Pliocene colonization of the Mediterranean by Great White Shark inferred from fossil records, historical jaws, phylogeographic and divergence time analyses

(Article begins on next page)

01 September 2024

Journal of Biogeography

Pliocene colonization of the Mediterranean by Great White Shark inferred from fossil records, historical jaws, phylogeographic and divergence time analyses

Complete List of Authors:	Leone, Agostino; Universita degli Studi di Bologna, Laboratory of Genetics & Genomics of Marine Resources and Environment (GenoDREAM); Dipartimento di Scienze Biologiche, Geologiche e Ambientali (BiGeA) Puncher, Gregory; University of New Brunswick, Department of Biological Sciences, Canadian Rivers Institute; Maurice-Lamontagne Institute, Fisheries and Oceans Canada Ferretti, Francesco; Stanford University, Hopkins Marine Station, Department of Biology Sperone, Emilio; Universita degli Studi della Calabria, Dipartimento di Biologia, Ecologia e Scienze della Terra (DiBEST) Tripepi, Sandro; Universita degli Studi della Calabria, Dipartimento di Biologia, Ecologia e Scienze della terra (DiBEST) Micarelli, Primo; Aquarium Mondo Marino, Sharks Studies Center (C.S.S.) Gambarelli, Andrea; Universita degli Studi di Modena e Reggio Emilia, Museo di Zoologia e Anatomia Comparata Sarà, Maurizio; Universita degli Studi di Palermo, Dipartmento di Scienze e Tecnologie Biologiche Chimiche e Farmaceutiche (STEBICEF) Arculeo, Marco; Universita degli Studi di Palermo, Dipartimento di Scienze e Tecnologie Biologiche Chimiche e Farmaceutiche (STEBICEF) Doria, Giuliano; Museo civico di storia naturale Giacomo Doria, Museo Civico di Storia Naturale "Giacomo Doria" Garibaldi, Fulvio; Universita degli Studi di Genova, Dipartimento di Scienze della Terra, dell'Ambiente e della Vita (DISTAV) Bressi, Nicola; Museo Civico di storia naturale, Museo Civico di Storia Naturale di Trieste Dall'Asta, Andrea; Museo Civico di storia naturale, Museo Civico di Storia Naturale di Trieste Minelli, Daniela; Universita degli Studi di Bologna, Museo di Anatomia Comparata Cilli, Elisabetta; Universita degli Studi di Bologna, Dipartimento di Beni Culturali Vanni, Stefano; Universita degli Studi di Firenze, Museo di Storia Naturale "La Specola" Sez. Zoologia Serena, Fabrizio; Consiglio Nazionale delle Ricerche, Istituto per le Risorse e le Biotecnologie Marine (IRBIM) Díaz-Jaimes, Píndaro; Universidad Nacional Autonoma de Mexico,

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

	Laboratorio de Genética de Organismos Acuáticos Instituto de Ciencias del Mar y Limnología
	Baele, Guy; Rega Institute, KU Leuven, Department of Microbiology and Immunology
	Cariani, Alessia; Universita degli Studi di Bologna, Laboratory of Genetics & Genomics of Marine Resources and Environment (GenoDREAM); Dipartimento di Scienze Biologiche, Geologiche e Ambientali (BiGeA)
	Tinti, Fausto; Universita degli Studi di Bologna, Laboratory of Genetics & Genomics of Marine Resources and Environment (GenoDREAM); Dipartimento di Scienze Biologiche, Geologiche e Ambientali (BiGeA)
Key Words:	Great White Shark, Mediterranean, historical DNA, divergence time, phylogeography, Carcharodon carcharias

Responses to Editor and Reviewers in Bold

Dear Dr. Leone,

Reference Number: JBI-19-0030

Title: Pliocene colonization of the Mediterranean by Great White Shark inferred from fossil records, historical jaws, phylogeographic and divergence time analyses

We have designated this paper as requiring Full Revision, indicating that significant concerns have been raised about the paper in the review process but that we are keeping the paper centre open for the submission of a revised version. Resubmitted versions of papers classified under this heading are normally sent out for a second round of peer review prior to an editorial decision being made. There is no certainty of eventual acceptance.

This email provides: (1) specific comments from the chief editor and from the handling editor, (2) any referee reports.

To upload a revised manuscript: Go into your author centre and click on "Manuscripts with decisions". You will then see a link allowing you to upload your revision to the correct number.

A copy of this letter is also available in your Corresponding Author Centre at https://mc.manuscriptcentral.com/jbi

Thank you again for submitting to the 'Journal of Biogeography'. We look forward to receiving the revised version of your paper, ideally within 60 days. You can contact the editorial office if you would like an extension.

With best wishes,

Peter Linder

Chief Editor, Journal of Biogeography

FURTHER COMMENTS FROM THE CHIEF EDITOR

EDITOR'S COMMENTS TO AUTHOR

Editor: Gaither, Michelle
Comments to the Author:
This is a well written and interesting paper and both reviewers agree that this is an interesting contribution to the field. However both also raised important questions about the fossil data used here. These independent reviewers raised similar concerns about the use of GWS records from the Paleobiology

Database....

One reviewed said
the authors need to be cautious "and use recent literature rather than data from PaleoDB that suffer from outdated taxonomic assignments"

We hope that authors will carefully consider the detailed comments by reviewers before resubmission.

Dr. Michelle Gaither

Dear Dr. Gaither,

First of all, we would like to thank you for finding our study interesting and worth reviewing, and for granting us the extension of the re-submission. We really appreciated that.

We are glad that our work was found very interesting by both reviewers, and we have improved the manuscript; which we hope you will find suitable for publication on the Journal of Biogeography. We have found the comments useful and have considered all of them in our revision. We have marked our modifications to the manuscript in word track changes mode and detail below our responses to each of the comments raised.

Specifically, we have integrated more recent literature about fossil history of the species, and we've faced the problems with the PaleoDB database, further investigating the literature on fossil occurences of the species. Furthermore, we have integrated some new divergence time analyses with alternative calibration priors (estimated divergence between the target species and an outgroup, inferred using molecular data and molecular clock calibration, rather than using fossil records only), trying to clarify any doubts about the reliability of our study, and showing that our work is solid.

Yours sincerely,

Agostino Leone

On behalf of all the other authors

REVIEWER COMMENTS TO AUTHOR

Referee: 1

Comments to the Author
In this paper, the authors propose that the Mediterranean (MED) population of great white sharks (GWS) originated in the Pacific Ocean based on multiple lines of evidence. I really enjoyed reading this well-written paper and found the results very interesting and though provoking. I applaud the great amount of data analysed and the fact the authors have considered different types of approaches (paleontological, historical, molecular). They also provide a very complete list of references which shows a good understanding of the literature on the different approaches used. The two main strengths of this paper, in my opinion, are the findings that the MED population originated in the Pacific, and not in the Atlantic and the interpretations on local extinctions. Nevertheless, I consider that this work, which I think has great potential, has a number of problems that should be addressed before its publication. Below I describe three main issues, and then I provide a line by line list.

We thank the reviewer for reviewing our manuscript and are glad of their overall interest and suggestions. We detail a point by point response below.

Main issues:

The fossil record of the GWS, and its origin: This paper needs a stronger support on the origin of the GWS, and therefore, it begs for a more detailed analyses of its fossil record. The authors use the Paleobiology Database (herein, PBDB) to assess the fossil record of this species, but they don't seem to have filtered the data using solid criteria (other than just the species name). It has been proposed that the earliest occurrence of this species is late Miocene (Ehret et al. 2012) and not middle Miocene as stated here. However, I don't think that the global fossil record of the GWS has ever been assessed. For instance, as shown in Table S4, there are records from as early as the early Miocene. Evaluating the global fossil record of the GWS may seem like a lot of work, but there aren't that many records in the PBDB (< 60) so it should not be so difficult considering the number of co-authors involved. I personally may have some of the references that would need to be evaluated (see detailed comments below) and would be happy to share the PDFs with the authors (through the editor, perhaps?), if accessibility is an issue. This should also be complemented by the addition of records that are not in the PBDB already, such a the various references cited in the paper on the MED Pliocene occurrences. Since the hypotheses here proposed are largely based on the fossil record, I consider critical to address this issue.

We thank the reviewer for offering the access to references. We strongly hope we've mentioned these references in the new version of the manuscript. A deeper scan of the available literature and related fossil records has been made in order to have all the available information on the origin of the species (See SI). However, we would like to point out that the origination time of the species is above the scope of this work, which is focused on the origin of the Mediterranean population using all the information available. The data from PaleoDB were filtered using the species name, the synonymous names, and all the records were subsequently checked and evaluated through literature. We have modified the statement in L143-150, in order to be clearer about that. The map where modified in order to remove dubious occurrences. We are aware that the origin of the species is a still-open debate in the scientific community. The earliest fossil attributed to the species is dated at 16 Mya from Calvert Formation, Maryland (National Museum of Natural History USNM 336204), but overall, following the literature, it is stated that the origin of the species is dated back to about 11 Mya, in the late Miocene (see also Chapter 4 of A. P. Klimley & D. G. Ainley (Eds.), Great white sharks: the biology of Carcharodon carcharias (pp. 19-36). San Diego, CA: Academic Press.). We have integrated some new divergence time analysis using a normal distributed prior for the first calibration point with mean at 11 Mya. Parallelly, we have done also a divergence time analysis using the estimated divergence between the *Carcharodon carcharias* and the *Lamna nasus* as first calibration point, rather than using a fossil calibration for the origin of the species. The outcomes were attached to the manuscript, showing no differences in the results on which we have built our discussion. At the moment, not all the records from the Mediterranean were included in the PaleoDB database. Our work aim is to reveal the origin of the Mediterranean population of GWS using all the information available, without detracting from the fact that the records can be submitted in the near future in PaleoDB database, even in the context of a new work dedicated to the fossil history of the species.

Calibration point and the Isthmus of Panama: It has been proposed that the Isthmus of Panama formed much earlier than here used. I think the authors need to discuss how an earlier formation of the isthmus of Panama would affect their results. Also, I encourage the authors to consider the intermittent shallow connections that could have allowed migrations of sharks across the isthmus, even if there was indeed an earlier closure (references provided under L240; but also see Pimiento et al. 2013. Journal of Paleontology 87: 755–774).

We thank the reviewer for the interesting point raised out. We would like to remark that the formation of the Isthmus of Panama has not been used as calibration point for our work, which rely on different calibrations from those used in Gubili, et al. 2010. We were aware of this discussion about a potential

earlier formation of the Isthmus of Panama, and we have discussed a possible effect of an earlier formation of the Isthmus of Panama (L357-364).

The statements on local extinctions lack of support: What were the potential mechanisms of those extinctions and in what basis do the authors suggest that there were extinctions at all? Keep in mind when addressing this that absence of fossils does not necessarily mean that the species went extinct, especially at a regional level. If this is based on the molecular-level analyses, then it needs to be better explained.

As explained in the manuscript, line 268-281, the analysis of RASP suggested that the population of the MED could be the result of a vicariance event. This fact, combined with the lack of an Atlantic lineage well structure, and a combined lineage NWA/SA, suggest the the NWA population represent a relatively young population which may have colonized the NWA from few South African individuals. The lack of a continuum NEP-NWA-MED, with the MED genetically similar to NEP, may indicate that the past GWS present in the NWA went through a local extinction or extirpation, leaving a phylogeographic gap filled subsequently by the SA lineage. Our discussion rely exclusively from these molecular analysis, since we are aware that the lack of fossils does not necessarily mean that the species went extinct in that specific area.

Line by line comments:

Background:
L74: Appropriate global reference: Myers et al (2007) Science 315, 1846-1850.

The reference has been integrated

L83: Missing reference on long distances: Domeier, M.L. & Nasby-Lucas, N. (2008) Marine Ecology: Progress Series, 370, 221–237.

The missed reference has been integrated

L125: The use of private collections represent a problem in terms of reproducibility. I am not sure what to recommend to do about this, as I understand how they can be essential, but they are often disregarded in the literature.

The private collection mentioned here are the samples LICC01LI, FACC01TI, FACC02TI, FACC03TI, FACC04TI, FACC05TI, from a private archive at the Regional Agency for the Protection of the Envinronment of Tuscany (ARPAT Livorno) the first, and from the tuna trap fishermen of Favignana the rest. The samples could be available under request, and some of these were used in past works (Gubili, et al. 2010). However, we think that given the difficulty of sampling the species in the Mediterranean area, it's essential to collect as many samples as possible with sufficient information on place and date of capture.

General: Through the paper, it is bit hard to grasp the aims of the different methodologies. It would therefore be useful to explain, perhaps in the last paragraph of the background, in a logical order, how the different types of data and approaches are answering the research questions (e.g., with fossils we aim to know x, with historical collections we aim to answer y, etc.).

Since the bibliographic research for fossil occurrences of the species have strengthened the results from molecular data, suggesting also good discussions and interpretations, we have modified the last paragraph of the background to "A deep evolutionary analysis of the newly generated Mediterranean sequences aligned with those from global GWS populations available in public repositories, and the scan of the fossil records occurrences in literature, with special emphasis on the Mediterranean area , provided an opportunity to better unravel the origin of the GWS Mediterranean population.". We strongly hope it is clearer now.

Methods:

L143: Specify what the filtering criteria was. Only the use of the species name? That would not be enough. The PBDB is great in that it presents all records of a species, but it has to be used carefully because some records are simply unreliable, i.e., some specimens are misidentified, or the age of some localities have been changed or refined. Unreliable records are still "correct" in the sense that they represent a record, but should not be included in the analyses. For an idea on how to filter the data you could see Pimiento and Clements (cited in the ms.), but basically, you would have to go to the primary reference that supports the PBDB occurrence and check if its accurate. As stated above, I may be able to help providing some PDFs (but would have to check).

The filtering criteria were the species name and relative synonyms. The Genus spp. records were discarded due to the doubts of wrong taxonomic assignments. Additionally, the related references associated to dubious fossil dating were checked for accuracy of species and dating assignment. The dubious records were then discarded. Additionally, a deep bibliographic research has been carried out, as mentioned above. All the info are attached in the SI.

L145: How was the bibliographic search made? Was it exhausted and comprehensive? What means were used (e.g., Google Scholar)? Were those papers added to the PBDB? Also, I am not sure if the map can be regarded as an analysis, and if the map is in fact an stratigraphic map.

We went through a deep and exhausted bibliographic research using public repository (NCBI, Scholar, PaleoDB, ecc.) and for each work, the references we subsequently consulted. Some papers were still not added to PaleoDB, since our aim was to retrieve as much information as possible on the stratigraphic distribution of the species and information regarding key fossils for our study (e.g. from the Mediterranean area), and the map summarize these information as well as graphically showing the events that affected the species during the formation of the different populations we know today.

L151: Why do you want to verify the constant rate of substitution, and the lack of a saturation effect? Please explain the aim of your methods (see above).

We agree with the reviewer that these analysis did not add anything more to the manuscript, and were removed. Also taking in mind that we did a parallel analysis using an outgroup. So, a relaxed clock was the best choice (also taking into account the big divergence between the two main lineages, MED-NEP-AUS vs SA-NWA).

General: The paper would benefit from a abbreviation sections as there are a lot throughout, some of them are not spelled out (e.g., CAS in L337).

The Central American Seaway has been spelled at L 254 of the revised manuscript, as well as all the acronyms. We think it is not necessary to add an abbreviation section, also due to problem with words limit.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Results:

L240-241: There has been some solid evidence that the isthmus of Panama formed much earlier (Farris et al. 2011. Geology, 39:1007–1010; Montes et al. 2012. Journal of Geophysical Research, 117:B0410; Montes et al. 2012. Geological Society of America

Bulletin, 124:780–799; also see Jaramillo et al. 2017 [comments on O'Dea et al. 2016] in Science Advances 3: e1602321).

A potential earlier formation of the Isthmus of Panama (Bacon et al., 2015; Lessions et al., 2015; Montes et al., 2015; O'Dea et al., 2016; Jaramillo et al., 2017) would not affect our results, since it is plausible that the Pacific lineage that gave rise to the MED population may already have been present in the Atlantic before and after the closure of the CAS and before its potential local extinction in the Atlantic. Also, past intermittent shallow connections are plausible and discussed in literature (Jaramillo et al., 2017).

L244-245: Why middle Miocene if Table S4 shows early Miocene occurrences? Also, consider that the oldest fossil of a species does not necessarily represent the time of origin because species can originate before they leave a record. Therefore, it is incorrect to say that the evolutionary history of the GWS dates to middle Miocene (or early Miocene in agreement with Table S4). I looked at one of the references supporting an early Miocene occurrence in the PBDB (collection number 42842) and it is based on a list of species, so it cannot be confirmed that that record was actually a GWS (there is also a note next to the species list in such collection warning cautious about that record).

L245-246: The earliest occurrence in the MED is based on Table S4? If is, it is not possible to track it there as no locality/region is provided. It is fundamental here to provide more details/references on what is the earliest occurrence of the GWS in the MED.

As explained in the manuscript, line xxx, the earliest occurrence in the MED, is the early Pliocene of South of Spain (see blabla), but it lack a defined dating. For the dating, we've used the fossil occurrences described and dated by xxx at around 3 Mya. There are also some Miocene records from the MED according to the work of XXx, but there is no consensus on the correct dating for these fossils (see XXX), and also, hypothetical Miocene individuals from the area would have faced local extinction due to the Messinian Salinity Crisis.

Discussion:

L275: I am not sure "evolutionary pathway" is the right term here. Also, what is the biological or ecological implication of the Messinian salinity crisis? It is only mentioned, but no mechanisms are suggested.

We think that "evolutionary pathway" fit well here, since we are highlighting intraspecies evolutionary processes mediated by historical dispersion, vicariance, ecc. We have included a statement on the biological and ecological implication of the Messinian Salinity Crisis. "entering the Mediterranean after the Messinian salinity crisis, which may have caused the local extinction of the marine megafauna due to the total or partial drying up of the Mediterranean with related ecological upheavals". L289-291.

L327: I can only see 4 records in the PBDB, 1 from Spain and 3 from Italy. If by abundant data the authors refer to are based on the references provided, are those included in the list of fossils? (I personally think should be added to the PBDB) and are they included in the map?

Actually, we have found several Mediterranean fossils of GWS, listed in the SI. In the PaleoDB and in the map, only the PaleoDB records are shown. A deeper analysis of the original location of these fossils is needed before the submission to PaleoDB (which is above the scope of this work).

L328: What do you mean by "origin of those specimens"? It seems to me that those are the dates of the fossils, so it is not necessarily the origin.

We agree with the reviewer. The statement was modified removing "origin of"

L332: How do you know it was gradual? See Pimiento et al. 2016. Journal of Biogeography 43(8): 1645–1655.

We thank the reviewer for the missed reference. We removed the statement "gradually" from the line

L356-362: Intriguing statement. It was proposed that the Pliocene marine megafauna extinction was related with sea level changes and consequent habitat loss. In this global event, which was studied at the genus level, the GWS was not affected. How does this fit with your hypothesis of a local extinction of great white shark populations in the MED?

