Journal of open archaeology data

Early African Hominin Localities from *Sahelanthropus to Australopithecus anamensis*: A Source-Audited, GIS-Ready Dataset

DATA PAPERS

IAN G. SMITH
SALLY C. REYNOLDS
*Author affiliations can be found in the back matter of this article

ubiquity press

ABSTRACT

Spatial information for fossil localities central to the earliest stages of human evolution is dispersed across publications, maps, supplementary material and open datasets, and many localities lack published numerical coordinates. This dataset reconciles those sources and derives usable positions where necessary by georeferencing peer-reviewed locality maps in QGIS. It brings together the published locality record identified within its defined scope: 43 records from *Sahelanthropus tchadensis* to *Australopithecus anamensis* across Chad, Ethiopia and Kenya. Each coordinate has a published source or transparent derivation and an assigned precision category, while taxonomic, chronological, spatial and palaeoenvironmental uncertainties are retained. The resulting source-audited, GIS-ready resource can be downloaded and imported directly into GIS software, where the locality data can be overlaid with geological, climatic, topographic and other spatial datasets.

CORRESPONDING AUTHOR:
Ian G. Smith
Independent Researcher at Bournemouth University, UK; Departments Health, Environmental and Medical Sciences (HEMS), Bournemouth University, Fern Barrow, Poole, Dorset BH12 5BB, UK
ismith1@bournemouth.ac.uk

KEYWORDS:
early hominins; palaeoanthropology; fossil localities; geospatial data; GIS; Africa; Late Miocene; Early Pliocene; open data

TO CITE THIS ARTICLE:
Smith IG, Reynolds SC 2026 Early African Hominin Localities from *Sahelanthropus to Australopithecus anamensis*: A Source-Audited, GIS-Ready Dataset. *Journal of Open Archaeology Data*, 14: 27, pp. 1–8. DOI: https://doi.org/10.5334/joad.239

(1) OVERVIEW

CONTEXT

Published locality information for the earliest African hominin record is dispersed across fossil descriptions [11], palaeoenvironmental studies [1], monographs [8], supplementary material [6], maps [12] and open-access online databases [4], and numerical coordinates are frequently absent. The difficulties extend beyond locating coordinates. Successive publications may revise locality names, specimen assignments, ages and geographical positions, while different sources may refer to individual fossil localities, grouped research areas, formations or broad regional points. Values that appear directly comparable may therefore represent different geographical entities, and later accounts may refine, correct or contradict earlier ones. Reuse requires reconstruction of the source history for each record, reconciliation where the evidence permits, and explicit retention of unresolved disagreement.

Although only 43 records fall within the defined scope, the evidential history behind many entries is disproportionately complex. That complexity, rather than the number of localities, helps explain the previous absence of a comparable source-audited, GIS-ready compilation. The resource is locality-centred: its basic units are fossil localities associated with the target taxa rather than individual hominins.

The dataset described here brings these records into a single source-audited, GIS-ready resource by following locality information through primary publications, later reassessments, published maps and relevant open datasets. Published coordinates were checked against their cited sources for copying or conversion errors and then standardised. Where no numerical coordinate was available, suitable locality maps were acquired through the wider peer-reviewed literature, including geological and later regional studies beyond the principal fossil-publication sequence. The relationship between each map and the target locality was checked, particularly where locality names, spatial units or research-area boundaries differed between publications. The maps were then georeferenced in QGIS [2] to derive usable positions. Where sources differed, conflicting values and interpretations were reconciled where possible; where no defensible resolution could be reached, the disagreement was retained. Each record preserves its source, coordinate derivation, precision category and relevant uncertainty.

This article describes the dataset boundary, source hierarchy, coordinate methods, treatment of uncertainty, file structure and reuse potential. It does not attempt to resolve wider debates over species attribution, fossil association, chronology or palaeoenvironmental reconstruction. Instead, it documents how disputed evidence was represented in the released dataset, the basis for those decisions and any uncertainty that remains. The dataset was initially developed during the spatial research underlying *Out of Afar: the first hominin migration? Long-term landscape changes in the Afar region and implications for hominin bipedalism* [14], but is published as an independent research resource.

