems in Northern Africa." *Ecosystems*, 6(6), pp. 524–32.

Franklin, J. and Miller, J.A. (2010). "Mapping species distributions: spatial inference and prediction." *Cambridge University Press*.

Freitas, I., Fahd, S., Velo-Antón, G. and Martínez-Freiría, F. (2018). "Chasing the phantom: biogeography and conservation of *Vipera latastei-monticola* in the Maghreb (North Africa)." *Amphibia Reptilia*, 39(2), pp. 145-161.

García-Cardenete, L., Flores-Stols, M.V. and Yubero, S. (2017). "New Case of syntopy between viperid snakes (Viperidae) in the Atlantic Sahara." *Go-South Bulletin*, 14, pp. 139-141.

Garrick, R.C., Sunnucks, P. and Dyer, R.J. (2010). "Nuclear gene phylogeography using PHASE: Dealing with unresolved genotypes, lost alleles, and systematic bias in parameter estimation." *BMC Evolutionary Biology*, 10(1), pp. 118.

Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Böttinger, M., Brovkin, V., Crueger, T., Esch, M., Fieg, K. and Glushak, K. (2013). "Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5." *Journal of Advances in Modeling Earth Systems*, 5(3), pp. 572-597.

Glennie, K.W. (1998) "The desert of southeast Arabia: a product of Quaternary climatic change." *Quaternary Deserts and Climate Change*, pp. 279–91.

Gonçalves, D.V., Brito, J.C., Crochet, P., Geniez, P., Padial, J.M., and Harris, D.J. (2012). "Phylogeny of north african *Agama* lizards (Reptilia: Agamidae) and the role of the Sahara Desert in vertebrate speciation." *Molecular Phylogenetics and Evolution*, 64(3), pp. 582-591.

Gonçalves, D.V., Pereira, P., Velo-Antón, G., Harris, D.J., Carranza, S. and Brito, J.C. (2018a). "Assessing the role of aridity-induced vicariance and ecological divergence in species diversification in North-West Africa using *Agama* lizards." *Biological Journal of the Linnean Society*, 124(3), pp. 363-380.

Gonçalves, D.V., Martínez-Freiría, F., Crochet, P.A., Geniez, P., Carranza, S. and Brito, J.C. (2018b). "The role of climatic cycles and trans-Saharan migration corridors in species diversification: biogeography of *Psammophis schokari* group in North Africa." *Molecular Phylogenetics and Evolution*, 118, pp. 64-74.

Guisan, A., Thuiller, W. and Zimmermann, N. E. (2017). "Habitat suitability and distribution models: with applications in R." *Cambridge University Press.*

Harris, N. (2004). "Atlas of the World's deserts." *Routledge.*

Harrison, R. G. and Larson, E. L. (2014). "Hybridization, introgression, and the nature of species boundaries." *Journal of Heredity*, 105(S1), pp. 795-809.

Hasumi, H. (2007). "K-1 coupled gcm (miroc) description." http://www.ccsr.u-tokyo.ac.jp/kyosei/hasumi/MIROC/tech-repo. pdf.

Hernandez, P. A., Graham, C. H., Master, L. L. and Albert., D. L. (2006). "The effect of sample size and species characteristics on performance of different species distribution modeling methods." *Ecography,* 29(5), pp. 773-785.

Hewitt, G. (1999) "Post-glacial re-colonization of european biota." *Biological Journal of the Linnean Society*, 68(1–2), pp. 87-112.

Hewitt, G. (2000). "The genetic legacy of the Quaternary ice ages." *Nature*, 405(6789), pp. 907-913.

Hewitt, G.M. (2004). "Genetic consequences of climatic oscillations in the Quaternary." *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359(1442), pp. 183-195.

Hickerson, M.J., Carstens, B.C., Cavender-Bares, J., Crandall, K.A., Graham, C.H., Johnson, J.B., Rissler, L., Victoriano, P.F. and Yoder, A.D. (2010). "Phylogeography's past, present, and future: 10 years after Avise, 2000." *Molecular Phylogenetics and Evolution,* 54(1), pp. 291-301.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. and Jarvis, A. (2005). "Very high resolution interpolated climate surfaces for global land areas*." International Journal of Climatology: A Journal of the Royal Meteorological Society*, 25(15), pp. 1965-1978.

Hoelzmann, P., Jolly, D., Harrison, S.P., Laarif, F., Bonnefille, R. and Pachur, H.J. (1998). "Mid-Holocene land-surface conditions in Northern Africa and the Arabian Peninsula: a data set for the analysis of biogeophysical feedbacks in the climate system." *Global Biogeochemical Cycles*, 12(1), pp. 35-51.

Holmes, J.A. (2008). "Ecology: how the Sahara became dry." Science, 320(5877), pp. 752-753.

Kapli, P., Lymberakis, P., Crochet, P.A., Geniez, P., Brito, J.C., Almutairi, M., Ahmadzadeh, F., Schmitz, A., Wilms, T., Pouyani, N.R. and Poulakakis, N. (2015). "historical biogeography of the lacertid lizard *Mesalina* in North Africa and the Middle East." *Journal of Biogeography*, 42(2), pp. 267-279.

Kearney, M., Shine, R. and Porter, W.P. (2009). "The potential for behavioral thermoregulation to buffer 'cold-blooded' animals against climate warming." *Proceedings of the National Academy of Sciences*, 106(10), pp. 3835-3840.

Koepfli, K. P., Pollinger, J., Godinho, R., Robinson, J., Lea, A., Hendricks, S., Schweizer, R.M., Thalmann, O., Silva, P., Fan, Z., Yurchenko, A.A., Dobrynin, P., Makunin, A., Cahill, J. A., Shapiro, B., Álvares, F., Brito, J.C., Geffen, E., Leonard, J. A., Helgen, K. M., Johnson, W.E., O'Brien, S.J., Valkenburgh, B.V. and Wayne, R.K. (2015). "Genome-wide evidence reveals that african and eurasian golden jackals are distinct species." *Current Biology*, 25(6), pp. 2158-2165.

Kotwicki, V. and Al Sulaimani, Z. (2009). "Climates of the Arabian Peninsula – Past, Present, Future." *International Journal of Climate Change Strategies and Management*, 1(3), pp. 297-310.

Laity, J.J. (2009) "Deserts and desert environments Vol. 3." *John Wiley & Sons.*

Le Berre, M. (1989) "Faune du Sahara. I. Poissons-Amphibiens-Reptiles" Lechevalier, R. Chabaud, Paris.

Le Houérou, H.N. (1997). "Climate, flora and fauna changes in the Sahara over the past 500 million years." *Journal of Arid Environments*, 37(4), pp. 619-647.

Leite, J.V., Álvares, F., Velo-Antón, G., Brito, J.C. and Godinho, R. (2015). "Differentiation of North African foxes and population genetic dynamics in the desert—insights into the evolutionary history of two sister taxa, *Vulpes rueppellii* and *Vulpes vulpes.*" *Organisms Diversity and Evolution*, 15(4), pp. 731-745.

Librado, P. and Rozas, J. (2009). "DnaSP v5: A software for comprehensive analysis of dna polymorphism data." *Bioinformatics*, 25(11), pp. 1451-1452.

Mace, G.M., Masundire, H. and Baillie, J.E.M. (2005). "Biodiversity. Ecosystems and human-well being: current state and trends" *Island Press, Washington*, DC., pp.77–122.

Machado, L., Šmíd, J., Mazuch, T., Sindaco, R., Al Shukaili, A.S. and Carranza, S. (2019). "Systematics of the Saharo-Arabian clade of the palearctic naked-toed geckos with the description of a new species of *Tropiocolotes* endemic to Oman." *Journal of Zoological Systematics and Evolutionary Research*, 57(1), pp. 159-178.

Maiorano, L., Amori, G., Capula, M., Falcucci, A., Masi, M., Montemaggiori, A., Pottier, J., Psomas, A., Rondinini, C., Russo, D., Zimmermann, N.E., Boitani, L. and Guisan, A. (2013). "Threats from climate change to terrestrial vertebrate hotspots in Europe." *PLoS ONE*, 8(9), e74989.