As explained in the statement, we don't think that the Mediterranean population of GWS has been affected by local extinction in the Pliocene. A Pliocene local extinction of the Pacific-derived GWS lineage in the North Atlantic Ocean and further replacement by Indian components, and the subsequent persistence of a Pacific-derived GWS in the Mediterranean, could explain such phylogeographic discontinuity. we have argued that a possible Mediterranean population settled before the Messinian Crisis of Salinity has gone extinct, but following suggestions from the reviewer 2 we removed this statement, since the presence of GWS in the Mediterranean Miocene is dubious, and it is irrelevant for our hypothesis.

Supplement material:
Table S4: I was not able to retrieve records using PaleoDB N° column. Those numbers do not correspond to PaleoDB collection numbers. So what are they? Provide more info, collection number is crucial, but also, country and or locality.

We thank the reviewer for pointing it out. We have replaced the column with the correct reference codes. All the necessary info were showed in the table. Contry and locality are implicit in the coordinates, since PaleoDB does not include country in the downloadable spreadsheet.

Referee: 2

Comments to the Author
Dear authors:

This study is well conceived and the manuscript is well and clearly written. There are only some minor corrections I'd like to ask you to consider. But there are also some inaccuracies concerning your usage of fossils, which might change your interpretations and discussion to some extend. Please consider the comments below.

We thank the reviewer for finding our study interesting and worth reviewing, and we found the comments and suggestions very useful. We detail a point by point response below.

Line 152: You consider that the substitution rate is constant and that you are intending to verify it. This

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

seems odd since substitution rates normally are considered not to be constant. Could you please include a short reference or support why you consider it constant?

We reconsidered this analyses using a relaxed clock for substitution rate, allowing for rate changes across branches, differently from past published work on the species, which used the constant rate. Using the relaxed clock we were also able to to a divergence time using the estimated divergence between the GWS and the outgroup.

Line 173: You are using the fossil shark described by Gottfried & Fordyce as oldest record of Carcharodon. This specimen is from the Oligocene. However, Gottfried is a proponent of traditional usage of taxonomic names and the assignment of this species to Carcharodon is considered not valid any more. The history of the taxonomic assignment of the species "megalodon", "angustidens", and auriculatus as several other megatooth lamniforms has been debated over decades. There is no common agreement that we are dealing with different clades, such as Otodus, Carcharocles, Megaselachus, Carcharodon, Carcharomodus, etc. The genus Carcharodon can only be traced back into the late Miocene, but normally it is a Pliocene species. The late Miocene occurrences are in South America - see Cione et al. (2012): Oldest record of the Great White Shark (Lamnidae, Carcharodon; Miocene) in the Southern Atlantic; Adnet et al. (2009): New fossil teeth of the White Shark (Carcharodon carcharias) from the Early Pliocene of Spain. Implication for its paleoecology in the Mediterranean; Cappetta (2012: Handbook of Palaeoichthyology. Especially the paper by Adnet might be useful for the discussion about the origin of Carcharodon carcharias. The Handbook summarizes our current knowledge about the systematics of the megatooth lamniform sharks.

In the mentioned Gottfried & Fordyce paper, there is an interesting paragraph on the origination time of *C.carcharias* with the description of a 16 My dated fossil, already described in past work (we additionally added the older reference), and we are referring just on that, and not on the *C.angustidens*. However we decided to use a more conservative calibration (11Ma) due to the fierce debate around the oldest fossil records of GWS. The line has been modified accordingly.
The work by Adnet et al. is extremely useful in terms of colonization of the MED by GWS, and we have considered all the info discussed in that work. We also proposed a new analysis taking into account the estimated divergence between the GWS and the *L.nasus* (see results and SI). Our hypothesis is still strongly supported by the results of the new analysis.

The discussion needs to be revised in the light of this information. I also would suggest to not use the Miocene records from PaleoDB, because this information is very outdated because of wrong taxonomic identifications and/or wrong stratigraphic placement. This is a general problem with this database. Rather refer to papers of specialists.

The reference records were deeply scanned for corrected species assignment to our target species, and the outdated records were removed. Also, we have considered the stratigraphic samples from middle Miocene, since we think they could add interesting information to discuss, and in any case, they don't affect our hypothesis since we tested it also considering different calibration priors (see divergence from outgroup in the results, as discussed above).

Line 176: The paper of Cigala-Fulgosi reports on indirect evidence, but not on an actuarl record. These traces also could have be caused by some other large sharks with serrated tooth crowns such as snaggle tooth or tiger sharks that were very large during the Miocene and also Pliocene. Rather use information from the paper of Adnet et al. or similar publications.

This dating is justified by well dated *C.carcharias* fossil records associated with ceataceans remains from the Pliocene. We have integrated the references with the right work. Also the paper from Adnet et al. has been taken into consideration, but it lacks precise dates that can be considered in divergence time

analysis, even if it is extremely useful in terms of discussion (They mention just the Pliocene epoch).

Line 200: Shouldn't it be "A MCC estimated" rather than "An MCC estimated)?

Corrected

Line 243: The statement that divergence between Mediterranean and Pacific GWS coincides with the closure of the Isthmus seems odd, because one would expect that this event separated the Pacific from Caribbean/Atlantic populations. Mediterranean GWS populations are quite far away. But Atlantic populations seemingly are different? How can this be explained. Especially if the oldest record comes from the south Aouth American Atlantic coast? The records of Carcharodon from the late Miocene of Chile are actually Pliocene in age and the remains from Sacao might represent a different species. There was also a record of a new species from the late Miocene of Chile but the stratigraphic age seems also ambiguous.

As explained in the manuscript, the closure of the CAS separated the Pacific populations from the Pacific-derived Atlantic and, consequently, Mediterranean populations. The local extinction of the Pacific-derived Atlantic population, allowing a more recent colonization of the North Atlantic by South African individuals, may explain the phylogeographic gap observed. We strongly think that any other phylogeographic pattern/hypothesis should have left a SA/NWA genetic signal in the MED population. We found unlikely the phylogeographic route proposed by past works, with single Pacific individuals crossing the entire Indian Ocean, then the Atlantic and finally the Mediterranean Sea without leaving a phylogeographic signal. Also, the Statistical Dispersal-Vicariance Analysis (S-DIVA), and the Dispersal-Extinction-Cladogenesis (DEC) analysis, suggest such vicariance hypothesis. Also, a local extinction in the Atlantic may have erased any local genetic signal, regardless of the fossil records, which, however, don't refute our hypothesis, especially concerning the Mediterranean.

In Line 320, you actually refer to the divergence of Pacific and Atlantic populations 3.5 Ma. Does this mean that you consider Atlantic and Mediterranean populations to be the synonymous?

Here we are referring to the method used by Gubili et al. to estimate the substitution rate of the species. As extensively discussed in the manuscript, we don't consider the contemporary Atlantic and Mediterranean populations to be synonymous.

Line 246: Your notion that the GWS extends back into the late Miocene in the Mediterranean area is wrong! The oldest records are from the Pliocene, well after the salinity crisis. This fossil record would support your interpretation that the GWS migrated into the Mediterranean Sea after the Mesinian crisis. Otherwise, you have to consider that the GWS was in the Mediterranean Sea before the crisis, became extinct and was replaced in the Pliocene by a new, different population. In this scenario, one could argue that the GWS migrated into the Mediterranean Sea during the Miocene via the open gate towards the Indian Pacific and after it was closed (in the Pliocene), GWS migrated into the Mediterranean Sea through the Gibraltar strait.

The presence of GWS in the Miocene of Mediterranean is still an open debate. Marsili (2006) has been able to catalog fossils from the Miocene of the area (Marsili, 2008), discussing how these fossils became much more common in the Pliocene. It seems that these fossils are problematic and to be re-date to Pliocene. In this case we agree with the reviewer in removing this confusing statement, since this is not the scope of our work, and more detailed paleontological analyses are needed to clarify the presence or not of the GWS in the Miocene of the MED.

Line 251: How do your estimates of divergences change when you exchange the incorrectly assigned fossil

records with correctly identified records? As I wrote above, don't rely to much on PaleoDB but make a through literature review using recent papers!

There are no different results that invalidate our hypothesis, also using different calibration points (see results). The calibration points were extrapolated from literature.

Discussion section: You have to reconsider your discussion in the framework of correctly used fossil records. This might change some of your interpretations.

We went through all the manuscript in the framework of the points suggested by the reviewers, and the interpretations are still solids.

Line 277: Is your interpretation still correct if the oldest fossil record of GWS is from the late Miocene of Argentina and not from Australia (the species "auriculatus" belongs to the genus Carcharocles)?

Yes, the interpretations are still correct. We also kept in mind that the oldest fossil record from Paraná Formation in Argentina (Cione et al. 2012. Oldest record of the Great White Shark (Lamnidae, Carcharodo n; Miocene) in the Southern Atlantic. Geobios, 45(2):167-172. doi: 10.1016/j.geobios.2011.06.002), could be a different species, as stated by the authors.

Line 288: The low estimated genetic diversity correlates well with the assumption that GWS colonized the Mediterranean Sea during the Pliocene and might have underwent a bottleneck then in relation to climatic changes during the Pliocene and Pleistocene. You should include a short discussion of climatic changes in the Mediterranean Sea during the Pliocene and Pleistocene in relation to your genetic findings.

We think it is well discussed, and a deeper analysis of the climate change in the paleo-Mediterranean and its effect on the genetic variation of the species (also more than our target species) need a dedicated work

Lines 345 following. Does this mean that the Mediterranean population diverged from Atlantic populations at 2.93 Ma? And does this date really correlate with oceanographic phenomena? This needs to be more elaborated because it is confusing and even contradictory as it stands momentarily.

We have retested the divergence time, which is now 3.23. That timing was an error. However, no. This divergence indicate that the Mediterranean population diverged from the common ancestors within the Pacific populations at 3.23 Mya.

Overall: There might be minor changes by adjusting oldest verified records of GWS but your data, especially concerning the colonization of the Mediterranean Basin by GWS is in good accordance with the fossil record.

We are glad of this, and the dubious oldest records were checked and eventually removed. For the divergence time analysis, we used only solid records.

Kind regards,
Jürgen Kriwet

1 **Pliocene colonization of the Mediterranean by Great White Shark inferred from fossil records,**

2 **historical jaws, phylogeographic and divergence time analyses**

3

4 Agostino Leone[1,π,*], Gregory Neils Puncher[2,3], Francesco Ferretti[4], Emilio Sperone[5], Sandro Tripepi[5],

5 Primo Micarelli[6], Andrea Gambarelli[7], Maurizio Sarà[8], Marco Arculeo[8], Giuliano Doria[9], Fulvio

6 Garibaldi[10], Nicola Bressi[11], Andrea Dall'Asta[11], Daniela Minelli[12], Elisabetta Cilli[13], Stefano

7 Vanni[14], Fabrizio Serena[15], Píndaro Díaz-Jaimes[16], Guy Baele[17], Alessia Cariani[1], Fausto Tinti[1]

8 [1]Dept. Biological, Geological & Environmental Sciences (BiGeA), Laboratory of Genetics &

9 Genomics of Marine Resources and Environment (GenoDREAM), University of Bologna, 48123,

10 Ravenna, Italy

11 [2]Dept Biological Sciences, Canadian Rivers Institute, University of New Brunswick, Saint John,

12 NB, E2L 4L5, Canada

13 [3]Genomics Laboratory, Maurice-Lamontagne Institute, Fisheries and Oceans Canada, Mont-Joli,

14 QC, G5H 3Z4, Canada

15 [4]Hopkins Marine Station, Department of Biology, Stanford University, 120 Oceanview Boulevard,

16 Pacific Grove, CA 93950, USA

17 [5]DiBEST Department of Biology, Ecology and Earth Science, University of Calabria, 87036,

18 Arcavacata di Rende, Italy

19 [6] Sharks Studies Center (C.S.S.), 58024, Massa Marittima (GR), Italy

20 [7]Museum of Zoology and Comparative Anatomy of Modena, University of Modena and Reggio

21 Emilia, 41121, Modena, Italy

22 [8]Dipartmento di Scienze e Tecnologie Biologiche Chimiche e Farmaceutiche (STEBICEF),

23 University of Palermo, 90123, Palermo, Italy

24 [9]Civic Museum of Natural History "Giacomo Doria", 16121, Genova, Italy

25 [10]Department of Earth Sciences, Environmental and Life, University of Genova, 16132, Genova, Italy

26 [11]Civic Museum of Natural History of Trieste, 34100, Trieste, Italy

27 [12]Museum of Comparative Anatomy, University of Bologna, 40126, Bologna, Italy

28 [13]Department of Cultural Heritage, University of Bologna, 48121, Ravenna, Italy

29 [14]Museum of Natural History of Firenze "La Specola", 50125, Firenze, Italy

30 [15]Institute for Biological Resources and Marine Biotechnology (IRBIM), National Research Council

31 – CNR, 91026, Mazara del Vallo (TP), Italy

32 [16]Laboratorio de Genética de Organismos Acuáticos Instituto de Ciencias del Mar y Limnología,

33 Universidad Nacional Autónoma de México, Apdo. Postal 70-305, 04510 Mexico, D.F., Mexico

34 [17] Department of Microbiology and Immunology, Rega Institute, KU Leuven, 3000 Leuven, Belgium

35 *Corresponding author: agostino.leone@unibo.it

36 π Present address: Marine Research Division, AZTI, 48395, Sukarrieta, Bizkaia, Spain

37

38

Abstract

Aim We aim to unravel the evolutionary origin of the contemporary, poorly characterized Mediterranean population of Great White Shark, *Carcharodon carcharias*, GWS, tracing back its paleo-migration pattern by phylogenetic and dispersal vicariance analyses at the global level.

Location Mediterranean Sea

Taxon *Carcharodon carcharias*

Methods We have built the largest control region mtDNA (CR) sequence dataset for the Mediterranean GWS from referenced historical jaws spanning 19th and 20th century. Mediterranean and global GWS CR sequences were analyzed for genetic diversity, phylogenetic relationship and divergence time. A Bayes factor approach was finally used to test two scenarios of divergence time and paleo-colonization of the Mediterranean by GWS, calibrated using fossil records and paleo-geographical events.

Results The results confirmed a closer evolutionary relationship between Mediterranean GWS and populations from Australia-New Zealand and the Northeastern Pacific coast rather than populations from South-African and Northwestern Atlantic. The Mediterranean GWS lineage showed the lowest genetic diversity at the global level, indicating its quite recent evolutionary origin. The divergence time of the Mediterranean GWS at 3.236 Ma by a potential dispersal-vicariance process involving Australian/Pacific paleo-populations was retrieved as the best divergence/paleobiogeographic scenario.

Main conclusion Based on the fossil records, phylogeographic patterns and divergence time, we revealed that the Mediterranean GWS population originated in the Pliocene following the Messinian Salinity Crisis. Colonization of the Mediterranean by GWS likely occurred via an eastward paleo-migration of Australian/eastern Pacific elements through Central America Seaway, before the complete closure of the Isthmus of Panama. The Pliocene origin of Mediterranean GWS by Pacific paleo-populations is in contrast with the previously ad hoc reconstructed scenario in which Australian GWS colonized the Mediterranean by antipodean northward migration caused by navigational errors from South Africa during Quaternary climatic oscillations.

Keywords Great White Shark, Mediterranean, historical DNA, divergence time, phylogeography, *Carcharodon carcharias*

Background

69
70   Large predatory shark populations in the Mediterranean Sea have declined dramatically over

71   the last century (Ferretti, Myers, Serena, & Lotze, 2008). The loss of apex shark predators, like sharks,

72   throughout the world's oceans over the past century is likely to have caused profound ecological

73   alterations and potentially large-scale trophic cascades (Myers et al., 2007; Ferretti, Worm, Britten,

74   Heithaus, & Lotze, 2010). Among the species that witnessed the most precipitous declines is the Great

75   White Shark (*Carcharodon carcharias*, L.1758, henceforth GWS; McPherson & Myers. 2009),

76   which is currently listed as Critically Endangered in European seas (Nieto et al., 2015). GWS are

77   widespread throughout the globe, aside from the polar regions, with hotspots of abundance located

78   off the coasts of South Africa, Australia, New Zealand, Japan, North and South America, and in the

79   Mediterranean (Compagno, 1984; Fergusson, 1996). Integrated ecological, genetic and tagging data

80   has revealed natal homing and philopatric behavior of GWS, with extraordinary trans-oceanic

81   migrations of both sexes between geographically distant populations in the Indian (Pardini et al.,

82   2001; Bonfil et al., 2005; Blower, Pandolfi, Bruce, Gomez-Cabrera, & Ovenden, 2012) and North

83   Pacific Ocean (Domeier & Nasby-Lucas, 2008; Jorgensen et al., 2010, 2012). In the Mediterranean,

84   GWS have long been observed and documented by the public, resource users and scientists, who have

85   provided opportunistic occurrence records, from direct sightings, fisheries catch, records of bite

86   marks found on prey, like dolphins and tuna, and museum specimens dating back to the early 19th

87   century (Mancusi et al., 2002; De Maddalena, 2006; Sperone et al., 2012). In the past, free-swimming

88   individuals and pairs were frequently recorded in areas where large pelagic fisheries were intense

89   (e.g. the Sicilian Channel, the Ligurian and Tyrrhenian seas; Fergusson, 1996; Storai, Vanni, Zuffa,

90   & Biagi, 2005). A few records of GWS pups from Turkey and Tunisia, as well as juveniles in the

91   Sicilian Channel, suggest that the Mediterranean may host GWS nursery areas (Fergusson, 1996;

92   Storai, Mojetta, Zuffa, & Giulian, 2000; Kabasakal, & Gedikoğlu, 2008). However, the natural history

93   of Mediterranean GWS still remains largely uncharacterized.

94   Efforts have been made in the past to resolve the phylogenetic relationships of the

95   Mediterranean GWS population using mitochondrial DNA (mtDNA) sequences (Gubili et al., 2010;

96   2015). Using up to five specimens, Gubili et al. (2010, 2015) concluded that the Mediterranean

97   population is more closely related to populations in the Pacific Ocean (Australia, New Zealand and

98   Northeastern Pacific) than to those from the western Indian Ocean (South-Africa) and northwestern

99   Atlantic Ocean (Florida). Based on a nucleotide substitution rate between the two major lineages

100 (Northeastern Pacific vs. North West Atlantic and Eastern Indian) calibrated by the formation of the

101 Isthmus of Panama (3.5 Ma) and the Sunda-Sahul Shelves (5 Ma) respectively, Gubili et al. (2010)

102 suggested that Mediterranean GWS are descendants of a few disoriented individuals who immigrated

103 from Australia/New Zealand during the Pleistocene (348-565 ka) by an antipodean route along the

104 western coast of Africa. The hypothesis of multiple relatively recent colonization events was also

105 considered given the haplotype relationships using few historical and contemporary Mediterranean

106 specimens (Gubili et al., 2015).