SPATIAL COVERAGE

Spatial coverage extends across present-day Chad, Ethiopia and Kenya. Records include the Toros-Menalla cluster in the Chad Basin; Gona, Middle Awash, Galili, Woranso-Mille and Fejej in Ethiopia; and the Tugen Hills, Baringo Basin, Kanapoi, Allia Bay and Ileret in Kenya.

Northern boundary: 16.2500 N
Southern boundary: 0.5792 N
Eastern boundary: 40.5450 E
Western boundary: 17.4875 E

TEMPORAL COVERAGE

Temporal coverage extends from approximately 7.0 to 3.8 Ma and includes *Sahelanthropus tchadensis*, *Orrorin tugenensis*, *Ardipithecus kadabba*, *Ardipithecus ramidus* and *Australopithecus anamensis*. Selected records of uncertain affinity are retained under cautious labels; later *Australopithecus* and early *Homo* localities fall outside Version 1.0. Ages are literature-based and may refer to a fossil horizon, locality, member, formation or a group of localities reported together. Published qualifiers and ranges are preserved, with conflicts explained in the Supporting Notes.

The principal characteristics of Version 1.0 are summarised in Table 1.

(2) METHODS

STEPS

Dataset boundary and candidate-record identification

The dataset boundary was fixed at the outset around five widely recognised early hominin taxa spanning approximately 7.0 to 3.8 Ma. The principal localities within this scope are well established in the palaeoanthropological literature; the purpose of the compilation was therefore not primarily to discover previously unrecognised sites, but to bring the dispersed locality record into a single source-audited, GIS-ready resource. Candidate localities were identified through a source-led review of primary fossil and geological publications, monographs, later reassessments, supplementary material, open datasets and established palaeontological databases, including the Human Fossil Record at Fossilized.org and the Paleobiology Database. No fixed systematic-review search string was used. Records were included where the published literature assigned or tentatively assigned them to one of the five target taxa; selected disputed records were retained

ELEMENT	VERSION 1.0 DATASET
Principal records	15
Lettered locality subrecords	28
Total identified records	43
Coordinate-bearing rows	38
Coordinate precision	22 high; 6 within 200 m; 10 within 2.5 km
Countries	Chad, Ethiopia and Kenya
Temporal coverage	ca. 7.0 to 3.8 Ma
Taxonomic scope	Five named target taxa, plus selected explicitly uncertain or mixed assignments

Table 1 Version 1.0 public workbook at a glance.

where later taxonomic reassessment continued to propose a direct association with one or more of those taxa. The dataset was not intended as a comprehensive catalogue of all indeterminate or generically assigned early hominin fossils within the same chronological interval. Each candidate record was then audited individually for locality identity, fossil association, chronology, taxonomic wording and spatial evidence before inclusion. Later *Australopithecus* and early *Homo* localities fell outside Version 1.0. Specimen counts may denote catalogue entries or reported fossils rather than biological individuals.

Principal records, subrecords and inherited information

Numbered rows represent principal sites or locality groups; lettered rows represent constituent localities with separate spatial evidence. Subrecords inherit parent references, landscape context and notes unless a separate source is given. Parent records may therefore lack coordinates where mapped subrecords provide the spatial positions, and subrecord specimen fields may remain blank where published material cannot be allocated reliably. Supporting Notes carry complex explanations without repeating them across subrows.

Source search, hierarchy and evidence classification

Primary fossil and geological publications formed the principal evidence base, supplemented by later reassessments, published maps, open datasets and established palaeontological databases. Later coordinates or interpretations were not accepted automatically but were checked against the earlier locality record to determine whether they referred to an individual fossil locality, a broader research area, a formation or a regional point. Where the source used to obtain or derive a coordinate differed from the principal publication supporting the locality, both sources were recorded separately.