Maley, J. (2010). "Climate and palaeoenvironment evolution in north tropical Africa from the end of the Tertiary to the Upper Quaternary." *Palaeoecology of Africa*, 30, pp. 227-278.

Martínez-Freiría, F., Crochet, P.A., Fahd, S. , Geniez, P., Brito, J.C. and Velo-Antón, G. (2017). "Integrative phylogeographic and ecological analyses reveal multiple pleistocene refugia for mediterranean *Daboia* vipers in North-West Africa." *Biological Journal of the Linnean Society,* 112(2), pp. 366-384.

Martinez-Freiria, F., Lizana, M., do Amaral, J.P. and Brito, J.C. (2010). "Spatial and temporal segregation allows coexistence in a hybrid zone among two mediterranean vipers (*Vipera aspis and V. latastei*)." *Amphibia-Reptilia* 31(2), pp. 195-212.

Martínez-Freiría, F., Tarroso, P., Rebelo, H. and Brito, J.C. (2016). "Contemporary niche contraction affects climate change predictions for elephants and giraffes." *Diversity and Distributions*, 22(4), pp. 432-444.

Martínez-Freiría, F., Velo-Antón, G. and Brito, J.C. (2015). "Trapped by climate: interglacial refuge and recent population expansion in the endemic iberian adder *Vipera seoanei.*" *Diversity and Distributions*, 21(3), pp. 331-344.

Merow, C., Smith, M.J. and Silander, J.A. (2013) "A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter." *Ecography,* 36(10), pp. 1058-1069.

Metallinou, M., Červenka, J., Crochet, P.A., Kratochvíl, L., Wilms, T., Geniez, P., Shobrak, M.Y., Brito, J.C. and Carranza, S. (2015). "Species on the rocks: Systematics and biogeography of the rock-dwelling *Ptyodactylus* geckos (Squamata: Phyllodactylidae) in North Africa and Arabia." *Molecular Phylogenetics and Evolution*, 85, pp. 208-220.

Miller, M.A., Pfeiffer, W. and Schwartz, T. (2010). "Creating the CIPRES Science Gateway for inference of large phylogenetic trees." in 2010 gateway computing environments workshop (GCE), pp. 1-8.

Mitchell, S. M., Muehlbauer, L.K. and Freedberg, S. (2016). "Nuclear introgression without mitochondrial introgression in two turtle species exhibiting sex-specific trophic differentiation." *Ecology and Evolution*, 6(10), pp. 3280-3288.

Mithen, S. (2004) "After the Ice: A Global Human History, 20,000–5000 BC." *Harvard University Press, Cambridge.*

Múrias dos Santos, A., Cabezas, M.P., Tavares, A.I., Xavier, R. and Branco, M. (2015). "tcsBU: a tool to extend TCS network layout and visualization." *Bioinformatics*, 32(4), pp. 627-628.

Murphy, D.D. and Weiss, S. B. (1992). "Effects of climate change on biological diversity in western North America: species losses and mechanisms." *Global Warming and biological diversity. Yale University Press, New Haven, Connecticut*, pp. 355-368.

Nagy, Z.T., Sonet, G., Glaw, F. and Vences, M. (2012) "First Large-Scale DNA Barcoding Assessment of Reptiles in the Biodiversity Hotspot of Madagascar, Based on Newly Designed COI Primers" PLoS one, 7(3), e34506.

Nicolas, V., Mataame, A., Crochet, P.A., Geniez, P., Fahd, S. and Ohler, A. (2018). "Phylogeography and ecological niche modeling unravel the evolutionary history of the African green toad, *Bufotes boulengeri boulengeri* (Amphibia: Bufonidae), through the Quaternary." *Journal of Zoological Systematics and Evolutionary Research*, 56(1), pp. 102-116.

Nicolas, V., Mataame, A., Crochet, P.A., Geniez, P. and Ohler, A. (2015). "Phylogeographic patterns in north African water frog *Pelophylax saharicus* (Anura: Ranidae)." *Journal of Zoological Systematics and Evolutionary Research*, 53(3), pp. 239–248.

Nogués-Bravo, D. (2009). "Predicting the past distribution of species climatic niches." *Global Ecology and Biogeography*, 18(5), pp. 521-531.

Noonan, B.P. and Chippindale, P.T. (2006) "Dispersal and vicariance: the complex evolutionary history of boid snakes." *Molecular Phylogenetics and Evolution*, 40(2), pp. 347-358.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V., Underwood, E.C., D'amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P. and Kassem, K.R. (2001). "Terrestrial Ecoregions of the World: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity." *BioScience*, 51(11), pp. 933.

Orme, A.R. (2002). "The Pleistocene legacy: beyond the ice front." *The Physical Geography of North America*, pp. 380-401.

Peterson, A.T. (2009). "Phylogeography is not enough: The need for multiple lines of evidence." *Frontiers of Biogeography*, 1(1).

Peterson, A.T. and Kluza, D.A. (2005). "Ecological niche modeling as a new paradigm for large-scale investigations of diversity and distribution of birds." in *Third International Partners in Flight Conference,* March 20-24, 2002. U.S. Department of Agriculture Forest Service, Asilomar, California.

Peterson, A.T., Soberón, J. and Sánchez-Cordero, V. (1999). "Conservatism of ecological niches in evolutionary time." *Science*, 285(5431), pp. 1265-1267.

Peterson, A.T., Soberón, J., Pearson, R.G., Anderson, R.P., Martínez-Meyer, E., Nakamura, M., and Araújo, M.B. (2012). "Ecological niches and geographic distributions." *Choice Reviews Online*, 56.

Phelps, T. (2010). " Old world vipers, A natural history of the Azemiopinae and Viperinae." *Edition Chimaira*

Phillips, S.J., Anderson, R.P. and Schapire, R.E. (2006). "Maximum entropy modeling of species geographic distributions." *Ecological Modelling*, 190(3-4), pp. 231-2559.

Pook, C.E., Joger,U., Stümpel, N. and Wüster,W. (2009). "When continents collide: phylogeny, historical biogeography and systematics of the medically important viper genus *Echis* (Squamata: Serpentes: Viperidae)." *Molecular Phylogenetics and Evolution*, 53(3), pp. 792-807.

Posada, D. (2008). "JModelTest: phylogenetic model averaging." *Molecular Biology and Evolution*, 25(7), pp. 1253-1256.

Pough, F.H. (1980). "Advantages of Ectothermy for Tetrapods." *The American Naturalist*, 115(1). pp. 92-112.

Pyron, R.A. and Burbrink, F.T. (2009) "Neogene diversification and taxonomic stability in the snake tribe *Lampropeltini* (Serpentes: Colubridae)." *Molecular Phylogenetics and Evolution*, 52(2), pp. 524-529.

Quante, M. (2010). "The Changing Climate: Past, Present, Future." In Relict species: phylogeography and conservation biology, pp. 9-56. *Springer*, Berlin, Heidelberg.

Rambaut, A. (2016). "FigTree v1.4.3 software." Molecular Evolution, Phylogenetics and *Epidemiology.*

Robinson, M., Arnold, N.E., and Carranza, S. (2009). "A preliminary analysis of phylogenetic relationships and biogeography of the dangerously venomous Carpet Vipers, *Echis* (Squamata, Serpentes, Viperidae) based on mitochondrial dna sequences." *Amphibia-Reptilia*, 30(2), pp. 273-282.

Santos, X., Brito, J.C, Sillero, N., Pleguezuelos, J.M, Llorente, G.A., Fahd, S. and Parellada, X. (2006). "Inferring habitat-suitability areas with ecological modelling techniques and GIS: a contribution to assess the conservation status of *Vipera latastei*." *Biological Conservation*, 130(3), pp. 416-425.

Sarnthein, M. (1978). "Sand deserts during glacial maximum and climatic optimum." *Nature*, 272(5648), p. 43.