107   The sampling of GWS specimens in the Mediterranean has proven to be difficult because of

108 their decreased abundance in recent decades. However, there is a great number of referenced and dry-

preserved GWS specimens in several Italian museums and private scientific archives, such as stuffed and mounted skins, jaws, vertebrae and teeth collected over the last two centuries in the Mediterranean (Mancusi et al., 2002; De Maddalena, 2006). Considering the recent advances in ancient DNA (aDNA) technologies for addressing questions of conservation genetics, this material represents a great opportunity for reconstructing the natural history of Mediterranean GWS (Willerslev, & Cooper, 2005; Wandeler, Hoeck, & Keller, 2007; Ahonen, & Stow, 2008; Riccioni et al., 2010). Using a suite of newly developed techniques (Hofreiter, Serre, Poinar, Kuch, & Pääbo, 2001), researchers have been able to analyze the spatio-temporal population dynamics and genetic diversity of several other marine fish species, using well-preserved historical samples (e.g. vertebrae, scales, shark teeth and jaws; Hauser, Adcock, Smith, Ramirez, & Carvalho, 2002; Hutchinson, van Oosterhout, Rogers, & Carvalho, 2003; Ahonen, & Stow, 2008; Riccioni et al., 2010). Unfortunately, most historical GWS specimens have been archived in sub-optimal conditions, thereby compromising their potential for DNA-based applications. Moreover, since the potential value of the preserved specimens can decrease considerably if they are subjected to damaging invasive sampling techniques, many collectors are reluctant to loan specimens for molecular studies.

By analyzing DNA sequences of preserved specimens of GWS caught over the last 195 years in the Mediterranean, from eight Italian museums and private collections, we have been able to deeply explore the evolutionary history of Mediterranean GWS. Using effective and affordable aDNA techniques widely used to extract and genotype DNA from historical specimens of marine fish (e.g. vertebrae, bones, skin), we have generated a publicly available mtDNA sequence dataset from 18 Mediterranean GWS individuals (see Table S1 in Appendix S3). A deep evolutionary analysis of the newly generated Mediterranean sequences aligned with those from global GWS populations available in public repositories, and the scan of the fossil records occurrences in literature, with special emphasis on the Mediterranean area, provided an opportunity to better unravel the origin of the GWS Mediterranean population.

Methods

Full details of the collected historical GWS specimens, sampling procedures, protocols for aDNA extraction, PCR amplification of Control Region (CR), sequencing and sequence analyses are provided in the Supplementary Methods in Appendix S1, Figures S1-S3 of Appendix S2 and Table S2 of Appendix S3.

Analysis of fossil evidences

The extensive catalogue of taxon-specific GWS fossils featured in the online and open access Paleobiology Database (https://paleobiodb.org/#/), and its R package 'paleoDB' (Varela et al., 2015), was used to create a distribution and stratigraphic map of global GWS fossils. The downloaded database was further filtered manually to avoid the use of homonym extinct taxa commonly named as "Great White Shark". Only records specifically classified as *Carcharodon carcharias* and relatives synonyms were retained (Table S4 in Appendix S3). Moreover, a deep bibliographic records research with most attention to those related to *C. carcharias* fossils from the Mediterranean area was carried

148 out (Table S5 in Appendix S3) (Cigala-Fulgosi, 1990; Applegate, & Espinosa-Arrubarrena, 1996;
149 Gottfried, & Fordyce, 2001; Bianucci et al., 2002; ~~Marsili, 2006;~~ Marsili, 2008; Adnet, Balbino,
150 Antunes, & Marín-Ferrer, 2009)~~. were selected and used to strengthen the analysis of fossil evidence.~~

Time of the Most Recent Common Ancestor (TMRCA) and estimation of evolutionary rate

152 ~~In order to verify the constant rate of substitution, and the lack of a saturation effect, a molecular~~
153 ~~clock test as implemented in MEGA7 v.7.0.14 software (Kumar, Stecher, & Tamura, 2016) and a~~
154 ~~correlation between transitions (Ts) and transversions (Tv) against the Tamura-Nei's distance model~~
155 ~~(Tamura, & Nei, 1993) using the 'ape' R package (as the HKY model is not available in ape) (Paradis,~~
156 ~~Claude, & Strimmer, 2004) were performed.~~

157 The divergence time analysis of the GWS lineages was carried out using Bayesian inference
158 through Markov Chain Monte Carlo (MCMC) as implemented in BEAST v1.10.0 (Suchard et al.,
159 2018). An initial analysis using calibration priors without sequence data was carried out to determine
160 if the calibration priors ~~were specified properly,~~ did not interact unexpectedly and to assess that the
161 data were informative~~.~~, ~~thus ensuring that the estimated divergence time was not the result of the~~
162 ~~priors alone~~ (Fulton, & Strobeck, 2010). ~~A~~ A ~~strict~~ relaxed molecular clock and ~~and~~ a constant
163 population size coalescent model, were used ~~as the clock and coalescent models of choice~~ to recover
164 time-stamped phylogenies ~~at the within-species level using a single locus~~ in BEAST. To ensure
165 convergence of the posterior distributions, three independent MCMC analyses were run (20 million
166 steps, sampled every 1k generation, burn-in 50%). Convergence and effective sample sizes (ESS)
167 were verified using Tracer v.1.7.0 (Rambaut, Drummond, Xie, Baele, & Suchard, 2018). A maximum
168 clade credibility (MCC) tree was summarized using TREEANNOTATOR V.1.10.0 and visualized in
169 FIGTREE V.1.4.3 (Rambaut, 2009). Since the molecular clock is sensitive to bias when a short
170 fragment with fewer polymorphisms is used, the TMRCA analyses were carried out using the two
171 CR sequence datasets of different length (516bp and 828bp) as previously used in the haplotype
172 network analysis. The divergence time of the GWS lineages was estimated using two alternative
173 combinations of calibration priors in BEAST v1.10.0 (Suchard et al., 2018). The alternative
174 combinations of calibration priors were built by integrating the estimated age of the earliest~~the oldest~~
175 GWS ~~fossils~~ fossil records (Applegate & Espinosa-Arrubarrena, 1996~~Gottfried, & Fordyce, 2001~~)
176 (~~Offset~~Mean: 1~~1~~5.0 Ma; SD: 1.0 ~~Ma - Quantiles: 2.5%=15.14; 5%=15.19; median=16.0; 95%=20.18;~~
177 ~~97.5%=22.10~~) with two different secondary calibrations: a) the best dated fossil record of GWS in
178 the Mediterranean Sea, ~~consisting of a GWS-dolphin trophic interaction~~ dated back to the Pliocene
179 (Cigala-Fulgosi, 1990; Bianucci et al., 2002) (Mean: 3.0 Ma; SD: 0.~~15~~30 Ma~~ - Quantiles: 2.5%=2.71;~~
180 ~~5%=2.75; median=3.0; 95%=3.25; 97.5%=3.29~~); b) the molecular divergence previously estimated
181 by Gubili et al. (2010) (Mean: 0.~~4~~5 Ma; SD: 0.15 ~~Ma - Quantiles: 2.5%=0.21; 5%=0.25; median=0.5;~~
182 ~~95%=0.75; 97.5%=0.79~~). ~~B~~In ~~b~~oth combinations of calibration, were implemented as normally
183 distributed ~~priors, the calibrations were implemented as lognormally and normally distributed,~~
184 ~~respectively~~. ~~The calibrations used were chosen to obviate undesirable effect of bias on node~~
185 ~~calibration (Ho, & Phillips, 2009; O'Reilly, & Donoghue, 2016).~~ Since the origination time of the
186 species is a still fiercely debated argument, the divergence time between the GWS and the outgroup

Lamna nasus (Genbank Acc. No. GU266755-GU266769) estimated at around 46 Mya (Martin, 1996) (Mean: 46.0 Ma; SD: 1.0), was used as alternative first calibration, rather than using debated fossil records from the middle Miocene (Gottfried & Fordyce, 2001).

Marginal likelihood estimation and testing divergence time hypotheses

After setting the first calibration for the earliest fossil attributed to the GWS, the two alternative secondary calibrations could be compared by estimating log marginal likelihoods using generalized stepping-stone (GSS) sampling, as implemented in BEAUti v1.10.0 and BEAST v1.10.0 (Suchard et al., 2018; Baele, Lemey, & Suchard, 2015). The log marginal likelihood values for the two different scenarios were first estimated using generalized stepping-stone sampling as 100 stepping stone of 1 million of iterations, logging Log likelihood every 1,000. A total of 101 power posteriors of one million iterations each were sampled using MCMC for the GSS approach. Subsequently, the log Bayes factor was calculated for both scenarios using the formula logBF = logPr (D | M1) - logPr (D | M2), where logPr (D | M1) is the log marginal likelihood for model 1, and logPr (D | M2) is the log marginal likelihood for model 2, in order to select the alternative calibrations that fits best with the principal timing information fixed for both scenarios. This analysis was performed on both sequence datasets (i.e. 516bp and 828bp).

Historical biogeography reconstruction

Two different approaches were implemented in order to reconstruct the historical biogeography of species: the Statistical Dispersal-Vicariance Analysis (S-DIVA), which is a parsimony method of historical biogeography, and the Dispersal-Extinction-Cladogenesis (DEC) analysis (Yu, Harris, & He, 2010; Ree, & Smith, 2008). An MCC estimated by using BEAST and TreeAnnotator on just haplotype sequences, and the specimen distribution through all biogeographical areas (A: Australia/New Zealand, AUS; B: Northeastern Pacific, NEP; C: Mediterranean, MED; D: South-Africa, SA; E: Northwestern Atlantic, NWA) was used to perform the S-DIVA and DEC analyses implemented in RASP v. 4.0 (reconstruct ancestral state in phylogenies) (Yu, Harris, Blair, & He, 2015).

Results

Partial CR sequences (516bp) were obtained from 18 GWS historical specimens (Acc. Num. XXXXXX-XXXXXX; provided after manuscript's acceptance).

The multiple sequence alignment obtained including sequences from all 18 historical samples and four Mediterranean homologous modern sequences deposited in GenBank (HQ540294-HQ540296; JF715925; Table S3 in Appendix S3) showed an extremely low nucleotide diversity for the CR fragment (< 0.1%), with only one variable site at position 244, in which the specimen FICC01LI showed a transition A > G. Once Mediterranean sequences were merged with the homologous sequences from the other global populations (N = 99), an indel at position 60 of the final alignment resulted in a final dataset of 117 CR sequences of 516bp. The global ML haplotype network revealed 31 haplotypes (Figure S4 in Appendix S2) that were clustered in two main haplogroups

differentiated by 28 mutations. As expected, based on the existing literature, the first haplogroup was formed by the individuals from the MED and the Pacific Ocean (AUS and NEP) and the second was composed of GWS from SA and the NWA. Three haplotypes of GWS collected in AUS (HQ414073, HQ414074 and AY026211; Table S3 in Appendix S3) clustered in the latter haplogroup and these individuals are SA−like individuals that likely migrated across the Indian Ocean (Pardini et al., 2001; Blower et al., 2012). The ML haplotype network built with 99 GWS sequences of 828bp revealed 68 haplotypes and a similar topology (Figure S4 in Appendix S2) to the one reconstructed using the shorter sequence dataset, with 2 main haplogroups (AUS-NEP-MED vs SA-NWA).

The cross-plot for the haplotype and nucleotide diversity of the Mediterranean and global populations of GWS revealed that the MED population showed lowest indexes values among the other global populations, with AUS showing the highest values for both indexes shown (Figure 1; Table S6~~5~~ in Appendix S3). The NEP and NWA populations showed high haplotype diversity and low nucleotide diversity while the SA population exhibited quite opposite positions in the plot depending on the reference study with high values for both estimates in Pardini et al. (2001) and O'Leary et al. (2015), and low values in Andreotti et al. (2016).

No polymorphisms were detected among the four complete contemporary CR sequences of GWS, demonstrating a low haplotype diversity among longer sequences as well, while the genetic diversity (e.g. nucleotide) is proportional to the length of the sequences.

Marginal likelihood estimation and test for divergence time hypotheses

The application of the log Bayes factor formula gave ~~decisive~~ significant support to the Pliocene calibration scenario ~~using the 828bp dataset and non-significant support using the 516bp dataset~~in every model tested (Table 1)~~, suggesting the lack of signal in the short sequence dataset~~. Based on the 828bp dataset, the GWS Mediterranean population diverged from the Pacific populations at 3.~~26~~ 23 Ma, a time that is congruent with the estimated closure of the Central American Seaway, CAS (□3.5 Ma)~~,~~ after the formation of the Isthmus of Panama (O'Dea et al., 2016). Interesting, we got a small mismatch between the "Pleistocene" calibration priors sampling and the priors posteriors, suggesting statistically that a Pleistocene second calibration is conflictual with the data, and a "Pliocenic divergence" scenario is preferable.

Time of the Most Recent Common Ancestor (TMRCA) and estimation of evolutionary rate

The analysis of fossil evidence suggested that GWS experienced a long evolutionary history ~~dating back to the middle Miocene~~ (Table S4;S5~~4~~ of Appendix S3). The occurrence of GWS fossils in the Mediterranean area ~~dated back to the Late Miocene and~~is rich in the Pliocene, after the Messinian Salinity Crisis.

~~The molecular clock test showed that the null hypothesis of equal evolutionary rate throughout the tree was not rejected (p = 0.14; Table S6 of Appendix S3). A saturation effect of transitions (Ts) and transversions (Tv) was not detected (Ts: r² = 0.99, P = < 0.0001; Tv: r² = 0.91, p < 0.0001).~~

The CR relaxed substitution rate range inferred from the TMRCA analysis was estimated at 0.38%-0.72%~~0.23%~~ substitutions/site/my ~~(divergence rate of 0.46%) and 0.33% (divergence rate of~~

264 0.66%) for the 3 Ma and 0.5 Ma calibrations, respectively for the best model following the Bayes

265 factor. Based on these estimates, the coalescence of the Mediterranean and Pacific GWS lineages

266 dated the divergence time of the Mediterranean population at 3.26 23 Ma (Figure 2) and 01.72 81 Ma

267 (Figure S5 of Appendix S2), respectively.

Historical biogeography reconstruction

269   The analysis of the ancestral ranges of GWS, using both S-DIVA and DEC approaches,

270 suggests that the Mediterranean population is the result of a potential dispersal-vicariance scenario.

271 The biogeographical reconstruction of the Mediterranean lineage estimated using S-DIVA resulted

272 in two principal ancestral ranges, AC (AUS-MED) and ABC (AUS-NEP-MED) with probabilities

273 (relative frequencies) for each range of 52.44% and 46.69% respectively. Two minor ancestral ranges

274 were detected with very low probability: AB (AUS-NEP) and BC (NEP-MED), with probabilities of

275 0.52% and 0.35% respectively. The DEC analysis resulted in concordant results with two ancestral

276 ranges, AC (AUS-MED); BC (NEP-MED), with probabilities of 50.01% and 49.99% respectively,

277 supporting that the origin of the Mediterranean population is likely among GWS of Australia and the

278 North Eastern Pacific. Both approaches suggest a route of dispersal-vicariance of route

279 AC>CAB>C|AB scenario. Interesting, for both S-DIVA and DEC analyses, Australian and North

280 Eastern Pacific lineages originated from a Pacific ancestral range AB (AUS-NEP) with a probability

281 of 100.00% through a vicariance phenomenon.

282

Discussion

284   Our phylogenetic analyses of contemporary and historical sequences indicated that the existing

285 population of GWS in the Mediterranean could be a relic of the Pacific GWS clade. This scenario

286 would suggest that Pacific-born ancestors migrated eastward across the ocean, before the complete

287 closure of the Central American Seaway, and colonized the North Atlantic Ocean before entering the

288 Mediterranean after the Messinian salinity crisis, which may have caused the local extinction of the

289 marine megafauna due to the total or partial drying up of the Mediterranean with related subsequent

290 ecological upheavals (Garcia-Castellanos et al., 2009). This evolutionary pathway (Figure 3) is

291 supported by Bayesian analyses of genetic diversity and divergence time estimates and it is coherent

292 with the age of the main paleo-geographical events, paleo-climatic patterns and fossil records. Our

293 proposed pathway represents an alternative hypothesis to that of an accidental antipodean long-

294 distance dispersal of a few Australian/New Zealand founder females that visited the Good Hope Cape

295 area and, because of Pleistocene climatic oscillations, navigated erroneously northward and entered

296 into the Mediterranean Sea (henceforth named "the antipodean dispersal hypothesis"; Gubili et al.,

297 2010).

298   The phylogenetic analyses of the CR sequences for the 18 Mediterranean GWS confirmed once

299 again a closer evolutionary relationship of this population with the Australian/New Zealand and

300 Northeastern Pacific lineage than with the South African and Northwestern Atlantic lineage (Pardini

301 et al., 2001; Jorgensen et al., 2010; Gubili et al., 2010; Blower et al., 2012). The Mediterranean GWS

302 exhibited a very shallow mtDNA genetic variation with only two haplotypes and extremely low

303 genetic diversity. Among the modern samples collected from throughout the world that have been
304 analyzed to date, the Mediterranean GWS have provided the lowest estimates of genetic diversity,
305 indicating that this population is relatively young and has likely originated by a founder event by a
306 single or a few mtDNA lineages or that the population has experienced a recent population bottleneck
307 (Group 1 of Grant, & Bowen (1998); Grant, & Waples, 2000). Within the Pacific/Mediterranean
308 mitochondrial clade, the Australian/New Zealand GWS population was identified as the most
309 ancestral with long evolutionary/demographic history and divergent haplotypes that have
310 accumulated over long periods of time (as indicated by the high values of haplotype and nucleotide
311 diversity; Group 3 of Grant, & Bowen (1998); Grant, & Waples, 2000). An intermediate evolutionary
312 position in the clade can be that of the Northeastern Pacific GWS population that possessed high
313 haplotype diversity and low nucleotide diversity which indicated a pattern typically associated with
314 population that have experienced a bottleneck event followed by rapid demographic growth and
315 accumulation of mutations (Group 2 of Grant, & Bowen (1998); Grant, & Waples, 2000).

316   The evolutionary pattern revealed by the plot analysis of haplotype and nucleotide diversity is
317 corroborated by the historical biogeography of the species, that suggests the closest relationship
318 between Mediterranean and both Pacific lineages as the best evolutionary scenario.

319   From those ancestral lineages from which the Mediterranean one originated, the Northeastern
320 Pacific GWS appears to have been an important contributor to the genetic composition of the
321 Mediterranean population. Furthermore, since both S-DIVA and DEC analyses suggest that the
322 Mediterranean GWS originated from a dispersal-vicariance event, an origin via westward dispersal
323 from a few Australian individuals during the Pleistocene. The founder event that gave origin to the
324 modern Mediterranean GWS was dated by the TMRCA analysis with high statistical support, to 3.26
325 23 Ma, during the late Pliocene, in the Piacenzian (3.60 – 2.58 Ma) and immediately after the
326 Zanclean (5.33 – 3.60 Ma). This estimate is greatly older than that reconstructed in the antipodean
327 dispersal hypothesis (348-565 ka; Calabrian, Pleistocene) by Gubili et al. (2010) based on an estimate
328 of mtDNA substitution rate of 1.19-0.74% of divergence between lineages per million year. These
329 rates of evolutionary change were calibrated by Gubili et al. (2010) using estimates of vicariance
330 events that separated GWS populations from the Northeastern Pacific and Northwestern Atlantic
331 oceans (i.e. the rising of the Isthmus of Panama dated at 3.5 Ma; O'Dea et al., 2016) and from the
332 Western Pacific and Indian oceans (i.e. the rising of the Sunda-Sahul shelves dated at 5 Ma; Haq,
333 Hardenbol, & Vail, 1987). However, all phylogenetic analyses carried out so far (Gubili et al., 2010;
334 Andreotti et al., 2016; present work) have revealed that the Northwestern Atlantic GWS are
335 phylogenetically linked to the South African population but not to the Northeastern Pacific GWS.
336 Therefore, the use of the vicariance event separating Atlantic and Pacific GWS at 3.5 Ma from a
337 common recent ancestor might have led to an overestimation of the mutation rate and the time of
338 divergence between Mediterranean and Pacific GWS at 348-565 ka in the Pleistocene. On the
339 contrary, all phylogenetic analyses consistently suggested that the Mediterranean and the Western
340 Atlantic populations belong to two different lineages, wherein the Mediterranean GWS are
341 descendants of populations found in the Australian/New Zealand and Northeastern Pacific, and

342 modern Northwestern Atlantic elements likely descended from South African GWS by a founder

343 event (Andreotti et al., 2016).