Coordinate determination and standardisation

Published coordinates were transcribed where they were clearly associated with a named locality and converted to decimal degrees where necessary. All released coordinates are stored as separate longitude and latitude fields in WGS84 decimal degrees.

Where no numerical coordinate was available, suitable locality maps were identified through searches of the wider peer-reviewed literature, including geological and regional studies outside the principal fossil-publication sequence. In QGIS, each published map was aligned to identifiable geographical features using the most appropriate available GIS basemap and, where available, other georeferenced maps from peer-reviewed publications. A coordinate was then assigned to the position shown for the locality on the georeferenced published map. The assigned precision reflected the map's scale and resolution and how securely the locality could be aligned to identifiable geographical features.

Public-database coordinates or proxy points, meaning approximate substitute locations representing a broader site or locality group, were used only when no more direct locality position could be established. They were checked against the published regional setting, identified as such in the Coordinate source / derivation field, and retained only as explicitly approximate points with reduced precision.

Coordinate precision and geographical meaning

Three coordinate-precision categories are used: high for explicit locality coordinates reported in a published paper or dataset, within 200 m for coordinates derived from detailed published maps, and within 2.5 km for broader map-derived positions, cluster midpoints, proxies and public-database points. These are source- and scale-based categories rather than statistical error estimates, and "high" does not imply metre-level accuracy.

Taxonomy, chronology and specimen wording

Taxonomic assignment required separate source tracing because published attributions could vary between research teams and change through time as material was reassessed. For each record, the history of its assignment was reviewed and taxonomic confidence was recorded independently of coordinate precision. Where uncertainty or disagreement remained, labels such as tentative,

probable, disputed or mixed were retained rather than forcing a firm species assignment. Ages and specimen information were standardised only to the level justified by the sources, preserving approximate values, ranges and uncertainty in the allocation of material among sublocalities. Record-level explanations are provided in the Supporting Notes worksheet.

Conflicting evidence, corrections and retained disagreement

Conflicts were investigated by returning to primary publications and relevant later reassessments. Where one source provided a better-supported coordinate, chronology or taxonomic wording, the selected value and its basis were documented. Where no defensible resolution was possible, uncertainty was retained through cautious wording, broader ranges, reduced coordinate precision or a Supporting Note rather than attempting to resolve the underlying scientific debate.

Illustrative examples of record construction

The following examples illustrate how the general procedures described above were applied to records requiring different forms of source tracing, spatial reconstruction and dataset structuring.

1. A coordinate derived from a published map
GWM67 illustrates the derivation of a usable locality coordinate for which no published numerical position was located. Simpson et al. [12] identified and mapped GWM67 in Figure 1C, while Semaw et al. [10] had published WGS84 coordinates for surrounding Gona Western Margin localities. The published figure was imported into QGIS and georeferenced through detailed comparison of the local landscape, the courses of the Busidima, Gawis and Sifi rivers, and the published coordinates of surrounding localities. The position of GWM67 was then extracted from the georeferenced map. The resulting coordinate, 40.322000° E, 11.158300° N, was not treated as equivalent to a directly published GPS position. It was assigned the precision category "within 200 m", and its map-derived basis was recorded in the Coordinate source / derivation field and Supporting Note B.
2. A coordinate requiring correction: Fejej / FJ-4
Fejej / FJ-4 illustrates why coordinates obtained from secondary databases must be treated with caution. Paleobiology Database collections 22563 and 22567 identify both records as "Fejej – FJ4" but report the rounded position 37.000000° E, 5.000000° N [9]. This point, used in an earlier working dataset, lies