Schield, D.R., Adams, R.H., Card, D.C., Perry, B.W., Pasquesi, G.M, Jezkova, T., Portik, D.M., Andrew, A.L., Spencer, C.L., Sanchez, E.E., Fujita, M.K., Mackessy, S.P. and Castoe, T.A. (2017). "insight into the roles of selection in speciation from genomic patterns of divergence and introgression in secondary contact in venomous rattlesnakes." *Ecology and Evolution*, 7(11), pp. 3951-3966.

Schuster, M., Duringer, P., Ghienne, J.F., Vignaud, P., Mackaye, H.T., Likius, A. and Brunet, M. (2006). "The Age of the Sahara Desert." *Science*, 311(5762, pp. 821-821.

Sindaco, R., Venchi, A. and Grieco, C. (2013). "The Reptiles of the Western Palearctic. 2. Annotated checklist and distributional atlas of the snakes of Europe, North Africa, Middle East and Central Asia, with an update to the 1." *Acta Herpetologica.*

Sinervo, B., Mendez-de-la-Cruz, F., Miles, D.B., Heulin, B., Bastiaans, E., Villagran-Santa Cruz, M., Lara-Resendiz, R., Martinez-Mendez, N., Calderon-Espinosa, M.L., Meza-Lazaro, R.N., Gadsden, H., Avila, L.J., Morando, M., De la Riva, I.J., Sepulveda, P.V., Rocha, C.F.D., Ibarguengoytia, N., Puntriano, C.A. , Massot, M., Lepetz, V., Oksanen, T.A., Chapple, D.G., Bauer, A.M., Branch, W.R., Clobert, J. and Sites, J.W. (2010). "Erosion of lizard diversity by climate change and altered thermal niches." *Science*, 328(5980), pp. 894-899.

Šmíd, J., Carranza, S., Kratochvíl, L., Gvoždík, V., Nasher, A.K. and Moravec, J. (2013). "Out of Arabia: A complex biogeographic history of multiple vicariance and dispersal events in the gecko genus *Hemidactylus* (Reptilia: Gekkonidae)." *PLoS one*, 8(5), p. e64018.

Šmíd, J. and Tolley, K.A. (2019). "Calibrating the tree of vipers under the fossilized birth-death model." *Scientific Reports*, 9(1), p. 5510.

Sommer, R.S. and Zachos, F.E. (2009). "Fossil evidence and phylogeography of temperate species: 'glacial refugia' and post-glacial recolonization." *Journal of Biogeography*, 36(11), pp. 2013-2020.

Sow, A.S., Martínez-Freiría, F., Dieng, H., Fahd, S. and Brito, J.C. (2014). "Biogeographical analysis of the Atlantic Sahara reptiles: Environmental correlates of species distribution and vulnerability to climate change." *Journal of Arid Environments*, 109, pp. 65-73.

Stephens, M. and Scheet, P. (2005). "Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation." *The American Journal of Human Genetics,* 76(3), pp. 449-462.

Stephens, M., Smith, N.J. and Donnelly, P. (2001). "A new statistical method for haplotype reconstruction from population data." *The American Journal of Human Genetics*, 68(4), pp. 978-989.

Stimpson, C.M., Lister, A., Parton, A., Clark-Balzan, L., Breeze, P.S., Drake, N.A., Groucutt, HS., Jennings, R., Scerri, E.M.L., White, T.S., Zahir, M., Duval, M., Grün, R., Al-Omari,

A., Al Murayyi, K.S.M., Zalmout, I.S., Mufarreh, Y.A., Memesh, A.M. and Petraglia, M.D. (2016). "Middle Pleistocene vertebrate fossils from the Nefud Desert, Saudi Arabia: implications for biogeography and palaeoecology." *Quaternary Science Reviews*, 143, pp. 13-36.

Stokes, S., Thomas, D.S.G. and Washington, R. (1997). "Multiple episodes of aridity in southern africa since the last interglacial period." *Nature*, 388(6638).

Stümpel, N., Rajabizadeh, M., Avci, A., Wüster, W. and Joger, U. (2016). "Phylogeny and diversification of mountain vipers (*Montivipera*, Nilson et al., 2001) triggered by multiple Plio-Pleistocene refugia and high-mountain topography in the Near and Middle East." *Molecular Phylogenetics and Evolution*, 101, pp. 336-351.

Tamar, K., Carranza, S., Sindaco, R., Moravec, J., Trape, J.F. and Meiri, S. (2016). "Out of Africa: Phylogeny and biogeography of the widespread genus *Acanthodactylus* (Reptilia: Lacertidae)." *Molecular Phylogenetics and Evolution*, 103, pp. 6-18.

Tamar, K., Metallinou, M., Wilms, T., Schmitz, A., Crochet, P.A., Geniez, P. and Carranza, S. (2018). "Evolutionary history of spiny-tailed lizards (Agamidae: *Uromastyx*) from the Saharo-Arabian region." *Zoologica Scripta*, 47(2), pp. 159-173.

Tarroso, P., Velo-Antón, G. and Carvalho, S.B. (2015). "PHYLIN: an R Package for phylogeographic interpolation." *Molecular Ecology Resources*, 15(2), pp. 349-357.

Townsend, T.M., Alegre, R.E., Kelley, S.T., Wiens, J.J. and Reeder, T.W. (2008). "Rapid development of multiple nuclear loci for phylogenetic analysis using genomic resources: an example from squamate reptiles." *Molecular Phylogenetics and Evolution*, 47(1), pp. 129-142.

Trabucco, A. and Zomer, R.J. (2009). "Global aridity index (global-aridity) and global potential evapo-transpiration (globalPET) geospatial database." *CGIAR Consortium for Spatial Information.*

Trauth, M.H., Larrasoaña, J.C. and Mudelsee, M. (2009). "Trends, rhythms and events in Plio-Pleistocene African climate." *Quat. Sci. Rev.* 28, 399-411.

Varela, S., Lima-Ribeiro, M.S. and Terribile, L.C. (2015). "A short guide to the climatic variables of the last glacial maximum for biogeographers*." PLoS one,* 10(6), p.e0129037.

Velo-Antón, G., Godinho, R., Harris, D.J., Santos, X., Martínez-Freiria, F., Fahd, S., Larbes, S., Pleguezuelos, J.M. and J. C. Brito. (2012). "Deep evolutionary lineages in a Western Mediterranean snake (*Vipera latastei/monticola* group) and high genetic structuring in Southern Iberian populations." *Molecular Phylogenetics and Evolution*, 65(3), pp. 965-973.

Velo-Antón, G., Parra, J.L., Parra-Olea, G. and Zamudio, K.R. (2013). "Tracking climate change in a dispersal-limited species: reduced spatial and genetic connectivity in a montane salamander." *Molecular Ecology*, 22(12), pp. 3261-3278.

Velo-Antón, G., Martínez-Freiría, F., Pereira, P., Crochet, P.A. and Brito, J.C. (2018). "Living on the edge: ecological and genetic connectivity of the spiny-footed lizard, *Acanthodactylus aureus*, confirms the Atlantic Sahara Desert as a biogeographic corridor and centre of lineage diversification." *Journal of Biogeography*, 45(5), pp. 1031-1042.

Verissimo, J., Znari, M., Stuckas, H., Fritz, U., Pereira, P., Teixeira, J., Arculeo, M., Marrone, F., Sacco, F., Naimi, M., Kehlmaier, C. and Velo-Antón, G. (2016). "Pleistocene diversification in Morocco and recent demographic expansion in the Mediterranean pond turtle *Mauremys leprosa*." *Biological Journal of the Linnean Society*, 119(4), pp. 943-959.

Wagner, P. and Wilms, T.M. (2010). "A crowned devil: new species of *Cerastes laurenti*, 1768 (Ophidia, Viperidae) from Tunisia, with two nomenclatural comments." *Bonn Zoological Bulletin*, 57(2), pp.297-306.