344   Abundant fossil data suggests that GWS have inhabited the Mediterranean Sea since the early

345 Pliocene, following the Messinian Salinity Crisis, with ~~origin of~~ numerous specimens estimated to be

346 between ~5 and ~2 Ma old, with peaks of abundance occurring during the Pliocene (Cigala-Fulgosi,

347 1990; Bianucci et al., 2002; Marsili, 2008; Adnet, Balbino, Antunes, & Marín-Ferrer, 2009). After

348 these paleoclimatic phases, GWS could have colonized the Mediterranean Sea occupying the

349 ecological niches ~~gradually~~ left empty by other apex marine predators, such as the giant megatooth

350 shark *Carcharocles megalodon*, which went extinct between 3.5 and 2.6 Ma (Pimiento, & Clements,

351 2014). Fossil evidence from Central America suggests that many species of sharks and marine

352 mammals that were part of the region's faunal assemblage may have migrated across the CAS

353 continuously before the formation of the Isthmus of Panama (Steeman et al., 2009; Pimiento et al.,

354 2013; Velez-Juarbe, Wood, De Gracia, & Hendy, 2015). Upon the closure of the CAS, the newly

355 formed Gulf Stream current could have facilitated an eastward trans-Atlantic migration by GWS

356 during the late Pliocene by way of the eastward dispersion of nutrients and, consequently, food

357 resources. There is evidence that intense fluctuations in the speed of the Gulf Stream occurred during

358 the formation of the Isthmus of Panama, reaching a height during the late Miocene and early Pliocene

359 (~6.1-4.8 Ma; Kaneps, 1979). The intensification of currents coincides with the end of the Messinian

360 Salinity Crisis (5.33 Ma), which culminated with the Zanclean inflow in the Mediterranean Sea

361 (Garcia-Castellanos et al., 2009), and it may have been the driver of eastward migrations by various

362 species of fish, and consequently, of their predators. Therefore, all these oceanographic phenomena

363 are concordant with the estimated divergence of the Mediterranean GWS population that we have

364 obtained ~~(2.93 Ma)~~.

365   A potential earlier formation of the Isthmus of Panama (Bacon et al., 2015; Lessions et al.,

366 2015; Montes et al., 2015; O'Dea et al., 2016) would not affect our results, since it is plausible that

367 the Pacific lineage that gave rise to the MED population may already have been present in the

368 Atlantic before and after the closure of the CAS and before its potential local extinction in the

369 Atlantic. Also, past intermittent shallow connections are plausible and discussed in literature

370 (Jaramillo et al., 2017).

371   The global phylogeography of contemporary GWS populations reveals a discontinuous

372 distribution of the Pacific/Mediterranean GWS evolutionary lineage in the North Atlantic being

373 interrupted by the recent penetration and colonization of the Northwestern Atlantic area by the Indian

374 lineage. The phylogenetic analyses highlighted that the GWS of Northwestern Atlantic population

375 does not exhibit a separated mtDNA cluster like the Pacific and Mediterranean populations, and it

376 form a unique lineage with the South-Africa (Figure S4 of Appendix S2). This evidence is suggestive

377 of a recent evolutionary history or a population bottleneck followed by population growth and

378 accumulation of mutations.

379   A local extinction of the past Pacific GWS lineage in the North Atlantic Ocean and further

380 replacement by Indian components could explain such a phylogeographic discontinuity. An extensive

381 analysis of the fossil records, highlighted an exceptional extinction rate of the marine fauna, including

sharks, during the Pliocene epoch (Pimiento et al., 2017). Furthermore, the highest extinction rates occurred in the late Pliocene, between 3.8 and 2.4 Ma, which coincides with the Mediterranean population divergence time, and may have led the local extinction of the ancient Atlantic population of GWS.

Additional genetic and tagging studies are required to determine if the Mediterranean GWS are ecologically and reproductively isolated from the adjacent populations occupying the North Atlantic Ocean. This issue is of high-priority for the development of robust conservation criteria and implementation of management strategies. In spite of this data poor situation, some information can be gleaned from the other populations in the Atlantic as well as other species. For example, several other epipelagic sharks migrate throughout the temperate waters of the Atlantic and several are suspected of having nursery areas in the mid-Atlantic Ridge (Kohler, Turner, Hoey, Natanson, & Briggs, 2002; Stevens, 2010; Vandeperre et al., 2014; O'Leary et al., 2015). Elsewhere, in the Indian Ocean GWS have been observed migrating between Australia and South Africa, while in the Pacific, GWS frequently migrate between the western coast of North America and Hawaii (Bonfil et al., 2005; Jorgensen et al., 2010, 2012; Blower et al., 2012). Clearly, the species is not averse to far reaching longitudinal movements. However, antipodean connections appear less frequent and have only been observed in the variation of the DNA sequences (O'Leary et al., 2015). Latitudinal range of GWS could be limited by thermal tolerance (Cliff, Dudley, & Davis, 1989), prey availability, social structure and fidelity to nursery areas.

Our results suggested that the Mediterranean GWS have a more ancient origins than previously thought, and that this population is genetically disconnected from the adjacent Atlantic population. Due to historical and widespread declines of sharks in the Mediterranean Sea (Coll, Palomera, Tudela, & Dowd, 2008; Lotze, Coll, & Dunne, 2011; Colloca et al., 2013; Ferretti, Osio, Jenkins, Rosenberg, & Lotze, 2013) as well as the ongoing over-exploitation of marine resources in the region, the current white shark population is at minimal levels of abundance and consequently may play a less prominent ecological role than in historical times. Yet this distinct genetic pool represents a valuable, albeit extremely precarious, component of the genetic diversity of a species that is endangered worldwide. Losing the Mediterranean GWS population would represent a hard blow to the global white shark. More effort for characterizing the ecology, biology and population structure of this critically endangered population (i.e. through tagging, genetic analyses and historical analyses of occurrence records) are now paramount to plan conservation and management actions to restore this important oceanic top-predator in the region.

1
2
3 422
4
5 423　**Figures and Tables**
6 424　**Figure 1.** Plot of the haplotype and nucleotide diversity values (expressed as percentage values)
7
8 425　estimated by the global populations of *Carcharadon charcarias* inferred using the 828bp dataset of
9 426　control region sequences. AUS: Australia/New Zealand; NEP: Northeastern Pacific; MED:
10
11 427　Mediterranean; SA: South-Africa; NWA: Northwestern Atlantic. The asterisk marks the values
12 428　reported by Andreotti et al. [29] for the South African population.
13
14 429

15
16
17
18
19
20
21
22
23
24
25
26
27
28

29
30
31
32 430
33
34 431
35 432
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
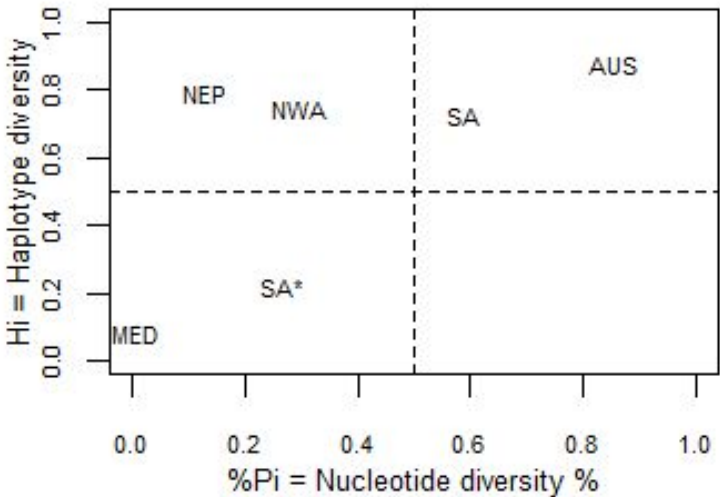
433 **Figure 2.** Bayesian divergence time tree of populations of *Carcharadon charcarias* inferred using

434 the 828bp dataset of control region sequences. High posterior density (HPD 95%) values are featured

435 as blue bars. Only information of node with at least 0.5 posterior value are shown. Abbreviations are

436 given in Figure 1.

437

438

439

Figure 3. Global dispersal and Pacific/Mediterranean vicariance hypothesis for *Carcharadon charcarias*. Ancient great White Sharks from the Pacific Ocean, namely Australia, dispersed via two routes: westward to South African coasts (1, light grey dotted line) and eastward to Northeastern Pacific Americas (2, black dotted line). The Pacific Great White Sharks were free to move eastward to the Atlantic, and in the Pliocene to an ancient Mediterranean Sea after the Messinian Salinity Crisis, that could have caused the local extinction of a great white shark population established in the middle Miocene (3-4). Past climatic oscillation due to the closure of the Central America Seaway and the formation of the Isthmus of Panama, could have caused a local extinction or an eastward mass migration of white shark from the North Atlantic Ocean, isolating the Mediterranean population from other ancestral populations. The North Atlantic was colonized, then, in relatively recent history, when the climate conditions were ideal to maintain a great white shark population, once again (5). Fossil records extrapolated from Paleobiology Database are shown by colored points (red: Miocene, blue: Pliocene, green: Pleistocene). The map was created using 'paleobioDB' package in R version 3.5.1 (Varela et al., 2015). Abbreviations are given in Figure 1.
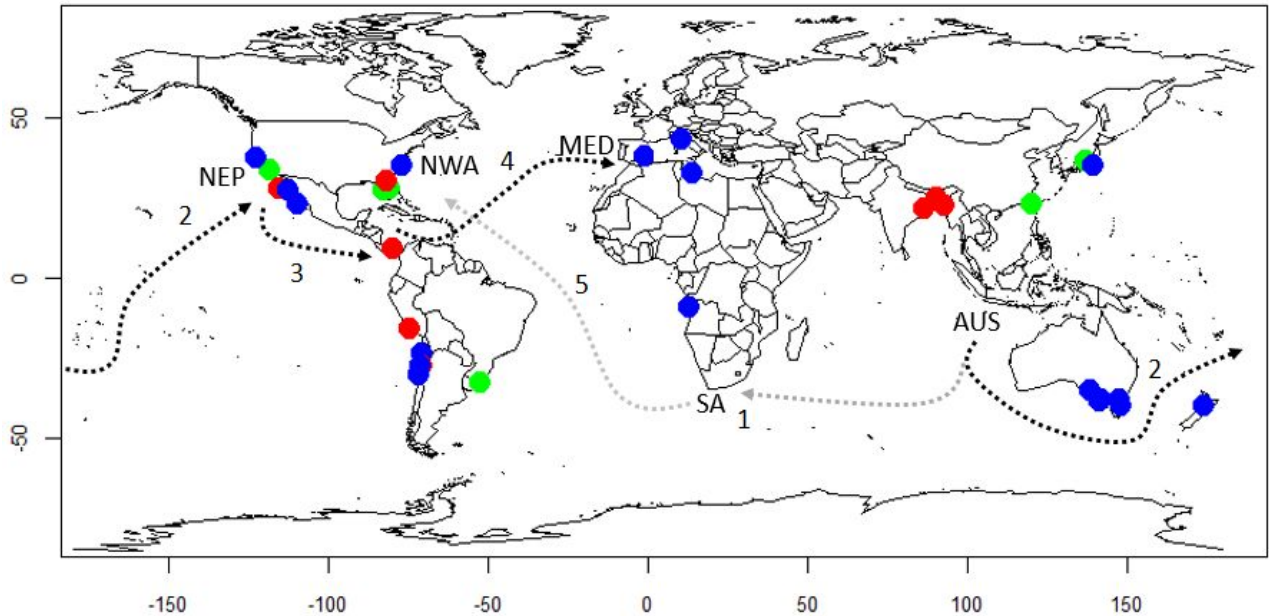
455
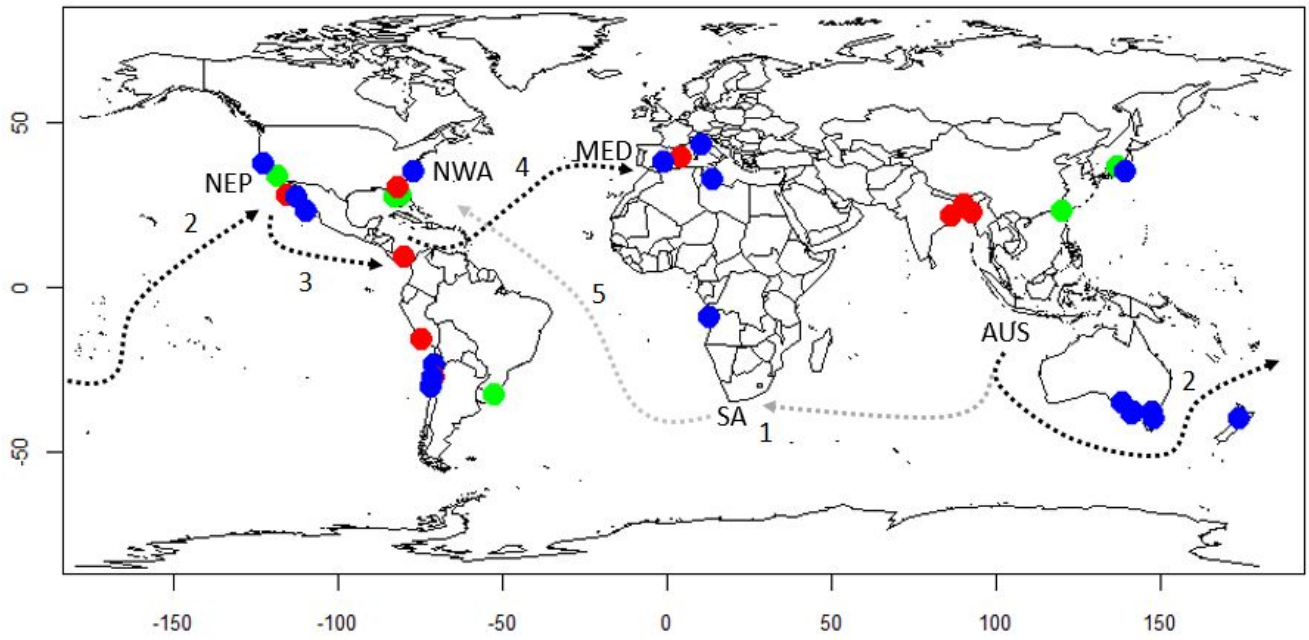
Table 1. Best model selection based on Bayes factors of tree topologies reconstructed with <u>the earliest fossil occurences and the estimated divergence between the GWS and the outgroup L.nasus (</u>the 828bp and 516bp sequence datasets, respectively. Node constraints and calibration parameters on the phylogeographical tree of *Carcharodon carcharias* for the two divergence models are also provided. logML_GSS: log marginal likelihood from generalized stepping stone model, BF_GSS: Bayes factors calculated using the logML_GSS. Models are ranked according to the logML values.

Model	Dataset	Node1	Node2	logML_GSS	logBF_GSS
MED Pliocene divergence	828bp	Ingroup	MED/Pacific		
		Mean: 11 Ma, SD: 1.0 Ma	Mean: 3.0 Ma, SD: 0.30 Ma	-2455.2796	0 (Best)
MED Pleistocene divergence	828bp	Ingroup;	MED/Pacific		
		Mean: 11 Ma, SD: 1.0 Ma	Mean: 0.4Ma, SD: 0.15 Ma	-2458.1264	2.8468
MED Pleistocene divergence	516bp	Ingroup;	MED/Pacific		
		Offset: 11 Ma, SD: 1.0 Ma	Mean: 3.0 Ma, SD: 0.30 Ma	-1440.6171	0
MED Pliocene divergence	516bp	Ingroup;	MED/Pacific		
		Offset: 11 Ma, SD: 1.0 Ma	Mean: 0.4Ma, SD: 0.15 Ma	-1443.0069	2.3898

References

Adnet, S., Balbino, A. C., Antunes, M. T., & Marín-Ferrer, J. M. (2009). New fossil teeth of the white shark (*Carcharodon carcharias*) from the early Pliocene of Spain. Implication for its paleoecology in the Mediterranean. *Neues Jahrbuch für Geologie und Paläontologie – Abhandlungen,* 256(1), 7-16. https://doi:10.1127/0077-7749/2009/0029

Ahonen, H., & Stow, J. (2008). Shark jaws and teeth: an unexploited resource for population genetic studies. *Journal of Fish Biology*, 73(2), 450-455. https://doi:10.1111/j.1095-8649.2008.01896.x

Andreotti, S., von der Heyden, S., Henriques, R., Rutzen, M., Meÿer, M., Oosthuizen, H., & Matthee, C. A. (2016) New insights into the evolutionary history of white sharks, *Carcharodon carcharias*. *Journal of Biogeography*, 43(2), 328-339. https://doi.org/10.1111/jbi.12641

Applegate, S. P., & Espinosa-Arrubarrena, L. (1996). The fossil history of *Carcharodon* and its possible ancestor, *Cretolamna*: a study in tooth identification. In A. P. Klimley & D. G. Ainley (Eds.), *Great white sharks: the biology of Carcharodon carcharias* (pp. 19-36). San Diego, CA: Academic Press.

Bacon, C. D., Silvestro, D., Jaramillo, C., Smith, B. T., Chakrabarty, P., Antonelli, A. (2015). Biological evidence supports an early and complex emergence of the Isthmus of Panama. *Proceedings of the National Academy of Sciences U.S.A*, 112(19), 6110-6115. doi: https://doi.org/10.1073/pnas.1423853112

Baele, G., Lemey, P., & Suchard, M. A. (2015). Genealogical working distributions for Bayesian model testing with phylogenetic uncertainty. *Systematic Biology*, 65(2), 250-264. https://doi.org/10.1093/sysbio/syv083

Bianucci, G., Bisconti, M., Landini, W., Storai, T., Zuffa, M., Giuliani, S., & Mojetta, A. (2002). Trophic interaction between white shark, *Carcharodon carcharias*, and cetaceans: a comparison between Pliocene and recent data from Central Mediterranean Sea. In M. Vacchi, G. Lamesa, F.

496   Serena, & B. Seret (Eds.), *Proceedings of the 4th European Elasmobranch Association Meeting 28-*
497   *30 November 2000* (pp. 33-48). Livorno, Italy: ICRAM, ARPAT & SFI.

498   Blower, D. C., Pandolfi, J. M., Bruce, B. D., Gomez-Cabrera, M.D., & Ovenden, J. R. (2012)
499   Population genetics of Australian white sharks reveals fine-scale spatial structure, transoceanic
500   dispersal events and low effective population sizes. *Marine Ecology Progress Series*, 455, 229-244.
501   https://doi: 10.3354/meps09659

502   Bonfil, R., Meÿer, M., Scholl, M. C., Johnson, R., O'Brien, S., Oosthuizen, H., … Paterson, M.
503   (2005). Transoceanic migration, spatial dynamics, and population linkages of white sharks. *Science*,
504   310(5745), 100-103. https://doi:10.1126/science.1114898

505   Canese, S. P., Giusti, M., Salvati, E., Angiolillo, M., Cardinali, A., & Consalvo, I. (2011). *Final*
506   *Report Project ISPRA PR0.3.452 Biodiversity of Sicily Channel*. Rome, Italy: Ministry of the
507   Environment, Territorial and Sea Protection.