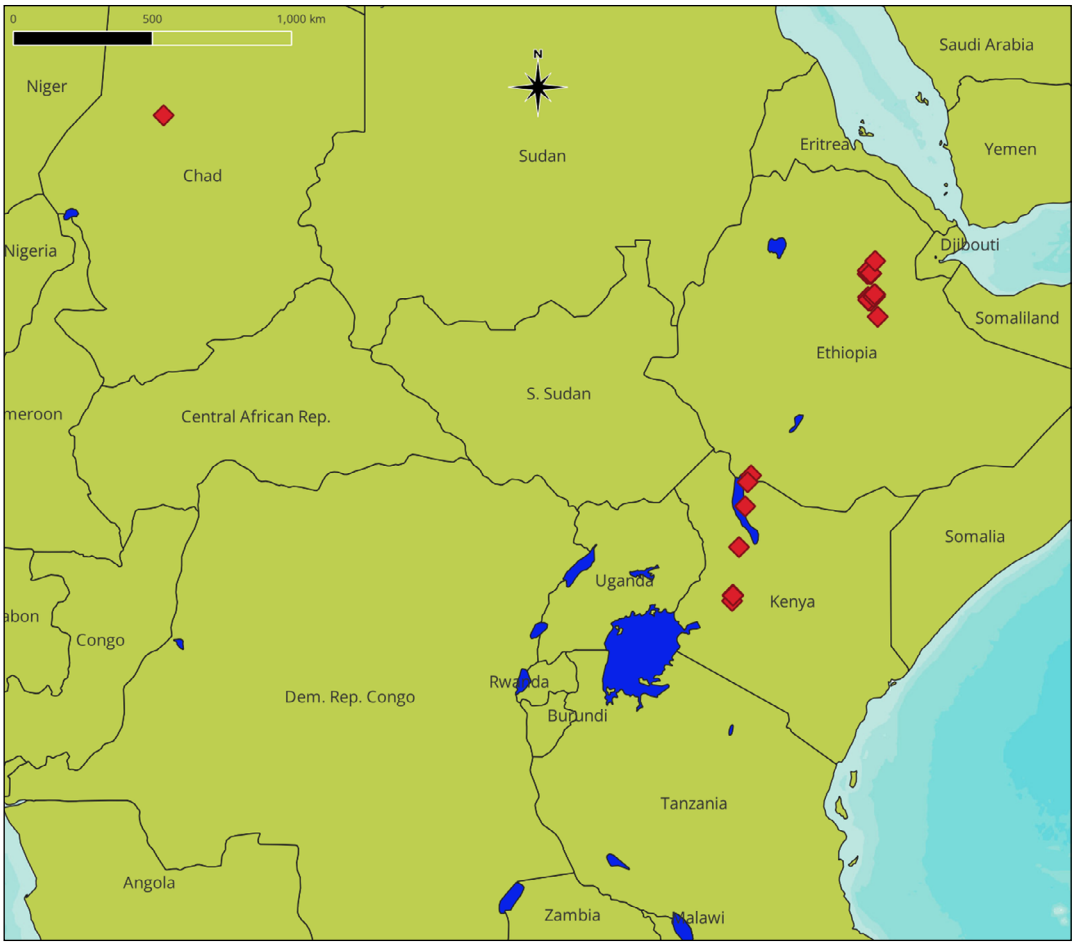

Figure 1 Geographic distribution of the early African hominin locality records included in the dataset. Locality coordinates are those reported or derived in the source-audited dataset. Base mapping from Natural Earth (public domain).

approximately 73 km from the final FJ-4 position. The final source audit then located an explicit published coordinate for FJ-4 in Table 4 of de Lumley et al. [3]: 4°37.6′ N, 36°27.5′ E. Conversion to decimal degrees produced 36.458333° E, 4.626667° N, which was entered in the released dataset with the precision category "high". This sequence demonstrates why apparently authoritative database coordinates still required checking against the specialist literature.

3. A principal record divided into constituent localities: Middle Awash Ardipithecus kadabba
   The Middle Awash *Ardipithecus kadabba* record comprised five constituent localities: Alayla, Digiba Dora, Saitune Dora, Amba East and Asa Koma Locality 3. Of these, only Alayla had a published numerical coordinate, reported by Haile-Selassie [5]. The positions of the other four localities were derived by georeferencing maps in Haile-Selassie and WoldeGabriel [8]. Five lettered subrecords were therefore created to preserve the separate spatial positions.