Walker, M., Johnsen, S., Rasmussen, S.O., Popp, T., Steffensen, J.P., Gibbard, P., Hoek, W., Lowe, J., Andrews, J., Björck, S., Cwynar, L.C., Hughen, K., Kershaw, P., Kromer, B., Litt, T., Lowe, D.J., Nakagawa, T., Newnham, R. and Schwander, J. (2009). "Formal definition and dating of the GSSP (Global Stratotype Section and Point) for the base of the Holocene using the Greenland NGRIP ice core, and selected auxiliary records." *Journal of Quaternary Science*, 24(1), pp. 3-17.

Waltari, E., Hijmans, R.J., Peterson, A.T., Nyári, A.S., Perkins, S.L. and Guralnick, R.P. (2007). "Locating Pleistocene Refugia: Comparing Phylogeographic and Ecological Niche Model Predictions." *PloS one*, 2(7), p. 2563.

Wan, Q.H., Wu, H., Fujihara, T. and Fang, S.G. (2004). "Which genetic marker for which conservation genetics issue?" *Electrophoresis*, 25(14), pp. 2165-2176.

Ward, D. (2016). "The Biology of Deserts." *Oxford University Press.*

Weiss, S. and Ferrand, N. (2007). "Phylogeography of southern European Refugia." *Dordrecht Springer.*

Welch K.R.G. (1982). "Herpetology of Africa: a checklist and bibliography of the orders Amphisbaenia, Sauria and Serpentes." *Malabar, Florida: Krieger.*

Werner, Y. L., Sivan, N., Kushnir, V., and Motro, U. (1999). "A statistical approach to variation in *Cerastes* (Ophidia: Viperidae), with the description of two endemic subspecies." *Kaupia*, 8, pp. 83-97.

Wiens, J.A., Stralberg, D., Jongsomjit, D., Howell, C.A. and Snyder, M.A. (2009). "Niches, models, and climate change: assessing the assumptions and uncertainties." *Proceedings of the National Academy of Sciences of the United States of America*, 106(Supplement 2), pp. 19729-19736.

Wiens, J.J. and Graham, C.H. (2005). "Niche conservatism: integrating evolution, ecology, and conservation biology." *Annual Review of Ecology, Evolution, and Systematics*, 36, pp. 519-539.

Wisz, M. S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H., Guisan, A., Elith, J., Dudík, M., Ferrier, S., Huettmann, F., Leathwick, J.R., Lehmann, A., Lohmann, L., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.C., Phillips, S.J., Richardson, K.S., Scachetti-Pereira, R., Schapire, R.S., Soberón, J., Williams, S.E. and Zimmermann, N.E. (2008). "Effects of sample size on the performance of species distribution models." *Diversity and Distributions*, 14(5), pp. 763-773.

Wüster, W., Peppin, L., Pook, C.E. and Walker, D.E. (2008). "A nesting of vipers: phylogeny and historical biogeography of the Viperidae (Squamata: Serpentes)." *Molecular Phylogenetics and Evolution*, 49(2), pp. 445-459.

Yackulic, C.B., Chandler, R., Zipkin, E.F., Royle, J.A., Nichols, J.D., Grant, E.H.C. and Veran, S. (2013). "Presence-only modelling using MaxENT: when can we trust the inferences?" *Methods in Ecology and Evolution*, 4(3), pp.236-243.

Zhang, Z., Ramstein, G., Schuster, M., Li, C., Contoux, C. and Yan, Q. (2014). "Aridification of the Sahara desert caused by Tethys Sea shrinkage during the Late Miocene." *Nature*, 513(7518), p. 401.

Zheng, Y. and Wiens, J.J. (2016). "Combining phylogenomic and supermatrix approaches, and a time-calibrated phylogeny for squamate reptiles (lizards and snakes) based on 52 genes and 4162 species." *Molecular Phylogenetics and Evolution*, 94, pp. 537-547.

Appendix

Appendix 1 Ancient DNA extraction protocol

Ancient DNA extraction protocol for tissue samples (modified from Rohland and Hofreiter, 2007 protocol created for extraction of ancient DNA from bone samples)

Laboratory conditions necessary:

- The extraction procedure must be conducted in a room specified only for non-invasive or historical DNA extractions and equipped with UV lights.
- Before starting the extraction procedure, UV lights should be turned on for at least 15 minutes to eliminate all DNA present and to sterilize the room. Posteriorly, all surfaces must be cleaned with bleach and ethanol.
- All buffers and materials necessary in this procedure must be prepared, washed and treated with UV lights at least 15 minutes to eliminate any contaminants.

1. Pipet PBS solution into each tube containing the tissue samples and incubate at room temperature overnight
2. Discard PBS solution and cut the tissue sample in small pieces
3. Place 50 to 150 mg of sample into 2.0mL Safelock Eppendorf tube containing the extraction buffer (Composition: ultrapure water, 0.5M EDTA, Tween 20 and Proteinase k (New England Biolabs))
3. Seal the tubes with parafilm and digest overnight at 37ºC
4. Prepare one 50mL falcon tube for each sample containing a mixture of binding buffer (Composition: ultrapure water, isopropanol, tween 20 and guanidine hydrochloride) and sodium acetate. An extra tube should be prepared for the negative control
5. Perform centrifugation of samples and transfer supernatant to the 50 mL falcon prepared previously
6. Assemble Zymo extension reservoirs to the minElute spin column from Quiagen and add the assemblage to a 50mL falcon tube without the collection tube
7. Transfer the mixture (step 5) to the assembly and centrifugate. Repeat centrifugation process with the assembly inverted. If needed, perform centrifugation for a few more minutes or increase the number of rotations per minute to ensure that all the liquid flows through the column.
8. Disassemble the column and attach it to the collection tube. Discard the remaining materials and the flow-through.
9. Dry spin the columns, add PE buffer (Quiagen), centrifugate and discard the flow-through. Repeat this step one more time
10. Perform first and second DNA elutions to a 1.5 mL Eppendorf Low retention tube with 25uL TET buffer per elution. If the concentration of the samples is low (<100 ng), perform another elution containing 50uL TET buffer to a different 1.5 mL Eppendorf Low retention tube or perform a second extraction

Appendix 2: PCR conditions for the sequenced genes of mtDNA and nuDNA

COI

MasterMix – 5 uL

Rep-COI-F – 0.4 uL

Rep – COI-R- 0.4 uL

H20 – 3.2 uL

DNA – 1 uL

Amplification step	Temperature (ºC)	Duration	Number of cycles
Initial denaturation	95	10 min	1
Denaturation	95	40 sec	
Annealing	50	45 sec	9
Extension	72	45 sec	
Denaturation	95	40 sec	
Annealing	48	45 sec	31
Extension	72	45 sec	
Final extension	72	10 min	1

NT3

MyTaq– 5 uL

NT3-F1 – 0.4 uL

NT3-R4 - 0.4 uL

H20 – 3.2 uL

DNA – 1 uL

Amplification step	Temperature (ºC)	Duration	Number of cycles
Initial denaturation	95	10 min	1
Denaturation	95	40 sec	
Annealing	57	30 sec	40
Extension	72	45 sec	
Final extension	72	7 min	1

PRLR

MasterMix– 5 uL

PRLR-F1 – 0.4 uL

PRLR-R3- 0.4 uL

H20 – 3.2 uL

DNA – 1 uL

Amplification step	Temperature (ºC)	Duration	Number of cycles
Initial denaturation	95	15 min	1
Denaturation	95	30 sec	
Annealing	56	30 sec	40
Extension	72	45 sec	
Final extension	60	10 min	1

VIM

MyTaq– 5 uL

VIM_Ex5_F2– 0.4 uL

VIM_Ex6_R2- 0.4 uL

H20 – 3.2 uL

DNA – 1 uL

Amplification step	Temperature (ºC)	Duration	Number of cycles
Initial denaturation	95	10 min	1
Denaturation	95	40 sec	
Annealing	60 (Touchdown:0.5)	45 sec	11
Extension	72	45 sec	
Denaturation	95	40 sec	
Annealing	55	30 sec	29
Extension	72	45 sec	
Final extension	72	7 min	1