508   Cigala-Fulgosi, F. (1990). Predation (or possible scavenging) by a great white shark on an extinct
509   species of bottlenosed dolphin in the Italian Pliocene. *Tertiary Research*, 12, 17-36.

510   Cliff, G., Dudley, S. F. J., & Davis, B. (1989). Sharks caught in the protective gill nets off Natal,
511   South Africa. 2. The great white shark *Carcharodon carcharias* (Linnaeus). *South African Journal*
512   *of Marine Science,* 8, 131-144. https://doi:10.2989/02577618909504556
513

514   Coll, M., Palomera, I., Tudela, S., & Dowd, M. (2008). Food-web dynamics in the South Catalan Sea
515   ecosystem (NW Mediterranean) for 1978–2003. *Ecological Modelling,* 217(1-2), 95-116.
516   https://doi:10.1016/j.ecolmodel.2008.06.013
517

518   Colloca, F., Cardinale, M., Maynou, F., Giannoulaki, M., Scarcella, G., Jenko, K., … Fiorentino, F.
519   (2013). Rebuilding Mediterranean fisheries: a new paradigm for ecological sustainability. *Fish and*
520   *Fisheries*, 14(1), 89-109. https://doi:10.1111/j.1467-2979.2011.00453.x)

521   Compagno, L. J. V. (1984). FAO species catalogue. Vol. 4. Sharks of the world. An annotated and
522   illustrated catalogue of shark species known to date. Part 1. Hexanchiformes to Lamniformes. *FAO*
523   *Fish Synopsis*, 125, 1-249.

524   De Maddalena, A. 2006 A catalogue of great white sharks *Carcharodon carcharias* (Linnaeus, 1758)
525   preserved in European museums. *Journal of the National Museum, Natural History Series*, 175(3-4),
526   109-125.

527   Domeier, M. L., Nasby-Lucas, N. (2008). Migration patterns of white sharks Carcharodon carcharias
528   tagged at Guadalupe Island, Mexico, and identification of an eastern Pacific shared offshore foraging
529   area. *Marine Ecology Progress Series*, 370, 221–237. doi: https://doi.org/10.3354/meps07628

530 Feretti, F., Worm, B., Britten, G. L., Heithaus, M. R., & Lotze, H. K. (2010). Patterns and ecosystem
531 consequences of shark declines in the ocean. *Ecology Letters*, 13(8), 1055-1071.
532 https://doi:10.1111/j.1461-0248.2010.01489.x

533 Fergusson, I. K. (1996) Distribution and autoecology of the white shark in the Eastern North Atlantic
534 and the Mediterranean Sea. In A. P. Klimley & D. G. Ainley (Eds.), *Great white sharks: the biology*
535 *of Carcharodon carcharias* (pp. 321-345). San Diego, CA: Academic Press.

536 Ferretti, F., Myers, R. A., Serena, F., & Lotze, H. K. (2008). Loss of large predatory sharks from the
537 Mediterranean Sea. *Conservation Biology*, 22(4), 952-964. https://doi:10.1111/j.1523
538 1739.2008.00938.x

539 Ferretti, F., Osio, G. C., Jenkins, C. J., Rosenberg, A. A., & Lotze, H. K. (2013) Long-term change
540 in a meso-predator community in response to prolonged and heterogeneous human impact. *Scientific*
541 *Reports*, 3, 1057. https://doi:10.1038/srep01057

542 Fulton, T. L., & Strobeck, C. (2010). Multiple fossil calibrations, nuclear loci and mitochondrial
543 genomes provide new insight into biogeography and divergence timing for true seals (Phocidae,
544 Pinnipedia). *Journal of Biogeography*, 37(5), 814-829. https://doi: 10.1111/j.1365-
545 2699.2010.02271.x

546 Garcia-Castellanos, D., Estrada, F., Jiménez-Munt, I., Gorini, C., Fernandéz, M., Vergés, J., & De
547 Vicente, R. (2009). Catastrophic flood of the Mediterranean after the Messinian salinity crisis.
548 *Nature*, 462, 778-781. https://doi:10.1038/nature08555

549 Gottfried, M. D., & Fordyce, R. E. (2001). An associated specimen of *Carcharodon angustidens*
550 (Chondrichthyes, Lamnidae) from the late Oligocene of New Zealand, with comments on
551 *Carcharodon* interrelationships. *Journal of Vertebrate Paleontology*, 21(4), 730-739.
552 https://doi:10.1671/0272-4634(2001)021[0730:AASOCA]2.0.CO;2

553 Grant, W. S., & Bowen, B. W. (1998). Shallow population histories in deep evolutionary lineages of
554 marine fishes: insights from sardines and anchovies and lessons for conservation. *Journal of Heredity*,
555 89(5), 415–426. https://doi.org/10.1093/jhered/89.5.415

556

557 Grant, W. S., & Waples, R. S. (2000). Spatial and temporal scales of genetic variability in marine and
558 anadromous species: implications for fisheries oceanography. In P. J. Harrison & T. R. Parsons (Eds.),
559 *Fisheries oceanography: an integrative approach to fisheries ecology and management* (pp. 63–93).
560 Oxford, UK: Blackwell Science.

561 Gubili, C., Bilgin, R., Kalkan, E., Karhan, S. Ü., Jones, C. S., Sims, D. W., … Noble, L. R. (2010).
562 Antipodean white sharks on a Mediterranean walkabout? Historical dispersal leads to genetic

563 discontinuity and an endangered anomalous population. *Proceedings of the Royal Society B:*
564 *Biological Sciences,* 278(1712), 1679–1686. https://doi:10.1098/rspb.2010.1856

565 Gubili, G., Robinson, C. E. C., Cliff, G., Wintner, S. P., de Sabata, E., De Innocentiis, S., … Jones,
566 C. S. (2015). DNA from historical and trophy samples provides insights into white shark population
567 origins and genetic diversity. *Endangered Species Research* 27(3), 233-241. https://doi:
568 10.3354/esr00665

569 Haq, B. U., Hardenbol, J., & Vail, P. R. (1987). Chronology of fluctuating sea levels since the
570 Triassic. *Science*, 235(4793), 1156-1167. https://doi:10.1126/science.235.4793.1156

571 Hasegawa, M., Kishino, H., & Yano, T. (1985). Dating of the human-ape splitting by a molecular
572 clock of mitochondrial DNA. *Journal of Molecular Evolution*, 22(2), 160-174.

573 Hauser, L., Adcock, G. J., Smith, P. J., Ramirez, J. H. B., & Carvalho, G. R. (2002). Loss of
574 microsatellite diversity and low effective population size in an overexploited population of New
575 Zealand snapper (*Pagrus auratus*). *Proceedings of the National Academy of Sciences of the United*
576 *States of America,* 99(18), 11742–11747. https://doi:10.1073/pnas.172242899

577 Ho, S. Y. W., & Phillips, M. J. (2009). Accounting for calibration uncertainty in phylogenetic
578 estimation of evolutionary divergence times. *Systematic Biology*, 58(3), 367-380.
579 https://doi.org/10.1093/sysbio/syp035

580 Hofreiter, M., Serre, D., Poinar, N. H., Kuch, M., & Pääbo, S. (2001). Ancient DNA. *Nature Reviews*
581 *Genetics*, 2, 353-359. https://doi:10.1038/35072071

582 Hutchinson, W. F., van Oosterhout, C., Rogers, S. I., & Carvalho, G. R. (2003). Temporal analysis
583 of archived samples indicates marked genetic changes in declining North Sea cod (*Gadus morhua*).
584 *Proceedings of the Royal Society B: Biological Sciences,* 270(1529), 2125-2132.
585 https://doi:10.1098/rspb.2003.2493

586 Jaramillo, C., Montes, C., Cardona, A., Silvestro, D., Antonelli, A., Bacon, C. D. (2017). Comment
587 (1) on "Formation of the Isthmus of Panama" by O'Dea et al. *Science Advances*, 3(6), e1602321. doi:
588 10.1126/sciadv.1602321

589 Jorgensen, S. J., Arnoldi, N. S., Estess, E. E., Chapple, T. K., Rückert, M., Anderson, S. D., & Block,
590 B. A. (2012). Eating or meeting? Cluster analysis reveals intricacies of white shark (*Carcharodon*
591 *carcharias*) migration and offshore behavior. *PLOS ONE* 7, e47819.
592 https://doi:10.1371/journal.pone.0047819

593 Jorgensen, S. J., Reeb, C. A., Chapple, T. K., Anderson, S., Perle, C., Van Sommeran, S.R., … Block,
594 A. B. (2010). Philopatry and migration of Pacific white sharks. *Proceedings of the Royal Society B:*
595 *Biological Sciences,* 277(1682), 679-688. https://doi:10.1098/rspb.2009.1155

596 Kabasakal, H., & Gedikoğlu, Ö. S. (2008). Two new-born great white sharks, *Carcharodon*
597 *carcharias* (Linnaeus, 1758) (Lamniformes; Lamnidae) from Turkish waters of the north Aegean Sea.
598 *Acta Adriatica*, 49, 125–135.

599 Kaneps, A. G. (1979). Gulf stream: velocity fluctuations during the late Cenozoic. *Science*,
600 204(4390), 297-301. https://doi:10.1126/science.204.4390.297

601 Kohler, N. E., Turner, P. A., Hoey, J. J., Natanson, L. J., & Briggs, R. (2002). Tag and recapture data
602 for three pelagic shark species: blue shark (*Prionace glauca*), shortfin mako (*Isurus oxyrinchus*), and
603 porbeagle shark (*Lamna nasus*) in the North Atlantic Ocean. *ICCAT Collective Volumes of Scientific*
604 *Papers,* 54, 1231-1260.

605 Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7 Molecular Evolutionary Genetics Analysis
606 Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, 33(7), 1870-1874.
607 https://doi.org/10.1093/molbev/msw054

608 Lessios, H. A. (2015). Appearance of an early closure of the Isthmus of Panama is the product of
609 biased inclusion of data in the metaanalysis. *Proceedings of the National Academy of Sciences USA,*
610 112(43), E5765-E5765. doi: 10.1073/pnas.1514719112

611 Lotze, H. K., Coll, M., & Dunne, J. A. (2011). Historical changes in marine resources, food-web
612 structure and ecosystem functioning in the Adriatic Sea, Mediterranean. *Ecosystems*, 14(2), 198-222.
613 https://doi: 10.1007/s10021-010-9404-8

614 Mancusi, C., Nicolosi, P., Arculeo, M., Barbagli, F., Carlini, R., Costantini, M., … Vacchi, M. (2002).
615 The presence of elasmobranchs in the collections of the main Italian natural history museums. In M.
616 Vacchi, G. Lamesa, F. Serena, & B. Seret (Eds.), *Proceedings of the 4th European Elasmobranch*
617 *Association Meeting 28-30 November 2000* (pp. 97-108). Livorno, Italy: ICRAM, ARPAT & SFI.

618 Marsili, S. (2008). Systematic, paleoecologic and paleobiogeographic analysis of the Plio-Pleistocene
619 Mediterranean elasmobranch fauna. *Atti della Società Toscana di Scienze Naturali, Memorie, Serie*
620 *A*, 113, 81-88.

621 Marsili, S. (2006). Analisi Sistematica, Paleoecologica e Paleobiogeografica della Selaciofauna plio-
622 pleistocenica del Mediterraneo. University of Pisa. PhD thesis

623 Martin, A. P. (1996). Systematics of the Lamnidae and the Origination Time of Carcharodon
624 carcharias Inferred from the Comparative Analysis of Mitochondrial DNA Sequences. In A. P.
625 Klimley & D. G. Ainley (Eds.), *Great white sharks: the biology of Carcharodon carcharias* (pp. 49-
626 53). San Diego, CA: Academic Press.

627 McPherson, J. M., & Myers, R. A. (2009). How to infer population trends in sparse data: examples
628 with opportunistic sighting records for great white sharks. *Diversity and Distributions* 15(5), 880-
629 890. https://doi: 10.1111/j.1472-4642.2009.00596.x

Montes, C., Cardona, A., Jaramillo, C., Pardo, A., Silva, J. C., Valencia, … Niño, H. (2015). Middle Miocene closure of the Central American Seaway. *Science* 348(6231), 226–229. doi: 10.1126/science.aaa2815

Myers, R. A., Baum, J. K., Shepherd, T. D., Powers, S. P., Peterson, C. H. (2007). Cascading Effects of the Loss of Apex Predatory Sharks from a Coastal Ocean. *Science*, 315(5820), 1846-1850. doi: 10.1126/science.1138657

Nieto, A., Ralph, G. M., Comeros-Raynal, M. T., Kemp, J., García Criado, M., Allen, D. J., … Williams, J. T. (2015) *European Red List of marine fishes*. Luxembourg, EU: Rosseels Printing.

O'Dea, A., Lessios, H. A., Coates, A. G., Eytan, R. I., Restrepo-Moreno, S. A., Cione, A. L., … Jackson, J. B. C. (2016). Formation of the Isthmus of Panama. *Science Advances*, 2(8), e1600883. https://doi:10.1126/sciadv.1600883

O'Leary, S. J., Feldheim, K. A., Fields, A. T., Natanson, L. J., Wintner, S., Hussey, N., … Chapman, D. D. (2015). Genetic diversity of white sharks, *Carcharodon carcharias,* in the Northwest Atlantic and Southern Africa. *Journal of Heredity*, 106(3), 258-265. https://doi:10.1093/jhered/esv001

O'Reilly, J. E., & Donoghue, P. C. J. (2016). Tips and nodes are complementary not competing approaches to the calibration of molecular clocks. *Biology Letters*, 12(4), 20150975. https://doi.org/10.1098/rsbl.2015.0975

Paradis, E., Claude, J., & Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20(2), 289-290. https://doi.org/10.1093/bioinformatics/btg412

Pardini, A. T., Jones, C. S., Noble, L. R., Kreiser, B., Malcolm, H., Bruce, B. D., … Martin, A. P. (2001). Sex-biased dispersal of great white sharks. *Nature*, 412(6843), 139-140. http://doi:10.1038/35084125

Pimiento, C., & Clements, C. F. (2014). When did *Carcharocles megalodon* become extinct? A new analysis of the fossil record. *PLoS ONE*, 9(10), e111086. https://doi.org/10.1371/journal.pone.0111086

Pimiento, C., Gonzalez-Barbad, G., Hendy, A. J. W., Jaramillo, C., MacFadden, B. J., Montes, C., … Shippritt, M. (2013). Early Miocene chondrichthyans from the Culebra formation, Panama: a window into marine vertebrate faunas before closure the Central American Seaway. *Journal of South American Earth Sciences*, 42, 159-170. https://doi:10.1016/j.jsames.2012.11.005

Pimiento, C., Griffin, J. N., Clements, C. F., Silvestro, D., Varela, S., Uhen, M. D., & Jaramillo, C. (2017). The Pliocene marine megafauna extinction and its impact on functional diversity. *Nature Ecology & Evolution*, 1, 1100-1106. https://doi:10.1038/s41559-017-0223-6

664 Puncher, G. N. (2015). *Assessment of the population structure and temporal changes in spatial*
665 *dynamics and genetic characteristics of the Atlantic bluefin tuna under a fishery independent*
666 *framework*. PhD Thesis, University of Bologna, Italy, and University of Ghent, Belgium.

667 Rambaut, A. (2009). FigTree v1.4.3. Retrieved from http://tree.bio.ed.ac.uk/software/figtree/

668 Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarisation
669 in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, 67(5), 901–904.
670 https://doi:10.1093/sysbio/syy032

671 Ree, R. H., & Smith, S. A. (2008). Maximum Likelihood inference of geographic range evolution by
672 dispersal, local extinction, and cladogenesis. *Systematic Biology*, 57(1), 4–14.
673 https://doi.org/10.1080/10635150701883881

674 Riccioni, G., Landi, M., Ferrara, G., Milano, I., Cariani, A., Zane, L., … Tinti, F. (2010). Spatio
675 temporal population structuring and genetic diversity retention in depleted Atlantic bluefin tuna of
676 the Mediterranean Sea. *Proceedings of the National Academy of Sciences of the United States of*
677 *America,* 107(5), 2102–2107. https://doi:10.1073/pnas.0908281107

678 Sperone, E., Parise, G., Leone, A., Milazzo, C., Circosta, V., Santoro, F., … Tripepi, S. (2012).
679 Spatiotemporal patterns of distribution of large predatory sharks in Calabria (Central Mediterranean,
680 Southern Italy). *Acta Adriatica*, 53, 13-24.

681 Steeman, M. E., Hebsgaard, M. B., Fordyce, R. E., Ho, S. Y. W., Rabosky, D. L., Nielsen, R., …
682 Willerslev, E. (2009). Radiation of Extant Cetaceans Driven by Restructuring of the Oceans.
683 *Systematic Biology*, 58(6), 573-585. https://doi:10.1093/sysbio/syp060
684

685 Stevens, J. D. (2010). Epipelagic Oceanic Elasmobranchs. In J. C. Carrier, J. A. Musick, & M. R.
686 Heithaus (Eds.), *Sharks and Their Relatives II: Biodiversity, Adaptive Physiology, and Conservation*
687 (pp. 3-37). Boca Raton, FL: CRC Press.

688 Storai, T., Mojetta, A., Zuffa, M., Giulian, S. (2000). Nuove segnalazioni di *Carcharodon carcharias*
689 (L.) nel Mediterraneo centrale. *Atti Società Toscana di Scienze Naturali*, 107, 139-142.

690 Storai, T., Vanni, S., Zuffa, M., & Biagi, V. (2005). Presenza di *Carcharodon carcharias* (Linnaeus,
691 1758) nelle acque toscane (Mar Ligure meridionale e Mar Tirreno settentrionale; Mediterraneo):
692 analisi e revisione delle segnalazioni (1839-2004). *Atti Società Toscana di Scienze Naturali*, 112,
693 153-166.

694 Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., & Rambaut, A. (2018).
695 Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution*, 4(1),
696 vey016. https://doi:10.1093/ve/vey016

697   Tamura, K, & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control

698   region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, 10(3),

699   512-526. https://10.1093/oxfordjournals.molbev.a040023

700   Vandeperre, F., Aires-da-Silva, A., Fontes, J., Santos, M., Serrão Santos, R., & Afonso, P. (2014).

701   Movements of blue sharks (*Prionace glauca*) across their life history. *PLoS ONE* 9(8), e103538.

702   https://doi:10.1371/journal.pone.0103538

703   Varela, S., González-Hernández, J., Sgarbi, L. F., Marshall, C., Uhen, M. D., Peters, S., &

704   McClennen, M. (2015). paleobioDB: an R package for downloading, visualizing and processing data

705   from the Paleobiology Database. *Ecography*, 38(4), 419–425. https://doi:10.1111/ecog.01154

706   Velez-Juarbe, J., Wood, A. R., De Gracia, C., & Hendy, A. J. W. (2015). Evolutionary patterns among

707   living and fossil kogiid sperm whales: evidence from the Neogene of Central America. *PLoS ONE*,

708   10(4), e0123909. https://doi:10.1371/journal.pone.0123909

709   Wandeler, P., Hoeck, P. E. A., & Keller, L. F. (2007). Back to the future: museum specimens in

710   population   genetics.   *Trends   in   Ecology   Evolution*,   22(12),   634-642.