Chronology required separate treatment. Haile-Selassie [5] reported an overall age range of 5.8–5.2 Ma for the grouped Middle Awash hominin material. The more detailed locality-specific ages were established principally from Haile-Selassie and WoldeGabriel [8]: approximately 5.7 Ma for Alayla and Digiba Dora, approximately 5.6 Ma for Saitune Dora, 5.63–5.57 Ma for Asa Koma Locality 3, and approximately 5.2 Ma for Amba East. The latter age is also reported by Haile-Selassie [5]. Recording the localities separately therefore preserved chronological differences that would have been obscured by the grouped 5.8–5.2 Ma range.

Taxonomic treatment also changed between publications. Haile-Selassie [5] grouped the 11 specimens from all five localities as a single taxon, then named *Ardipithecus ramidus kadabba*. In the 2009 monograph, Haile-Selassie, Suwa and White [7] identified the material from Alayla, Digiba Dora, Saitune Dora and Asa Koma Locality 3 as *Ardipithecus kadabba*, but listed the Amba East specimen separately as *Ardipithecus* cf. *kadabba*. The dataset therefore records four firm locality assignments and one tentative assignment, with the parent record summarised as mixed, mostly firm. Recording the localities separately preserves a taxonomic distinction that would have been obscured if all five had been represented under a single unqualified species label.

Specimen allocation required a further reconstruction. Because the 2009 monograph grouped much of the material by geological member and taxon rather than by locality, the catalogue numbers had to be traced back to their individual localities before cautious specimen counts could be established for each site. The released dataset records four specimens at Alayla, one at Digiba Dora, four at Saitune Dora, one at Amba East and eight at Asa Koma Locality 3. Recording the localities separately therefore allowed the published fossil material to be associated with the correct site rather than remaining within one undifferentiated Middle Awash record.

QUALITY CONTROL

Each record was checked against its cited sources for locality identity, taxonomy, chronology, specimen information and coordinate provenance. All coordinates were plotted in QGIS to check coordinate conversion and consistency with published mapping and known geography. The final workbook and CSV exports were compared to confirm that the released files preserved the dataset and Supporting Notes.

CONSTRAINTS

Coordinate precision varies among records. Published coordinates, positions derived from georeferenced maps, cluster midpoints, proxy points and open-dataset coordinates may differ in geographical meaning despite similar decimal precision. The three precision categories are source- and scale-based, not statistical error estimates; high-resolution analyses or modelling should therefore use appropriate buffers, exclusions or sensitivity tests.

Numbered parent records and lettered subrecords represent different levels of the same hierarchy and should not be counted together as independent occurrences. Blank coordinate and specimen fields are intentional where the published evidence does not justify a value.

Taxonomic assignments, ages, fossil associations, locality definitions and palaeoenvironmental interpretations may change. Where the published evidence is uncertain or disputed, this is retained in the relevant dataset field and, where necessary, explained in the Supporting Notes. Landscape descriptions are broad source syntheses rather than uniform specimen-level habitats, and the cited primary sources should be consulted.

The dataset represents the published fossil-locality record, not the complete former distribution of these taxa. That record is shaped by preservation, exposure, accessibility, collection history and research intensity. Version 1.0 is therefore an updateable compilation that can incorporate newly published localities, revisions to existing evidence, and more precise coordinates where these become available from published sources or site researchers and can be released openly.

USE OF AI-ASSISTED TECHNOLOGIES

During preparation of the dataset documentation and manuscript, Ian G. Smith used ChatGPT (OpenAI) to assist with extracting, comparing and summarising

information from source PDFs supplied by the author, organising material, and refining the text. The authors retained responsibility for interpreting the sources and for all final dataset and manuscript decisions.