Appendix 3 - List of sequenced samples used in this study and respective haplotype number from Fig. 5 and Fig. 6

CODE_coll	Code	Species	Latitude	Longitude	COI	NT3	PRLR	VIM
9166	14CC001	*Cerastes cerastes*	22.62	-14.60	1	-	-	-
5136	5136	*Cerastes cerastes*	20.15	-16.14	2	-	-	-
13693	17CC005	*Cerastes cerastes*	25.10	-11.52	3	-	-	-
6270	11CC022	*Cerastes cerastes*	18.48	-16.02	3	-	-	-
2898	09CC003	*Cerastes cerastes*	21.49	-11.33	4	-	-	-
3232	09CC010	*Cerastes cerastes*	17.38	-11.82	5	-	-	-
11984	11984	*Cerastes cerastes*	20.25	-13.09	6	-	-	-
3045	09CC005	*Cerastes cerastes*	18.08	-11.87	7	-	-	-
2908	09CC006	*Cerastes cerastes*	21.06	-11.42	7	8,14	1,11	8,9
5852	11CC016	*Cerastes cerastes*	21.78	-12.88	8	-	-	-
5920	11CC018	*Cerastes cerastes*	20.79	-12.20	8	-	-	-
ABR17 - FER2	17CC028	*Cerastes cerastes*	27.10	-11.06	8	-	-	-
9076	9076	*Cerastes cerastes*	26.51	-11.99	8	8,16	1,15	-
	17CC002	*Cerastes cerastes*	27.80	-11.09	9	6,15	1,1	19,20
1621	08CC001	*Cerastes cerastes*	21.19	-13.58	10	-	-	-
5205	10CC004	*Cerastes cerastes*	20.28	-15.08	10	-	-	-
5220	10CC005	*Cerastes cerastes*	20.41	-14.96	10	-	-	-
5251	10CC007	*Cerastes cerastes*	21.06	-14.37	10	-	-	-
6427	11CC023	*Cerastes cerastes*	20.74	-16.42	10	-	-	-
SPM002639	SPM002639	*Cerastes cerastes*	30.00	12.00	11	7,16	1,16	-
SPM002345	SPM002345	*Cerastes cerastes*	35.22	2.32	12	-	-	-
9089	9089	*Cerastes cerastes*	32.48	-1.60	13	-	-	-
8288	8288	*Cerastes cerastes*	32.36	-2.85	13	-	-	-
9059	9059	*Cerastes cerastes*	32.28	-3.32	13	6,14	1	-
BEV.8502	BEV.8502	*Cerastes cerastes*	32.15	-1.26	13	-	-	-
325	325	*Cerastes cerastes*	31.30	10.62	14	-	-	-

T31-37 Ceracera1	T31-37 Ceracera1	*Cerastes cerastes*	34.15	8.29	14	8,16	14	-
10405	10405	*Cerastes cerastes*	31.53	-4.18	15	-	-	-
set/16	17CC015	*Cerastes cerastes*	29.11	-8.66	15	-	-	-
6602	6602	*Cerastes cerastes*	31.24	-4.57	15	-	-	-
9061	9061	*Cerastes cerastes*	30.80	-6.75	15	-	-	-
S1339	S1339	*Cerastes cerastes*	31.20	-4.92	15	13,14	1,6	-
SPM002638	SPM002638	*Cerastes cerastes*	30.94	-7.21	15	-	-	-
BEV.10313	BEV.10313	*Cerastes cerastes*	30.75	-6.09	16	8,16	1	19,20
	15CC067	*Cerastes cerastes*	30.55	-7.15	17	-	-	-
	16CC008	*Cerastes cerastes*	28.56	-9.80	18	8,16	8,9	1,2
17CC040	17CC040	*Cerastes cerastes*	28.98	-9.85	18	-	-	-
11504	11504	*Cerastes cerastes*	28.84	-10.31	19	-	-	-
	15CC001	*Cerastes cerastes*	30.09	-6.88	19	-	-	-
8370	8370	*Cerastes cerastes*	29.76	-7.74	19	-	-	-
9065	9065	*Cerastes cerastes*	29.76	-8.52	19	-	-	-
BEV.8704	BEV.8704	*Cerastes cerastes*	29.84	-7.21	19	7,14	1,14	18,20
TAU.R16430	TAU.R16430	*Cerastes cerastes*	30.93	34.41	21	12,17	7	3,6
SPM002585	SPM002585	*Cerastes cerastes*	29.08	34.23	22	-	-	-
12953	12953	*Cerastes cerastes*	18.92	20.91	23	-	-	-
12959	12959	*Cerastes cerastes*	18.92	20.91	23	11,16	7	4,5
6734	6734	*Cerastes cerastes*	15.86	11.45	23	8,16	7	4,6
92	92	*Cerastes cerastes*	25.29	32.55	24	10,16	7	4,7
25132/2	25132/2	*Cerastes cerastes*	13.66	30.02	25	-	-	-
SPM000783	SPM000783	*Cerastes cerastes*	36.91	7.76	25	8,16	1,12	-
38248	ZFMK38248	*Cerastes cerastes*	24.37	9.53	26	6,14	1	-
AO141	AO141	*Cerastes gasperettii*	17.96	53.75	27	-	-	-
CN2672	CN2672	*Cerastes gasperettii*	20.78	58.31	28	-	-	-
CN3768	CN3768	*Cerastes gasperettii*	21.76	59.49	28	-	-	-