711   https://doi:10.1016/j.tree.2007.08.017

712   ~~Willerslev, E., & Cooper, A. (2005). Ancient DNA. *Proceedings of the Royal Society B: Biological~~

713   ~~Sciences,* 272(1558), 3-16. https://doi:10.1098/rspb.2004.2813)~~

714   Yu, Y., Harris, A. J., & He, X. (2010). S-DIVA (Statistical Dispersal-Vicariance Analysis): A tool

715   for inferring biogeographic histories. *Molecular Phylogenetics and Evolution*, 56(2), 848-850.

716   https://doi.org/10.1016/j.ympev.2010.04.011

717

718   Yu, Y., Harris, A. J., Blair, C., & He, X. J. (2015) RASP (Reconstruct Ancestral State in Phylogenies):

719   a tool for historical biogeography. *Molecular Phylogenetics and Evolution,* 87, 46-49.

720   https://doi:10.1016/j.ympev.2015.03.008

721

722

Biosketch

724

725   **Agostino Leone** is a researcher of the marine research division at fundación AZTI

726   (https://www.azti.es/join-azti-tecnalia/team/name/agostino-leone/), and past PhD student at the

727   University of Bologna in the Tinti Lab. He has broad interests in researches related to marine

728   conservation, ranging from biogeography to population genomics, investigating marine populations

729   structure and their response to different pressures. AL, FT conceived the study; AL, GNP, EC, AC

730   carried out molecular work and sequences analysis; AL, GNP, FF, PDJ, GB carried out statistical

731   analyses; ES, ST, PM, AG, MS, MA, GD, FG, ADA, DM, SV, FS collected specimens; AL, GNP,

732   MA, EC, FS, PDJ, GB, AC, FT drafted the manuscript.

1
2
3 733
4
5 734   **Supporting Information**
6 735
7
8 736   Additional Supporting Information may be found in the online version:
9 737   **Appendix S1** Supplementary methods for the tissue sampling, historical DNA extraction, species-
10
11 738   specific primers design, DNA amplification, sequencing and genetic diversity analysis.
12 739   **Appendix S2** Supplementary figures
13
14 740   **Appendix S3** Supplementary tables
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Supporting Information

Pliocene colonization of the Mediterranean by Great White Shark inferred from fossil records, historical jaws, phylogeographic and divergence time analyses

Agostino Leone, Gregory Neils Puncher, Francesco Ferretti, Emilio Sperone, Sandro Tripepi, Primo Micarelli, Andrea Gambarelli, Maurizio Sarà, Marco Arculeo, Giuliano Doria, Fulvio Garibaldi, Nicola Bressi, Andrea Dall'Asta, Daniela Minelli, Elisabetta Cilli, Stefano Vanni, Fabrizio Serena, Píndaro Díaz-Jaimes, Guy Baele, Alessia Cariani, Fausto Tinti

Appendix S1 Supplementary methods for the tissue sampling, historical DNA extraction, species-specific primers design, DNA amplification, sequencing and genetic diversity analysis.*Sampling*

Supplementary methods

Tissue samples from 18 historical specimens identified as *Carcharodon carcharias*, captured in the Italian Seas from 1823 to the 1980s, were collected from museums and private archives (table S1; figure S1). Due to the cultural importance of the GWS museum specimens, sampling operations were carried out with utmost care to avoid extensive and unsightly damage (e.g. collecting tissue samples from the inner surface of jaws, internal dental pulp of teeth and dried skin debris). When jaws were available, the internal point of attachment between the lower hemi-arches was used for sampling, as it is concealed from public view and is the thickest part of the jaw. In this way, it was possible to drill deep into the jaw and avoid the use of surface materials that could act as a source of contaminants (figure S2). Exogenous DNA was removed by saturating sample surfaces with a 3.0% v/v sodium hypochlorite solution for approximately 10 minutes according to the protocol of Kemp and Smith (Kemp & Smith, 2005). All instruments were sterilized with bleach and UV irradiation between samplings. Holes of 5 mm in diameter were drilled into each jaw using an electric drill set to minimum speed to avoid damage to DNA due to thermal stress (Gibbon, Penny, Štrkalj, & Ruff, 2009). As the drill bit was removed from the cartilaginous jaws and vertebrae, all dust was carefully collected. The amounts of collected tissue ranged from 16 to 409.2 mg. The powdered tissue was then transported to a sterilized laboratory dedicated to aDNA analysis. The resulting holes left in specimens were filled with a low temperature restorative paste commonly used in anthropological studies (figure S2).

Historical DNA extraction

Extraction of DNA and polymerase chain reaction set-up were conducted in the "pre-PCR" clean-room of the Laboratory of Genetics & Genomics of Marine Resources and Environment (GenoDREAM) of the University of Bologna, dedicated to the analysis of degraded and low copy number DNA. The most stringent criteria to minimize and detect exogenous contamination in aDNA

analysis (Cooper & Poinar, 2000) were followed (e.g. the use of extraction and amplification blanks as negative controls in each reaction).

DNA was extracted from all samples using a protocol modified from Riccioni et al. (2010), whereby homogenized tissue powder was incubated in EDTA buffer (0.5M, pH 8.0) in a shaker overnight at room temperature. Samples were then precipitated and incubated in an extraction buffer (0.1M EDTA, 0.5% N-laurylsarcosine-Na) and Proteinase K (20 mg/mL) for an additional 24 hours at 44°C. After spinning, 250 μL supernatants were transferred to tubes containing 3.5 μL of 1 μg/μL Dextran Blue, 250 μL of 4M NH_4-acetate, and 500 μL of 96% v/v ethanol. After precipitation, washing and drying, the genomic DNA was re-suspended in 50 μL of distilled sterile water and stored at -20°C.

Design of primer pairs, PCR amplification and sequencing of historical DNA

A fragment of the highly variable fragment of mitochondrial control region (CR) was targeted for analysis, since it is the most commonly used marker for reconstructing white shark phylogeography, and because there are available several sequences and haplotypes deposited in the public repositories (e.g. GenBank). Since historical DNA can be extensively fragmented (Allentoft et al. 2012), species-specific primers suitable for amplifying short fragments (167-221bp) of overlapping and adjacent DNA sequence fragments were developed (figure S3). Due to the intrinsic characteristic of ancient genetic material to be highly damaged, mainly for deamination processes, resulting in transitions from C to T and G to A (Hansen, Willerslev, Wiuf, Mourier, & Arctander, 2001), at least two independent amplifications were performed for each sample, in order to improve the detection of the damaged sites. Control region sequences/haplotypes of contemporary GWS specimens were retrieved from GenBank (table S3) and aligned with MEGA v.7.0.14 (Kumar, Stecher, & Tamura, 2016). using the ClustalW algorithm (Thompson, Higgins, & Gibson, 1994). From the alignment, five CR primer pairs (table S2) were designed with the software PRIMER3 v.4.0.0 (Untergrasser et al. 2012). These were subsequently tested *in silico* (figure S3) using AmplifX software, version 1.44 (©Nicolas Jullien 2004-2013; CNRS, Aix-Marseille Université, http://crn2m.univ-mrs.fr/pub/amplifx-dist). PCR conditions for all gene fragments consisted of 3 minutes of denaturation at 94°C, followed by 35 cycles of 30s at 94°C, 30s at 50°C, 30s at 72°C, and a final extension period of 7 minutes at 72°C. All PCR reactions were performed in a volume of 50 μL containing approximately 10-20ng of template DNA, 1X Tris-HCl, 200mM of each dNTP, 3mM $MgCl_2$, 0.5μM of forward and reverse primers, and 1.25 units of *Taq* DNA Polymerase (Invitrogen). PCR amplicons were cycle-sequenced from both strands by a commercial sequence service provider (Macrogen Europe, Amsterdam, Netherlands).

Sequence Analysis

The mitochondrial Control Region (CR) partial sequences obtained from the historical samples were checked and edited using MEGA v.7.0.14 (Kumar, Stecher, & Tamura, 2016) and aligned with homologous sequences deposited in the GenBank (table S3) using the ClustalW algorithm (Thompson, Higgins, & Gibson, 1994). The total number of haplotypes and haplotype and nucleotide

diversities of the Mediterranean sequences (with associated standard deviations) were estimated using DnaSP v.5.1 (Librado, & Rosaz, 2009). The genetic diversity data of the global populations were taken from literature (see results). The relationship between haplotype and nucleotide diversity of each population was plotted using the R software package (R Core Team, 2008) to investigate GWS population diversity history. Phylogenetic relationships and haplotype genealogies were inferred using HapView (Salzburger, Ewing, & von Haeseler, 2011). Maximum likelihood (ML) clustering was constructed using the DNAML program in PHYLIP v.3.695 (Felsenstein, 2005), run in HapView. The best evolutionary model used in the phylogenetic analyses was inferred with JModelTest 2.1.1 (Darriba, Taboada, Doallo, & Posada, 2012), according to the Akaike Information Criteria (AIC; Akaike, 1974). Subsequently, the Hasegawa-Kishino-Yano model (Hasegawa, Kishino, & Yano, 1985) with the discrete Gamma distribution (0.8) and allowing for a proportion of invariant sites (0.4) to exist (HKY85+G+I) was selected as the best-fit model.

Since the historical Mediterranean sequence alignment (515bp) was shorter than most of the CR sequences deposited in the GenBank, a ML haplotype network was also reconstructed using a longer sequence alignment of 828bp based on haplotypes available in GenBank obtained from global contemporary GWS populations and specimens, including four Mediterranean sequences previously reported (table S3). A comparison between the topologies of the two haplotype networks permitted a test of the potential loss of informative sites in the shorter sequence alignment and the possible effects this might have on the reconstructed phylogenetic relationships.

References of the supplementary methods

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. https://doi:10.1109/TAC.1974.1100705

Allentoft, M. E., Collins, M., Harker, D., Haile, J., Oskam, C. L., Hale, M. L., … Bunce, M., (2012). The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proceedings of the Royal Society B: Biological Sciences,* 279(1748), 4724-4733. https://doi:10.1098/rspb.2012.1745

Cooper, A, & Poinar, H. N. (2000). Ancient DNA: do it right or not at all. *Science*, 289(5482), 1139. https://doi:10.1126/science.289.5482.1139b

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9, 772. https://doi:10.1038/nmeth.2109

Felsenstein, J. (2005). *PHYLIP (Phylogeny Inference Package) version 3.6*. University of Washington, Seattle. Retrieved from http://evolution.genetics.washington.edu/phylip.html.

Gibbon, V. E., Penny, C. B., Štrkalj, G., & Ruff, P. (2009). Brief Communication: Minimally Invasive Bone Sampling Method for DNA Analysis. *American Journal of Physical Anthropology*, 139(4), 596-599. https://doi:10.1002/ajpa.21048

Hansen, A. J., Willerslev, E., Wiuf, C., Mourier, T., & Arctander, P. (2001) Statistical Evidence for Miscoding Lesions in Ancient DNA Templates. *Molecular Biology and Evolution,* 18(2), 262–265. https://doi:10.1093/oxfordjournals.molbev.a003800.

Hasegawa, M., Kishino, H., & Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, 22(2), 160-174.

Kemp, B. M., & Smith, D. G. (2005). Use of bleach to eliminate contaminating DNA from the surface of bones and teeth. *Forensic Science International*, 154(1), 53-61. https://doi:10.1016/j.forsciint.2004.11.017

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7 Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, 33(7), 1870-1874. https://doi.org/10.1093/molbev/msw054

Librado, P., & Rosaz, J. (2009). DNAsp v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25(11), 1451-1452. https://doi:10.1093/bioinformatics/btp187

R Core Team. (2008). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Riccioni, G., Landi, M., Ferrara, G., Milano, I., Cariani, A., Zane, L., … Tinti, F. (2010). Spatio temporal population structuring and genetic diversity retention in depleted Atlantic bluefin tuna of the Mediterranean Sea. *Proceedings of the National Academy of Sciences of the United States of America,* 107(5), 2102–2107. https://doi:10.1073/pnas.0908281107

Salzburger, W., Ewing, G. B., & von Haeseler, A. (2011). The performance of phylogenetic algorithms in estimating haplotype genealogies with migration. *Molecular Ecology*, 20(9), 1952-1963. https://dx.doi.org/10.1111/j.1365-294X.2011.05066.x

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22(22), 4673-4680. https://doi:10.1093/nar/22.22.4673

Untergrasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012 Primer3 - new capabilities and interfaces. *Nucl. Acids Res*. **40**, e115. https://doi:10.1093/nar/gks596)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16

Appendix S2 Supplementary figures

Figure S1. a) Jaws of the Great White Shark archived in the Museum of Comparative Anatomy of the University of Bologna. b, c) The cover and original illustrative note from Antonio Alessandrini (1854) "Catalogo degli Oggetti e Preparati più interessanti del Gabinetto D'Anatomia Comparata della Pontificia Università di Bologna dalla sua Fondazione all' Ottobre del 1852". The jaws were prepared from a GWS individual (TL = 473 cm) collected in the Adriatic Sea in 1827 and displayed to the public at the University of Bologna. d) Original cover from the publication of Ricciardi (1721) Pontificia Università di Bologna, Italy.
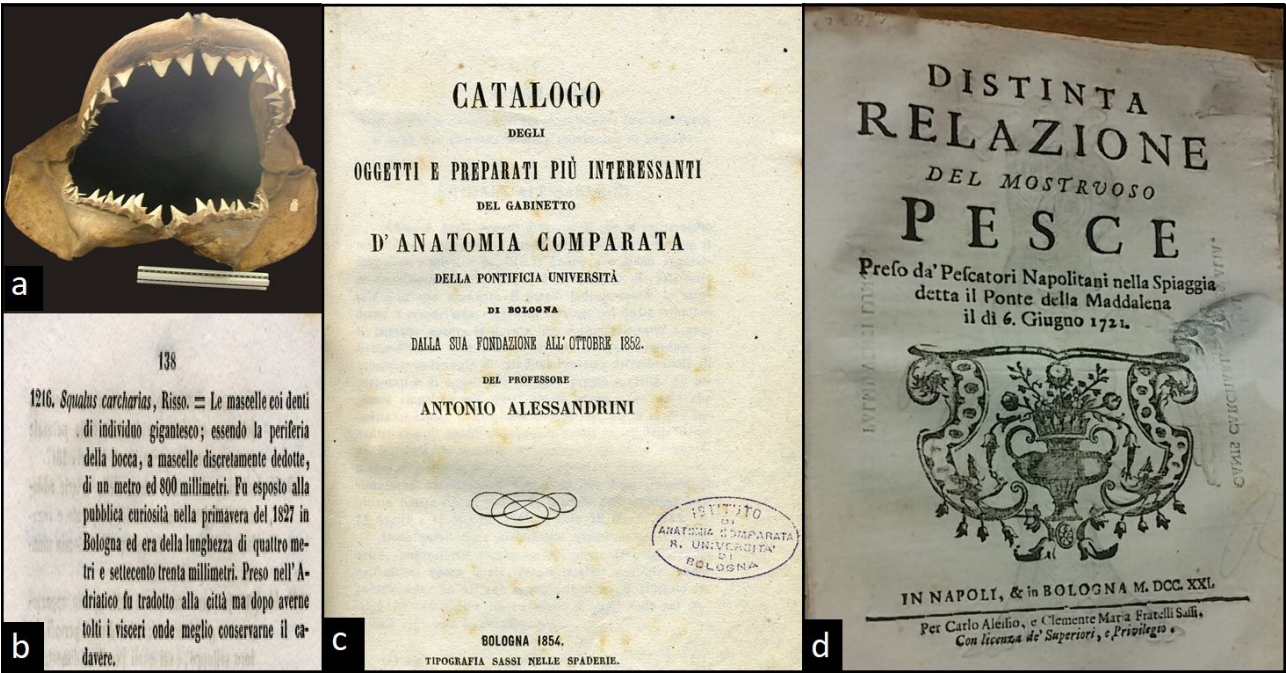
Figure S2. Illustration of the sampling procedures for the museum specimens of Great White Shark. a: Tooth pulp extraction from the crown. b-d: Drill sampling from dried jaw and restoration of the holes.

1
2
3
4
5
6
7
8
9
10

Figure S3. Results of the *in silico* PCR simulation with primer pairs designed for the amplification of the Control Region (CR) of *Carcharadon charcarias* (see Table S2). The annealing positions of the primers with respect to the starting position of the gene on a reference mitochondrial genome (GenBank accession number NC_022415) was shown in brackets. The length in base pairs (bp) of the amplified fragments are marked in yellow.