(3) DATASET DESCRIPTION

OBJECT NAME

- Early_African_Hominin_Localities_v1.0.xlsx: authoritative workbook containing the Hominin Localities and Supporting Note worksheets.
- Early_African_Hominin_Localities_v1.0.csv: CSV export of the Hominin Localities worksheet.
- Early_African_Hominin_Localities_Supporting_Notes_v1.0.csv: CSV export of the Supporting Note worksheet.
- README.txt: guidance on dataset scope, structure, limitations, citation and versioning.

DATA TYPE

Secondary data compiled from published sources, with processed spatial coordinates and documentation.

FORMAT NAMES AND VERSIONS

Microsoft Excel (.xlsx), CSV (.csv) and plain text (.txt); Version 1.0.

CREATION DATES

Compiled and source-audited through July 2026; published 3 August 2026.

DATASET CREATORS

Ian G. Smith, Independent Researcher at Bournemouth University, ORCID 0009-0003-6687-0818.

LANGUAGE

English.

LICENSE

Creative Commons Attribution 4.0 International (CC BY 4.0).

REPOSITORY LOCATION

Zenodo. Version 1.0 is openly available under a Creative Commons Attribution 4.0 International licence at https://doi.org/10.5281/zenodo.21773300. The Concept DOI, https://doi.org/10.5281/zenodo.21773299, identifies the continuing dataset series.

PUBLICATION DATE

3 August 2026.

The 14 fields in the Hominin Localities worksheet are described in Table 2.

The Supporting Note worksheet contains three fields: the lettered note identifier, the public explanation of the relevant source or dataset decision, and any additional bibliographic reference required.

(4) REUSE POTENTIAL

The 38 coordinate-bearing records can be imported directly into QGIS and other GIS software from the released CSV file. The five parent rows without coordinates are hierarchical summary records and do not generate point geometry when the CSV is imported using the longitude and latitude fields. For counting or summarising localities, numbered parent records and lettered subrecords should be treated as different levels of the same hierarchy and should not be counted together as independent occurrences. Where the main table contains a Note identifier, users should consult the corresponding entry in the separate Supporting Notes CSV, which records source decisions, uncertainties and qualifications relevant to interpretation.

A principal reuse is as a ready-to-use locality layer that can be overlaid with other spatial evidence in GIS. In the research underlying Out of Afar [14], the locality data were compared with elevation, geological formations, tectonic structures, drainage and reconstructed palaeolakes. Published palaeogeographic and palaeoenvironmental maps from different sources can also be georeferenced and layered with the locality data. When these spatial sources are viewed together, relationships not apparent in the original publications can become visible, helping researchers identify patterns and develop or test hypotheses about landscape change, environmental context and hominin locality distribution.

The same framework can be extended to later *Australopithecus*, early *Homo* and other hominin taxa, allowing comparisons across species and time periods. Researchers can also compare the geological, topographic and hydrological settings surrounding known localities with comparable landscape conditions elsewhere.

In university teaching, the dataset can add a geographical dimension to discussions of early hominin species, their chronological relationships and possible evolutionary connections. Localities can be mapped by taxon and age and displayed over a digital elevation model, allowing students to see where each species is known to have occurred and to appreciate the surrounding landscape. Maps and terrain visualisations can therefore complement anatomical and chronological evidence and make relationships among species, localities and environments easier to understand.