CN7622	CN7622	*Cerastes gasperettii*	18.44	55.27	29	1,2	2,4	21,25
	BEV.10077	*Cerastes gasperettii*	29.39	47.07	30	2,3	2,3	22,25
S7071	S7071	*Cerastes gasperettii*	22.45	58.68	30	-	2	23,24
TAU.R17182	TAU.R17182	*Cerastes gasperettii*	32.07	34.79	30	-	-	-
553	553	*Cerastes vipera*	25.76	-14.60	31	8,16	1,5	10,11
2888	09CV002	*Cerastes vipera*	21.50	-11.62	32	-	-	-
503	503	*Cerastes vipera*	18.98	-16.21	32	-	-	-
7312	12CV031	*Cerastes vipera*	23.58	-15.23	32	-	-	-
5284	10CV009	*Cerastes vipera*	20.47	-15.61	33	-	-	-
11616	11616	*Cerastes vipera*	24.78	-14.85	33	-	-	-
11964	11964	*Cerastes vipera*	19.11	-14.94	33	-	-	-
5763	11CV011	*Cerastes vipera*	21.20	-14.22	33	-	-	-
5784	11CV014	*Cerastes vipera*	21.52	-12.85	33	9,15	1	-
13775	13775	*Cerastes vipera*	23.20	-11.95	33	-	-	-
ABR17 - FER4	17CV030	*Cerastes vipera*	26.87	-11.75	33	-	-	-
BEV.10867	BEV.10867	*Cerastes vipera*	22.74	-14.55	33	-	-	-
15CV165	15CV165	*Cerastes vipera*	27.74	-11.46	34	-	-	-
OCT16 - 4	17CV017	*Cerastes vipera*	28.30	-9.34	35	8,15	1	14,15