GWCRA1F (130)
GWCRA2F (133)
GWCRB1F (295)
GWCRB2F (299)
GWCRC1F (442)

GWCRA1R (299) GWCRB1R (443) GWCRB2R (496) GWCRC1F (624)
GWCRA2R (299)

191bp
188bp
167bp
221bp
202bp

100 bp
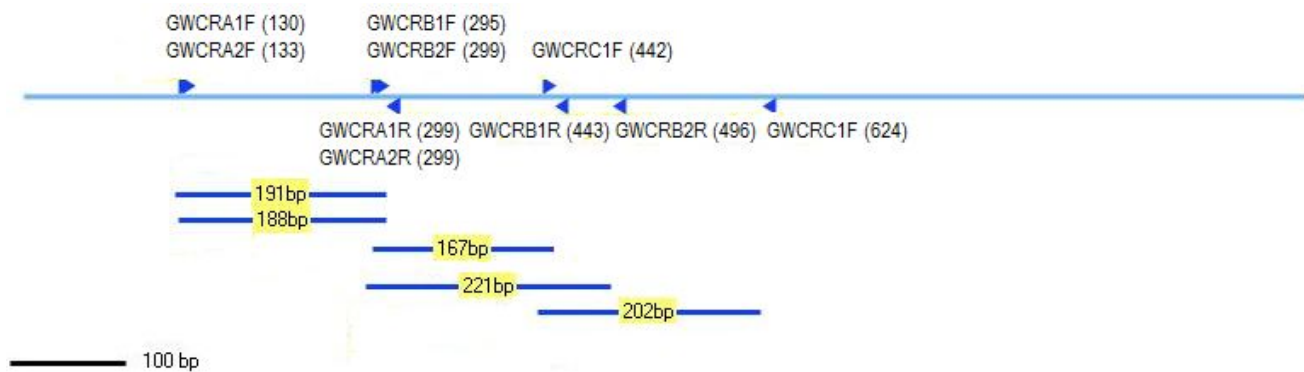
Figure S4. Maximum likelihood haplotype networks of worldwide *Carcharadon charcarias* population samples reconstructed using the 516bp (a) and the 828bp (b) CR sequence alignments, respectively. Acronyms are given in table S2. Numbers inside circles indicates the number of GWS individuals bearing the CR haplotype. Small blue dots indicate single nucleotide substitutions. Numbers outside the circles indicates the gross number of nucleotide substitutions separating the two haplogroups.
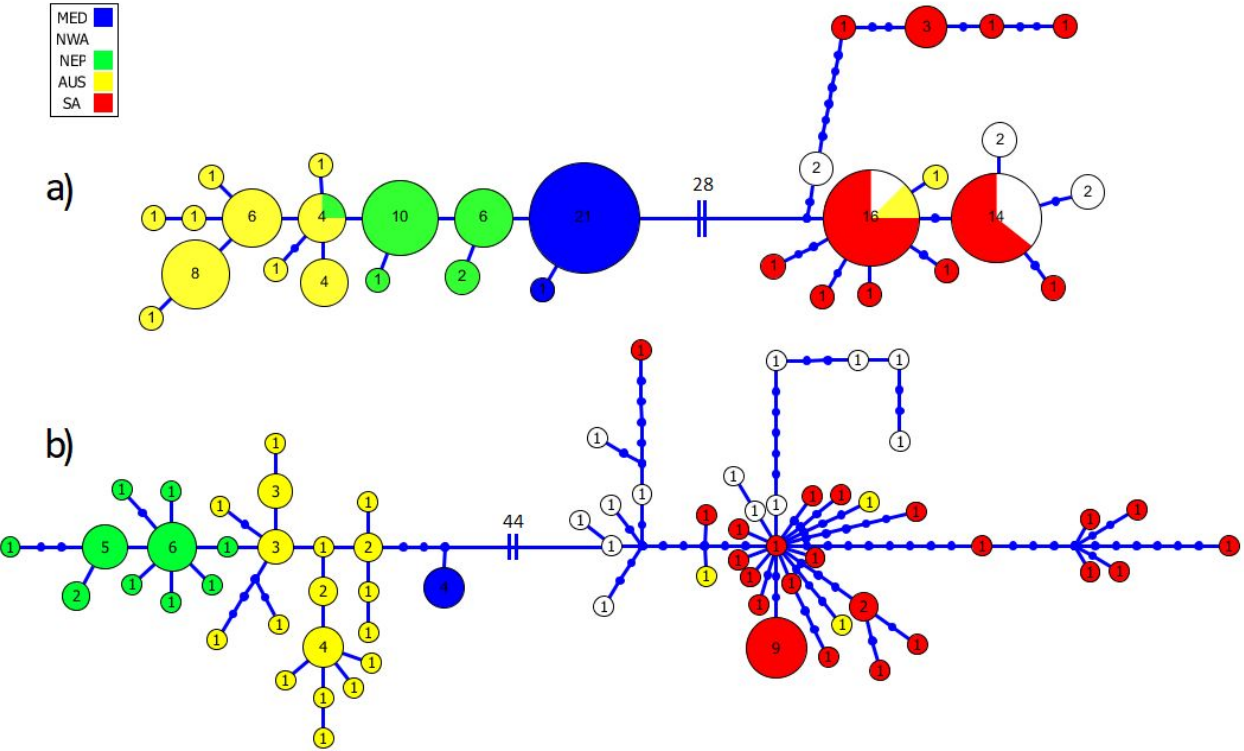
Figure S5. Bayesian divergence time tree of populations of *Carcharadon charcarias* inferred using the 516bp dataset of control region sequences. High posterior density (HPD 95%) in blue bars. Abbreviations are given in table S2.

0,51
0,33 NEP
1
0,72
0,09
0,53 AUS
0,07
0,62
0,59 MED
15,91
1
1,25 SA
1
3,91
1
1,15 SA/NWA

2.0
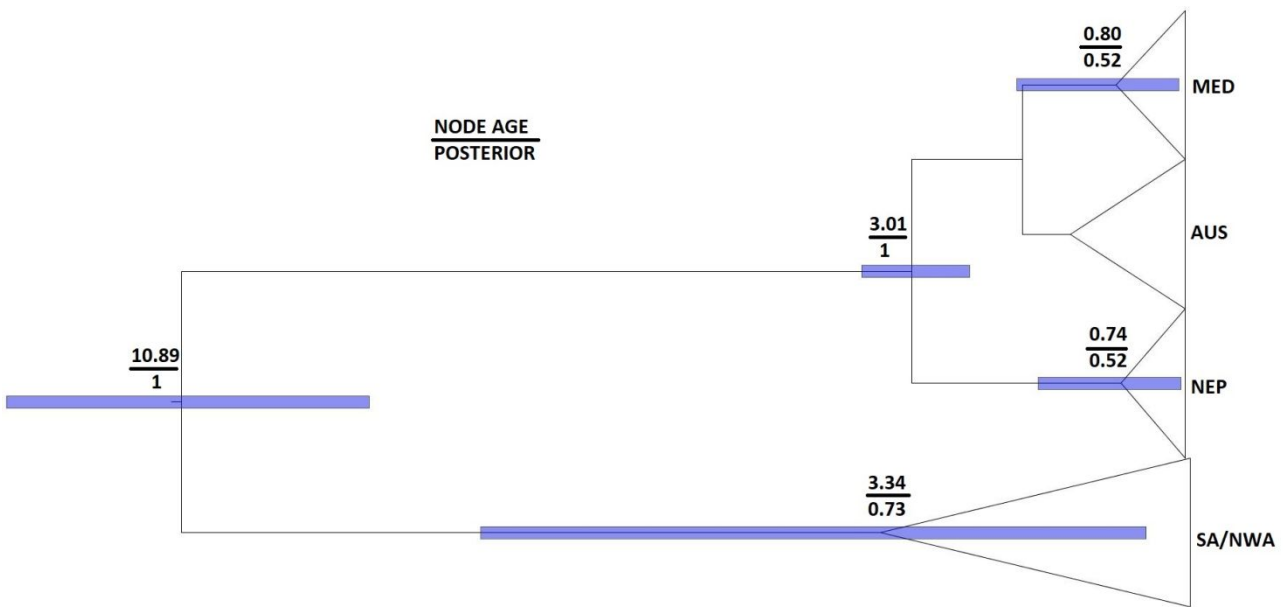
Appendix S3 Supplementary tables

Table S1. List of the 18 historical specimens of Mediterranean Great White Sharks, *Carcharodon carcharias* collected from museum and private archives.

Specimen code	Geographic area	Sampling year/period	Tissue type*	Source§	Museum/Archive, Biological and Collecting Data
TRCC01AD	Adriatic Sea	1906	V, T	1	Cat. No. 1182. Sex: female. Collecting site: Quarnero Gulf (Istria, Croatia). Collecting date: 29 May 1906. Donor: A. Morin
GECC01LI	Ligurian Sea	1935	J	2	Cat. No. C.E. 32695. Sex: unknown. Collecting site: Riva Trigoso, Sestri Levante (Genoa, Italy). Collecting date: 03 July 1935
GECC02LI	Ligurian Sea	1930s	J	2	Cat. No. C.E. 31916. Sex: unknown. Collecting site: unknown. Collecting date: 17 March 1933. Donor: E. Olivieri
GECC03LI	Ligurian Sea	1958	J, T	2	No detailed data are available
LICC01LI	Ligurian Sea	1950s	J	3	No detailed data are available
BOCC01AD	Adriatic Sea	1823	J	4	Cat. No. 811, Catalogue Alessandrini. 1823 Sex: unknown. Collecting site: unknown. Collecting date: unknown. Additional info: mouth extension of 1.15 m
BOCC02AD	Adriatic Sea	1827	J	4	Cat. No. ACP 114*; 1216 Catalogue Alessandrini. Sex: unknown. Collecting site: unknown. Collecting date: unknown. Additional info: mouth extension of 1.80 m
FICC01LI	Ligurian Sea	1891	J	5	Cat. No. 6032; *Carcharodon rondeletii* M.H., 2775. Sex: female. Collecting site: Monterosso (Spezia, Italy). Collecting date: 10 December 1891. Additional info: Total length ~600 cm; Weight ~600 Kg. Donor: S.H. Giglioli.
FICC02LI	Ligurian Sea	1879	V, T	5	Cat. No. 5983. Sex: unknown. Collecting site: Viareggio (Lucca, Italy). Collecting date: unknown.
MOCC01LI	Ligurian Sea	1883	J, T	6	Cat. No. 50; *Carcharodon rondeletii* M.H. Sex: Male. Collecting site: Portofino (Genoa, Italy). Collecting date: January 1883.
PACC01TI	Tyrrhenian Sea	1980s	J	7	No detailed data are available
PACC02TI	Tyrrhenian Sea	1980s	J	7	No detailed data are available
PACC03TI	Tyrrhenian Sea	1980s	J	7	No detailed data are available
FACC01TI	Tyrrhenian Sea	1980s	V, M	8	Sex: unknown. Collecting site: Favignana (Trapani, Italy). Collecting date: unknown.
FACC02TI	Tyrrhenian Sea	1980s	M	8	Sex: unknown. Collecting site: Favignana (Trapani, Italy). Collecting date: unknown.
FACC03TI	Tyrrhenian Sea	1980s	V	8	Sex: unknown. Collecting site: Favignana (Trapani, Italy). Collecting date: unknown.
FACC04TI	Tyrrhenian Sea	1980s	V	8	Sex: unknown. Collecting site: Favignana (Trapani, Italy). Collecting date: unknown.

| FACC05TI | Tyrrhenian Sea | 1980s | M | 8 | Sex: unknown. Collecting site: Favignana (Trapani, Italy). Collecting date: unknown. |

* V: cartilage, vertebrae; T: osteodentine, tooth; J: cartilage, jaws; M; dried skeletal muscle

§ 1: Civic Museum of Natural History of Trieste; 2: Civic Museum of Natural History of Genova "Giacomo Doria"; 3: Regional Agency for Environmental Protection of Tuscany, ARPAT Livorno; 4: Museum of Comparative Anatomy, University of Bologna; 5: Museum of Natural History of Firenze "La Specola"; 6: University Museum of Natural History and Instrumentation of Modena; 7: Museum of Zoology of Palermo "P. Doderlein"; 8: Favignana Tuna Trap

Table S2. List of primer pairs (F: forward primer; R: reverse primer) designed for the PCR amplification and sequencing of the CR gene fragments from the historical DNA of Great White Shark.

Primer	5'>3' sequence
GWCRA1F:	TGACCTTCACCTAATGGTATCACA
GWCRA1R*:	AAGTCTCTGTGAGTGGAAGGAA
GWCRA2F:	CCTTCACCTAATGGTATCACACTC
GWCRA2R*:	AAGTCTCTGTGAGTGGAAGGAA
GWCRB1F:	TTCCTTCCACTCACAGAGACTT
GWCRB1R:	CAAGGACTGAAGTGTTACAAGCA
GWCRB2F:	TTTATTCCTTCCACTCACAGAGAC
GWCRB2R:	GACGGAAATGCTGTTAAAGG

* these two primers have identical sequences

Table S3. The mtDNA Control Region (CR) haplotypes and/or sequences of modern Great White Shark deposited and retrieved from the GenBank.

Geographic Origin	Acronym	N	Reference	GenBank Accession Number
Australia	AUS	14	[9]	HQ414073 - HQ414086
Australia	AUS	12		
New Zealand	AUS	4	[7]	AY026196 - AY026224
South Africa	SA	13		
Northeastern Pacific	NEP	20	[10]	GU002302 - GU002321
Florida	NWA	2		
Mediterranean	MED	3	[17]	HQ540294 - HQ540298
Mediterranean	MED	1	[27]	JF715925
Northwestern Atlantic	NWA	11		
South Africa	SA	15	[28]	KC511601 - KC511626
South Africa*	SA	4*	[29]	KP058665 - KP058902*

*During the analyses, the 238 unique sequences from Andreotti et al. [29], were collapsed in the unique four haplotypes observed and added to the final dataset. AUS: Australia/New Zealand; NEP: Northeastern Pacific; MED: Mediterranean; SA: South-Africa; NWA: Northwestern Atlantic.

Table S4. Fossil data downloaded from PaleoDB (https://paleobiodb.org/#/) using the species-specific taxon "*Carcharodon carcharias*" and synonims. References downloadable from the PaleoDB database.

PaleoDB N°	coll_no	match_no	early_interval	Epoch	early_age	late_age	Ref_no	longitude	latitude
248318	24237	83174	Tortonian	Miocene	11.62	7.246	6853	-79.800.003	9.35
248318	24237	83174	Tortonian	Miocene	11.62	7.246	6853	-79.800.003	9.35
465463	45458	83174	Messinian	Miocene	7.246	5.333	12282	-115.175.552	28.114.721
465463	45458	83174	Messinian	Miocene	7.246	5.333	12282	-115.175.552	28.114.721
465581	13079	83174	Messinian	Miocene	7.246	5.333	12182	-74.719.719	-15.580.833
465581	13079	83174	Messinian	Miocene	7.246	5.333	12182	-74.719.719	-15.580.833
465597	46083	83174	Messinian	Miocene	7.246	5.333	12182	-74.690.277	-15.557.222
465597	46083	83174	Messinian	Miocene	7.246	5.333	12182	-74.690.277	-15.557.222
504614	51304	83174	Calabrian	Pleistocene	1.806	0.781	13658	139.516.663	35.383.331
504614	51304	83174	Calabrian	Pleistocene	1.806	0.781	13658	139.516.663	35.383.331
506112	51328	83174	Zanclean	Pliocene	5.333	3.6	13672	141.603.058	-38.361.942
506112	51328	83174	Zanclean	Pliocene	5.333	3.6	13672	141.603.058	-38.361.942
506149	51335	83174	Zanclean	Pliocene	5.333	3.6	13672	141.944.443	-37.728.333
506149	51335	83174	Zanclean	Pliocene	5.333	3.6	13672	141.944.443	-37.728.333
506450	46068	83174	Zanclean	Pliocene	5.333	3.6	13672	14.796.666	-37.854.721
506450	46068	83174	Zanclean	Pliocene	5.333	3.6	13672	14.796.666	-37.854.721
506468	51405	83174	Piacenzian	Pliocene	3.6	1.806	13672	138.609.726	-34.833.332
506468	51405	83174	Piacenzian	Pliocene	3.6	1.806	13672	138.609.726	-34.833.332
506557	51414	83174	Early Pliocene	Pliocene	5.333	2.588	13672	148.083.328	-39.983.334
506557	51414	83174	Early Pliocene	Pliocene	5.333	2.588	13672	148.083.328	-39.983.334
518397	52644	83174	Calabrian	Pleistocene	1.806	0.781	14149	-52.326.389	-32.388.332
518397	52644	83174	Calabrian	Pleistocene	1.806	0.781	14149	-52.326.389	-32.388.332
520948	28039	83174	Waipipian	Pliocene	3.6	3	14399	17.428.334	-39.583.332
520948	28039	83174	Waipipian	Pliocene	3.6	3	14399	17.428.334	-39.583.332
533364	20400	83174	Irvingtonian	Pleistocene	1.8	0.3	1960	-82.5	27.700.001
533364	20400	83174	Irvingtonian	Pleistocene	1.8	0.3	1960	-82.5	27.700.001
558936	58089	83174	Irvingtonian	Pleistocene	1.8	0.3	15601	-80.811.386	27.830.278
558936	58089	83174	Irvingtonian	Pleistocene	1.8	0.3	15601	-80.811.386	27.830.278
593653	52582	83174	Zanclean	Pliocene	5.333	3.6	18094	-76.817.497	35.359.722
593653	52582	83174	Zanclean	Pliocene	5.333	3.6	18094	-76.817.497	35.359.722

634589	68271	83174	Zanclean	Pliocene	5.333	3.6	19640	-71.5	-30.333.332
634589	68271	83174	Zanclean	Pliocene	5.333	3.6	19640	-71.5	-30.333.332
639047	55535	83174	Messinian	Miocene	7.246	5.333	19852	-70.841.667	-27.080.833
639047	55535	83174	Messinian	Miocene	7.246	5.333	19852	-70.841.667	-27.080.833
645725	69730	83174	Tortonian	Miocene	11.62	7.246	23394	-7.087.944	-27.139.168
645725	69730	83174	Tortonian	Miocene	11.62	7.246	23394	-7.087.944	-27.139.168
645727	69731	83174	Zanclean	Pliocene	5.333	3.6	23394	-7.087.944	-27.139.168
645727	69731	83174	Zanclean	Pliocene	5.333	3.6	23394	-7.087.944	-27.139.168
706800	45480	83174	Piacenzian	Pliocene	3.6	2.588	24838	10.737.222	43.636.665
706800	45480	83174	Piacenzian	Pliocene	3.6	2.588	24838	10.737.222	43.636.665
716083	76844	83174	Early Pleistocene	Pleistocene	2.588	0.781	25983	120.407.219	23.048.889
716083	76844	83174	Early Pleistocene	Pleistocene	2.588	0.781	25983	120.407.219	23.048.889
732439	78614	83174	Middle Miocene	Miocene	15.97	11.608	26436	9.273.333	22.883.333
732439	78614	83174	Middle Miocene	Miocene	15.97	11.608	26436	9.273.333	22.883.333
789236	20646	83174	Late Pleistocene	Pleistocene	0.126	0.0117	28773	-118.199.997	34
789236	20646	83174	Late Pleistocene	Pleistocene	0.126	0.0117	28773	-118.199.997	34
803512	42842	83174	Early Miocene	Miocene	23.03	15.97	29458	8.672.139	21.934.999
803512	42842	83174	Early Miocene	Miocene	23.03	15.97	29458	8.672.139	21.934.999
807796	88328	83174	Piacenzian	Pliocene	3.6	2.588	29650	10.888.611	43.672.222
807796	88328	83174	Piacenzian	Pliocene	3.6	2.588	29650	10.888.611	43.672.222
899319	100174	83174	Late Pliocene	Pliocene	3.6	2.588	34371	-70.534.447	-233.575
899319	100174	83174	Late Pliocene	Pliocene	3.6	2.588	34371	-70.534.447	-233.575
981089	117471	83174	Zanclean	Pliocene	5.333	3.6	37795	-122.446.381	3.747.015
981089	117471	83174	Zanclean	Pliocene	5.333	3.6	37795	-122.446.381	3.747.015
984829	118104	83174	Zanclean	Pliocene	5.333	3.6	38036	-0.676944	38.084.999
984829	118104	83174	Zanclean	Pliocene	5.333	3.6	38036	-0.676944	38.084.999
1087691	136597	83174	Zanclean	Pliocene	5.333	3.6	43697	14.004.444	32.732.498
1087691	136597	83174	Zanclean	Pliocene	5.333	3.6	43697	14.004.444	32.732.498
1192384	154111	83174	Pliocene	Pliocene	5.333	2.588	49963	13.3	-8.75
1192384	154111	83174	Pliocene	Pliocene	5.333	2.588	49963	13.3	-8.75
1192476	154117	83174	Pliocene	Pliocene	5.333	2.588	49968	-112.291.389	27.368.055
1192476	154117	83174	Pliocene	Pliocene	5.333	2.588	49968	-112.291.389	27.368.055
1192505	154118	83174	Pliocene	Pliocene	5.333	2.588	49969	-109.616.669	23.166.668