For field-survey planning, the locality layer can be displayed over a digital elevation model and combined with evidence for fossil-bearing geology, exposure, preservation and accessibility. This allows researchers

FIELD	DESCRIPTION
Record ID	Whole numbers identify principal records; lettered identifiers identify constituent locality subrecords. Parent and subrecords are not independent occurrences.
Specimens	Published fossil material or conservative specimen-count wording. Entries may be descriptive, minimum catalogue counts or intentionally blank where material cannot be allocated to sublocalities.
Age range, Ma	Published age or range in millions of years, retaining qualifiers such as "ca." where required.
Longitude	East longitude in WGS84 decimal degrees; blank for parent records represented by mapped subrecords.
Latitude	North latitude in WGS84 decimal degrees; blank for parent records represented by mapped subrecords.
Coordinate precision	High, within 200 m or within 2.5 km, describing source basis and geographical meaning rather than displayed decimal places.
Country	Modern country containing the locality: Chad, Ethiopia or Kenya.
Site name	Published locality, site, formation or grouped research-area name, including useful alternative names or codes.
Hominin species	Taxonomic label used for the record, retaining cautious or indeterminate wording where necessary.
Taxonomic confidence	Published assignment recorded as firm, tentative, probable, disputed, mixed or probable / contested; independent of coordinate precision.
Note	Pointer to Supporting Notes A to I where further explanation is required.
Landscape and rift context	Brief source-based summary of geological, depositional and wider landscape setting, not a specimen-level habitat assignment.
Original or most relevant reference	Principal publication or publications supporting the locality record; subrecords may inherit the parent reference.
Coordinate source / derivation	Source from which the coordinate was obtained or derived where this differs from the principal locality reference, including published coordinates, georeferenced published maps, proxy points or public databases.

Table 2 Fields in the Version 1.0 Hominin Localities worksheet.

to examine the landscape surrounding known localities remotely, compare established sites with similar landscape settings elsewhere, and prioritise areas for investigation before committing resources to field reconnaissance.

DATA ACCESSIBILITY STATEMENT

The dataset described in this article is openly available on Zenodo as Early African Hominin Localities from Sahelanthropus to Australopithecus anamensis: A Source-Audited, GIS-Ready Dataset, Version 1.0 [13], at https://doi.org/10.5281/zenodo.21773300. The Concept DOI for the continuing dataset series is https://doi.org/10.5281/zenodo.21773299.

ETHICS AND CONSENT

The dataset contains only previously published scientific information and no personal data or privately supplied coordinates. Where coordinates were derived from published maps, only openly published scientific sources were used; no private, restricted or confidential locality information was incorporated. Ethics approval and participant consent were not required.

FUNDING INFORMATION

This research received no specific funding.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

Conceptualization: Ian G. Smith; Methodology: Ian G. Smith; Validation: Ian G. Smith, Sally C. Reynolds; Investigation: Ian G. Smith; Resources: Ian G. Smith; Data curation: Ian G. Smith; Writing – original draft: Ian G. Smith; Writing – review & editing: Ian G. Smith, Sally C. Reynolds; Project administration: Ian G. Smith.

AUTHOR AFFILIATIONS

Ian G. Smith orcid.org/0009-0003-6687-0818
Independent Researcher at Bournemouth University, UK; Departments Health, Environmental and Medical Sciences (HEMS), Bournemouth University, Fern Barrow, Poole, Dorset BH12 5BB, UK
Sally C. Reynolds orcid.org/0000-0001-6313-6405
Departments Health, Environmental and Medical Sciences (HEMS), Bournemouth University, Fern Barrow, Poole, Dorset BH12 5BB, UK