83340	ZFMK83340	*Cerastes vipera*	27.10	-13.39	36	-	-	-
BEV.10185	BEV.10185	*Cerastes vipera*	24.56	10.89	38	8,16	1,13	16,17
6658	6658	*Cerastes vipera*	14.68	13.14	39	8,16	1	10,11
NMPR72142/6	NMPR72142/6	*Cerastes vipera*	27.58	29.86	39	-	-	-
TAU.R17180	TAU.R17180	*Cerastes vipera*	30.93	34.42	40	18,19	1	12,13
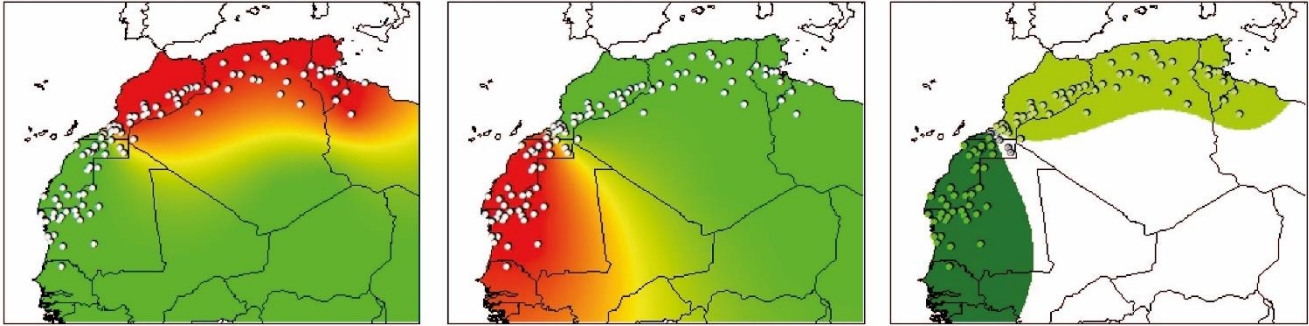
Appendix 4 - Delimitation of *Cerastes cerastes* lineages and sub-lineages (West-North and West-South) occurrences using spatial interpolations of mtDNA
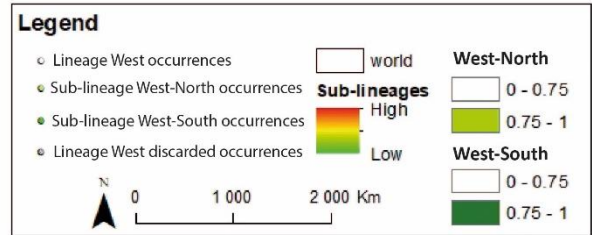
Appendix 5 - Probabilistic models of *Cerastes cerastes* for the present and the past climatic periods with standard deviation
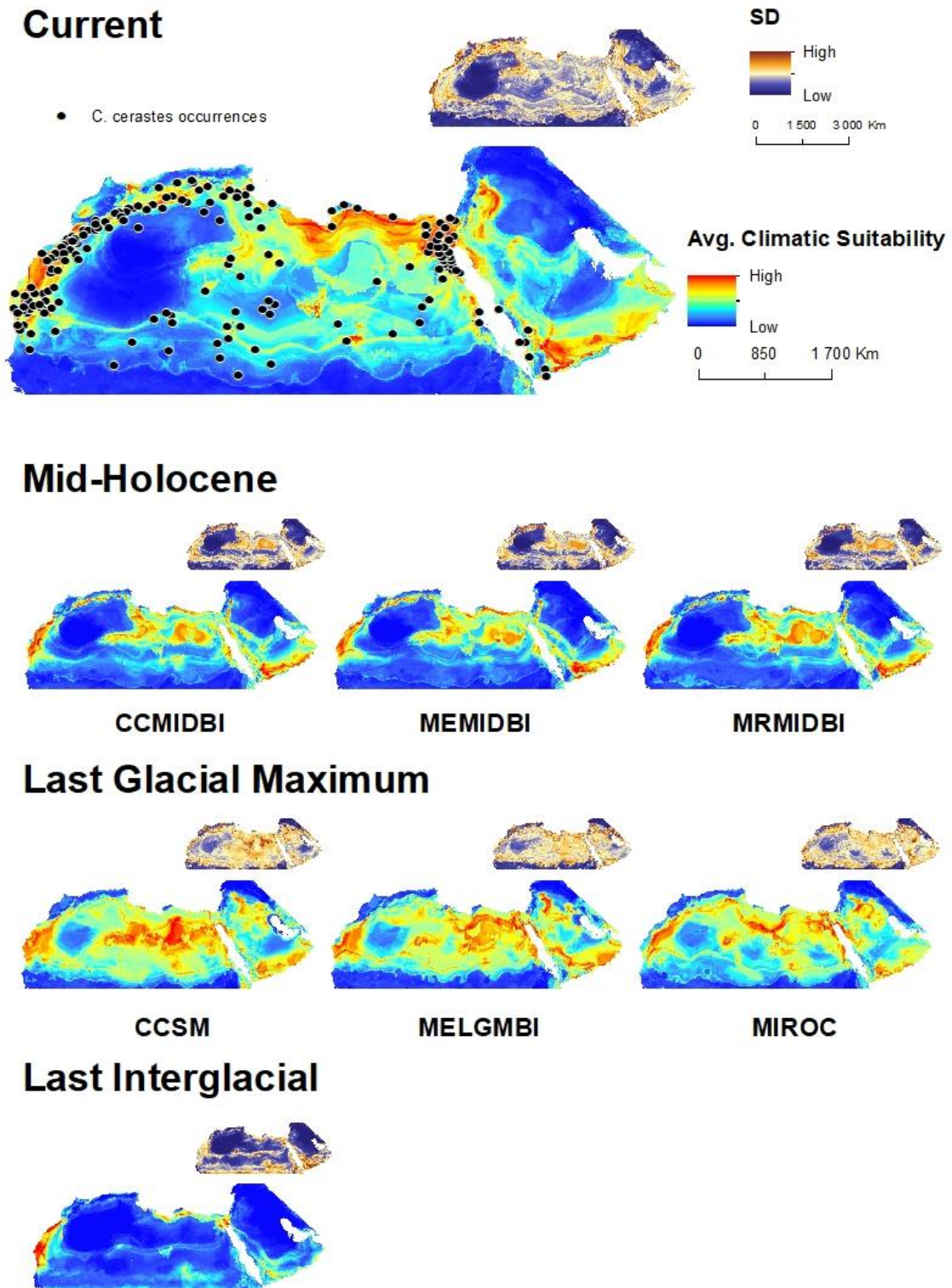
Appendix 6 - Probabilistic models of *Cerastes vipera* for the present and the past climatic periods with standard deviation
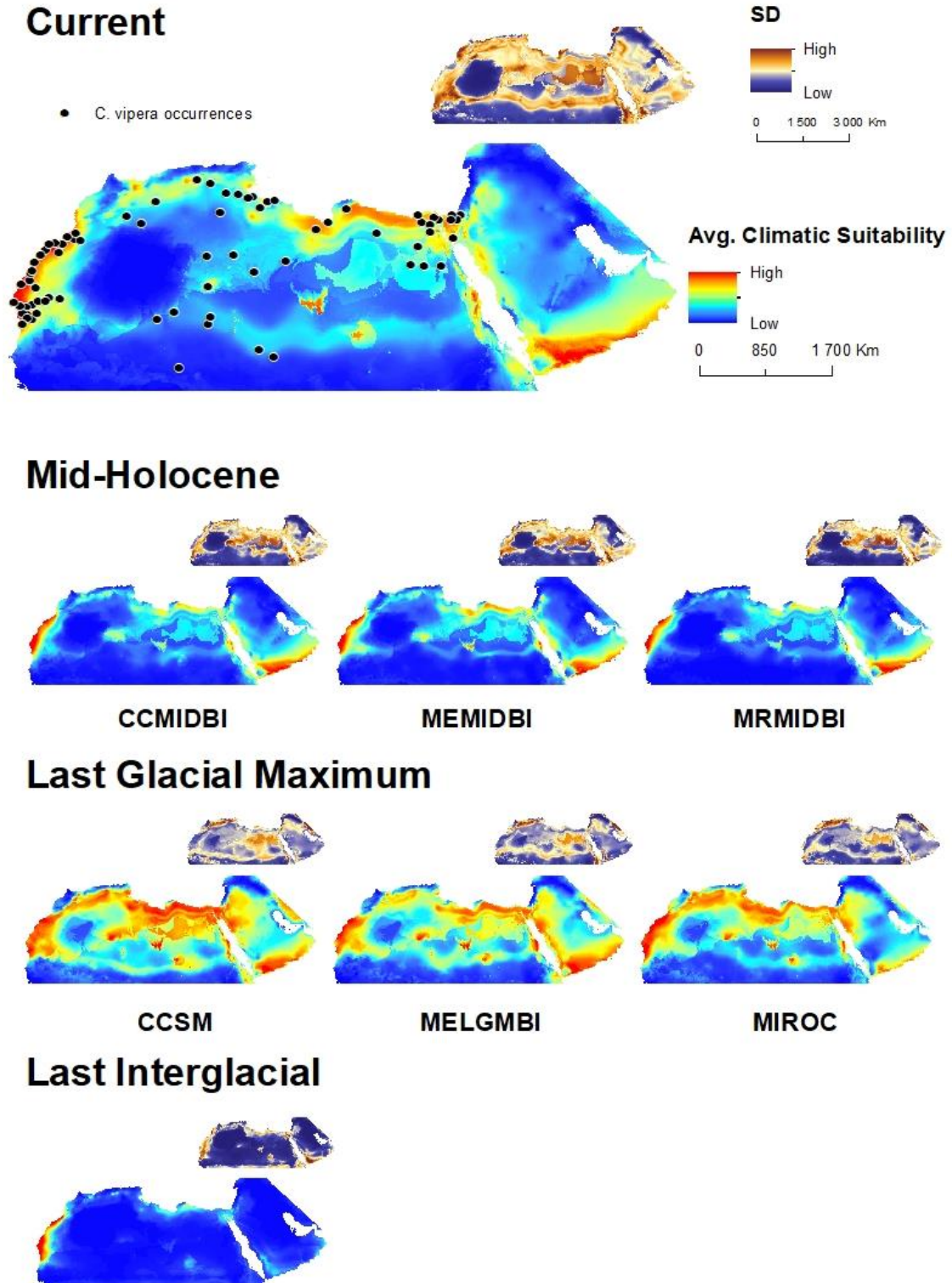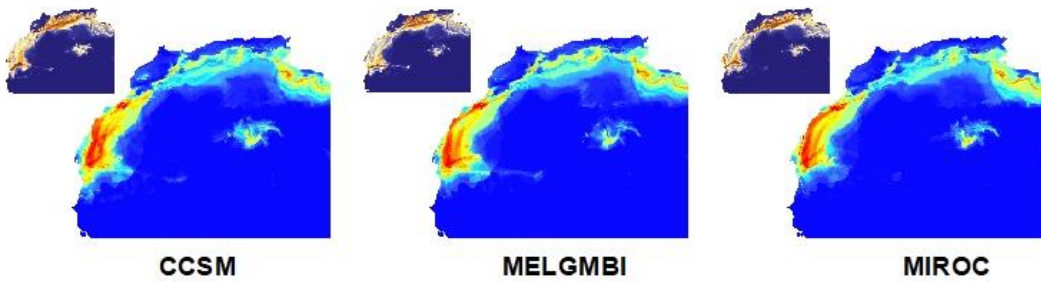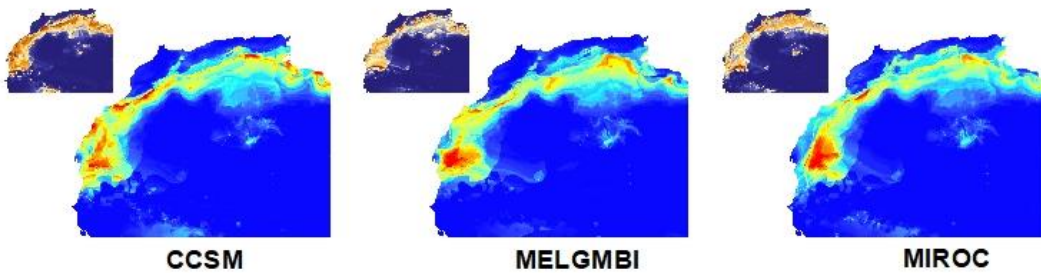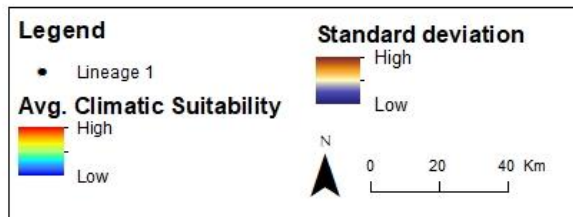