1192505	154118	83174	Pliocene	Pliocene	5.333	2.588	49969	-109.616.669	23.166.668
1221552	159914	83174	Late Pliocene	Pliocene	3.6	2.588	52184	138.600.006	-34.916.668
1221552	159914	83174	Late Pliocene	Pliocene	3.6	2.588	52184	138.600.006	-34.916.668
1227809	161451	83174	Pliocene	Pliocene	5.333	2.588	52598	139.645.554	35.447.498
1227809	161451	83174	Pliocene	Pliocene	5.333	2.588	52598	139.645.554	35.447.498
1228090	161503	83174	Early Pleistocene	Pleistocene	2.588	0.781	52618	13.696.666	36.763.889
1228090	161503	83174	Early Pleistocene	Pleistocene	2.588	0.781	52618	13.696.666	36.763.889
1234295	162457	83174	Late Miocene	Miocene	11.608	5.333	52569	-81.650.002	30.33
1234295	162457	83174	Late Miocene	Miocene	11.608	5.333	52569	-81.650.002	30.33
1236423	162703	132707	Late Miocene	Miocene	11.608	5.333	52975	4.304.722	39.850.555
1236423	162703	132707	Late Miocene	Miocene	11.608	5.333	52975	4.304.722	39.850.555

Table S5. Fossil data from the Mediterranean area. For museal fossil teeth catalogued from 1 to 87 see S. Marsili (2006).

n	Location	Age	Reference	Stratigraphic References
1a	**Salsomaggiore Terme, Parma, Emilia-Romagna (Italy)**	**Pliocene**	**Bianucci et al. 2002**	**Bianucci et al. 1998**
2a	**Guardamar del Segura, Alicante (Spain) Pliocene**	**Pliocene**	**Adnet et al. 2009**	
1	Terreti, Reggio Calabria: Calabria(Italy)	Pliocene sup./Pleistocene ?	Seguenza, 1901	Gaetani et al., 1986.
2	Nasiti e S.Agata, Reggio Calabria: Calabria(Italy)	Pleistocene inf.	Seguenza, 1901	Lombardo
3	Cetona (vicinanze), Siena: Toscana (Italy)	Pliocene	Principi, 1920	
4	Castiglione del Lago (a Ovest di), Perugia: Umbria (Italy)	Pliocene	Principi, 1920	
5	Città della Pieve, Perugia: Umbria (Italy)	Pliocene	Principi, 1920	
6	Imola (varie località): Emilia Romagna (Italy)	Pliocene	Collection Scarabelli (De Stefano, 1911)	
7	Imola (varie località): Emilia Romagna (Italy)	Pliocene	Collection Scarabelli (De Stefano, 1911)	
8	Castell'Arquato, Piacenza: Emilia Romagna (Italy).	Pliocene	De Stefano, 1912	
9	Orciano, Pisa: Toscana (Italy)	Pliocene inf.-middle	Landini, 1977	
10	Punta Ristola, Capo di Leuca, Lecce: Puglia (Italy)	Pliocene middle	Menesini, 1968	
11	Torre del Lago, Lucca: Toscana (Italy)	Pleistocene	Ghelardoni, 1956	
12	Calanna, Reggio Calabria: Calabria (Italy)	Pliocene sup.	De Stefano, 1901	Barrier et al., 1986.
13	Castell'Arquato, Piacenza: Emilia Romagna (Italy).	Pliocene	Carraroli, 1897	

14	Rometta, Messina: Sicilia (Italy)	Pliocene sup-Pleistocene inf.	Seguenza, 1900	For Violanti, 1989
15	Milazzo, Messina: Sicilia (Italy)	Pliocene sup ?	Seguenza, 1900	
16	Guardamar del segura, Alicante: Valezia (Spain)	Pliocene inf.	Mora Morote, 1996	
17	Ruvo del Monte, Potenza: Basilicata (Italy)	Pliocene	Bassani, 1901	
18	Taranto: Puglia (Italy)	Pleistocene	Bassani, 1905	
19	Volpedo, Alessandria: Piemonte (Italy)	Pliocene middle-sup	De Alessandri, 1895	Gabba, 82
20	Ruvo del Monte, Potenza: Basilicata (Italy)	Pliocene	Pasquale, 1903	
21	S.Agata , Reggio Calabria: Calabria(Italy)	Pleistocene inf.	Pasquale, 1903	Lombardo
22	Terreti, Reggio Calabria: Calabria (Italy)	Pliocene inf.-middle	Pasquale, 1903	Gaetani et al., 1986.
23	Reggio, Nasiti, Reggio Calabria: Calabria (Italy)	Pliocene	Pasquale, 1903	
24	Capo di Leuca, Lecce: Puglia (Italy)	Pliocene middle	Pasquale, 1903	Bossio et al., 2001
25	San Demetrio Corone, Cosenza: Calabria (Italy)	Pleistocene	Pasquale, 1903	
26	Tabiano, Piacenza: Emilia Romagna (Italy);.	Pliocene inf.	De Stefano, 1912	
27	Calanna, Reggio Calabria: Calabria (Italy)	Pliocene sup.	Pasquale, 1903	Barrier et al., 1986. Iannone et al., 1979; Cherubini et al., 1996; Pomar et al., 2001.
28	Matera: Basilicata (Italy)	Pliocene sup.-Pleistocene inf.	Pasquale, 1903	
29	Sestano e Medano, Siena: Toscana (Italy)	Pliocene inf.	Manganelli & Spadini, 2003	
30	San Quirico d'Orcia, Giustrigona, Terre Rosse e I Sodi, Siena: Toscana (Italy)	Pliocene middle	Manganelli & Spadini, 2003	
31	Allerona, Terni: Umbria (Italy)	Pliocene middle	Bellocchio et al., 1991	
32	Guardamar del segura, Alicante: Valezia (Spain)	Pliocene inf.	Mora Morote, 1996	
33	Colline Toscane (Italy)	Pliocene	De Stefano, 1910: Collection Lawley di Bologna	
34	Colline Toscane (Italy)	Pliocene	De Stefano, 1910: Collection Lawley di Bologna	
35	Orciano, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection Lawley di Pisa	
36	Orciano, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze	
37	Volterra, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze	
38	Val d'Orcia: Toscana (Italy)	Pliocene	Collection fossil teeth di Firenze	
39	Orciano, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze	
40	Orciano, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze	
41	Orciano, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze	
42	Orciano, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze	
43	Orciano, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze	
44	Orciano, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze	

45		Pliocene	Collection fossil teeth di Firenze
46	Orciano, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze
47	Orciano, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze
48	San Frediano, Pisa: Toscana (Italy)	Pliocene middle	Collection fossil teeth di Firenze
49	Volterra, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze
50	Colline Senesi: Toscana (Italy)	Pliocene	Collection fossil teeth di Firenze
51		Pliocene	Collection fossil teeth di Firenze
52	Orciano, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze
53	Orciano, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze
54	San Vivaldo, Firenze: Toscana (Italy)	Pliocene	Collection fossil teeth di Firenze
55	San Quirico d'Orcia, Siena: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze
56	Colline Senesi: Toscana (Italy)	Pliocene	Collection fossil teeth di Firenze
57	Chiusi, Siena: Toscana (Italy)	Pliocene	Collection fossil teeth di Firenze
58		Pliocene	Collection fossil teeth di Firenze
59	Orciano, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze
60	Orciano, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze
61	Orciano, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze
62	Volterra, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze
63	Siena: Toscana (Italy)	Pliocene	Collection fossil teeth di Firenze
64	San Quirico d'Orcia, Siena: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze
65	Bagni di Casciana, Pisa: Toscana (Italy)	Pliocene	Collection fossil teeth di Firenze
66	Orciano, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze
67	San Quirico d'Orcia, Siena: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze
68	Volterra, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze
69	Volterra, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection fossil teeth di Firenze
70	Santa Luce, Pisa: Toscana (Italy)	Pliocene	Collection fossil teeth di Firenze
71	Rione Castellana, Palermo: Sicilia (Italy)	Pleistocene	Gemellaro, 1913
72	Pradalbino (varie loc.), Bologna: Emilia Romagna (Italy)	Pliocene inf.-middle	Vinassa de Regny, 1899
73	Pradalbino (varie loc.), Bologna: Emilia Romagna (Italy)	Pliocene inf.-middle	Vinassa de Regny, 1900
74	Orciano; Volterra, Pisa: Toscana (Italy)	Pliocene inf.-middle	Bassani, 1901
75	San Quirico d'Orcia (dintorni), Siena: Toscana (Italy)	Pliocene inf.-middle	Simonelli, 1880
76	unknown locality	Pliocene	Accademia Fisiocritici di Siena
77	unknown locality	Pliocene	Accademia Fisiocritici di Siena

78	? Monte Follonico(1), Siena: Toscana (Italy)	Pliocene	Accademia Fisiocritici di Siena
79	Monte Follonico, Siena; ? Volterra(1), Pisa: Toscana (Italy)	Pliocene	Accademia Fisiocritici di Siena
80	Pod. Casabianca (Trequanda), Siena: Toscana (Italy)	Pliocene	Accademia Fisiocritici di Siena
81	Medane (Asciano), Siena: Toscana (Italy)	Pliocene	Accademia Fisiocritici di Siena
82	Volterra, Pisa: Toscana (Italy)	Pliocene inf.-middle	Collection Lawley di Pisa
83	Pontedera, Pisa: Toscana (Italy)	Pliocene	Collection Lawley di Pisa
84	Piacentino: Emilia Romagna (Italy)	Pliocene	Collection Lawley di Pisa
85	Terricciola, Pisa: Toscana (Italy)	Pliocene	Collection Lawley di Pisa
86	Val di Cecina: Toscana (Italy)	Pliocene	Collection Lawley di Pisa
87	Peccioli, Pisa: Toscana (Italy)	Pliocene	Collection Lawley di Pisa

PaleoDB N°	early_age (Ma)	late_age (Ma)	early_age_interval	epoch	longitude	latitude
49	0.126	0.0117	Late Pleistocene	Pleistocene	-118.199.997	34
50	0.126	0.0117	Late Pleistocene	Pleistocene	-118.199.997	34
25	1.8	0.3	Irvingtonian	Pleistocene	-82.5	27.700.001
26	1.8	0.3	Irvingtonian	Pleistocene	-82.5	27.700.001
27	1.8	0.3	Irvingtonian	Pleistocene	-80.811.386	27.830.278
28	1.8	0.3	Irvingtonian	Pleistocene	-80.811.386	27.830.278
9	1.806	0.781	Calabrian	Pleistocene	139.516.663	35.383.331
10	1.806	0.781	Calabrian	Pleistocene	139.516.663	35.383.331
21	1.806	0.781	Calabrian	Pleistocene	-52.326.389	-32.388.332
22	1.806	0.781	Calabrian	Pleistocene	-52.326.389	-32.388.332
43	2.588	0.781	Early Pleistocene	Pleistocene	120.407.219	23.048.889
44	2.588	0.781	Early Pleistocene	Pleistocene	120.407.219	23.048.889
77	2.588	0.781	Early Pleistocene	Pleistocene	13.696.666	36.763.889
78	2.588	0.781	Early Pleistocene	Pleistocene	13.696.666	36.763.889
17	3.6	1.806	Piacenzian	Pliocene	138.609.726	-34.833.332
18	3.6	1.806	Piacenzian	Pliocene	138.609.726	-34.833.332
41	3.6	2.588	Piacenzian	Pliocene	10.737.222	43.636.665
42	3.6	2.588	Piacenzian	Pliocene	10.737.222	43.636.665

53	3.6	2.588	Piacenzian	Pliocene	10.888.611	43.672.222
54	3.6	2.588	Piacenzian	Pliocene	10.888.611	43.672.222
55	3.6	2.588	Late Pliocene	Pliocene	-70.534.447	-233.575
56	3.6	2.588	Late Pliocene	Pliocene	-70.534.447	-233.575
73	3.6	2.588	Late Pliocene	Pliocene	138.600.006	-34.916.668
74	3.6	2.588	Late Pliocene	Pliocene	138.600.006	-34.916.668
23	3.6	3	Waipipian	Pliocene	17.428.334	-39.583.332
24	3.6	3	Waipipian	Pliocene	17.428.334	-39.583.332
19	5.333	2.588	Early Pliocene	Pliocene	148.083.328	-39.983.334
20	5.333	2.588	Early Pliocene	Pliocene	148.083.328	-39.983.334
65	5.333	2.588	Pliocene	Pliocene	13.3	-8.75
66	5.333	2.588	Pliocene	Pliocene	13.3	-8.75
69	5.333	2.588	Pliocene	Pliocene	-112.291.389	27.368.055
70	5.333	2.588	Pliocene	Pliocene	-112.291.389	27.368.055
71	5.333	2.588	Pliocene	Pliocene	-109.616.669	23.166.668
72	5.333	2.588	Pliocene	Pliocene	-109.616.669	23.166.668
75	5.333	2.588	Pliocene	Pliocene	139.645.554	35.447.498
76	5.333	2.588	Pliocene	Pliocene	139.645.554	35.447.498
11	5.333	3.6	Zanclean	Pliocene	141.603.058	-38.361.942
12	5.333	3.6	Zanclean	Pliocene	141.603.058	-38.361.942
13	5.333	3.6	Zanclean	Pliocene	141.944.443	-37.728.333
14	5.333	3.6	Zanclean	Pliocene	141.944.443	-37.728.333
15	5.333	3.6	Zanclean	Pliocene	14.796.666	-37.854.721
16	5.333	3.6	Zanclean	Pliocene	14.796.666	-37.854.721
29	5.333	3.6	Zanclean	Pliocene	-76.817.497	35.359.722
30	5.333	3.6	Zanclean	Pliocene	-76.817.497	35.359.722
31	5.333	3.6	Zanclean	Pliocene	-71.5	-30.333.332
32	5.333	3.6	Zanclean	Pliocene	-71.5	-30.333.332
37	5.333	3.6	Zanclean	Pliocene	-7.087.944	-27.139.168
38	5.333	3.6	Zanclean	Pliocene	-7.087.944	-27.139.168
59	5.333	3.6	Zanclean	Pliocene	-122.446.381	3.747.015
60	5.333	3.6	Zanclean	Pliocene	-122.446.381	3.747.015

61	5.333	3.6	Zanclean	Pliocene	-0.676944	38.084.999
62	5.333	3.6	Zanclean	Pliocene	-0.676944	38.084.999
63	5.333	3.6	Zanclean	Pliocene	14.004.444	32.732.498
64	5.333	3.6	Zanclean	Pliocene	14.004.444	32.732.498
3	7.246	5.333	Messinian	Late Miocene	-115.175.552	28.114.721
4	7.246	5.333	Messinian	Late Miocene	-115.175.552	28.114.721
5	7.246	5.333	Messinian	Late Miocene	-74.719.719	-15.580.833
6	7.246	5.333	Messinian	Late Miocene	-74.719.719	-15.580.833
7	7.246	5.333	Messinian	Late Miocene	-74.690.277	-15.557.222
8	7.246	5.333	Messinian	Late Miocene	-74.690.277	-15.557.222
33	7.246	5.333	Messinian	Late Miocene	-70.841.667	-27.080.833
34	7.246	5.333	Messinian	Late Miocene	-70.841.667	-27.080.833
79	11.608	5.333	Late Miocene	Late Miocene	-81.650.002	30.33
80	11.608	5.333	Late Miocene	Late Miocene	-81.650.002	30.33
81	11.608	5.333	Late Miocene	Late Miocene	4.304.722	39.850.555
82	11.608	5.333	Late Miocene	Late Miocene	4.304.722	39.850.555
1	11.62	7.246	Tortonian	Late Miocene	-79.800.003	9.35
2	11.62	7.246	Tortonian	Late Miocene	-79.800.003	9.35
35	11.62	7.246	Tortonian	Late Miocene	-7.087.944	-27.139.168
36	11.62	7.246	Tortonian	Late Miocene	-7.087.944	-27.139.168
67	15.97	3.6	Middle Miocene	Middle Miocene	-70.833.336	-27.133.333
68	15.97	3.6	Middle Miocene	Middle Miocene	-70.833.336	-27.133.333
39	15.97	11.608	Middle Miocene	Middle Miocene	90.666.664	25.166.668
40	15.97	11.608	Middle Miocene	Middle Miocene	90.666.664	25.166.668
45	15.97	11.608	Middle Miocene	Middle Miocene	9.273.333	22.883.333
46	15.97	11.608	Middle Miocene	Middle Miocene	9.273.333	22.883.333
57	23.03	2.588	Miocene	Early Miocene	-70.966.667	-27.633.333
58	23.03	2.588	Miocene	Early Miocene	-70.966.667	-27.633.333
47	23.03	15.97	Early Miocene	Early Miocene	93	23
48	23.03	15.97	Early Miocene	Early Miocene	93	23
51	23.03	15.97	Early Miocene	Early Miocene	8.672.139	21.934.999
52	23.03	15.97	Early Miocene	Early Miocene	8.672.139	21.934.999

Table S65. Control Region diversity of the Mediterranean and global populations of *Carcharodon carcharias*.

Population	Acronym	N	Nh	h ± SD			π ± SD		Reference
Mediterranean	MED	22*	3	0.091	±	0.124	0.0002	±0.0007	This work
Australia-New Zealand	AUS	94	14	0.880	±	0.015	0.0085	±0.0045	[7,9]
Northeastern Pacific	NEP	59	20	0.790		na	0.0013	±0.0009	[10]
South Africa	SA	34	15	0.720		na	0.0059	na	[28]
South Africa	SA*	238	4	0.205			0.0027		[29]

Acronyms are given in table S3.

N: number of individuals analysed; Nh: number of haplotypes; h: haplotype diversity; π nucleotide diversity; SD: Standard deviation

* this sample included sequence records from GenBank (Acc. Num. JF715925, HQ540294, HQ540295, HQ540296).

Table S6. Best model selection based on Bayes factors using alternative primary calibration. logML_GSS: log marginal likelihood from generalized stepping stone model, BF_GSS: Bayes factors calculated using the logML_GSS. Molecular clock test performed by comparing the Maximum Likelihood values for the given topology with and without the molecular clock constraints under Hasegawa-Kishino-Yano (1985) model (+G+I) [39]. The null hypothesis of equal evolutionary rate throughout the tree was not rejected at a 5% significance level. Significance: P = 0.14 (P-value less than 0.05 is used to reject the hypothesis of equal evolutionary rate throughout the tree). The analysis involved 56 nucleotide sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There was a total of 1,080 nucleotide positions in the final dataset. Evolutionary analyses were conducted with MEGA7 v.7.0.14 [31].

Constraint	ln$L$	Parameters	+G	+I
With Clock	-2887.856	61	0.868	0.48
Without Clock	-2826.106	115	0.91	0.47

Model	Dataset	Node1	Node2	logML_GSS	logBF_GSS
MED Pliocene divergence	828bp	*C.carcharias*/*L.nasus* divergence	MED/Pacific		
		Mean: 46 Ma, SD: 1.0 Ma	Mean: 3.0 Ma, SD: 0.3 Ma	-3210.6566	0 (Best)
MED Pleistocene divergence	828bp	*C.carcharias*/*L.nasus* divergence	MED/Pacific		
		Mean: 46 Ma, SD: 1.0 Ma	Mean: 0.4Ma, SD: 0.15 Ma	-3215.5172	4.8606
MED Pleistocene divergence	516bp	*C.carcharias*/*L.nasus* divergence	MED/Pacific		
		Offset: 46 Ma, SD: 1.0 Ma	Mean: 3.0 Ma, SD: 0.3 Ma	-1905,6997	0

MED Pliocene divergence	516bp	*C.carcharias*/*L.nasus* divergence	MED/Pacific		
		Offset: 46 Ma, SD: 1.0 Ma	Mean: 0.4Ma, SD: 0.15 Ma	-1914,4034	8.7037