REFERENCES

1. **Cerling TE, Levin NE, Quade J, Wynn JG, Fox DL, Kingston JD, Klein RG, Brown FH.** Comment on the paleoenvironment of Ardipithecus ramidus. *Science*. 2010;328(5982):1105–1105. DOI: https://doi.org/10.1126/science.1185274
2. **Dawson N, Fischer J, Kuhn M, Pasotti A, Rouzaud D, Bruy A, Sutton T, Dobias M, Pellerin M, Rouault E.** qgis/QGIS: 3.44.5. *Zenodo*; 2025.
3. **de Lumley H, Bahain J-J, Beyene Y, Marchal F, Pleurdeau D.** Les sites paléontologiques et préhistoriques de la région de Fejej. In: de Lumley H, Beyene Y, editors. *Les sites préhistoriques de la région de Fejej, Sud-Omo, Éthiopie, dans leur contexte stratigraphique et paléontologique*. Paris: ADPF, Éditions Recherche sur les Civilisations. 2004;109–128.
4. **Gilbert H.** n.d. *Human Fossil Record* [online]. Fossilized.org. Available from: https://www.fossilized.org/Human_paleontology/ [Accessed August 2026].
5. **Haile-Selassie Y.** Late Miocene hominids from the Middle Awash, Ethiopia. *Nature*. 2001;412(6843):178–181. DOI: https://doi.org/10.1038/35084063
6. **Haile-Selassie Y, Melillo SM, Vazzana A, Benazzi S, Ryan TM.** A 3.8-million-year-old hominin cranium from Woranso-Mille, Ethiopia. *Nature*. 2019;573(7773):214–219. DOI: https://doi.org/10.1038/s41586-019-1513-8
7. **Haile-Selassie Y, Suwa G, White T.** 7. Hominidae. In: Haile-Selassie Y, WoldeGabriel G, editors. *Ardipithecus kadabba: Late Miocene Evidence from the Middle Awash, Ethiopia*. Berkeley: University of California Press. 2009;159–236. DOI: https://doi.org/10.1525/9780520942509-011
8. **Haile-Selassie Y, WoldeGabriel G.** *Ardipithecus kadabba: Late Miocene Evidence from the Middle Awash, Ethiopia*. Berkeley: University of California Press; 2009. DOI: https://doi.org/10.1525/9780520942509
9. **Paleobiology Database.** *Fejej – FJ4, collections 22563 and 22567* [online]; 2026. Available from: https://paleobiodb.org/classic/basicCollectionSearch?collection_no=22563 [Accessed 6 August 2026].
10. **Semaw S, Simpson SW, Quade J, Renne PR, Butler RF, McIntosh WC, Levin N, Dominguez-Rodrigo M, Rogers MJ.** Early Pliocene hominids from Gona, Ethiopia. *Nature*. 2005;433(7023):301–305. DOI: https://doi.org/10.1038/nature03177
11. **Simpson SW, Kleinsasser L, Quade J, Levin NE, McIntosh WC, Dunbar N, Semaw S, Rogers MJ.** Late Miocene hominin teeth from the Gona Paleoanthropological Research Project area, Afar, Ethiopia. *Journal of Human Evolution*. 2015;81:68–82. DOI: https://doi.org/10.1016/j.jhevol.2014.07.004
12. **Simpson SW, Levin NE, Quade J, Rogers MJ, Semaw S.** Ardipithecus ramidus postcrania from the Gona Project area, Afar Regional State, Ethiopia. *Journal of Human Evolution*. 2019;129:1–45. DOI: https://doi.org/10.1016/j.jhevol.2018.12.005
13. **Smith IG.** *Early African Hominin Localities from Sahelanthropus to Australopithecus anamensis: A Source-Audited, GIS-Ready Dataset* (version 1.0) [Data set]. Zenodo; 2026. DOI: https://doi.org/10.5281/zenodo.21773300
14. **Smith IG, Reynolds SC.** *Out of Afar: the first hominin migration? Long-term landscape changes in the Afar region and implications for hominin bipedalism [preprint]* [online]. [Accessed August 2026]; 2025. DOI: https://doi.org/10.31223/X5TT85

TO CITE THIS ARTICLE:
Smith IG, Reynolds SC 2026 Early African Hominin Localities from *Sahelanthropus to Australopithecus anamensis*: A Source-Audited, GIS-Ready Dataset. *Journal of Open Archaeology Data,* 14: 27, pp. 1–8. DOI: https://doi.org/10.5334/joad.239

Submitted: 22 September 2026 **Accepted:** 05 November 2026 **Published:** 25 November 2026

COPYRIGHT:
© 2026 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See https://creativecommons.org/licenses/by/4.0/.

Journal of Open Archaeology Data is a peer-reviewed open access journal published by Ubiquity Press.