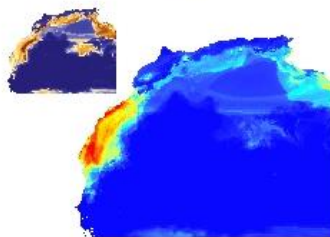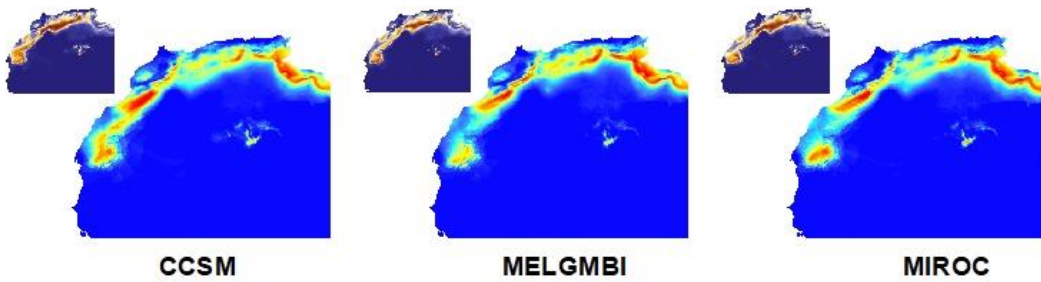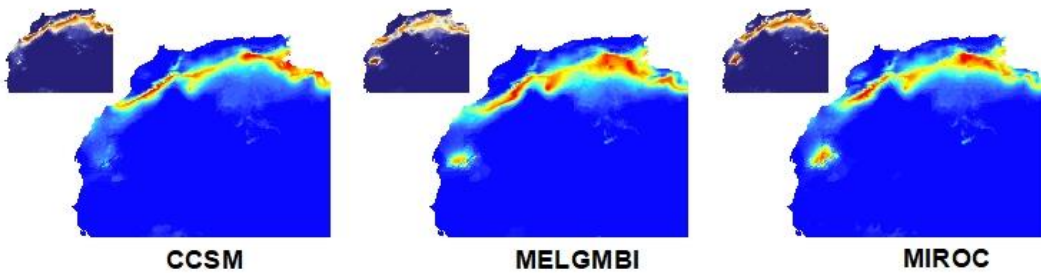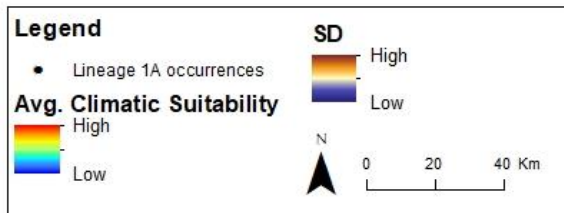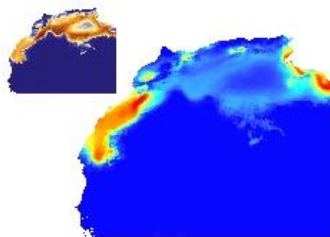
Appendix 7 - Probabilistic models of *Cerastes gasperettii* for the present and the past climatic periods
with standard deviation
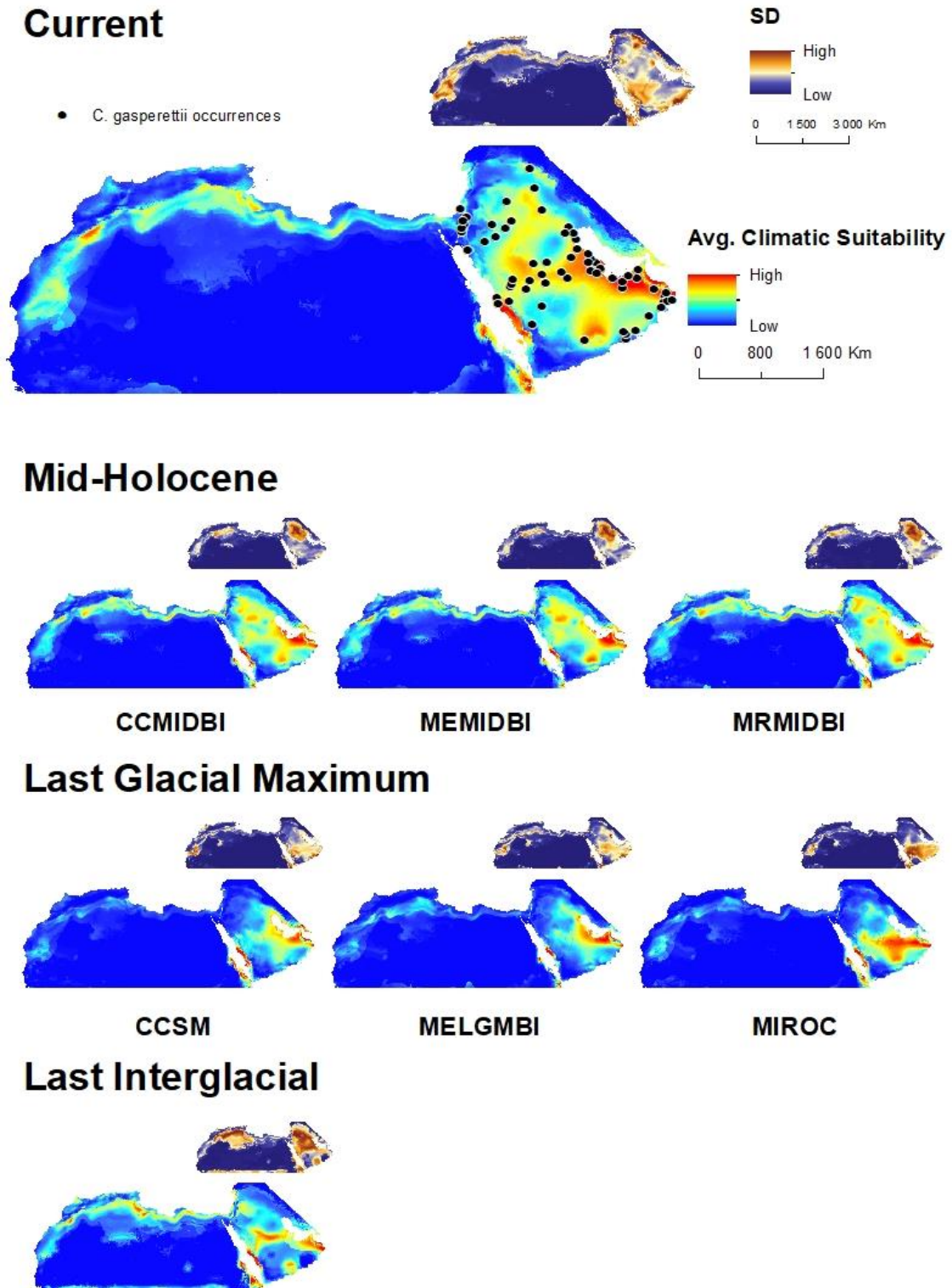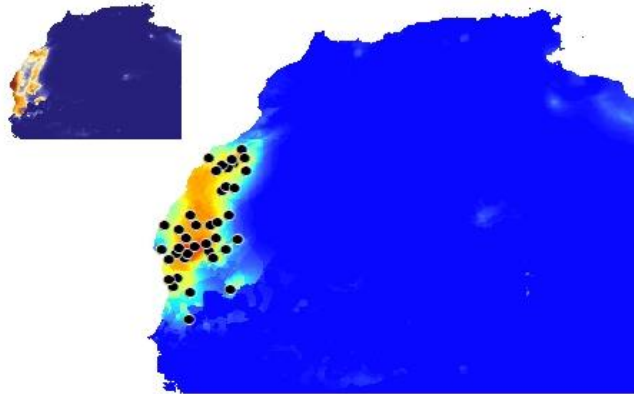
Appendix 8 - Probabilistic models of *Cerastes cerastes* West lineage for the present and the past climatic periods with standard deviation
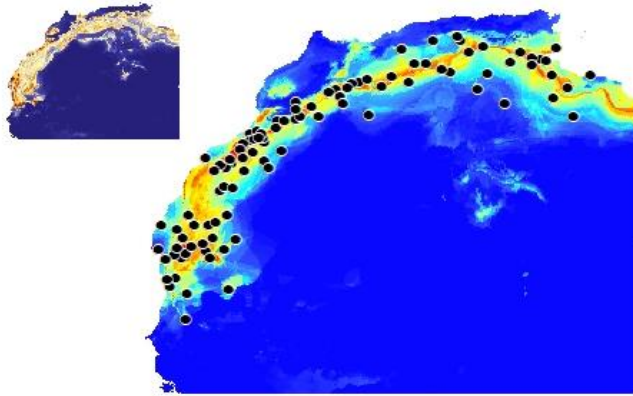
Appendix 9 - Probabilistic models of *Cerastes cerastes* West-North sub-lineage for the present and the past climatic periods with standard deviation
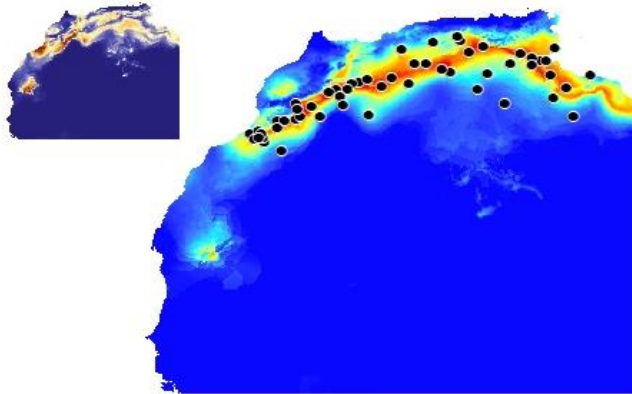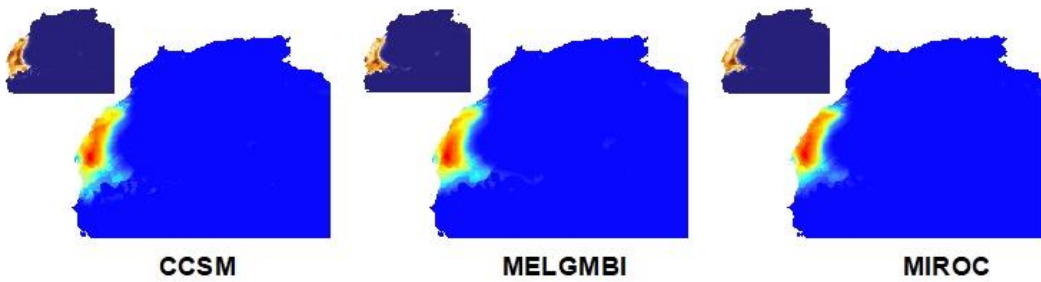
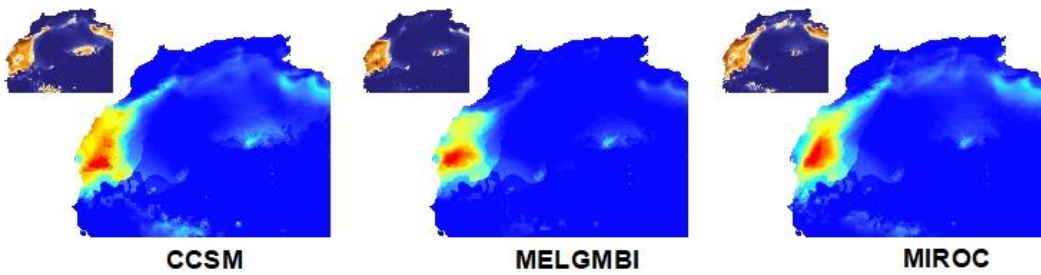*Appendix 10* - Probabilistic models of *Cerastes cerastes* West-South sub-lineage for the present and the past climatic periods with standard deviation